PROTO MALAYIC:

THE RECONSTRUCTION OF ITS PHONOLOGY AND PARTS OF ITS LEXICON AND MORPHOLOGY

K. Alexander Adelaar

TABLE OF CONTENTS

PREFACE

This publication is a revised version of my PhD thesis of the same title which I defended in October 1985 at Leiden State University. Most of the revisions are editorial, and I made only minor corrections and additions regarding the subject matter of the original thesis.

A number of publications relevant to the history of the Malayic isolects have appeared since 1985. The more important among these are Asmah Haji Omar (1981, 1985), Collins (1985, 1986a, 1986b, 1987), Hogan (1988) and a bundle of contributions to the International Workshop on the History of Malay held in Kuala Lumpur in 1986 (Mohd. Thani Ahmad and Zaini Mohamed Zain eds 1988). Nothofer's contribution to this workshop is of particular interest, as it is a critical evaluation of the definitions and subclassification of Proto Malayic that have been made by me and others (Nothofer 1988). Elsewhere I will treat his evaluation at greater length (Adelaar in press a and forthcoming). Here I limit myself to reacting to three of his most important criticisms, which are as follows.

Firstly, Nothofer observes that my definition of 'Malayic' is different from that of Blust (1981), a difference which I do not discuss, although I do refer to Blust (1981) in another context. This is correct: in the introduction to the thesis (and to this publication) I quote Hudson (1970, 1978), who coined the term, and to whose definition I clearly adhere.

Secondly, my thesis does not provide any quantitative or qualitative evidence supporting the existence of a Malayic subgroup. It is true that I did not give explicit evidence for this subgroup. In Chapter 7, however, I listed the phonological and lexical changes that have taken place between PAN and PM, and these changes as a whole are critical for the definition of the Malayic subgroup. But I readily admit that my thesis would have been the better for it if it had contained an explicit diagnostic device for the definition of Malayic and for the identification of Malayic isolects. As such a device, I propose a set of phonological criteria which I have added to Chapter 1 of this publication. In reference to Nothofer (1988), this phonological criterion excludes Embaloh (or Maloh, spoken in West Kalimantan, Indonesian Borneo) and Rejang (spoken in South Sumatra) from the Malayic subgroup (see Adelaar in press a and Blust 1984 for further details).

Finally, Nothofer criticises the fact that I treat the six isolects forming the basis of my thesis as if they all continued directly from Proto Malayic, and that only in the last chapter of my thesis do I conclude that "...it seems opportune to make a provisory bipartite division of the Malayic isolects, with IBN [Iban] in one branch, and all five other isolects in the other". Nothofer is not so much concerned about this conclusion, but rather about the way in which it was reached. He himself gives supportive lexical evidence for a separate Iban branch in the Malayic subgroup. I agree with him that my statement was rather offhand. In this publication I make no attempt at any internal classification with regard to the six isolects forming the base of my reconstruction. Historically Iban and other 'Malayic Dayak' isolects underwent a separate development from other Malayic isolects, but that does not necessarily

imply that they form a separate branch. There are, at any rate, no sufficient linguistic criteria to define such a putative branch. Nor does the fact that Iban is considered a language in its own right, rather than a Malay dialect (Hudson 1970), imply that it is genetically more remote from Standard Malay than other Malayic isolects: many of the shared similarities between other Malayic isolects seem to be the result of language convergence. The question of the internal classification of the Malayic subgroup, and hence the question of the difference between 'Malay' and 'Malayic' as discussed by Blust (1988) and Nothofer (1988), remains unanswered, and it is doubtful whether sound solutions will be obtained from the comparative method alone.

ACKNOWLEDGEMENTS

LIST OF ABBREVIATIONS

(adv)	adverb	pers.	person
a.i.	all isolects (SM, MIN, BH, SWY, IBN, and JKT)	pl.	plural
		S	subject
excl.	exclusive	sg.	singular
IC	irregular correspondence(s)	s.o.	someone
incl.	inclusive	s.th.	something
k.o.	kind of	UIC	unexplained irregular correspondence(s)
(n)	noun		
N	nasal	(v)	verb
N-	nasalisation, active verbal prefix	VDI	dynamic intransitive verb
n.c.	no cognate available	Vl	intransitive verb
O	object	VSI	stative intransitive verb
o.i.	other isolects (forming basis of this study)	VTR	transitive verb

SYMBOLS USED

*....	form belongs to one of the reconstructed proto-languages
+....	form occurred in a stage between Proto Malayic and the present
[...]	indicates the phonetic realisation of a sound or a word
'...'	1. indicates the orthography of a sound or word
	2. indicates the meaning of a label
'(..)'	indicates the function of a label, or the meaning of a precategorial
<	'originated from'
>	'became'
~	'is a variant of, and similar in meaning to'
≠	'is different from'
/	(within a word) indicates a historical morpheme boundary between a fossilised affix and a lexeme, or between two former lexemes
#	indicates a word boundary
-	indicates a morpheme boundary, or, with phonemes, indicates whether this phoneme occurs initially (...-), medially (-...-), or finally (-...)
l...._....	'with a position (_) in the environment (....)'

Languages and dialects most frequently referred to in this study, with abbreviations and primary lexical sources are as follows. When other sources are used, this will be indicated.

Achehnese	(Hoesein Djajadiningrat 1934)
AN	Austronesian
Angkola-Mandailing	Angkola- and Mandailing-Batak (Eggink 1936)
AR	Arabic (Wehr 1976)
BAC	Bacan Malay (Collins 1986a)
Balinese	(Kersten 1978)
BH	Banjarese Malay, Hulu isolect (Abdul Jebar 1977)
BK	Banjarese Malay, Kuala isolect (Abdul Jebar 1977)
BRU	Brunei Malay (Wilkinson 1959; Prentice pers.comm.)
BSM	Middle Malay, Besemah isolect (Helfrich 1904)
Cham	(Moussay 1972; Lee 1966)
CHI	Chinese (Wilkinson 1959, Klinkert 1916)
Dairi	Dairi Batak (Tindi Radja Manik 1977)
DU	Dutch
ENG	English
IBN	Iban (Scott 1956)
Jarai	(Lafont 1968)
JKT	Jakartanese Malay (Abdul Chaer 1976)
JV	Javanese (Jansz 1913; Pigeaud 1938)
Karo	Karo-Batak (Neumann 1951)
KCI	Kerinci (Hakim Usman 1985)
KD	Kendayan(-Dayak) (Dunselman 1949-1950)
Madurese	(Penninga & Hendriks 1937)
Manobo	(Elkins 1968)
Mansaka	(Svelmoe & Abrams 1955)
MIN	Minangkabau (Thaib 1935)
Malagasy	(Dempwolff 1938)
Mualang	Mualang-Dayak (Dunselman 1955)
Ngaju	Ngaju-Dayak (Hardeland 1859)
Old Javanese	(Zoetmulder 1982)
OM	Old Malay (Çoedès 1930, De Casparis 1956)
Paiwan	(Ferrell 1982)
PAN	Proto Austronesian
Proto Batak	(Adelaar 1981)
Proto Chamic	(Lee 1966)
Persian	(Steingass 1930)
PHF	Proto Hesperonesian-Formosan (Zorc 1982)
PHN	Proto Hesperonesian (Zorc 1982)
PM	Proto Malayic
PMJ	Proto Malayo-Javanic (Nothofer 1975)
PMP	Proto Malayo-Polynesian (Dempwolff 1938)
POR	Portuguese (Wilkinson 1959; Klinkert 1916)
PWMP	Proto Western-Malayo-Polynesian (Blust 1980a, 1984a, 1986)

Rhade	(Lee 1966)
Roglai	(Lee 1966)
SAR	Sarawak Malay (Collins 1987)
Sasak	(Goris 1938)
SD	Selako(-Dayak) (Ina Anak Kalom & Hudson 1970)
Simalungun	Simalungun-Batak (Saragih 1989)
SKT	Sanskrit (Gonda 1973)
SM	Standard Malay (Wilkinson 1959)
SUN	Sundanese (Coolsma 1913)
SWY	Middle Malay, Seraway isolect (Helfrich 1904)
Tagalog	(Panganiban 1966)
TAM	Tamil (Wilkinson 1959; Klinkert 1916; Van Ronkel 1902)
Toba	Toba-Batak (Warneck 1977)
Timugon Murut	(Prentice forthcoming)
Tioman-Malay	(Collins 1985)
Urak Lawoi'	(Hogan 1988)
WMP	Western-Malayo-Polynesian

MAP OF AREAS WHERE MALAYIC ISOLECTS ARE SPOKEN
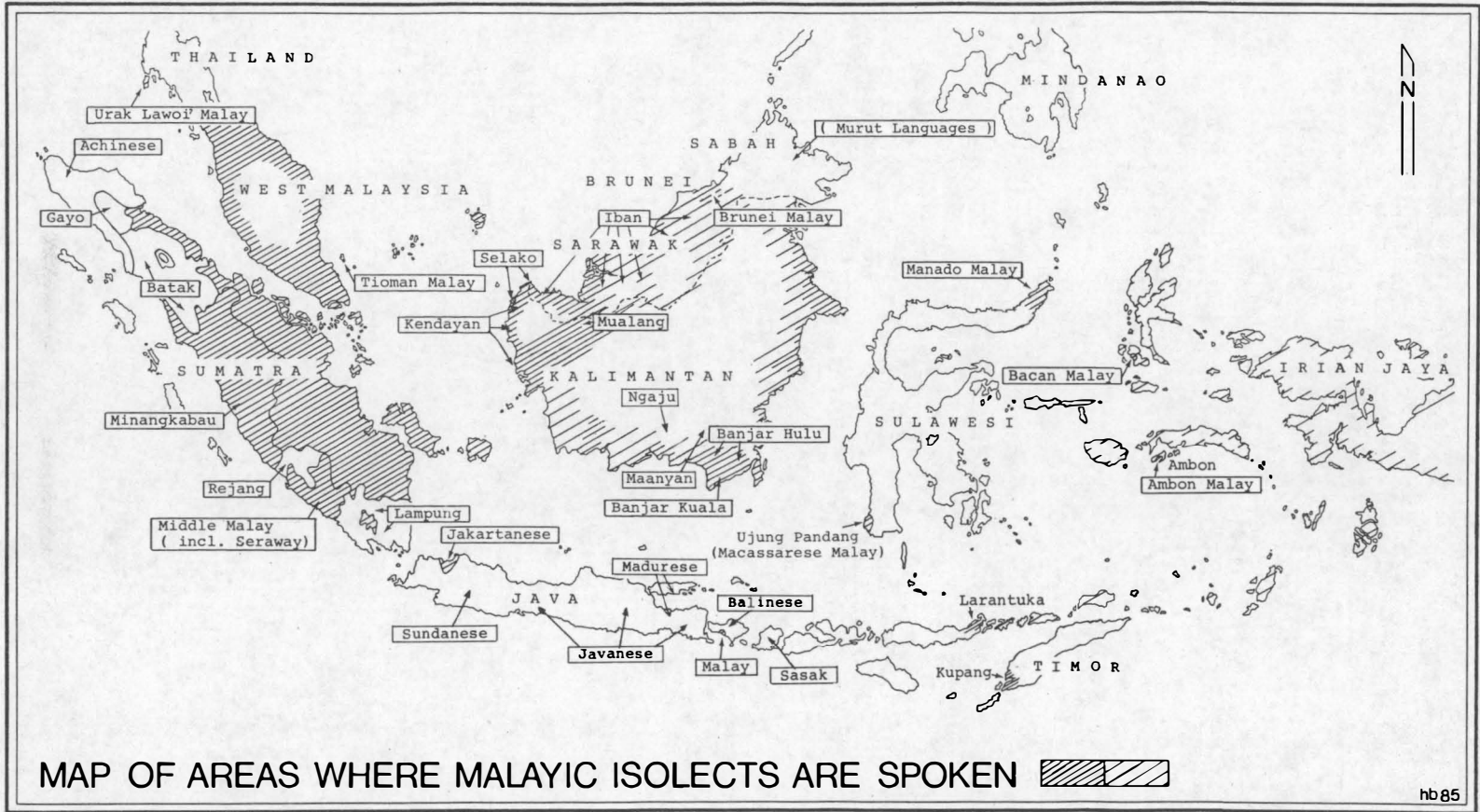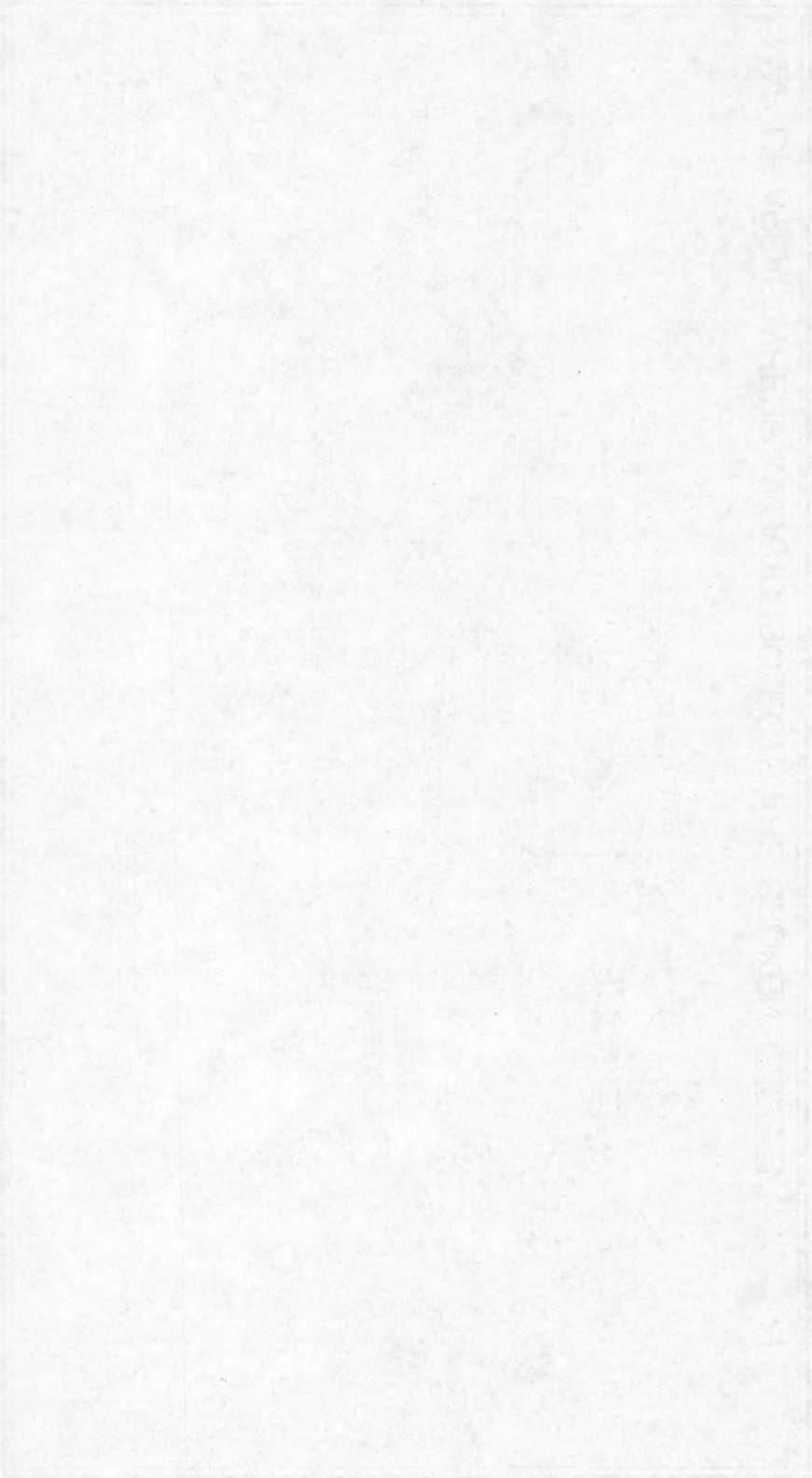
CHAPTER 1

INTRODUCTION

Malay is spoken in a number of dialects in coastal areas of the Malay Peninsula, in Borneo,[1] in South and South-east Sumatra, and in nearly all major trade centres of the Indonesian archipelago. The Malay language belongs to the group of Malayic languages, which also includes Minangkabau and Kerinci[2] in Sumatra, and various languages and dialects of inland western Borneo, of which the most important is Iban.[3] In what follows I will use the term 'isolect' to refer to a speech form without respect to its status as a language or a dialect.[4] Hudson (1970:302-303) applies the name 'Malayic' to all isolects which seem, with Malay, to be "immediately related through descent from a common ancestor".

In this study I compare the Malayic isolects, and attempt to reconstruct their common ancestor, Proto Malayic. There are several motives for this study. First, little is known about the linguistic evolution of the Malayic isolects, or the nature of Proto Malayic (PM). The reconstruction of PM may also provide more insight into the relative position of the Malayic isolects within the Austronesian (AN) language family, and more particularly, within the loosely defined branch of Western-Malayo-Polynesian (WMP) languages belonging to this family.[5] An internal classification of Malayic isolects may give some indication as to the homeland of the PM speakers. The PM lexicon will ultimately yield more insight into PM culture and social organisation. Knowledge about the relative position of PM within the

[1]For the sake of convenience and in line with international usage, I use the term 'Borneo' to refer to the whole island which is nowadays divided into four administrative parts: Sabah and Sarawak (which are part of Malaysia), Kalimantan (which is part of Indonesia) and Brunei.

[2]Van der Toorn (1891: VI- VII) calls Minangkabau a Malay dialect, but according to Dyen (1965a:18,26), its shared lexicostatistical percentage with Standard Malay (68.8) falls below the language limit (set by Dyen at 70.0). According to Dyen (1965a:26) Kerinci has a standard cognate percentage of 66.0 with Standard Malay, but Prentice and Hakim Usman counted a maximum of 86.2, and a minimum of 79.8; they argue that from a lexical point of view one should call Kerinci a Malay dialect, but that "the phonological changes...are of such a striking nature, and, in combination with other changes such as the loss of all suffixes, have had such far-reaching effects on the Kerinci sound system and morphology...that there appears ample justification for regarding Kerinci as a separate language" (Prentice & Hakim Usman 1978:123. See also Steinhauer & Hakim Usman for the morphology of Kerinci).

[3]Other Malayic languages/dialects in this area reported by Hudson (1970, 1978) are: Selako, Banana', Kayung, Semitau, Ambawang, Kendayan, Suhait, Keninjal, Delang, and the Ibanic group, which includes apart from Iban: Sebuyau, Mualang, Kantu', and Air Tabun. Hudson opposes the statement of several scholars reiterated by Cense and Uhlenbeck (1958) that IBN is a Malay dialect. He considers it a language in its own right, a "close relative of Malay, one more like a first cousin than a delinquent child" (Hudson 1970:302-303).

[4]The term isolect denotes "any language unit that is accorded a separate name by its speakers, regardless of whether it is, technically, a dialect or a language"; its use is "connotationally neutral in r gard to language-dialect identification" (Hudson 1967:12).

[5]Blust (1980a:11-12) makes the following classification of Austronesian lan uages: there is a primary split into four groups, three of which are exclusively Formosan (viz. Atayalic, Tsouic and Paiwanic), while the fourth, Malayo-Polynesian, includes all the other Austronesian languages. Within Malayo-Polynesian Blust distinguishes a Central-Eastern-Malayo- Polynesian subgroup, and labels the remaining languages as Western-Malayo-Polynesian. He divides the Central-Eastern-Malayo-Polynesian languages further into Central-Malayo-Polynesian (including languages of the Lesser Sunda Islands and languages of the southern and central Moluccas (including the Aru islands and the Sula archipelago)) and Eastern-Malayo-Polynesian (including languages of South Halmahera and West New Guinea, and the Oceanic languages). He does not define the Western-Malayo-Polynesian group as a subgroup of languages sharing specific innovations: it is a residual group which did not undergo changes characteristic to the languages of the Central- and Eastern-Malayo-Polynesian subgroup.

1

2

WMP language family and knowledge of its lexicon are relevant to the study of related fields such as anthropology, archaeology and prehistory. Finally, the reconstruction of PM will, it is hoped, contribute to the improvement of higher order reconstructions viz. Proto Austronesian (PAN), Proto Malayo-Polynesian (PMP) and Proto Malayo-Javanic (PMJ).

The Malayic isolects are characterised by a set of phonological developments which took place in their history. Most of these developments are not decisive in themselves (as, for instance, Chamic languages, Achehnese and Balinese each turn out to share more than half of them). It is the co-occurrence of the following developments which defines the members of the Malayic subgroup as such:

(1) devoicing of final stops;
(2) PAN *j > PM *d, *-t;
(3) *Z (and *z) > PM *j;
(4) *R (and *r) > r;
(5) reduction of consonant clusters to their last component;
(6) heterorganic nasal + stop clusters became homorganic nasal + stop clusters;
(7) *w- > ø;
(8) *i, *-ey, *-uy, *-iw > PM *i;
(9) *u, *-ew > PM *u;
(10) *q > PM *h;
(11) *h, *ʔ > PM *ʔ or ø.

PM morphology and lexicon do not provide strong criteria for a Malayic subgroup. Malayic morphology is admittedly quite different from what has been reconstructed as PAN morphology, but the differences in question seem to be shared with other western Indonesian languages.

There are many vocabulary items which are well attested and inherited in the Malayic subgroup as a whole, but there always seems to be a non-Malayic language which has a corresponding form. It is therefore all but impossible to collect a body of vocabulary items which are critical for a subgrouping argument. This fact reflects an important sociolinguistic reality: several forms of lingua franca Malay have had a tremendous influence on many insular Southeast Asian languages and on Malagasy, to the extent that non-Malayic languages have borrowed too many vocabulary items from all domains of the Malayic lexicon.

SM is one of the isolects Dempwolff used for the reconstruction of PMP (as explained below), and this is a good reason for comparing SM with the isolects with which it forms an exclusive subgroup. In a language family, lower order reconstructions are indispensable for a better understanding of higher order proto-languages, including the language ancestral to all members belonging to the family. This is not to say that an interim reconstruction of the highest order proto-language is not warranted: rather, lower order and higher order reconstructions should constantly be tested against one another. This may sound like a commonplace statement to anyone familiar with the comparative method, but the fact is that in the context of AN comparative linguistics lower order reconstructions have until recently remained rather neglected. It is only since the 70s that more than a tiny number of them have been made, and there is still much that remains to be done.[6]

---

[6]The only pre-1970 subgrouping attempts based on qualitative evidence and lower order reconstructions known to me are: Niemann (1891, on the relation of Cham to Achehnese); Stresemann (1927, on Ambonese languages); Dahl (1938, on Proto Malagasy; 1951, a comparison of Malagasy and Maanyan Dayak); Grace (1959, on Proto ceanic); Lee (19 , on Proto Chamic); and Hudson (1967, on Proto Ba ito). Other attempts

I make this reconstruction on the basis of SM, and on the basis of five other Malayic isolects which show important phonological retentions from PAN/PMP which are not found in SM, and for which there is a sufficient lexical and grammatical corpus available. Isolects such as Kerinci, Manado Malay, Moluccan Malay, and the Malayic isolects spoken by the Orang Laut, Orang Darat, Orang Akit and Orang Utan, are not systematically incorporated in the comparison because, although there is sufficient material available for them, they show no structural retentions not also present in SM, and thus will not alter the overall picture of the reconstruction of PM.[7] It also follows that some isolects with considerable retentions from an older stage (e.g. isolects from the Malay Peninsula, Bacan, Old Malay) will only occasionally be drawn into the comparison, because the lack of material makes a systematic use of them impossible.

The isolects fulfilling the above requirements are Minangkabau, Banjarese, Middle Malay, Iban and Jakartanese: together with SM they will form the basis of this study.

SM is the isolect on which Bahasa Indonesia and Bahasa Malaysia are based, and by which is meant "...the literary Malay which represents the direct descendant of the language used in the court of the Malacca sultanate...and which continued to be used in the court of the Sultans of Riau and Johore" (Prentice 1978:23). SM itself is based on Classical Malay which is the Malay of literary works from the sixteenth till the nineteenth century.

Minangkabau (MIN) is very close to SM in its structure and vocabulary. It is spoken in the province of West Sumatra, in some adjacent regencies in Jambi (along the Batang Hari river), and in the regency of Kampar in the province of Riau. It is also spoken by immigrant groups in Jamèë (West Acheh), and in Negeri Sembilan (Malaysia). The isolect of Koto Gadang (near Bukittinggi, in the Agam regency) is the base of the dictionary and grammar which are used for MIN in this study. It has a threefold vowel distinction *(a, i, u)* in the first syllable of trisyllabic lexemes, whereas the other isolects in this study (except Banjarese) only permit a schwa in this position. In a few cases it also retains heterorganic consonant clusters, which are reduced to a single medial consonant in other isolects, for example, *saŋla, sala* 'broiling; cooking at an open fire' (Wilkinson 1959), from PAN *saŋəlaR, cf. SM *səlar* 'branding', IBN *sala*ʔ 'smell of roasting flesh'.

Banjarese is spoken on the south coast of Borneo, and on the east coast up to Kutai. Its centre is the city of Banjarmasin (Abdul Jebar 1977:1). Banjarese has two main variants: Bahasa Banjar Hulu (BH), and Bahasa Banjar Kuala (BK).[8] BH is the isolect that I use in this comparison. Like MIN (and also like BK), it has a threefold vowel distinction *(a, i, u)* in antepenultimate syllables. It also consistently reflects PAN/PMP *q as *h*, whereas in initial position and between unlike vowels this proto-phoneme is largely lost in other isolects.

---

were made from 1974 onwards: Blust (1974b, on Proto North Sarawak); Mills (1975, on Proto South Sulawesi); Nothofer (1975, on Proto Malayo-Javanic); Tsuchida (1976, on Proto Tsouic); Dahl (1977, on Proto Barito); Sneddon (1978, on Proto Minahasan); and various reconstructions on the basis of Philippine and Oceanic subgroups.

[7] The phonological and morphological differences between Kerinci and the other Malayic isolects are innovative (Prentice & Hakim Usman 1978:passim); Kerinci has two reflexes *(h* and *r)* for SM *r*, but Kerinci *r* tends to be restricted to loans. Manado Malay and Moluccan Malay do not show any phonological retentions that are not also present in SM, and have a very reduced affix system. The phonological differences between SM and the Malayic isolects spoken by the Orang Laut, Orang Darat, Orang Akit and Orang Utan (cf. Kähler 1966), are due to innovations in the latter.

[8] Bahasa Banjar Kuala is spoken in the neighbourhood of Banjarmasin, Martapura and Palaihari; Bahasa Banjar Hulu is spoken in the Hulu Sungai area, and more specifically, in the regencies Tapin, Hulu Sungai Selatan, Hulu Sungai Tengah, Hulu Sungai Utara and Tabalong. (Abdul Jebar 1977:1).

Middle Malay is the name given to a number of closely related Malay isolects spoken in Bangka Hulu (Bencoolen) and in the Palembang highlands (Voorhoeve 1955:18).[9] For two of these isolects, Besemah (BSM) and Seraway (SWY), Helfrich collected a wordlist and wrote a grammatical outline (Helfrich 1904; also Helfrich 1915, 1921, 1927, 1933). BSM and SWY both oppose a uvular (or velar)[10] x to a trilled r. Whether this x and r reflect PAN/PMP *R and *r respectively will be discussed in Chapter 3 (3.7). SWY is the Middle Malay isolect that is used as one of the six basic isolects in this study.

Iban, or Sea Dayak (henceforth IBN), is an isolect widely spoken by non-Moslem peoples in the western part of Borneo. IBN has b (from PAN/PMP *b) in the environment $a \_ a$, whereas the other isolects have w, for example, IBN *laban* 'against' from PMP *$laban$, cf. other isolects (o.i.) *lawan* 'against'. It also has a final glottal stop, which apparently reflects PAN *S, PAN *H (> PMP *h), and PAN *$?$ (PMP *$?$).

Jakartanese (JKT) is the isolect of Jakarta. It has many variants;[11] the variant of Mester is used in Abdul Chaer's dictionary (see below) and will also be the one used in this study. JKT a and schwa before final consonants reflect PAN/PMP *a and *e (schwa) respectively; in other isolects these proto-phonemes are merged in this position.

Table 1 summarises those phonological retentions in MIN, BH, SWY, IBN and JKT, that have been lost in SM.

My main sources for the six isolects are:

for SM: D. Gerth van Wijk: *Spraakleer der Maleische taal* (1889).[12]

              D.J. Prentice: Malay (Indonesian and Malaysian) (1987) and lecture-notes on Indonesian and Malaysian grammar (n.d.).

              R.J. Wilkinson: *A Malay - English dictionary* (1959).

for MIN: J.L. van der Toorn: *Minangkabausche spraakkunst* (1899).

              M. Thaib gl. St. Pamoentjak: *Kamoes bahasa Minangkabau - bahasa Melajoe-Riau* (1935).[13]

for BH: Abdul Jebar Hapip: *Kamus Banjar - Indonesia* (1977).

              M. Asfandi Adul: Morfologi Bahasa Banjar (1976).

for SWY: O.L. Helfrich: *Bijdragen tot de kennis van het Midden Maleisch (Besemahsch en Serawajsch dialect)* (1904, + supplements and additions).

              Z.N. Aliana et al.: *Bahasa Serawai* (1979).

for IBN: N.C. Scott: *A dictionary of Sea Dayak* (1956).[14]

              Asmah Haji Omar: The Iban language (1977).

---

[9]Voorhoeve (1955:18) credits J.L.A. Brandes with first suggesting the name (i.e. Dutch 'Midden Maleisch'). Its customary translation into English as 'Middle Malay' is actually ill chosen, as it suggests an earlier stage of Malay, and not a geographical variant. A more appropriate translation would have been 'Central Malay' (Prentice & Hakim Usman 1978:158 n.13). SM has a shared cognate percentage of 87.1 with Besemah, and 87.7 with Seraway; Besemah and Seraway have a shared cognate percentage of 99.3 (Dyen 1965a:28).

[10]The exact phonetic realisation is not clear from the sources.

[11]Van der Tuuk calls this isolect "basically Low Balinese" (cf. Teeuw 1961:45), but Kähler (1966:I) and Abdul Chaer (1976:XVII) treat it as a Malay dialect. Abdul Chaer (p.XVIII) further distinguishes four main variants: Karet, Tanah Abang, Mester and Kebayoran.

[12]I decided to use this work because it provides a large number of examples from Classical Malay.

[13]I prefer Thaib's dictionary to Van der Toorn's (1891) because, although largely based on the latter, it contains additional information.

[14]A new *Iban-English dictionary* by Richards appeared in 1981. Although it contains much more information than Scott (1956), it does not distinguish between long and short vowels. I will, however, frequently use it as a supplement to Scott.

for JKT: Abdul Chaer Mad'ie: *Kamus dialeck Melayu Jakarta - Bahasa Indonesia* (1976);[15]

Muhadjir: *Morphology of Jakarta dialect, affixation and reduplication* (1981).

	TABLE 1: PHONOLOGICAL RETENTIONS					
	SM	MIN	BH	SWY	IBN	JKT
1)	ə	a	a	ə	ə	ə		C_CVCVC
	ə	i	i	ə	ə	ə		
	ə	u	u	ə	ə	ə	
2)	C	CC	C	C	C	C		V_V (in a few cases)
3)	h, ø	ø	h	ø	ø	ø		
								V
4)	r	r	r	x	r	r	
	r	r	r	r	r	r	
5)	w	w	w	w	w	w		a_a
	w	w	w	w	b	w		
6)	ø	ø	ø	ø	ʔ	ø		_#
7)	a	a	a	a	a	a		_C#
	a	a	a	a	a	ə	

Isolects that are not systematically used but that are often drawn into the comparison are Kendayan Dayak (KD), Selako Dayak (SD), Kerinci (KCl) (see above), Bacan (BAC), and Old Malay (OM).

According to Hudson (1967), KD and SD belong to the Malayic Dayak subgroup of the Malayic isolects. BAC is a Malayic isolect spoken on the island of Bacan (near Buru, eastern Indonesia). It differs from other eastern Indonesian Malayic isolects in that it is more archaic. Apart from many innovations it also shows a number of lexical retentions that have been lost in SM (Collins 1986a).

OM is the language of the inscriptions of South Sumatra and Kedah. It belongs to the Malayic group. According to Teeuw (1959:140-146) most scholars have been using the term Old Malay without careful critical checking of the data against other forms of Malay. One exception should be made for Aichele (1942-43), who made a study of the differences between Old Malay and Classical Malay. He was of the opinion that the language of the old inscriptions was rightly termed Old Malay, and attributed a number of the differences to influence from (early) Batak and (to a lesser extent) from Old Javanese. Teeuw criticises some of the views and methodological weaknesses in Aichele's work. But summarising the

---

[15]For JKT there is also an older dictionary by Hans Kähler (1966) which, however, contains less information than that of Abdul Chaer.

phonological, morphological and lexical data, he nonetheless concludes that "All in all it must be admitted...that it is not related to any other present-day language so closely as to Malay" (Teeuw 1959:146). In addition to Teeuw's views it is worth mentioning that in this study a number of the peculiarities of OM vis-à-vis Classical Malay are shown to have correspondences in other Malayic isolects, so that they must be retentions from PM (cf. 5.3 on numerals, and Chapter 6 on the affixes *di-* and *-a*).

Although there are no systematic comparisons of Malayic isolects, SM has been used in several comparative and historical studies with wider aims.

Dempwolff's three-volume work, *Vergleichende Lautlehre des austronesischen Wortschatzes* , was the first systematic attempt to reconstruct the phonology of a hypothetical Malayo-Polynesian proto-language.[16] He based his reconstruction primarily on three non-Oceanic Malayo-Polynesian languages: Tagalog, Javanese and Toba-Batak, and tested it against three other non-Oceanic Malayo-Polynesian languages (SM, Ngaju-Dayak and Malagasy), two Melanesian languages (Fijian and Sa'a) and three Polynesian languages (Tongan, Futuna and Samoan). The third volume of his work contains a list of 2,215 proto-lexemes, which has served as a starting point for later scholars of AN comparative linguistics, who have improved and extended it.

In a footnote Dyen (1949) used evidence primarily from SM to split Dempwolff's PMP *-ay* into *-ay* (on the basis of Tagalog *-ay*, SM *-ay*, and Tongan *-e*) and *-ey* (on the basis of Tagalog *-ay*, SM *-i* and, by extension, Tongan *-e*).

In his *Proto-Malayo-Polynesian laryngeals* (1953) Dyen improved Dempwolff's reconstruction by the introduction of two 'laryngeals', namely *q* and *h.* His inferences rest on evidence from the so-called 'Tagalic' languages (Tagalog, Bisayan and Bikol), SM, Javanese and Tongan.

Blust (1970, 1980, 1984a, 1986) gives a large number of PAN (and lower order) proto-lexemes which are to a large extent based on IBN and SM.

In 1975 Nothofer published *The Reconstruction of Proto-Malayo-Javanic* . This is a phonological and lexical reconstruction based on Javanese, SM, Sundanese and Madurese, four languages which, according to Dyen (1965:26) belong with several others in a relatively close-knit subgroup (the Javo-Sumatran Hesion of the West Indonesian Cluster of the Hesperonesian Linkage).

In an article on vocative forms Blust (1979) interprets the irregular final glottal stop (phonemically analysed as *k*) in SM kinship terms and titles as a petrified vocative suffix.

In a later publication (1982a) he examines the loss of medial vowels (usually schwa) and the subsequent reduction of heterorganic consonant clusters in historically trisyllabic SM lexemes.

Finally Zorc (1982) discusses the reflexes of PAN laryngeals in more than 100 AN isolects. He argues that IBN final glottal stop[17] reflects:

(1) PAN *H* when a corresponding *h* is found in Formosan and Philippine languages;

---

[16]Dempwolff called his reconstruction Proto Austronesian, but he made it without Formosan evidence. Nowadays a reconstruction on the basis of Austronesian languages minus the Formosan ones would be labelled Proto Malayo-Polynesian (cf. fn.5).

[17]This glottal stop is represented as 'q' in Scott's Systematic Spelling, a phonemic spelling for IBN which he uses in his dictionary between brackets besides the official spelling (cf. Scott 1956:VII and passim; Scott 1957).

(2) PAN *ʔ when a corresponding ʔ is found in various Philippine languages (and probably Takituduh Bunun and Ami in Formosa);

(3) PAN *S when a corresponding sibilant is found in Formosan languages (especially Paiwan and Ami), and an *h* in most Philippine languages.

The topics treated in this study are organised in the following way. Following this introductory chapter, Chapter 2 presents a phonological description of the six Malayic isolects treated here: for each of them a short outline is given of the phonemes, the morphophonemic changes, and the phonological constraints. Chapter 3 is a reconstruction of PM phonemes. These are arranged as follows: (1) vowels, (2) diphthongs, (3) semivowels, (4) voiceless stops, (5) voiced stops, (6) nasals, (7) liquids, (8) a sibilant, (9) a glottal spirant, and (10) intervocalic ∅. Only phonemes as such are treated: their distribution within a lexeme is treated in Chapter 4. (There are two exceptions to this: for the sake of convenience, the reconstruction of antepenultimate vowels (3.1.3), and IBN changes in antepenultimate syllables and in adjacent penultimate consonant clusters (3.11) are dealt with in Chapter 3.) Chapter 4 treats PM word structure: PM phonotactic constraints are given, and tendencies to phonotactic constraints are discussed. Chapter 5 deals with PM lexicon: it includes a basic vocabulary and lexicon pertaining to well-defined semantic fields. In Chapter 6 an attempt is made to reconstruct the PM affixes. Chapter 7 follows the developments from PAN to PM; an account is given of the sound changes and of the phonotactic and lexical changes that took place between PAN and PM. Chapter 8 is a concluding chapter, which includes a summary of the most important findings, suggestions for further research and some ideas on subclassification of Malayic isolects and on the PM homeland.

I will represent phonemes from other proto-languages as I find them in the lexical instances in the linguistic literature. Although there is good reason to question some of the PAN/PMP proto-phonemes proposed by Dempwolff (and maintained by Dyen and later Austronesianists), their reconstruction does not affect the interpretation of PM phonology.

CHAPTER 2

PHONOLOGICAL DESCRIPTION OF THE MALAYIC ISOLECTS

In this chapter a brief outline is given of the phoneme systems, morphophonemic changes, and phonotactic constraints in the inherited vocabulary of the isolects compared. Deviations from this pattern in loanwords are noted separately.

2.1 THE STANDARD MALAY PHONEME SYSTEM

2.1.1 SM PHONEMES

The SM phonemes are as follows:

VOWELS

	front	central	back
high	*i*		*u*
mid	*e*	*ə*	*o*
low		*a*	

(diphthongs: *-ay, -aw*)

CONSONANTS

	labial	supra-dental	alveolar	palatal	velar	glottal
stops voiceless	*p*	*t*		*c*	*k*	
voiced	*b*		*d*	*j*	*g*	
nasals	*m*		*n*	*ñ*	*ŋ*	
fricatives			*s*			*h*
liquids			*l*		*r*	
semivowels	*w*			*y*		

d is always alveolar, whereas for most speakers *t* is supradental.

n is basically alveolar, but it becomes homorganic to a directly following *t*.

k syllable finally is realised as a glottal stop.

h is a glottal fricative; in the speech of some speakers it is heard only between like vowels, and in final position.[18]

r is pronounced as a velar or uvular fricative by some speakers (and then usually elided word finally), and as an apical flap by others. The apical flap is dominant outside the traditional Malay areas and in official Indonesian.

-ay and *-aw* (written 'ai' and 'au' in the official spelling) are actually sequences of *a* + a semivowel.

---

[18]*h* between unlike vowels is usually not pronounced, but there are minimal pairs of intervocalic *-h-* with *-ø-*, e.g. *liat* 'clay' vs *lihat* 'look!', and *tuan* 'sir' vs *Tuhan* '(the) Lord, God'. A minimal pair of *h* and *ø* in initial position is *arus* 'current' and *harus* 'have to, must'.

8

Stress falls on the penultimate syllable of a lexeme, unless this has a schwa followed by a singular consonant: then it falls on the final syllable. Stress is non-phonemic, and there are no other supra-segmental features with phonemic relevance.

2.1.2 SM MORPHOPHONEMIC ALTERNATIONS

A. The following changes are brought about by the active verbal prefix *mə(N)-*, the actor/instrument prefix *pə(N)-* and the nominal circumfix *pə(N)- an*:

(1) homorganic nasal substitution for *p, t,* and *k*, e.g.

(pukul):	*məmukul* 'hit'	*pəmukul* 'hammer'
(tolak):	*mənolak* 'refuse'	*pənolakan* 'refusal'
(kirim):	*məŋirim* 'send'	*pəŋirim* 'sender'

(2) palatal nasal substitution for *s,* e.g.

| *(salin):* | *məñalin* | exchange, translate | *pəñalin* | translator |

(3) homorganic nasal accretion[19] before voiced stops and *c*, e.g.

(bəli):	*məmbəli* 'buy'	*pəmbəli* 'purchaser'
(dəŋar):	*məndəŋar* 'hear'	*pəndəŋar* 'hearer'
(jual):	*məñjual* 'sell'	*pəñjualan* 'sale'
(gali):	*məŋgali* 'dig'	*pəŋgali* 'shovel, spade'
(cari):	*məñcari* 'look for'	*pəñcarian* '(means of) subsistence'

(4) velar nasal accretion before vowels and *h,* e.g.

(aku):	*məŋakui* (+ transitivising suffix *-i*) 'confess'	*pəŋakuan* 'confession'
(ukur):	*məŋukur* 'measure'	*pəŋukuran* 'measurement'
(hituŋ):	*məŋhituŋ* 'count'	*pəŋhituŋan* 'counting, count'

Before other phonemes only *mə-/pə-* is prefixed.

B. With some speakers of SM, suffixation of *-an* to lexemes with final *h* preceded by a high or a mid-vowel causes loss of *h* and the emergence of a non-phonemic glide (*-an* here denotes collectivity, or is part of the nominal circumfix *pə(N)- an*), e.g.

puluh [puluh] 'ten' *puluhan* [puluwan] 'tens'
məlatih [məlatih] 'exercise (v)' *pəlatihan* [pəlatiyan] 'exercise (n)'

but cf.
məŋalahkan [məŋalahkan] (+ transitivising suffix *-kan*)
'defeat' *pəŋalahan* [pəŋalahan] 'victory'

2.1.3 SM PHONOTACTIC CONSTRAINTS

The canonical form of the SM lexemes is C V C V C. A few monosyllables and trisyllables also occur. Each C can be ø, and medial C can also be a cluster (see below).

Constraints on the distribution of vowels in the non-borrowed lexicon are as follows:

---

[19] The palatal homorganic nasal in *-ñj-* and *-ñc-* clusters is written 'n' in official spelling.

(1) Schwa does not occur   (a) in final syllables;

                             (b) before *h,* a semivowel or a vowel;

                             (c) in antepenultimate syllables preceded by *h* and followed by a single consonant.

(2) Before a cluster of *r* + a consonant only schwa is permitted, e.g. *tərbaŋ* 'fly (v)', *bərsih* 'clean'.

(3) As a rule, only schwa occurs in antepenultimate syllables, e.g. *bəlakaŋ* 'back', *bənua* 'land, continent'; exceptions are *binataŋ* 'animal' and lexemes where an antepenultimate non-schwa has been prevented from becoming schwa by constraint (1), e.g. *kuala* 'mouth of a river', *baharu* 'new', *harimaw* 'tiger'.

(4) Schwa preceded by *h* or *ø* is followed by *l, r, s* or a nasal, e.g. *(h)əmpədu* 'gall', *əsaŋ* 'blow one's nose', *məŋ/əram* 'brood', *əlaŋ* (also *həlaŋ, laŋ* 'bird of prey', *həndak* 'will (aspect marker)'. It is only followed by a stop in a few exceptional cases.[20]

(5) In final open syllables only *a, i, u* or diphthongs are permitted.

(6) Diphthongs only occur in lexeme-final position, e.g. *baŋkay* 'corpse', *pulaw* 'island'.

(7) Vowels *e* and *o* occur in penultimate syllables and in closed final syllables, but only in penultimate syllables are they in phonemic contrast with *i* and *u* respectively.[21]

(8) There is vowel harmony of high and mid-vowels such that the last-syllable vowel agrees in height with the penultimate vowel, cf. *tipis* 'thin', *kurus* 'slim', *pohon* 'tree', *leher* 'neck', *kulit* 'skin, bark', *hituŋ* 'count (v)', *belok* 'turn (v)', *boleh* 'be allowed, obtain'; CeCiC-, CiCeC-, CoCuC-, and CuCoC- sequences do not occur.

---

[20]viz. *əjan* 'squeeze out by pressure', which is a variant of *rəjan* 'painful straining (in coughing or in easing the bowels)', and *ədap* 'suck up moisture', a variant of *p ap*; *ətam,* the base of *pəŋətam* 'reaping knife', is due to false analogy (cf. *kətam*) – *ədap* and *ətam* are not found in Iskandar.

[21]The distribution of high and mid-vowels is problematic. Wilkinson followed Winstedt' spelling rules for peninsular Malay (Winstedt 1927:48-49), and wrote only mid-vowels before final *h* and *k*, and only *o* before *-ŋ*, except if the last syllable was immediately preceded by another vowel. Furthermore, in the last syllable only mid-vowels (or *a*) were found if the preceding syllable had a mid-vowel, and in final open syllables only high vowels (or *a*) were found. Exceptions to these rules were loanwords. However, this was a convention, and did not reflect the linguistic reality of all peninsular Malay isolects. In Sumatra Van Ophuysen' convention was followed. In this convention high vowels before *-h*, *-k* and *-ŋ*, were not lowered, cf. Wilkinson's *jatoh* 'fall (v)', *gunoŋ* 'mountain', *bəneh* 'seed', which occur as *jatuh, gunuŋ,* and *bənih* in Klinkert. On the other hand, Van Ophuysen's spelling agreed with that of Wilkinson and Winstedt in lowering final syllable high vowels wherever mid-vowels occur in the penultimate syllable. Emeis (1955:199) pointed out the rules for height harmony in SM as it is spoken in Sumatra:
- penultimate *a, i, u* and *ə*, co-occur with *a, i, u* and with diphthongs in the final syllable, and
- penultimate *e* and *o* co-occur with *a, e* and *o*, in the final syllable.
Emeis drew this pattern from 4,548 inherited lexemes found in Poerwadarminta' (Emeis 1955:192). This pattern differs only in detail from that given by Van Ophuysen.
As to SM mid-vowels in penultimate syllables, there is no conditioning factor for their occurrence in lexemes like *pohon* 'tree', *leher* 'neck', *ekor* 'tail', *oleh* 'by', *oraŋ* 'human being', *tembak* 'shoot'. In cases like this mid-vowels are compulsory, and they are sometimes in contrast with high vowels, cf. *boroŋ* 'wholesale, by the gross', and *buruŋ* 'bird' (respectively *boroŋ* and *buroŋ* in Wilkinson). Dempwolff (1937:21-22) saw in the occurrence of SM high and mid-vowels a tendency to sound shift. He used SM *i/e* as a base for reconstructing PMP **i*, and SM *u/o* as a base for reconstructing PMP **u*. These correspondences are generally accepted among other Austronesianists. Dyen and Nothofer attributed the origin of SM mid-vowels to "secondary developments including dialectal and interlinguistic borrowing" (Nothofer 1975:50).
The relevant facts involved in the distribution of SM high and mid-vowels can be summarised as follows. Firstly, there is in various degrees a non-phonemic lowering of high vowels occurring in final closed syllables beginning with a consonant. This lowering is represented in Wilkinson before *-k* and *-h* and, for *o*, before *-ŋ*, whereas it is ignored in Van Ophuysen' orthography. Secondly, in penultimate syllables (and only here) high and mid-vowels can be in phonemic contrast. Thirdly, the vowel of the last syllable is never higher than the vowel of the penultimate syllable. The Ejaan Yang Disempurnakan (The Perfected Spelling, the uniform spelling for Bahasa Indonesia and Bahasa Malaysia which has been generally accepted since 1972) follows Van Ophuysen in the representation of high and mid-vowels, as I do in this study.

Constraints on the distribution of consonants in the non-borrowed lexicon are as follows:

(1) No voiced stops or palatals occur in final position.[22]

(2) The only consonant clusters that occur are:

(a) nasals + homorganic stops, e.g. *hantu* 'evil spirit', *baŋga* 'proud', *(h)əmpədu* 'bile, gall';

(b) velar nasal + *s*, e.g. *laŋsuŋ* 'straight, direct(ly)';

(c) *r* + any consonant but *h, r, y* or *w*, e.g. *bərsih* 'clean', *tərbaŋ* 'fly (v)'.[23]

(3) Semivowels do not occur initially, except in *yaitu* 'that is, namely', and in *yaŋ* '(the relative clause marker)', which are analysable as ⁺*ia* + ⁺*itu,* and ⁺*ia* + ⁺*ŋ* respectively. As has been said before, *y* has developed from an earlier ⁺*i.*[24] It does not occur adjacent to schwa or *i/e,* nor does *w* occur adjacent to schwa or *u/o,* although non-phonemic glides are heard (and sometimes orthographically represented) when *i/e* and *o/u* are adjacent to vowels of a different colouring (e.g. *cium ~ ciyum* 'kiss (v, n); sniff with the nose'; *duit ~ duwit* 'cent; money'; *baur ~ bawur* 'mixing up, confusing'; *əmpuan ~ əmpuwan* 'k.o. tray'). Exceptions are loanwords, and *kəyaŋan* 'heaven, fairyland' which is actually morphologically complex (< ⁺*kə* + ⁺*(h)iaŋ* + ⁺*an*).

(4) Two *r*'s rarely occur within one lexeme.[25]

N.B. I have modified Wilkinson's spelling (which was the basis of the pre-1972 spelling of SM in Malaya) in the following ways:

Wilkinson	present spelling
ch	*c*
e	*ə*
ng	*ŋ*
ny	*ñ*
-ai	*-ay*
-au	*-aw*

Furthermore, I follow the conventions of the Ejaan Yang Disempurnakan for the representation of high and mid-vowels (see fn. 21).

---

[22]Final voiced stops however are written in loanwords and in a very few sound-symbolic words e.g. *səbab* [səbap] 'reason; because', *abyad* [abyat] 'white' (both loanwords, from AR), *ləmbab* (also *ləmbap*) 'damp, humid'.

[23]In Wilkinson a few *-rh-* clusters occur in loanwords, in onomatopoeic forms, on morpheme boundaries, in *gərhaw* 'agape', and in *gərham* (with variants *gəraham* and *gaham* 'molar tooth'). Lexemes with *-rC-* clusters often have variants with *-rəC-,* e.g. *kərñih, kərəñih* 'grin (v)', *bərkas, bərəkas* 'bundle'.

[24]Prentice (pers. comm.) and others (cf. Wilkinson for *ya tu*).

[25]Of the few lexemes in Wilkinson having more than one *r,* most are loanwords from MIN and JV (*rərawitan* 'emotional melody' is based on a JV lexeme *rawit,* and only occurs in the poem 'Panji Semirang', Wilkinson 1959). Others have a variant with one *r,* or they are onomatopoeic. The form *raras* 'a large tree, *Radermachera gigantea*' is unexplained (it does not occur in Iskandar).

2.2 THE MINANGKABAU PHONEME SYSTEM

2.2.1 MIN PHONEMES

The MIN phonemes are as follows:

VOWELS

	front	central	back
high	*i*		*u*
mid	*e*		*o*
low		*a*	

(diphthongs: *iᵊ, uᵊ, uy, ay, aw*)

CONSONANTS

		labial	dental	palatal	velar	glottal
stops	voiceless	*p*	*t*[26]	*c*	[k ~ ʔ]	
	voiced	*b*	*d*	*j*	*g*	
nasals		*m*	*n*	*ñ*	*ŋ*	
fricatives			*s*			*h*
liquids			*l,r*			
semivowels		*w*	*y*			

k has two allophones: [ʔ], which occurs syllable finally, and [k], which occurs elsewhere.

h is a glottal spirant; in initial position it is not phonemic.[27]

ay and *aw* usually occur in word-final position, where they are actually sequences of *a* + a semivowel. *ay* sometimes occurs before a final glottal stop in ill-adapted loanwords, e.g. *siayʔ* 'healthy' (< ML *sehat* 'healthy' < AR *ṣiḥḥa(t)* 'health').

Stress falls on the penultimate syllable.[28] It is not phonemic, nor are there any other suprasegmental features with phonemic relevance.

2.2.2 MIN MORPHOPHONEMIC ALTERNATIONS

A. The following changes occur with affixation of the active verbal prefix *ma(N)-*, the actor/instrument prefix *pa(N)-* and the nominal circumfix *pa(N)- an:*

(1) homorganic nasal substitution for initial *p, t* and *k*, e.g.

(piliᵊh):	*mamiliᵊh* 'choose'	*pamiliᵊh* 'chooser'
(turuyʔ):	*manuruyʔ* 'follow'	*panuruyʔ* 'follower, docile person'
(kayuᵊh):	*maŋayuᵊh* 'row with short oars'	*paŋayuᵊh* 'k.o. agricultural implement'

(2) palatal nasal substitution for initial *s*, e.g.

(sudah):	*mañudah* 'end, finish, terminate'	*pañudahan* 'termination, end'

---

[26]It is likely that *t* is a supradental whereas *d* and *n* are alveolars (as in SM and JKT). Such a difference in articulation place between *t* and *d/n* is quite common in Southeast Asian languages (Henderson 1965:420-421). However, it is usually not pointed out in the linguistic literature, as *t/d/n* are phonemically in the same structural relationship as *p/b/m, c/j/ñ* and *k/g/ŋ*.

[27]See Van der Toorn (1891:X); in Thaib's dictionary initial *h* does not occur.

[28]Van der Toorn (not specifying whether this applies to lexemes or to words) speaks of a slightly longer penult mate vowel (Van der Toorn 1899:XIV). Moussay (1981) does not comment on MIN stress.

(3) homorganic nasal accretion[29] before initial voiced stop and *c*, e.g.

(bunuᵊh):	*mambunuᵊh* 'kill'	*pambunuᵊh* 'instrument to kill'	
(daŋa):	*mandaŋa* 'hear'	*pandaŋa* '(sense of) hearing'	
(jamu):	*mañjamu* 'treat, entertain'	*pañjamu* 'host'	
(gandin):	*maŋgandin* 'thrash, flog'	*paŋgandin* 's.th. to thrash or beat with'	
(cari):	*mañcari* 'look for, seek, search'	*pañcarian* 'livelihood, trade, business'	

(4) before other lexemes only *ma-/pa-/pa- an* is affixed, e.g.

maliᵊŋ 'thief':	*mamaliᵊŋ* 'steal'
	pamaliᵊŋ '(s.o.) prone to stealing'
itam 'black':	*maitam* 'become black'
anam 'six' (+ transitivising suffix *-kan*):	*maanamkan* 'make (s.th.) six'

B. The following changes take place when the transitivity marker *-i* or the nominaliser *-an* are suffixed to a lexeme:

(1) if the lexeme ends in *-a, -iᵊ,* or *-uᵊ,* a liquid appears between the lexeme ending and the suffix,[30] e.g.

kapuᵊ 'chalk':	*maŋapuᵊri* 'plaster, whitewash'
(pikiᵊ):	*pikiᵊran* 'thought'
badiᵊ 'gun':	*sapambadiᵊlan* (also *sapambadiᵊran*) 'the distance of a gunshot'
(sasa):	*mañasali* 'be sorry about s.o. or s.th.'

(2) if the lexeme ends in a glottal, this stop is sometimes followed and/or replaced by *p* or *t*, e.g.

rambuyʔ 'hair of head':	*rambuyʔtan* 'k.o. fruit, the rambutan'
sakiʔ 'ill':	*pasakitan* (Van der Toorn *pasakiʔan, pasakitan*) 'difficulty, impediment'
ikuyʔ 'follow':	*uraŋ paikuyʔtan* 's.o. without an opinion of his/her own'
tutuyʔ 'closed':	*tutuyʔpan* 'prison'

(3) if the lexeme ends in *-h*, this is sometimes replaced by *s*, e.g.

manih 'sweet':	*manisan labah* (Thaib), *manihan labah* (Van der Toorn) 'honey'

N.B. (1) The exact realisation (as preglottalised stops?) of the clusters *-ʔp-* and *-ʔt-* is uncertain.

(2) The morphophonemic changes described in (2) and (3) are exemplified in the MIN dictionaries, but they are not treated in Van der Toorn (1899).

---

[29] See fn. 19.
[30] Thaib assumes that all lexemes with final *a, iᵊ,* and *uᵊ* really end in an underlying liquid; he writes (.) wherever he expects a final liquid which is not attested through suffixation, e.g. *pitata(.)* 'k.o. tree', *garigiᵊ(.)* 'shiver (v)'.

2.2.3 MIN PHONOTACTIC CONSTRAINTS

The phonotactic shape of MIN lexemes is C V C V C. Monosyllables, trisyllables and tetrasyllables also occur. Each C can be *ø*, and intervocalic C can also be a cluster (see below).

Constraints on the distribution of vowels:

(1) Only *a*, *i* and *u* are found in antepenultimate syllables,[31] for example, *kulambu* 'mosquito net', *tilañjaŋ* 'naked', *jambatan* 'bridge'.

(2) Diphthongs only occur in the last syllable of a lexeme (as in *gunuᵊŋ* 'mountain', *pulaw* 'island', *piliᵊh* 'choose'); all other vowels are also found in the penultimate syllable.

(3) *a* occurs in final syllables before   *h, ?, m, n, ŋ* or *ŋ*;
   i   "   "   "   "   "   *h, ?, n* or *ŋ*;
   u   "   "   "   "   "   *ŋ, n*;
   o   "   "   "   "   "   *h, ?, ŋ* or *ŋ*;
   e   "   "   "   "   "   *h, ?, n, ŋ* or *ŋ*;
  uy   "   "   "   "   "   *h, ?*;
   iᵊ, uᵊ occur in final syllables before   *h, ?, ŋ* or *ŋ*;
   ay and *aw* only occur lexeme finally (except for *ay* in some loanwords).

Constraints on the distribution of consonants:

(1) Clusters consist of a homorganic nasal + stop, or *n + s*.[32]

(2) The only final consonants that occur are *h, ?, m, n* and *ŋ*; final *m* only occurs when preceded by *a*.

N.B. I have modified Thaib's spelling in the following ways:

Thaib	present spelling
dj	*j*
tj	*c*
j	*y*
ng	*ŋ*
nj	*ñ*
è	*e*
oe	*u*
-ai	*-ay*
-au	*-aw*
ie	*iᵊ*
oee	*uᵊ*
oei	*uy*
'	*?*
-(l), -(r)	(omitted)

---

[31]Exceptions to this rule are a few loanwords, a few forms with petrified affixes, and the lexeme *lemala?* (also *mala?*) 'k.o. bookmaker'.

[32]The only other cluster that occurs in Thaib is *-rg-* in *targutu* and *turgutu*, both onomatopoeic forms symbolising the sound of turtle-doves; Wilkinson gives a doublet *sala, saŋla* 'broiling, cooking at an open fire', where the velar nasal is a retention from PAN (cf. 4.6).

I have applied the same modifications to Van der Toorn's spelling, with the addition of the following:

oea, oee (1891)	*uᵊ*
ia, ie	*iᵊ*
ô, ò	*o*
é, è	*e*
(hamza)	*ʔ*

2.3 THE BANJAR HULU PHONEME SYSTEM[33]

2.3.1 BH PHONEMES

The BH phonemes are as follows:

VOWELS

	front	central	back
high	*i*		*u*
low		*a*	

(diphthongs: *-ay, -uy, -aw*)

CONSONANTS

		labial	dental	palatal	velar	glottal
stops	voiceless	*p*	*t*[34]	*c*	*k*	
	voiced	*b*	*d*	*j*	*g*	
nasals		*m*	*n*	*ñ*	*ŋ*	
fricatives			*s*			*h*
liquids			*l,r*			
semivowels		*w*		*y*		

Abdul Jebar does not give any explicit pronunciation rules for the BH phonemes.

Stress is not phonemic, and there appear to be no other phonemic suprasegmental features in BH.

A non-phonemic glottal stop is heard between like vowels, and when *-i* (the transitive verbal suffix) is suffixed to a verb ending in a vowel.

Diphthongs occur only lexeme finally,[35] and they actually consist of *a* or *u* + a semivowel. The occurrence of *-uy* seems to be restricted to loanwords.

---

[33]The two main sub-dialects of Banjarese, BH and BK, differ mainly in their lexicon and phonology: BK adds a pair of mid-vowels to the vowels it shares with BH. (In his orthography Abdul Jebar uses two different letters 'e' and 'é'; he does not explain the difference between them, and, as there is only one phoneme *e* according to Abdul Jebar's own analysis of BK, one presumes that 'e' and 'é' refer to the same phoneme. Another possibility is that 'e' refers to a schwa, and 'é' to a mid-front vowel, but again, Abdul Jebar does not mention the occurrence of BK schwa in his phonological outline).
A problem with the treatment of BH and BK lexicon is that Abdul Jebar does not distinguish between them consistently: some of his entries are marked with BH or BK, but most are not; one might suppose that unidentified lexemes belong to BK if they contain a mid-vowel, but even then one is left with many lexemes belonging either to BH, to BK, or to both. I cannot but treat all Abdul Jebar's material as belonging to BH, unless it contains a mid-vowel or is explicitly marked as BK.
[34]See fn. 26.
[35]According to Durdje and Djantera there are also interconsonantal diphthongs, e.g. *kaina* 'wait; soon', *sauda* 'no, not', *kuitan* 'aged person'.

2.3.2 BH MORPHOPHONEMIC ALTERNATIONS

The following alternations are brought about by prefixation of the active verbal prefix *ma(N)-* and the nominalising affix *pa(N)- (an)*:

(1) homorganic nasal substitution for *p, t* and *k,* e.g.

(turut):	*manurut* 'follow'	*panurut* 'who/which likes to follow'
(paŋkuŋ):	*mamaŋkuŋ* 'hit'	*pamaŋkuŋ* 'hammer'
(kuluh):	*maŋuluh* 'monopolise'	*paŋuluhan* 'always trying to monopolise, greedy'

(2) palatal nasal substitution for *s,* e.g.

(saŋat):	*mañaŋat* 'sting (v)'	*pañaŋat* 'wasp'

(3) homorganic nasal accretion[36] before voiced stops and *c,* e.g.

(banam):	*mambanam* 'burn'	*pambanam* 'burning place'
(dulaŋ):	*mandulaŋ* 'look for diamonds'	*pandulaŋ* 'place where diamonds are found'
(jajah):	*mañjajah* 'oppress'	*pañjajah* 'coloniser, colonising power'
(gantih):	*maŋgantih* 'twine, twist, spin'	*paŋgantih* 'spinning-wheel'
(cucuk):	*mañcucukkan* 'pin up'	*pañcucuk* 'pin'

Before bases with other initial phonemes, only *ma-* or *pa- (an)* are affixed.

2.3.3 BH PHONOTACTIC CONSTRAINTS

The canonical shape of BH lexemes is C V C V C. Monosyllables, trisyllables and tetrasyllables also occur. Each C can be ø, and intervocalic C can also be a cluster (see below).

Constraints on the distribution of consonants:

(1) Voiced stops, *c* and *ñ* occur in initial and medial position only.

(2) The consonant clusters which occur are:

(a) homorganic nasal + stop, e.g. *jambatan* 'bridge', *hantu* 'spirit, ghost', *kañciŋ* 'button', *ampat* 'four', *baŋkay* 'corpse';
(b) velar nasal + *s,* e.g. *buŋsu* 'youngest', *laŋsuŋ* 'direct(ly), straight';
(c) (in a few cases, mostly loanwords) stop + *l, r,* or *s; r* or *s* + stop or nasal; and *ŋ + h, l,* e.g. *baksa* 'dance (v)' (< JV *bəksa*), *ruksuy* 'bad (quality), ugly' (? < DU *rotzooi* 'mess; things of bad quality'), *ubrak* 'waste, sell-out', *ciprat* 'stain, spot', *hablur* 'crystal' (< Persian); *gargaji* 'saw' (< SKT), *garbaŋ* 'spread out, hanging down', *marma* 'horrible, terrible', *kasturi* 'musk, civet' (< SKT), *tarkam* 'pounce'; *ba-tiŋhuy* 'whistle (v)', *siŋhaja* 'purposely' (< SKT), *laŋlam* 'disappeared, submerged'; other consonant clusters are found only in loanwords.

Other constraints: diphthongs only occur in lexeme-final position.

---

[36]See fn. 19.

N.B. I have made the following alterations to Abdul Jebar's (and Asfandi's) orthography:

Abdul Jebar	present spelling
ng	ŋ
ny	ñ
-ai	-ay
-au	-aw
-ui	-uy
'	(omitted)[37]

2.4 THE SERAWAY PHONEME SYSTEM

2.4.1 SWY PHONEMES

The SWY phonemes are as follows:

VOWELS

	front	central	back
high	*i*		*u*
mid	*e*	*ə*	*o*
low		*a*	

(diphthongs:[38] *iᵊ, uᵊ, oy, ey, (-ay), (-aw)*)

CONSONANTS

		labial	alveolar	palatal	velar	glottal
stops	voiceless	*p*	*t*	*c*	*k*	*ʔ*
	voiced	*b*	*d*	*j*	*g*	
nasals		*m*	*n*	*ñ*	*ŋ*	
fricatives			*s*[39]		*x*	*(h)*
liquids			*l,r*			
semivowels		*w*		*y*		

e does not occur in Helfrich; where Aliana et al. has *e*, Helfrich has a corresponding *i*. Diphthongs occur in the final syllable of a lexeme. In Aliana et al. *iᵊ* and *uᵊ* are in phonemic contrast with final *i* and *u* respectively. *-ay* and *-aw* are phonemically sequences of *a* + a semivowel.

x, a (uvular or velar?) fricative, is in free variation with *r* in a number of lexemes.[40]

---

[37]Abdul Jebar sometimes uses an apostrophe to represent the glottal stop which is heard between like vowels, and before *-i* which is suffixed to verbs ending in a vowel. He does so for the sake of beginning students of Banjarese, and (before *i*) to indicate that the two vowels belong to different syllables. There are, however, many lexemes with sequences of like vowels where he does not use the apostrophe. As the BH glottal stop is non-phonemic, I do not represent it in my orthography.

[38]*iᵊ* and *uᵊ* only occur before *-h* and *-ʔ*. In Helfrich's analysis they are allophones of *i* and *u* respectively, as *i* and *u* do not occur before *-h* or *-ʔ*. But in Aliana et al. *-h* is not represented, and here *iᵊ* and *uᵊ* are in phonemic contrast with lexeme-final *i* and *u*. Examples of minimal pairs: *ali* 'sperm' and *aliᵊ(h)* 'change place, move', *aŋgu* 'use, wear' and *aŋguᵊ(h)* 'k.o. bird's nest', *bəli* 'buy' and *bəliᵊ(h)* 'wild, not tame'.
ey and *oy* do not occur in Helfrich's description. In Aliana et al. they are listed as diphthongs, but their occurrence seems restricted to the exclamations *ey* and *oy* (Aliana et al. p.12). In Aliana et al.'s orthography the symbol *uə* is found in two lexemes: *kukuəx* 'scratch (v)' and *təluəx* 'egg' (what I write as *iᵊ* and *uᵊ* in this study is represented as *ia* and *ua* in Aliana et al.). Their *uə* monophthongises under the same conditions as *iᵊ* and *uᵊ*.

[39]Aliana et al. also list *z* as a SWY phoneme, but I found only one instance (*zaman* 'time, period' (< AR)) and therefore do not consider it an inherited phoneme.

[40]The exact place of articulation (velar or uvular?) of *x* is uncer in. Helfrich describes it as a "burred *r*"; he represents it as 'gr', and Aliana et al. as 'gh' and 'g', which points to a velar articulation.

h is only inherited in final position.[41] In this position it occurs in Helfrich, but is not represented in Aliana et al.; hence it is written between brackets in this study.

According to Helfrich stress falls on the penultimate syllable of a word, unless this contains a schwa (in which case it falls on the final syllable); stress is not phonemic, and there are no other phonemic suprasegmental features.

2.4.2 SWY MORPHOPHONEMIC ALTERNATIONS

A. Nasalisation of actor-oriented verbs, affixation of the nominalising affixes *pəN-* and *pəN-an*, yield the following alternations:

(1) velar nasal accretion for initial vowels, e.g.

(adil):	*ŋadili* 'give judgement', *pəŋadilan* 'law court'
(ikat):	*ŋikat* 'bind', *pəŋikat* 's.th. to bind'
(ukur):	*ŋukur* 'measure', *pəŋukur* 'measurer, gauger'

(2) homorganic nasal substitution for initial stops, e.g.

(poti²(h)):	*moti²(h)* 'make white'
(pikul):	*pəmikul* 's.o. carrying a yoke'
(bəli):	*məli* 'buy'
(buka²):	*pəmuka²* 'entrepreneur'
(dəŋax):	*nəŋax* 'hear'
(dapat):	*pənapat* 'opinion'
(timbo):	*nimbo* 'bale (water)'
(tuñju²):	*pənuñju²* 'pointer'
(cukur):	*ñukur* 'shave', *pəñukur* 'barber'
(jəliŋ):	*ñəliŋ* 'look sidelong, squint'
(jait):	*pəñait* 'dressmaker, needle'
(kapur) (in Helfrich *kapux*):	*ŋapur* 'whitewash'
(kikir):	*pəŋikir* 'a file'
(gundal):	*ŋundali* 'throw at (s.th.)'
(gantuŋ):	*pəŋantuŋan* 'place to hang s.th.'

N.B. As the above examples show, in Aliana et al. stops become homorganic nasals. Helfrich, however, is vague about the exact nature of nasalisation of stops, but from his examples it appears that voiceless stops usually become homorganic nasals, and that voiced stops usually have homorganic nasal accretion. Apparent exceptions are *(baix)*, *mbaix*, *maix* 'pay', *(bəli)*, *mbəli*, *məli* 'buy', (with nasal accretion as well as nasal substitution), *(dəŋax)*, *nəŋax* 'hear' (Helfrich specifies that *(dəŋax)* is the only lexeme with initial *d* undergoing nasal substitution), *(pampo)*, *(mə)mpampo* 'carry on both hands'.

Examples of *x* in free variation with *r*: *xilaw* = *rilaw* 'k.o. eel', and *kikix* = *kikir* 'a file'; but cf. also *kuxo* 'milt, spleen' vs *kuro* 'turtle', and *raŋas* 'bare, leafless' vs *xaŋas, mə--* 'swing (of lianas)'.

[41] Helfrich has *h-* in only ten lexemes, five of which are AR loans (viz. *haji* 'pilgrim to Mecca', *hakim* 'judge', *hasil* 'product, result', *hukum* 'law, sentence', and *hormat* 'respect, honour'), two have variants without *h-* (*habis, abis* 'finished, used up; completely', and *hati, ati* 'heart, core, mind'), and one is an exclamation (*hay* 'hey'); the remaining two lexemes are: *haño* 'only', and *himpit* 'closely fitting'. (*hutan* is only found in BSM in *ni² hutan*, ('grandparent of the forest') a taboo term for 'tiger'; in other (BSM and SWY) contexts *utan* 'forest' is used).

Intervocalic *h* occurs in Helfrich in a very few Middle Malay lexemes which are not marked for SWY or BSM: I presume that these are BSM (BSM *h* ~SWY ø), e.g. *daha²* 'phlegm', *kahar* 'cart' (< DU *kar*). (In one case it occurs between unlike vowels: *laho* 'hissing sound of the *ulax muxo* (a very poisonous snake)').

(3) palatal nasal substitution for initial *s*, e.g.

(subur):	*ñuburka* 'make fertile'
(sədut):	*pəñədut* 'lazybones'

(4) if the initial consonant is a liquid or nasal, *mə-* or *ŋ-*, *pə-*, and *pə- an* are affixed, e.g.

(lupɯᵊ(h)):	*məlupuə(h)* 'flatten bamboo', *pəlupɯᵊ(h)* 'flattened bamboo'
maxaʔ 'flare up':	*məmaxaʔ* 'fan/feed (a fire)', 'light s.th.'
	pəmaxaʔan damax 'between 6.30 and 7 p.m. (i.e. the time to light a torch)'
(ŋayaw):	*məŋayaw* 'make *dudul* (a k.o. cake)'
	pəŋayaw 'k.o. spoon to stir dough for making *dudul*'
nixis 'sieve (v)':	*pənixisan* 'sieve (n)'
(nalo):	*mənalo* 'bark (of a dog)'
rəkat 'stick (v)':	*pərəkat* 'glue (n)'
(ximbo):	*məximbo* 'clear a forest for cultivation'

B. Before lexemes with initial vowel, the allomorphs *so-* or *s-* of the clitic *sə-* occur (*sə-* denotes singularity, or the circumstance of belonging to the same category) occur, e.g.

axi 'day':	*saxi* 'one day, per day'
alus 'fine, refined':	*soalus* 'as fine/refined as'
ikuᵊʔ 'tail':	*sikuᵊʔ* 'one tail; one piece per item'
ixup 'draught':	*soixup* 'a draught, a pull'
uxaŋ (*oxaŋ* in Helfrich) 'person, human being':	*suxaŋ* 'someone, a person, per person'
umur 'age':	*soumur* '(same) age; as old as'

C. The intransitive verbal prefix *bə-* is realised:

(1) as *bo-* before initial *a*, e.g.

anaʔ 'child':	*boanaʔ* 'have children, give birth'

(2) as *bo-* or *bəx-* before initial *i* or *u*, e.g.

isi 'content(s)':	*boisi* 'contain'
(ixis):	*boixis* '(be) carved with'
(iŋgut):	*bəxiŋgut* 'move, stir'
(uba(h)):	*bəxuba(h)* 'change (v)'
umur 'age':	*boumur* 'having the age of, old, mature, last (v)'

D. The prefix *tə-* (denoting involuntariness, or superlative degree) is realised:

(1) as *to-* or *təx-* before initial *a*, e.g.

(ambiᵊʔ):	*toambiᵊʔ* 'taken away'
alap 'beautiful':	*toalap* 'very beautiful'
(aŋkat):	*təxaŋkat* 'raised, lifted'

(2) as *təx-* before initial *i*, e.g.

(iŋgut):	*təxiŋgut* 'touched'

(3) as *t-* before initial *u*, e.g.

(untap):	*tuntap* 'bump up against'

E. The nominal circumfix *kə- an* (which forms nouns, or verbs with the meaning 'be affected by [base]') is realised as follows:

(1) as *ko- an* before *a* and *i*, e.g.

adil 'just':	*koadilan* 'justice'
alus 'fine, refined':	*koalusan* 'refinement'
idup 'live, life':	*koidupan* 'life, living'
iluᵖʔ 'beautiful, good':	*koiluʔan* 's.th. that can be improved, goodness'

(2) as *k- an* before *u*, e.g.

ujan 'rain':	*kujanan* 'be caught in the rain'

F. Before the clitics *-ño* (referring to the third person as an object or, if the verb is patient-oriented, to the third person as an agent) and *-la(h)* (emphatic particle), the transitive verbal suffix *-ka* is realised as *-kə-*, e.g.

diambiəʔka 'taken away':	*diambiəʔkəño* 'be taken away by her/him'
	diambiəʔkəla(h)! 'take (it) away!'

G. The following sandhi rules apply word internally or within a word group:[42]

(1) *i + i > ii*, which is pronounced as a long vowel, e.g. *di + ixup + o > diixupo*[43] [di:xupo] 'sucked in, absorbed by him/her'

(2) *n + ñ > ñ*, e.g. *dagaŋan + ño > dagaŋaño* 'her/his merchandise'

(3) *a(h) + a > a(h)a* which is pronounced as a long vowel, e.g. *bəbua(h) + an > bəbua(h)an*[44] [bəbua:n] 'various fruits'

(4) *o + a > ua*, e.g. *mañco + an > mañcuan* 'reflection, consideration'

(5) *ə + a > a*, e.g. *kə + aiᵖʔ > kaiᵖʔ* 'to the water'

(6) *ə + u* becomes *u*, e.g. *kə + ulu > kulu* 'to the beginning/source; upriver'

(7) *ə + i > i*, e.g. *kə + ilix > kilix* 'downstream'

H. The diphthongs *iᵖ* and *uᵖ* are monophthongised when the lexeme to which they belong is suffixed with *-i /-an*, or cliticised with *-yo*, e.g.

jauᵖ + i + la(h) > jauila(h) 'keep (it) away (from s.th./s.o.)!';
di + iluᵖʔ + i + ño + la(h) > diiluʔiñola(h) [di:luʔiñola][45] 'embellished by him/her (+emphasis)';
kə+iluᵖʔ+an > koiluʔan 'beauty, goodness';
di+ambiᵖʔ+i > diambiʔi 'fetched, taken away';
kə+kəciᵖʔ+an > kəkəciʔan 'smallness';

---

[42]Only the general principles of sandhi are given here (see Aliana et al. p.25).
[43]Aliana et al. analyse the vowels in these forms as single vowels which are phonetically long, e.g. *bəbuan* [bəbua:n], *kəbərsian* [kəbərsiya:n], *dixupo* [di:xupo] (in the examples given in Aliana et al. pp. 24-25 they are indicated as long vowels (with a bar to indicate length), but elsewhere in Aliana et al. they are written as 'a', 'i', so apparently they are phonemically interpreted as allophones of *a* and *i*).
[44]See fn. 43.
[45]See fn. 43.

kə+bərsiᵊ(h)+an > kəbərsi(h)an (with long [a:][46]: [kəbərsiya:n]) 'cleanness';
di+tagiᵊ(h)+yo > ditagi(h)yo[47] [ditagiyo] 'dunned, pressed for payment (by a 3rd person)'.

N.B. This monophthongisation process is not mentioned or exemplified in Helfrich.

2.4.3 SWY PHONOTACTIC CONSTRAINTS

The canonic shape of SWY lexemes is C V C V C. Each C can be ø, and intervocalic C can also be a cluster (see below). Monosyllables, trisyllables and tetrasyllables also occur.

Constraints on the distribution of vowels:

(1) Schwa does not occur in final syllables.

(2) Initial schwa is always followed by a nasal or, in one case, a liquid (cf. *əlaŋ, laŋ* 'kite (bird)'). There are two exceptions, both of which are SKT loanwords, viz. *ərti* 'meaning', *ərto* 'goods'.

(3) As a rule trisyllabic bases have an antepenultimate schwa.

(4) Diphthongs only occur in the last syllable of a lexeme:

-ay and *-aw* only occur lexeme finally (as in *baŋkay* 'corpse' and *pulaw* 'island');
iᵊ and *uᵊ* only occur before final *ʔ*, and before Helfrich's final *h* (which does not occur in Aliana et al., see 2.4.1), e.g. *kəciᵊʔ* 'small', *iluᵊʔ* 'beautiful', *bərsiᵊ(h)* 'clean', *jauᵊ(h)* 'far'.

(5) There are two mid-vowels, one of which, *e*, does not occur in Helfrich (2.4.1); *o* does occur in Helfrich, but not in final closed syllables.

Constraints on the distribution of consonants:

(1) Inherited *h* occurs only lexeme finally after *a*, *iᵊ* and *uᵊ*.

(2) According to Aliana et al., semivowels occur initially, but in Helfrich no lexemes with initial *y* are found, and only a few with initial *w* (sixteen, of which at least seven are loanwords: *wajib* 'obligation; obligatory', *wakil* 'representative, agent', *waris* 'heir' (all from AR), and *watas* 'frontier', *waraŋ* '(term of address to parents of child-in-law)', *wajik* 'k.o. cake', *walaŋ (ati)* 'depressed because of a misfortune or because of the refusal of a request' (all from JV). *y* does not occur adjacent to *ə* or *i* (the [y] heard in this position is a non-phonemic glide, as in *aiᵊʔ* [ayiᵊʔ] 'water', or *siaŋ* [siyaŋ] '(after)noon').[48]

(3) *ʔ* does not occur initially; in inherited lexemes it occurs only in final position.

(4) Voiced stops, *c* and *ñ* never occur in final position.

---

[46]See fn. 43.
[47]According to Aliana et al. (pp.24-25), the form *ditagiyo* is the result of the contraction of ᵊ ('a' in Aliana et al.) + *y* to *y*.
[48]In Helfrich these non-phonemic glides are usually represented in the spelling.
w occurs adjacent to *ə* or *o/u* in the following cases: *(pisan) kəwali* 'k.o. banana'; *kəwarasan* (< JV) 'recovered'; *kəwikkəwik* '(onomatopoeia for the sound of a kite)'; *rawuᵊ(h)* 'yell while calling s.o.'. In three other cases, *səwidaʔ ~ suidaʔ* 'sixty' (< JV), *nawon ~ naon* 'fruits of the *kəmiliŋ* tree that fell a long time ago', and *mə-lawuŋ~ mə-lauŋ* 'yell while calling s.o.', the semivowel is also a non-phonemic glide.
BSM also has *rauh* 'at, to(wards)' (< JV), and *xawuh* 'long and slim (of people)'.

(5) Clusters only occur intervocalically. The occurring combinations are:

    (a) homorganic nasal + stop, e.g. *xumput* 'grass', *kəntal* 'thick, coagulated';
    (b) ŋ + *l* or *s*, e.g. *buŋsu, boŋsu* 'youngest child in a family', *baŋlay* 'k.o. plant';
    (c) *s* + *t* (only in loanwords, like *məsti* 'certainly; have to, must' (< JV);
    (d) *r, x* + C[49] (this combination is always preceded by schwa) e.g. *kəxbay* 'married woman', *gəman* 'molar tooth', *sərpat* 'together'.

N.B. I have modified Helfrich's spelling in the following ways:

Helfrich	present spelling
ă, *a*	*a*
dj	*j*
ě	*ə*
gʳ	*x*
i, *i*	*i*
i'ă	*iə*
j	*y*
ng	ŋ
nj	ñ
ô, o	*o*
o'ă	*uᵊ*
oe, *o*	*u*
tj	*c*
-aj	*-ay*
-aw	*-aw, -o*[50]

Glottal stop and the non-phonemic glides -(w)- and -(j)- are omitted from the present spelling.

The modifications I have made to the spelling of Aliana et al. are as follows:

Aliana et al.	present spelling
ia	*iə*
ua	*uᵊ*
ng	ŋ
ny	ñ
ė	*e*
gh, ġ	*x*
k̇	ʔ

2.5 THE IBAN PHONEME SYSTEM

2.5.1 IBN PHONEMES

    The IBN phonemes are as follows:

---

[49]Aliana et al. do not mention *x* as member of a cluster, but Helfrich gives ample evidence for it.
[50]For Helfrich's SWY *-aw* (to which correspond BSM *-aw* and *-o*) Aliana et al. have *-aw* and *-o*. I will follow the orthography of the latter.

VOWELS

	front	central	back
high[51]	*i*		*u*
mid[52]	*e*	*ə*	*o*
low		*a*	

(diphthongs: *-ay, -aw, -uy*)

CONSONANTS

		labial	dental	palatal	velar	glottal
stops	voiceless	*p*	*t*[53]	*c*	*k*	*ʔ* .
	voiced	*b*	*d*	*j*	*g*	
nasals		*m*	*n*	*ñ*	*ŋ*	
fricatives			*s*			*h*
semivowels		*w*		*y*		

a, i and *u* each have long counterparts which are phonemically analysed as sequences of like vowels, e.g. *maaʔ* [ma:ʔ] 'k.o. fruit', *liiŋ* [li:ŋ] *(ləñiŋ)* 'grease',[54] *təmbuuŋ* [təmbu:ŋ] in *təmbuuŋ padi* 'paddy that is more advanced than the rest' (cf. *maʔ* 'a load, burden', *təmbuŋ* 'staff, cudgel', *liŋ* 'term of address for youths').

ʔ and *h* occur only in final position.

k is a voiceless stop; in final position it is realised as a glottal stop, and is phonetically identical with *ʔ* .[55]

-ay, -aw and *-uy* are phonemically sequences of *a/u* + a semivowel.

Stress falls on the penultimate syllable, unless this contains a schwa, in which case it falls on the last syllable.[56] It is not phonemic, nor are there any other phonemically relevant suprasegmental features.

2.5.2 IBN MORPHOPHONEMIC ALTERNATIONS

A. The active verbal prefix *N-* brings about the following changes to a base:

---

[51] *i* is an unrounded front vowel (between high and mid); before a final consonant other than *ʔ* it is realised as "a diphthong with movement from an unrounded front vowel closer than Cardinal [e] to an unrounded central vowel", e.g. *pətik* 'a spot' is pronounced [pəteᵊʔ] (Scott 1957:510).
u is an unrounded back vowel (between high and mid); before a final consonant other than *ʔ* it is realised as "a diphthong with movement from a rounded back vowel closer than Cardinal [o] to an unrounded central vowel", e.g. *buluh* 'bamboo' is pronounced [boloᵊh] (Scott 1957:510).

[52] Scott (1956, 1957) treats *ə* and *e* as allophones, and, as a matter of fact, they are in complementary distribution. I consider them as separate phonemes, because there is a structural harmony between *e* and *o*, both mid-vowels occurring in the same environment. Furthermore, there is a historical reason for treating *ə* and *e* as two phonemes, as will be seen in Chapter 3.

[53] See fn. 26.

[54] The contrast between long and short vowels is ignored in the orthography (and also in Richards, see fn. 14). The advantage of analysing the long vowels as sequences of two identical vowels is that the resulting syllable structure conforms to the general word pattern of IBN (as a long vowel is either the only vowel in a lexeme, or, in a few cases, it occurs in a lexeme of which the preceding syllable has a schwa).

[55] The difference between *-ʔ* and *-k* is distributionally determined: *-ʔ* is preceded by an unrounded open vowel [ɑ], by monophthongic allophones of *i* and *u*, or by *e* and *o*. *-k* (like other final consonants) is preceded by *a* realised as an open front vowel [a], or by *i* or *u* realised as diphthongs.
A description distinguishing between *-ʔ* and *-k* is more economical than one distinguishing only *-ʔ* besides two phonemic diphthongs *iᵊ* and *uᵊ*. Moreover, it reflects the historical development of the sounds in question.

[56] cf. Richards (1981:XIV-XV), who furthermore specifies that stress falls on the last syllable if the penultimate vowel is a schwa, unless this schwa is followed by a consonant cluster containing *r* + a consonant, or a homorganic nasal + stop.

24

(1) homorganic nasal substitution for initial stops, e.g.

(puaŋ):   *muaŋ* 'empty out'
(bunuh):   *munuh* 'kill'
(tusu):   *nusu* 'suck the breast (babies)'
(dədat):   *nədat* 'beat'
(kirum):   *ŋirum* 'send'
(gagay):   *ŋagay* 'chase'
(campur):   *ñampur* 'mix'
(jəmbuy):   *ñəmbuy* 'expose to the sun'

(2) palatal nasal subsitution for *s-*, e.g.

(səpu):   *ñəpu* 'blow (on s.th.)'

(3) velar nasal accretion before initial vowel, e.g.

(ayun):   *ŋayun* 'swing'
(idar):   *ŋidar* 'change, transfer, move'
(undaŋ):   *ŋundaŋ* 'visit, attend'

(4) prefixation of *ŋə-* before *l* or *r*, or before a monosyllabic base, e.g.

(laban):   *ŋəlaban* 'oppose'
(riŋat):   *ŋəriŋat* 'be angry'
(ñut):   *ŋəñut* 'shake'

Non-monosyllabic verbs with initial nasal do not undergo any change.

B. The nominalising prefix *pə-* is added to active verb forms (i.e. to nasalised verb forms as described in A), e.g.

(saup):   *ñaup* 'help', *pəñaup* 'helper'
(təmu):   *nəmu* 'know', *pənəmu* 'knowledge'
(riŋat):   *ŋəriŋat* 'be angry', *pəŋəriŋat* 'anger'
(untuŋ):   *ŋuntuŋ* 'give a share', *pəŋuntuŋ* 'profit'

C. The allomorphs of the intransitive verbal prefix *əN-* are homorganic to the following (voiceless) stop or *s*. (From Asmah's examples as well as from the examples in the dictionary, it seems that *əN-* is always prefixed to bases with an initial voiceless stop or *s*.) e.g.

(cabaw):   *əncabaw*[57] 'cut, mow'
(tacan):   *əntacan* 'skim'
(səput):   *ənsəput* 'breathe'
(puak):   *əmpuak* 'be surrounded by'
(kəpal):   *əŋkəpal* 'be pressed down'

D. The intransitive verbal prefix *bə-* has the following allomorphs:[58]

(1) usually *bə-*, and sometimes *ba-* (in free variation with the former) before a consonant, e.g.

---

[57]See fn. 19.
[58]The grammatically determined allomorph *pə-* is not dealt with here (Asmah 1977:86).

(puñca):	*bəpuñca* 'coloured'
(bəris):	*bəbəris* 'drizzle'
(titi?):	*bətiti?*, *batiti?* 'visit one another'

(2) before vowels the usual allomorphs are *bə-* and *bər-* ; these are in many cases in free variation with each other, and also with *ba-* and *b-* ; an allomorph *bəl-* also occurs, which sometimes freely alternates with *bər-*, e.g.

(ulun):	*bəulun* 'regard as a slave'
(anak):	*bəranak* 'give birth'
(ai?):	*bərai?*, *bəlai?* '(be) watery'
(idar):	*bəlidar* 'be transferred, move'
(umay):	*bəumay*, *bumay* 'farm'
(utaŋ):	*bərutaŋ* 'be in debt'
(ikan):	*bəikan* 'contain fish', and (with different meaning)
	bərikan 'fishing'
(acuk):	*bəacuk*, *bəracuk* 'keep poking'

E. *tə-* (denoting non-volition, or the possibility of an action) has two alternates: *tə-* occurs before any phoneme, and *t-* occurs only before vowels, e.g.

(ulih):	*tulih* 'acquire by accident; be able to acquire'
(indik):	*təindik* 'tread on unintentionally; be able to tread on'
(pəda?):	*təpəda?* 'notice; be able to see'

F. The passive voice marker *di-* has two allomorphs: *di-* occurring before any verb, and *d-*, which may be found before vowels,[59] e.g.

(əmpa?):	*diəmpa?*, *dəmpa?* 'be eaten'[60]
(diŋa):	*didiŋa* 'be heard'
(kə-datay-ka):	*dikədatayka* 'be brought'

N.B. When *bə-*, *tə-* and *di-* are prefixed to a base with an initial vowel and the vowel of the prefix is maintained, a glottal stop is heard between the prefix and the following vowel, e.g. *bəulun* [bə?ulun], *bəumay* [bə?umay], *təindik* [tə?indi?], *diəmpa?* [di?əmpa?].

2.5.3 IBN PHONOTACTIC CONSTRAINTS

The canonic shape of IBN lexemes is C V C V C . Each C can be ø, and medial C can also be a cluster.[61] Monosyllables, trisyllables, and tetrasyllables also occur.

Constraints on the distribution of vowels:

(1) Schwa does not occur in final syllables.

(2) Initial schwa is always followed by a nasal.

(3) As a rule, antepenultimate syllables contain a schwa.

---

[59]Instances with initial vowels other than schwa are not found in Asmah (1977).

[60]In Scott's Systematic Spelling (Scott 1956:VII; 1957) there are also clusters in initial position (consisting of a stop or *s* preceded by a homorganic nasal). In my analysis these are *əNC*-sequences. This analysis is favoured by the case of *əmpa?* (Scott: *mpa?*): Asmah gives *dəmpa?* (along with *diəmpa?* [di?əmpa?] as a derived form). Now *d-* is only prefixed to initial vowels, and a *-dmp-* consonant cluster would at any rate be very improbable.

[61]See fn. 60.

(4) In penultimate syllables only *a, ə, i,* and *u* occur; *o* and *e* only occur lexeme finally or before *-ʔ*.[62]

(5) The 'long vowels', *aa, ii* and *uu* only occur in closed final syllables.

(6) Diphthongs only occur in final position.

Constraints on the distribution of consonants:

(1) Semivowels do not occur adjacent to schwa or to a vowel of the same colouring (the [y] heard between *i* and *a/u,* and the [w] heard between *u* and *a/i,* are non-phonemic glides).[63]

(2) Voiced stops, *c* and *ñ,* do not occur in final position.

(3) *h* and *ʔ* only occur in final position.

(4) Clusters consist of a stop or *s* preceded by an homorganic nasal.

N.B. The orthography I use for IBN is Scott's Systematic Spelling,[64] with a few modifications:[65]

Scott's Systematic Spelling	present spelling
e	e (in final syllables)
	ə (elsewhere)
ch	c
ng	ŋ
ny	ñ
cch	cəc
nng	ŋəŋ
nny	ñəñ
q	ʔ

2.6 THE JAKARTANESE PHONEME SYSTEM

2.6.1 JKT PHONEMES

The JKT phonemes are as follows:

VOWELS

	front	central	back
high	*i*		*u*
	é	*ə*	*ó*
mid			
	è		*ò*
low		*a*	

(diphthongs: *-ay, -aw, -ey, -oy*)

---

[62]See fn. 52.
[63]An apparent exception in Scott is 'wong' ('rapids') which, however, he analyses phonemically as *uuŋ*.
[64]cf. Scott (1956:VII; 1957).
[65]In Scott 'cch', 'nng' and 'nny', are intended to represent geminated versions of *c, ŋ* and *ñ*. In Scott's Systematic Spelling there are geminated initial consonants: the schwa which is sometimes heard between initia-like consonants is considered to be non-phonemic. I consider this schwa as phonemic, and analyse Scott's geminated consonants as *CəC*- sequences.

CONSONANTS

		labial	supra-dental	alveolar	palatal	velar	glottal
stops	voiceless	*p*	*t*[66]		*c*	*k*	*ʔ*
	voiced	*b*		*d*	*j*	*g*	
nasals		*m*		*n*	*ñ*	*ŋ*	
fricatives				*s*			*h*
liquids				*r,l*			
semivowels		*w*			*y*		

ó and *ò* are in phonemic opposition, e.g. *ŋulón* 'weep, cry heavily', and *ŋulòn* 'go west'. So are *é* and *è*, e.g. *gulé* 'k.o. curry soup', and *gulè* 'sugar'; but generally speaking the distinction between *ó* and *ò* is rather weak.[67]

As a rule, diphthongs only occur in lexeme-final position.[68] They are actually sequences of vowels + semivowels.

h occurs initially only in loanwords, and finally only in exclamations and particles. In inherited lexemes it occurs intervocalically between like vowels. This intervocalic *h* sometimes alternates with *ʔ*, and is never in contrast with it (Muhadjir 1981:20).

ʔ occurs intervocalically and finally, and does not contrast with *h*.

2.6.2 JKT MORPHOPHONEMIC ALTERNATIONS

A. The active verbal prefix *N-* has the following allomorphs:

(1) homorganic nasal substitution for initial voiceless stop, e.g.

(kirim):	*ŋirim* 'send'
(pukul):	*mukul* 'beat, hit'
(tusuk):	*nusuk* 'stab'
(cari):	*ñari* 'look for'

N.B. According to Abdul Chaer (pp.XXVI-XXVII), forms with initial *c* can also have nasal accretion,[69] e.g. *ñcari*, and *(cium)*: *ñcium, ñium* 'kiss'; they can also have *ŋə-* prefixed, but Abdul Chaer does not give examples;

(2) palatal nasal substitution for *s*, e.g.

(simpən):	*ñimpən* 'save, put away'

C.D. Grijns informs me that in JKT *t* is a supradental, whereas *d* and *n* are alveolars. This observation is not made in Abdul Chaer, Ikranagara or Muhadjir.

[67]Muhadjir (1981:15) treats *ó* and *ò* as one phoneme in a seven-vowel system which looks as follows: low-central *a;* low-front *ε;* mid-front *e;* schwa *ə;* mid-back *o,* high-front *i;* high-back *u.*
Ikranagara (1980:113) distinguishes a system with six underlying vowels: low *a;* mid *e, ə* and *o;* high *i* and *u.* These underlying vowels generate 14 surface vowels, i.e. [ə], [i], [e], [u], [o], [a]; four lax counterparts of the non-low tense vowels: [I], [E], [U], [O]; and four non-low tense vowels with off-glides: [iy], [ey], [uw], [ow]. She does not distinguish between *ó* and *ò* in the way Abdul Chaer does, and she treats Abdul Chaer's *è* as a surface vowel representing underlying *a.*

[68]In the Introduction to his dictionary, Abdul Chaer (p.XXII) gives *eyt* as an example of a diphthong in non-final position. But this is an exclamation, and there are no other known examples of non-final diphthongs.

[69]See fn. 19.

(3) homorganic nasal accretion before initial voiced stops, e.g.

(bə ʔəŋ):	*mbə ʔəŋ* 'lie, tell untruths'
(dəkət):	*ndəkətin* 'approach, get near'
(gali):	*ŋgali* 'dig'
(jual):	*ñjual*[70] 'sell'

N.B. Verbal bases with initial voiced stops can also be prefixed by *ŋə-* in free variation with the above device, e.g. *ŋədəkətin, ŋəgali, ŋəjual, ŋəbòʔòŋ.*

(4) velar nasal substitution for *h,* e.g.

(hargè (< SKT)):	*ŋargain* 'give a price to, evaluate, bargain for'

(5) velar nasal accretion before initial vowel, e.g.

(atur):	*ŋatur* 'arrange, organise'
(ikut):	*ŋikut* 'follow'
(òbròl):	*ŋòbròl* 'chatter'

In other cases either *ŋə-* is prefixed (i.e. before initial *l, r, w,* and *y*), or *ø* is prefixed (i.e. if the initial consonant is a nasal). Prefixation of *ŋə-* also takes place before monosyllabic bases (which are usually loanwords), e.g.

(kir):	*ŋəkir* 'examine (medically), test (mechanically)'
	(< DU *keuren* 'id.')
(pèl):	*ŋəpèl* 'wash floors, mop' (< DU *dweilen* 'mop')

B. The intransitive verbal prefix *bə(r)-* has six allomorphs:

(1) *bə-* occurs before any consonant; it sometimes also occurs before a vowel; e.g.

(bisik):	*bəbisik* 'whisper'
(darè):	*bədarè* 'bleed'
(tarò):	*bətarò* 'bet'
(jalan):	*bəjalan* 'walk'
(cərέ):	*bəcərè* 'divorce'
(abaŋ):	*bəabaŋ* 'have an older brother'

(2) *b-* occurs before liquids (and alternates with *bə-*), e.g.

(lagaʔ):	*blagaʔ, bəlagaʔ* 'behave'
(rasè):	*brasè, bərasè* 'feel'

(3-4) *bər-,* and *br-* are in free variation before vowels, e.g.

(alaŋan):	*bəralaŋan, bralaŋan* 'prevented'
(untuŋ):	*bəruntuŋ, bruntuŋ* 'lucky'
(əmpat):	*bərəmpat, brəmpat* 'be four'

---

[70]See fn. 19.

(5) *bəʔ-* is mainly used before base forms which begin with a vowel in a closed syllable (and is sometimes in free variation with *bə-*), e.g.

(arti):	*bəʔarti* 'mean'
(ampròk):	*bəʔampròk* 'meet'
(əmpòʔ):	*bəʔəmpòʔ* 'have an older sister'
(abaŋ):	*bəʔabaŋ, bəabaŋ* 'have an older brother'

N.B. Whether there is any semantic difference between *bə-* and *bəʔ-* before initial vowels is not clear from Muhadjir's description.

(6) *bl-* only occurs with *ajar* viz. *blajar* 'learn'.

C. Final *è* of a base becomes *a* when *-an* or *-in* is suffixed, e.g.

(lupè):	*kəlupaan, lupaan* 'forgotten', *lupain* 'forget'
(lamè):	*lamaan* 'later', *lamain* 'prolong, extend; delay'

2.6.3 JKT PHONOTACTIC CONSTRAINTS

The canonical shape of JKT lexemes is C V C V C . Monosyllables, trisyllables and, to a lesser extent, tetrasyllables, also occur. Each C can be ø or (except morpheme finally) a cluster (see below).

Constraints on the distribution of vowels:

(1) As a rule, only schwa occurs as an antepenultimate vowel.

(2) Initial schwa only occurs before a nasal or a liquid, except for *əcóəcóan* 'at random'; (Abdul Chaer also has 'eyang' which is erroneous for 'éyang', cf. JV (polite) and SUN *éyaŋ* 'grandparent').

(3) *a* and schwa do not occur in word-final position (this rule applies specifically to the isolect of Mester; there are other JKT isolects with final *a* or *ə*).

(4) Diphthongs only occur in lexeme-final position (see above).

Constraints on the distribution of consonants:

(1) Palatals do not occur in final position, nor do voiced stops. Although Adbul Chaer includes many entries with orthographical final 'b', 'd', and 'g', and gives some potential minimal pairs for final *g* and *k* (e.g. *dədəg* 'terrain, surface', and *dədək* 'bran'), according to Muhadjir there are only a few lexemes with final *p, t,* or *k,* which undergo voicing of the final consonant when they take the suffixes *-in* or *-an.* The examples (six in number) given by Muhadjir appear to be all loanwords.[71]

---

[71]viz. *gəbuk* 'hit with s.th.', *gəbugin* 'hit repeatedly'; *krukup* 'blanket', *krukuban* 'covered with a blanket'; *uruk* 'put in the ground', *(k)urugan* 'buried' (all from JV); *jawap* 'answer (n)', *jawabin* 'answer (v)'; *mulut* 'anniversary of Mohammed's birth', *muludan* 'celebrate the birthday of Mohammed' (both from AR); *parut* 'grater', *parudin* 'grate' (< SUN).

(2) Clusters occur initially and intervocalically; many combinations of consonants are found. The common patterns are:

(a) for initial clusters:[72]

– homorganic nasal + stop, e.g. *mbòk* 'brother's wife', *ŋgaʔ/ŋgèʔ* 'no, not', *ŋkòŋ* 'grandparent', *ŋkali* (short for *baraŋkali*) 'perhaps';

– stop (other than *j* or *c*) + *r*, and stop (other than *d, t, j* or *c*) + *l*, e.g. *kriukkriuk* 'rumblings of the stomach (from hunger)', *trèm (tərèm)* 'tram, streetcar' (< DU), *tròtòl-an* 'full of sunspots', *klik* 'clique' (< DU), *ŋ-grèmbèŋ* 'swaying (of walk)'.

(b) for intervocalic clusters:

– homorganic nasal + stop, e.g. *rambut* 'hair', *pundak* 'shoulder', *baŋké* 'corpse', *nanti* 'wait; soon', *əmpat* 'four'; and

– velar nasal + *s*, e.g. *baŋsè* 'people', *taŋsi* 'barracks'.

(c) for intervocalic clusters (less common clusters, mainly occurring in loanwords and onomatopoeia):[73]

– stop or *s* + liquid, e.g. *gòblòk* 'dumb, stupid', *ŋ-gəbrak* 'hit hard with both hands; threaten', *cəplòk* I. 'stamp'; II. 'fried egg', *pòʔlòt* 'pencil' (< DU *potlood*), *bəslè* 'seizure' (< DU *beslag*), *pòkròl* 'solicitor' (< DU *prokureur*), *məsrè* 'close, intimate' (< SKT), *gèsrèk, ikan --* 'k.o. dried fish of inferior quality';

– nasal + homorganic stop + liquid, e.g. *santrònin* 'rush on s.th.', *səlòmprèt* 'trumpet' (< DU), *ŋ-òmplòk* 'gather';

– velar nasal + *s* + *r*, *saŋsraŋ (saŋsaŋ)* 'caught in, hooked on';

– *r* + stop or *s*, e.g. *pərban* 'bandage', *pərsèn* 'present' (both from DU), *gərtak* 'noise made to frighten someone';

– velar nasal + liquid, e.g. *buŋlòn* 'chameleon', *raŋraŋ* 'k.o. red ant';

– velar nasal + heterorganic stop, e.g. *dəŋdè* 'punishment, fine' (< SKT), *aŋpaw* 'k.o. gift in an envelope' (< CHI), *tòŋtòn* 'see, watch (a performance etc.)', *òŋji* 'issuing of a licence' (< CHI?);

– *r* + *l*, e.g. *bərlian* 'brilliant, diamond' (< DU), *pərlu* 'have to, need' (< AR);

– *s* + stop, e.g. *məsti* 'must' (< SKT), *pistòl* 'pistol' (< DU), *miskin* 'poor' (< AR), *aspal* 'asphalt' (< DU).

(3) In inherited lexemes *h* only occurs intervocalically between like vowels. It sometimes alternates with *ʔ*, as in *puhun, puʔun* 'tree'.

---

[72]Grijns says that most bases with an initial cluster have a variant form with an epenthetic schwa between the consonants. Whether a group of consonants is realised as a cluster or acquires an epenthetic schwa depends among other things on the number of syllables in the lexeme. Disyllabism is favoured in the canonical shape of JKT lexemes, and the insertion of *ə* would often result in more than two syllables. Nevertheless, there are inherited JKT lexemes which always occur with an initial consonant cluster, and which can form a contrasting minimal pair with lexemes containing a schwa. They may have a sound-symbolic value, in which case they have an initial cluster consisting of a stop or *s* + a liquid. Sometimes they are the result either of extension of a monosyllabic lexeme, or of reduction of another (non-monosyllabic) lexeme, in which case they have an initial cluster consisting of a nasal + a homorganic stop. (Grijns 1981:19-28). In the word-structural interpretation underlying Abdul Chaer's dictionary, initial clusters of *s* + a liquid do not occur.

[73]This list of clusters is not exhaustive (Abdul Chaer 1976:passim).

(4) In initial position semivowels occur in some lexemes, but these lexemes are recognised as loans. In intervocalic position *y* is found preceding final *è* (as in *bahayè* 'danger') but it is a non-phonemic glide when following a front-vowel. (This glide is orthographically represented in Abdul Chaer in one lexeme, *réyòt* 'run down, ruined, broken' (a variant of *rèòd, rèòt*, and a SUN loan). Adjacent to *ə, y* is probably not inherited.[74] *w* occurs adjacent to *a, ə, è, é* or *i*; following *u, ó* or *ò* it is a non-phonemic glide [w].

---

[74]Adbul Chaer gives 17 lexemes with *y* adjacent to *ə*. Of these, 12 can be explained as loanwords, i.e.
(a) (probably from JV) *ayəm* 'calm, at ease', *ayəŋ-ayəŋ-an* 'turn around, walk around aimlessly, doing things without sense, like s.o. insane', *bayəm* 'k.o. vegetable, amaranth', *əmpəyak* (also *əmpiak*) 'penthouse, lean-to', *gayəm* 'ruminate', *ŋ-gayəŋ* 'weep long and silently', *uyəg-uyəg (ŋuyəg-ŋuyəg)* 'massage the knee with palm of hand', *puyəŋ* 'dizzy, having a headache';
(b) (probably borrowed from SUN) *ri[y]əp* (actually an orthographic variant of *riəp*) 'become dark' (SUN *rɨip* 'id.');
(c) (borrowed from DU) *bəyonèt* 'bayonet', *puyər* 'powder';
(d) *əyaŋ*, which is an erroneous spelling of *éyaŋ* 'grandparent' (borrowed from JV or SUN, cf. JV, SUN *éyaŋ*).
In one case *y* is an orthographically represented non-phonemic glide: *kiyek-kiyek* 'the peeping of young chickens' (an onomatopoeic form). I analyse this form phonemically as *kiək-kiək*. The four remaining cases are: *dəmpəyak* 'overgrown, with branches low to the ground', *gayəm* 'calm, quiet' (~ *ayəm?*), *puyən* 'dirt', *kələyəŋan, kəliyəŋan* 'dizzy, having a headache'.

THE RECONSTRUCTION OF PROTO MALAYIC PHONEMES

In this chapter a reconstruction is made of PM phonemes on the basis of correspondences found in the six isolects. The following reconstructions (with different reflexes in the isolects) are treated, and are reflected in the proto-lexemes occurring in the following pages:

*a > SM, BH, IBN a, MIN a/o/e, SWY a/-o, JKT a/-è (3.1.1, 3.1.1.1, 3.1.1.3, 3.1.1.4);
penultimate *ə > SM, SWY, IBN, JKT ə, MIN, BH a (3.1.1, 3.1.1.2);
final-syllable *ə > SM, BH, SWY, IBN a, MIN a/o/e, JKT ə (3.1.1.5);
*i > BH, IBN i, SM i/e, MIN i/iᵃ/e, SWY i/iᵃ, JKT i/é/è (3.1.2, 3.1.2.1, 3.1.2.3);
*u > BH, IBN u, SM u/o, MIN u/uᵃ/uy/o, SWY u/uᵃ/o, JKT u/ó/ò (3.1.2, 3.1.2.2, 3.1.2.4);
*A > MIN, BH a, o.i. ə (3.1.3, 3.1.3.1);
*-ay > JKT -é, o.i. -ay (3.2, 3.2.1);
*-aw > JKT -ò, o.i. -aw (3.2, 3.2.2);
*-p > MIN -ʔ, o.i. -p (3.4.2, 3.4.2.1);
*-t > MIN -ʔ, o.i. -t (3.4.2, 3.4.2.2);
*-k > MIN, SWY -ʔ, o.i. -k (3.4.2, 3.4.2.3);
*-ʔ > IBN -ʔ, o.i. ø (3.4.2, 3.4.2.4);
*b I *a _ *a > IBN b, o.i. w (3.5, 3.5.1);
*-m I *(i,u)_# > MIN -n, o.i. -m (3.6.3, 3.6.3.1);
*r > SWY x/(r), MIN r/-ø, o.i. r (3.7, 3.7.3, 3.7.4, 3.7.5);
*-s > MIN -h, o.i. -s (3.8, 3.8.2);
*h (non-final) > BH h, SM h/ø, o.i. ø (3.9, 3.9.1);
*h (between like vowels, or between V₁ and *ə) > IBN, SWY ø, o.i. h (3.9, 3.9.2);
*ø > BH, SWY, IBN ø, o.i. h between like vowels (3.10).

3.1 THE PM PLAIN VOWELS

3.1.1 PM *a AND *ə

In all isolects a occurs in both the penultimate and final syllable. In JKT ə occurs in all syllables, while in SM, SWY, and IBN, it occurs only in non-final syllables.

Penultimate a and ə agree in SM, SWY, IBN, and JKT, whereas MIN and BH show a for both. As the distinction between a and ə in SM, SWY, IBN, and JKT penultimate syllables is not conditioned and reflects a PMP distinction, I reconstruct PM *a and *ə.

In closed final syllables a agrees in SM, BH, SWY and IBN, and this a corresponds to MIN a, e, o, and JKT a, ə or è: MIN has a corresponding e before -ʔ which developed from PM *t (cf. 3.4.2, 4.3.2.2) and before -h which developed from PM *s (3.8, 3.8.2); it has a corresponding o before -ʔ which developed from PM *p (3.4.2, 3.4.2.1), see also (C) last N.B. below; otherwise it has a.

JKT has *-è* corresponding to SWY *-a(h),* other isolects *-ah* (PM *-h* was lost in JKT (3.9, 3.9.3). Before other final consonants it has *a* or *ə,* and there is no conditioning factor for the distinction between these vowels.

In open final syllables, SM, BH, IBN *a* corresponds to MIN, SWY *o,* JKT *è.*

TABLE 2: CORRESPONDENCES OF *a* AND *ə* FOUND IN THE SIX ISOLECTS

	SM	MIN	BH	SWY	IBN	JKT
antepenultimate syllable	ə	a	a	ə	ə	ə
	ə	i	i	ə	ə	ə
	ə	u	u	ə	ə	ə
penultimate syllable	ə	a	a	ə	ə	ə
	a	a	a	a	a	a
final closed syllable	a	a/e/o	a	a	a	ə
	a	a/e/o	a	a	a	a
	a	a	a	a	a	è
final open syllable	a	o	a	o	a(ʔ)	è
(more specified)	-ap	-oʔ	-ap	-ap	-ap	-əp
	-ap	-oʔ	-ap	-ap	-ap	-ap
	-at	-eʔ	-at	-at	-at	-ət
	-at	-eʔ	-at	-at	-at	-at
	-ak	-aʔ	-ak	-aʔ	-ak	-ək
	-ak	-aʔ	-ak	-aʔ	-ak	-ak
	-al	-a	-al	-al	-al	-əl
	-al	-a	-al	-al	-al	-al
	-ar	-a	-ar	-ax,-ar	-ar	-ər
	-ar	-a	-ar	-ax,-ar	-ar	-ar
	-as	-eh	-as	-as	-as	-əs
	-as	-eh	-as	-as	-as	-as
	-ah	-ah	-ah	-a(h)	-ah	-è
	-ay	-ay	-ay	-ay	-ay	-é (3.2.1)
	-aw	-aw	-aw	-aw	-aw	-ó (3.2.2)
	-a	-o	-a	-o	-a	-è
	-a	-o	-a	-o	-aʔ	-è

Since the occurrence of JKT *a* and *ə* in final closed syllables is not conditioned, I will take JKT as a test language for the reconstruction of PM **a* and **ə* in final syllables (before a final C other than **h* or **ʔ*).

The interpretation of JKT final-syllable *ə* however needs further investigation. JKT has borrowed heavily from JV, SUN, Balinese, and Sasak, which are closely related to the Malayic group and which maintain schwa in final syllables. One could argue that lexemes with final-syllable *ə* came into JKT through borrowing. There are, however, several reasons

for assuming that JKT final-syllable ə is inherited and reflects PMP *e (schwa), and for reconstructing PM *ə on the basis of it.

A. In Blust's 200-item basic wordlist for PMP (see.5.7), there are 32 reconstructions with final-syllable *a or *e and a final C other than *q which have a JKT reflex (PMP *q became h or ø in the Malayic isolects, cf. 7.1). PMP and JKT agree in their distribution of *a/a and schwa, as can be seen from the following list (if the meaning of the JKT reflex differs from PMP, this meaning is given next to the JKT reflex):

	PMP		JKT	meaning
3.	*ka-wanan	1.	k/anan	right side
6.	*Zalan	2.	jalan	road, path
15.	*tuqelan/tuqelaŋ	3.	tulaŋ	bone
21.	*DemDem	4.	dəndəm 'grudge, revengeful feeling'	think, meditate, brood
25.	*liqeR	5.	lèhèr	neck
37.	*kaʔen	6.	ma/kan	eat
42.	*isep	7.	isəp	suck
44.	*deŋeR	8.	dəŋər	hear
47.	*(ma)-huab	9.	uap	yawn
49.	*inep	10.	inəp 'spend night'	lie down (sleep)
56.	*anak	11.	anak	child
62.	*qatep	12.	atəp	roof, thatch
66.	*iket	13.	ikət	tie up, fasten
81.	*tazem	14.	tajəm	sharp
89.	*gemgem	15.	gəŋgəm	hold (in the fist)
104.	*meñak/miñak	16.	miñak 'oil'	fat, grease
111.	*ihekan	17.	ikan	fish
115.	*akaR/*wakaR, *uRat	18.	akar	root
		19.	urat 'vein, tendon'	root
127.	*qutan	20.	utan	woods, forest
129.	*bulan	21.	bulan	moon
133.	*quZan	22.	ujan	rain (n)
135.	*kilat	23.	kilat	lightning
138.	*(ma-)panas	24.	panas 'warm, hot'	warm, of weather
142.	*(ma-)beReqat	25.	bərat	heavy
147.	*(ma-)qitem	26.	itəm	black
165.	*Zaqat/*Zaqet	27.	jahat	bad, evil
166.	*(ma-)bener	28.	bənər	correct, true
174.	*qi Dalem	29.	daləm	in, inside
175.	*a taʔas	20.	atas	above
196.	*bilaŋ	21.	bilaŋ (also 'say')	count (v)
200.	*hepat	22.	əmpat	four

N.B. two apparent exceptions, *kaʔen > JKT ma/kan and *liqeR > JKT lèhèr are explained in sections 4.5 and 3.1.1.5 IC respectively; see also 3.1.1.5 IC for *zaqat/zaqet > JKT jahat.

A broader picture of this correspondence is obtained by comparing JKT with the vocabulary of Dempwolff (1938). This comparison yields 324 lexemes[75] which have JKT reflexes, and of these the following do not match:

	PMP	JKT	meaning
1.	*ancam	añcəm 'threaten'	plan, what one has in mind
2.	*bulat	bulet	round, circular
3.	*deRas	dərəs 'fast, rapid'	be in a hurry
4.	*deŋan	dəŋən 'with'	companion
5.	*DeDak	dədək	bran
6.	*zinak	jinek	tame
7.	*asap	asəp	smoke
8.	*h(ei)(N)Zam	p/iñjəm	borrow
9.	*i(n)zak	iñjək 'step (v)'	pace, step, dance
10.	*cecak	cəcək, cəkcək	gecko, house lizard
11.	*seDaŋ	sədəŋ 'medium, enough'	average, medium
12.	*raŋkap	raŋkəp 'a pair'	unite, become one
13.	*pantas	pantəs 'capable, correct'	be successful
14.	*pesan	pəsən 'order, direction, command'	carry on the shoulder
15.	*putat	putət, putat	name of a tree
16.	*cepat	cəpət	quick, skilful
17.	*tukar	tukər	exchange
18.	*tambal	tambəl	mend, repair
19.	*taŋkap	taŋkəp	seize, hold
20.	*tegaŋ	təgəŋ	firm, steadfast
21.	*teman	təmən 'companion'	be used to
22.	*tikar	tikər	mat
23.	*tikam	tikəm	stab
24.	*baDer	badar	k.o. fish
25.	*zelateŋ	jəlatəŋ 'stinging nettle'	name of a tree, a nettle
26.	*Re(n)teŋ	rəntaŋ	stretch out
27.	*puket	pukat	drag-net, trawl-net
28.	*qu(n)tek	otak	brain
29.	*tempet	təmpat	abode
30.	*tilem	tilam	mattress
31.	*iheq	kəm/ih	urine

Several of these reconstructions can now be reinterpreted (all with final-syllable *e instead of final *a):

2. *bulat must be *buled on account of PMJ *buləd (Nothofer 1975:138).

---

[75]Dempwolff's reconstructions are given in Dyen's orthography for PAN/PMP, as modified by Blust. The representation of these reconstructions also follows their reinterpretation by Dyen and later Austronesianists. However, I present Zorc's evidence for PAN *-S, *-H, *-ʔ > IBN -ʔ in the orthography used in Zorc (1982). When the difference between PMP and PAN is not relevant for the argument, I refer to PMP. The main difference between PAN and PMP phonology is constituted by the following mergers:
PAN *S, *H> PMP *h,
PAN *t and *C > PMP *t, and
PAN *n and *N > PMP *n.
As I already stated in the introduction, I will represent proto-phonemes as I find them in the lexical instances in the linguistic literature.

3. *deRas must be *deRes on account of Ngaju dehes (Dyen 1956: passim).

4. *deŋan is attested in Toba and Ngaju (Dempwolff 1938) and furthermore in Sasak (dəŋan 'accompany'), but in SUN one finds diŋin 'all that is eaten with rice; side-dish', which makes a reconstruction of PMP *deŋen at least as warranted as *dəŋan (it is not unlikely that the Toba, Ngaju and Sasak correspondences are borrowed from SM).

5. *DeDak is not a valid reconstruction. Dempwolff made it on the basis of Tagalog lalak, Toba dodak, JV dədak, and SM dədak 'bran'. Lalak is not found in the dictionaries, but there is darák 'bran', which is a Malay loanword according to Wolff (1976:366). In JV dədək occurs as well as dədak, but Old Javanese only has dədək. Although there is at first sight no reason to consider Toba dodak a loan,[76] corresponding forms in other languages justify this assumption, cf. Balinese dədək, dəkdək, Madurese dhəkdhək (and Old Javanese dədək). The shape of a P(W)MP proto-lexeme for 'bran' should therefore be *DekDek.

7. *asap should be *qasep on account of PMJ *hasəp (Nothofer 1975:171).

8. *h(ei)(N)Zam has two correspondences which reflect PMP *e: SUN iñjim and Kroë-Lampong ŋ-iñjəm (Helfrich 1891), both meaning 'borrow', which warrants the reconstruction of a doublet *inzem.

9. *i(n)zak has also two other correspondences reflecting PMP *e: Mansaka indig (Philippines) and Tagalog t-indíg 'stand (v)' point to a doublet *inzeg.

10. *cecak is not a valid reconstruction: Toba sosak is probably a loan from SM,[77] and Ngaju tasak, Malagasy tsatsaka have an irregular vowel (cf. Dempwolff 1938); on the other hand, Balinese and Sasak have cəkcək which is regular. Balinese, Sasak cəkcək, Timugon Murut sosok, Proto Philippine *cekcek (Zorc 1971) and SM cəcak point to PMP *cekcek; JKT cəcak/cəkcək may be a loan from Balinese or Sasak on account of the intervocalic cluster in cəkcək.

11. *seDaŋ must be *seDeŋ on account of PMJ *səddəŋ (Nothofer 1975:154).

16. *cepat is not valid: Ngaju capat is a Malay loan (on account of its first vowel), and Sa'a toha 'rejoice' does not agree very well semantically. On the basis of Toba sopot 'untimely, premature' and Balinese cəpət 'precise, accurate; fast, accurate (mind)', PWMP *cepet is more plausible.

17. *tukar must be *tukeR on account of PMJ *tukər₂ (Nothofer 1975:164).

19. *taŋkap should be *ta(ŋ)kep on account of Toba tahop,[78] Ngaju takep, takep/an, kep/an 'put on, pulled on (clothes); be attached, fastened, tied to', SUN taŋkip 'put one's arm around s.th.; wrap its roots, around s.th. (of plants); entwine, clasp', and JV taŋkəp 'caught, seized, arrested'. Toba taŋkap [takkap] and Ngaju taŋkap (in Dempwolff 1938) must be loans.

---

[76] Dodak is the regular term for 'bran' in Toba as well as in other Southern Batak isolects (viz. Angkola-Mandailing and Simalungun). But the Southern Batak isolects also have uhut (Toba- and Simalungun), ut (Angkola-Mandailing) with this meaning, and Karo (which is Northern Batak) does not have a form corresponding to dodak (it has kədəp for 'bran').

[77] Sosak is found in Van der Tuuk but not in either Tambunan (1977) or Warneck (1977); the usual Toba term for 'gecko' is ilik. Toba-Batak sosak is not compatible with PMP *cekcek (⁺soksok would have been a regular reflex), and it is probably borrowed from SM.

[78] cf. Van der Tuuk (tahop) manahop 'put one's arm around s.o., as in wrestling; seize s.th. in its claws, as a vulture'; also taŋkop [takkop] 'fit, of clothes'; Tambunan tahop 'fishing net'; Warneck tahop/tahop ni bodil 'cock of a gun'.

22. *tikar was reconstructed on the basis of SM *tikar* and Malagasy *tsihy*. But it should be reinterpreted as *tikeR* on account of Bisaya and Bikol *tikog*, Ilokano *tiker* 'reed-mace, cattail' (cf. Conant 1911:80). This is not inconsistent with Malagasy *tsihy* 'mat' which could reflect either *tikaR* or *tikeR* (but not *tikar*).

Apart from these interpretations there is also the irregular JKT *kəmih*, which must be a loan from another Malayic isolect because of its *-h*.

Of the 31 correspondences which do not match, 14 can be explained through erroneous reconstructions in Dempwolff (1938) or through borrowing (as with *kəmih*), which leaves 17 irregular correspondences, or 5.55 per cent of the 324 tested ones. This quantitative evidence makes it difficult to believe that JKT final-syllable ə does not reflect PMP *e*.

B. If PMP *e* in fact became JKT *a* in final syllables, one would have to assume that almost all inherited JKT lexemes containing *a* (< *ə*) in this position were replaced by cognate loanwords from JV, SUN, Balinese, or Sasak, and that no Malayic isolect interfered in this process. But Malayic isolects (especially SM) have had considerable influence on the JKT lexicon. Furthermore, one would then have to explain a number of JKT lexemes with final-syllable ə which have cognates only in other Malayic isolects, or which have cognates in JV, SUN, Balinese, and/or Sasak which agree far better with cognates in other Malayic isolects in other (phonological, semantic, and formal) aspects. For instance, the following cognate sets appear not to have cognates in isolects outside the Malayic group:[79]

JKT *səmpət, ŋərəm*: see (C) below;
JKT *diəm*, IBN *diaw* (3.2.3), o.i. *diam* 'be quiet, silent; dwell';
JKT *əŋgən*, SM *əŋgan*, MIN *aŋgan*, IBN *əŋgay* (3.2.3) 'unwilling, reluctant';
JKT *kəjəm*, SM, SWY *kəjam*, MIN, BH *kajam* 'close the eyes';
JKT *ləŋkəp*, SM *ləŋkap*, MIN *laŋkoʔ*, BH *laŋkap*, SWY *laŋkap* (penultimate a unexplained) 'complete, having its parts or requisites';
JKT *rəndəm*, SM, IBN *rəndam*, MIN, BH *randam*, SWY *rəndam, xəndam* 'soak, steep';
JKT *təgəp*, SM, IBN *təgap*, MIN *tagoʔ*, BH *tagap* 'strong, firm'.

The following JKT lexemes agree much better with cognates within the Malayic group than with cognates in JV, SUN, Balinese, and/or Sasak:

JKT *garəm*, SM, MIN, IBN *garam*, SWY *garam, gaxam* 'salt' (cf. Sasak *garəm* 'grain');
JKT *antəm*, SM *(h)antam*, MIN *antam*, BH *hantam* 'fight, beat violently' (Sasak *antəm* 'remorse (n)');
JKT *ənam*, SM, SWY, IBN *ənam*, MIN, BH *anam* 'six' (Balinese *nəm, nəmnəm*, JV, Sasak *nəm* 'six');
JKT *kəñcəŋ*, SM *kəñcəŋ* 1. 'fast, rapid' 2. 'tight' (JV *kəñcəŋ*, SUN *kèñcaŋ* (with irregular vowel correspondences) 'tight');
JKT *m/urəm*, SM, MIN, BH *m/uram* 'cloudy, overcast', IBN *uraw* 'cloudy, of indefinite colour' (SUN *hiɨm* 'shade, shadow', Nothofer 1975:161; Balinese *urəm* 'sombre, overcast');

---

[79]In Pigeaud JV *diəm, kəjəm, ləŋkəp, ŋərəm, rəndəm/rəndəm, səmpət* (with similar meanings as in SM), and *təgəp* 'quick' are found. But *diəm, ləŋkəp, ŋərəm, rəndəm/rəndəm*, and *səmpət* are labelled as local, and *kəjəm, təgəp* are labelled as bookish. Moreover, *diəm* and *rəndəm* have a (retroflex) *ḍ* instead of an alveolar *d* (which points to borrowing).
In Jansz none of these forms is given (except for *səmpət*, which is labelled as SM). They must be considered as not inherited: Pigeaud gives many loanwords that entered JV after Jansz (1913, 2nd edition) was published, and does not systematically indicate these loanwords as such (Pigeaud 1938 was only intended to be a practical and concise dictionary preliminary to a more complete edition, see his Introduction, sections 4-6).

JKT *pərəm,* SM *pəram,* MIN, BH *param* 'store fruit for artificial ripening' (SUN *piyim* 'ketan which has been sweetened by ragi', Nothofer 1975:81, 161);

JKT *təkən,* SM, SWY, IBN *təkan,* MIN, BH *takan* 'press down' (JV *təkən* 'walking-stick; (archaic) a support').

Further, in the case of JKT *kəbəl* (SM, SWY, IBN *kəbal,* MIN *kabal*) 'invulnerable', there are languages reflecting PMP final-syllable **e,* but these are unlikely to be a source for borrowing: Toba *hobol,* Karo *kəbəl* 'invulnerable'.

C. Internal evidence for PM **ə* is sometimes obtained from other Malayic isolects.

IBN and Mualang show internal evidence for a final schwa in some originally complex forms with a petrified suffix *-i?.* This suffix became petrified before the change of PM final-syllable **e* to *a,* but it has a still productive cognate *-i* in SM, MIN, BH, and SWY (cf. 6.1.1). It follows that final-syllable **ə* (which changed into *a* after loss of the suffix) still occurred at least in morphologically complex forms of PM. For example:

IBN *əmpət/i?* 'use spare moments' (loss of **s-* unexplained), and IBN *səmpat,* SM *səmpat,* BH *sampat,* JKT *səmpət,* 'having sufficient time';

IBN *ŋ-ərəm/i?* 'press down', Mualang *ŋ-arəm/i?* 'brood',[80] and SM *əram,* JKT *ŋ-ərəm* 'brood';

IBN *sələt/i?* 'fill a gap', and IBN *səlat* 'alternation, stripe, streak', SM *səlat* 'strait, narrows' (cf. also *səlat/an* 'south', 5.2.1), MIN *sale?,* SWY *səlat* 'intervening space, narrow space between two objects'.

Furthermore, there is internal evidence for a final-syllable schwa in some SM, SWY, IBN, and Mualang forms with a petrified suffix *-/an.* This suffix was lost in IBN and Mualang (before the change of final-syllable **ə* to *a*), but it is still productive in the other isolects. For example:

SWY *kələm/an* 'obscure, dark', IBN *kələm/an* 'the moon on the wane', Mualang *kaləm/an* 'set (of sun, v)',[81] and SM, SWY *kəlam,* MIN *kalam* 'obscure, dark', IBN *kəlaam,* JKT *kələm* 'go under';

SWY *dələm/an* 'be pregnant', *daxa(h) ən/dələm/an* 'black and blue, of a contusion', and JKT *daləm* 'inside, inner, in', IBN *alam,* o.i. *dalam* 'deep; depth';

SM, JKT *dəp/an* 'before; next, coming' (with apocope of the first syllable, see 3.11c), and SM *(h)adap,* BH *hadap,* SWY *adap,* JKT *adəp* '(be) in front of, before', MIN *ma/ado?* 'face, be in front of'.

N.B. All the originally complex forms given above underwent antepenultimate neutralisation (Mualang to *a,* the other isolects to *ə,* see also 3.1.2).

Tioman Malay (spoken on Tioman Island, off the south-east coast of the Malay Peninsula) reflects PMP **e* in final syllables if the penultimate vowel equally reflects **e,* thus PMP **e* > Tioman Malay *ə* | *CəC_ C#* (Collins 1985). For example:

---

[80] *ŋ-arəm/i?* is found in verse 3003, and *kaləm/an* in verse 2709, of the *Kana Sera* (Dunselman 1955). Dunselman gives *a* and *ə* (usually *a* in prefixes, and *ə* in lexemes) for the neutralised penultimate vowel in Mualang.
[81] See fn. 80.

	PMP	Tioman	meaning
	*tebel	təbəl	thick
	*peZem	pəjəm	close (eyes)
	*gemgem	gəŋgəm	squeeze in the fist
	*demdem	dəndəm	hold a grudge
whereas:	*Dalem	dʌlam	in(side)
	*habaRat	bʌɣat	west
	*qalesem > PM *m-asəm >	masam	sour

Finally, internal evidence is found in the correspondence SWY *liax,* SM *leher,* MIN *li(h)iᵊ,* JKT *lèhèr* 'neck', and SWY *ən/tuat,* SM *lu/tut,* MIN *lu/tuyⁿ,* BH *tuut* 'knee', where the different development of the SWY final vowel is a reflex of PMP *e*: see 3.1.1.5 IC.

A later change from JKT +a to ə is unconvincing, since one must then explain why this phonological change did not affect those final *a*'s which represent PMP *a*. Therefore I conclude that JKT final ə is directly inherited from PMP *e*. This conclusion does not rule out the occurrence of ə from +a through false analogy, which seems to be the case in loanwords like *kuləm* 'pond, pool, reservoir, tank' (from TAM *kulam*) and *imət* 'solicitude, care, attention' (from AR *himmat*).[82] Nor does it rule out the occurrence of loanwords from other Malay isolects with final syllable *a* from PM *e*: borrowing from SM is probably the origin of the irregular JKT reflexes 24-31 (see Table 4).

N.B. In MIN, the reflexes of PM *e* and *a* undergo identical changes in final closed syllables. Consequently, the merger of PM *a* and *ə* took place prior to other changes in the last syllable. The following rule order can be established for changes in MIN last syllables with *-eⁿ, -eh,* and *-oⁿ.*

TABLE 3: THE DEVELOPMENT OF PM *ə AND *a IN MIN FINAL CLOSED SYLLABLES

I	PM *ə and *a	>	+a ǀ _ C#
II	+a	>	+o ǀ _ +p#
		>	+e ǀ _ +t#, +s#
		>	+a elsewhere
III	-+p, -+t	>	-ⁿ
	-+s	>	-h

It is possible that *a and *ə first merged to +ə: a schwa would more easily give rise to vowel differentiation. But in some MIN isolects where no raising took place, final *a and *ə are reflected as *a*. Note that the merger of *a and *ə took place in the last syllable (rule I) before it did in other syllables. Tamsin Medan's dialect study of MIN shows a merger of final syllable *a and *ə in all MIN isolects, whereas a merger in other syllables took place only in a few of them (among which Koto Gadang). In other isolects non-final syllable *ə became *e* or *o* (Tamsin Medan 1980:78). Tamsin Medan also shows that in some other isolects rules II and III did not apply at all, or not to the same extent.

---

[82]But even here there is another possible explanation: JKT *imət* may be a loan from SUN (SUN *imit* 'diligent, industrious' < AR). (The TAM lexeme *kulam* is pronounced as [kulɔ̃]; it is not impossible that it was borrowed into JKT at some earlier stage of TAM (before nasalisation of the last vowel took place), and that its last vowel was perceived as a schwa).

3.1.1.1 PM PENULTIMATE *a > ALL ISOLECTS a

Examples:

*bara? 'live coal'; SM, BH bara, MIN baro, SWY baxo, IBN bara?, JKT barè;

*hari 'day'; SM (h)ari, BH hari, MIN, IBN, JKT ari, SWY axi;

*salah 'at fault, amiss'; SWY sala(h), JKT salè, o.i. salah.

3.1.1.2 PM PENULTIMATE *ə > SM, SWY, IBN, JKT ə, MIN, BH a

Examples:

*pənuh 'full'; SM pənuh, MIN panuᵊh, BH panuh, SWY pənuᵊ(h), JKT pənu;

*tənun 'weave'; MIN, BH tanun, o.i. tənun;

*səŋət 'sting of (venomous) insect'; SM, SWY, IBN səŋat, MIN saŋe?, BH saŋat; JKT n.c., but cf. PMP *seŋet 'sting'.

UNEXPLAINED IRREGULAR CORRESPONDENCES

(1) SWY laŋkap has a penultimate a corresponding to SM and JKT ə, cf. SM ləŋkap, JKT ləŋkəp 'complete' < *ləŋkəp (3.1.1, 3.1.1.5).

(2) BH i reflecting PM *ə. BH sometimes reflects i for PM *ə. This i often co-occurs with a preceding or following i, and may also occur in both the final and penultimate syllable. For example:

BH cilup (also culup) 'soak, dye', SM, JKT cəlup 'id.' (SWY cəlup 'drop in, pass by' may be a cognate; Wilkinson (1959) also gives MIN colop, JKT cələp 'id.');

BH kilim 'k.o. small seam on a rug', SM, SWY kəlim 'id.', JKT kəlim, ŋ-- 'fold, sew an edge' (this must be borrowed from a West-Asiatic language (through Persian or AR?), cf. Turkish kilim, which refers to (1) a flatweave carpet, and (2) the woven edge of a carpet).

BH pijim (also pajam) 'close the eyes', SM, IBN pəjam, MIN pajam, pijam (cf. PMP *peZem 'id.', Dyen 1951:536);

BH rikit 'stick, glue', SM, SWY rəkat, MIN rake? (cf. PMP *reket and *riket 'id.', Blust 1970);

BH imit 'forethought, judgement; saving, careful', SM hemat 'id.', JKT imət 'saving, thrifty' < AR;

BH tikin 'sign (v)', SM tekən, JKT tèkən < DU.

3.1.1.3 PM *a |*_ C# (≠ *?) > MIN o | _ ? (< *p)#, OTHER ISOLECTS a
MIN e | _ ? (< *t)#, OTHER ISOLECTS a
MIN e | _ h (< *s)#, OTHER ISOLECTS a
MIN a ELSEWHERE, OTHER ISOLECTS a
JKT è | _ ø (< *h),
JKT a ELSEWHERE, OTHER ISOLECTS a

Examples:

*akar 'root, creeper'; SWY akax, MIN aka, o.i. akar;

*hutan 'jungle, wilderness'; SM, BH hutan, o.i. utan;

*uraŋ 'outsider'; SM, JKT oraŋ 'human being', SWY oxaŋ (Helfrich), uraŋ (Aliana et al.) 'id.', o.i. uraŋ 'id.' (see 5.7 lemma 53 for the meaning of this etymon);

*anak 'child'; SWY ana?, o.i. anak;

*sayap 'wing'; SM, SWY, IBN, JKT sayap, MIN sayo?;

*bəras 'uncooked rice'; MIN bareh, BH baras, SWY bəxas, o.i. bəras;

*rumah 'house'; SWY xuma(h), JKT rumè, o.i. rumah;
*surat 'letter, written paper, book'; MIN sure?, SWY suxat, o.i. surat;
*salah 'at fault, amiss' (3.1.1.1).

3.1.1.4 PM FINAL SYLLABLE *a | _ *?/ø > MIN, SWY o, JKT è, OTHER ISOLECTS a

Examples:

*apa 'what (interrogative)'; MIN, SWY apo, JKT apè, o.i. apa;
*dua(?) 'two'; MIN, SWY duo, JKT duè, o.i. dua (cf. 5.3);
*bara? 'live coal' (3.1.1.1).

3.1.1.5 PM *ə > MIN e (|_ ?# < *t, h# < *s), o (|_ ? < *p), a ELSEWHERE

Examples of cognate sets are given in 3.1.1 (B) and (C); here follow the PM reconstructions
that can be made on the basis thereof:

*sədəp 'pleasant, tasty';   *(ŋ-)ərəm 'brood';
*kəjəm 'close the eyes';   *pərəm 'ripen fruit artificially';
*ləŋkəp 'complete';   *əŋgən 'unwilling, reluctant';
*təgəp 'strong, firm';   *diəm 'be quiet, silent; dwell';
*təkən 'press down';   *səmpət 'having sufficient time';
*rəndəm 'soak, steep';   *hantəm 'fight, beat violently';
*ənəm 'six';   *kə-lə(hø)əm 'obscure, dark; go under, wane'
*garəm 'grain' (cf. 5.7 lemma 125)   (see 4.5 for subsequent vowel contraction);
*kəbəl 'invulnerable';   *hadəp '(be) in front of, before';
*urəm 'cloudy, overcast';   *(d-)aləm 'deep, depth; in(ner)';
                                           *sələt 'narrows, intervening space'.

IRREGULAR CORRESPONDENCES: ASSIMILATION OF *ə TO THE PRECEDING VOWEL

In two cases SWY exhibits final-syllable a corresponding to a high vowel in other isolects:

SWY ən/tuat 'knee', SM, JKT lu/tut, MIN lu/tuy?, BH tuut, IBN tu(u)t (Richards 1981); cf.
BRU tuhut (Wilkinson 1959), PMP *tuhed 'knee';
SWY liax 'neck', SM leher, MIN li(h)iə,[83] JKT lèhèr; cf. PMP *liqeR 'neck'.

In these cases SWY reflects PMP *e. The other isolects assimilated PMP *e to the vowel of
the preceding syllable whenever the two vowels were separated by ø or *h. In SWY
assimilation did not take place, and PMP *e became a (following the general rule PMP *e >
a |_ C#).

There are no other examples of this retention: the only other cognate sets with a final-
syllable vowel reflecting PMP *e which was separated from the preceding vowel by ø or *h,
are cases with penultimate a like SM, BH, JKT jahat, MIN jae?, SWY jaat, IBN jai? (with
unexplained -i?, see 3.2.3) 'bad, evil' (< PMP *Zaqat/*Zaqet), and SM g/ər/aham, MIN
g/ar/am/an, SWY g/əx/m/an, IBN gaam 'molar tooth' (< PMP *Raqem; the initial consonant
in the Malayic isolects is unexplained). In both examples it is not clear whether final-syllable
a came about through assimilation, or through the general rule that PMP *e > a in all isolects
except JKT. Since PMP final-syllable *e > JKT ə if the assimilation of *e to a preceding
vowel separated by PMP *q, *h, or ø, does not apply, and since the SWY developments

---

83Thaib gives liə as well as lihiə.

show that this assimilation had not yet taken place in PM, PM final-syllable *ə must be reconstructed in the sets SM *jahat* etc. and SM *g/er/aham* etc. Hence:

*gahəm 'molar tooth';
*jahət 'bad, evil';
*lihər 'neck'; and
*tuøət 'knee' (see 3.10 for ø); (see also *tahən and *pahət, 3.9.2).

N.B. It is also possible that PM still maintained the distinction between PMP *-aq and *-eq (which was lost in the Malayic isolects, including JKT). Cases like *lu/tut, ən/tuat*, etc., and *leher, liax*, etc. as well as other lexemes deriving from proto-lexemes containing *ə adjacent to *h (cf. 3.1.3.3) show that in many positions the Malayic isolects have a strong avoidance of ə adjacent to h, which is not inherited from PM. But as there is no stronger evidence against the merger of PMP *-eq and *-aq (which is also testified in many AN languages outside the Malayic group), and it is reflected in all Malayic isolects, I presume that this merger had already taken place at the PM level.

3.1.2 THE PM HIGH VOWELS

The following correspondences of high and mid-vowels are found in the isolects:

(a) BH: In BH two high vowels, *i* and *u*, are found; mid-vowels do not occur.

(b) IBN: On the phonemic level, IBN has a pair of high vowels corresponding to BH *i* and *u*; it has also *o* and *e*, but these phonemes do not occur in the inherited vocabulary.

(c) SM: Corresponding to BH, IBN *i* and *u*, SM has respectively *i/e* and *u/o*. These high and mid-vowels are in phonemic contrast only in penultimate syllables followed by closed final syllables with an initial consonant (i.e. they are in contrast in CVCVC-structures as opposed to CVCV-, and CVV(C)-structures, see 2.1.3 and fn. 21). SM *i* and *e* reflect PMP *i*, and SM *u* and *o* reflect PMP *u* (fn. 22).

(d) MIN: Corresponding to BH, IBN *i*, SM *i/e*, MIN usually has *i* or *iə*.
   (i) It has *i* in non-final syllables, and in final syllables before -ʔ from PM *p or *t, e.g.
       sisip > sisiʔ, and *jahit > jaiʔ* (see 3.1.2.1 - 3.1.2.4) for the reconstruction of this
       and the following examples);
       -n from PM *m and *n, e.g. *kirim > kirin*, and *aŋin > aŋin*;
       -h from PM *s, e.g. *habis > abih*;
       -ø, e.g. *lagi, tiŋgi*.
   (ii) It has *iə* in final syllables before -h from PM *h, e.g. *putih > putiəh*;
       -ʔ from PM *k, e.g. *tarik > tariəʔ*;
       -ŋ, e.g. *kəniŋ > kaniəŋ*;
       -ø from PM *l and *r, e.g. *paŋgil > paŋgiə*, *lihər > li(h)iə*.

Corresponding to BH, IBN *u*, SM *u/o*, MIN usually has *u, uə*, or *uy*.
   (i) It has *u* in non-final syllables, and in final syllables before
       -n from PM *m and *n, e.g. *bəlum > (b)alun*, and *puhun > puhun*;
       -ø, e.g. *tuŋku, abu*.
   (ii) It has *uə* in final syllables before
       -h from PM *h, e.g. *jatuh > jatuəh*;
       -ʔ from PM *k, e.g. *duduk > duduəʔ*;

-ŋ, and -ø from PM *l or *r, e.g. *sambuŋ > sambuᵊŋ, *pukul > pukɩᵊ, and *təlur > taluᵊ.

(iii) it has *uy* in final syllables before

-h from PM *s, e.g. *tərus > taruyh;

-ʔ from PM *p and *t, e.g. *tutup > tutuyʔ, and *turut > turuyʔ.

N.B. (1) In some MIN isolects diphthongisation of final-syllable high vowels took place only to a very limited extent, cf. the MIN of Padang Sibusuk (Kabupaten Sawah Lunto Sijunjung, central-east of the MIN area) and Tapan (Kabupaten Pesisir Selatan, extreme south); Tapan also maintains -p and -t (Tamsin Medan 1980:70-75).

(2) The following rule ordering can be applied to the changes that MIN underwent lexeme finally:

I   merger of final labial and dental consonants in ⁺t and ⁺n if the preceding vowel was high;

II   diphthongisation of high vowels preceding final velars and liquids; diphthongisation of *u preceding ⁺t (< *-p, *-t) and ⁺s;

III merger of final stops in -ʔ; merger of final *s and *h to h; loss of final liquids.

Examples:

	I		II		III	
*bəlum		⁺bəlun		⁺bəlun		(b)alun
*puhun		⁺puhun		⁺puhun		puhun
*tutup		⁺tutut		⁺tutuyt		tutuyʔ
*turut		⁺turut		⁺turuyt		turuyʔ
*kirim		⁺kirin		⁺kirin		kirin (but cf. 4.4)
*sisip		⁺sisit		⁺sisit		sisiʔ (but cf. 4.4)
*jahit		⁺jahit		⁺jahit		jaiʔ
*duduk		⁺duduk		⁺duduᵊk		duduᵊʔ
*jatuh		⁺jatuh		⁺jatuᵊh		jatuᵊh
*hituŋ		⁺hituŋ		⁺hituᵊŋ		hituᵊŋ
*təlur		⁺təlur		⁺təluᵊr (/⁺taluᵊr?)		taluᵊ
*pukul		⁺pukul		⁺pukuᵊl		pukɩᵊ
*tarik		⁺tarik		⁺tariᵊk		tariᵊʔ
*putih		⁺putih		⁺putiᵊh		putiᵊh
*kəniŋ		⁺kəniŋ		⁺kəniᵊŋ (/⁺kaniᵊŋ?)		kaniᵊŋ
*lihər > ⁺lihir >		⁺lihir		⁺lihiᵊr		li(h)iᵊ
*paŋgil		⁺paŋgil		⁺paŋgiᵊl		paŋgiᵊ
*habis		⁺habis		⁺habis		abih
*tərus		⁺tərus		⁺teruys (/⁺taruys?)		taruyh

MIN also has mid-vowels in the penultimate syllable. In a few cases these mid-vowels correspond to PMP high vowels. Lexemes containing them may be loans from SM or other languages, since they are relatively few in number, and some of them are of marginal use (cf. Van der Toorn 1891:passim), for example,

goreŋ 'fry'; the usual MIN term for frying is sala or saŋla (4.6); the usual term in SM is goreŋ;

merah 'red'; the usual MIN term is sirah, whereas in SM it is merah;

potoŋ 'cut'; the usual term is kareʔ (in SM it is potoŋ);

tembo? 'wall'; this term is rarely used; SM also has *tembok*, "...which does not occur in old books" (Wilkinson 1959).

Some of these lexemes have a doublet with high vowels, cf. *teñju ~ tiñju* 'fist', *goso? ~ gusuᵊ?* 'rub' (cf. 3.4.1.4 UIC), *gele ~ giliᵊ* 'turn, turning', *lompe? ~ lumpe?* 'jump' (according to Van der Toorn (1899:XII), the dialect of Koto Gadang often lowers original high vowels in the penultimate syllable (cf. *teñju, lompe?*). Most of the MIN mid-vowels correspond to SM mid-vowels, whereas the contrary (i.e. SM mid-vowels corresponding to MIN ones) is often not the case.

(e) SWY: SWY has *i, u,* and *o* corresponding to high and mid-vowels in the other isolects. In final syllables, *i* and *u* were diphthongised to *iᵊ* and *uᵊ* respectively when preceding -? or -(h), e.g. *paŋgil, bibix, kəniŋ, kirim/kixim, jait,* and *bini,* maintained final-syllable *i,* but in *bəniᵊ(h), boliᵊ(h), tariᵊ?,* it was diphthongised to *iᵊ*; furthermore *tuxut, tutup, təlux, puun, tuluŋ, əmbun,* etc. vs *bunuᵊ(h), jatuᵊ(h), duduᵊ?* (see 3.1.2.1 - 3.1.2.4 for the meaning of these lexemes, and their PM ancestral forms). *o* is rare, and does not occur in final syllables; e.g. *obat, oxaŋ* (*uxaŋ* in Aliana et al.), *potiᵊ(h)* (see 3.1.2.2).

(f) JKT: JKT has *i, é, è,* and *u, ó, ò.*

é occurs word finally and in penultimate syllables preceding final *é* or *ó,* and *è* occurs in non-final position (lexeme-final *é* reflects PM *a). *é* is sometimes in free variation with *i*; it does not occur very often; e.g. *lagi ~ lagé* 'again'.

ó is often in free variation with *u,* e.g. *anu ~ anó* 'so-and-so'; *dapur ~ dapór* 'kitchen'. It does not occur very often (cf. fn. 57), and I found only four instances where it contrasts with *ò*.[84] *é* and *ó* often alternate with *è* and *ò* in variants with final -? (e.g. *légó ~ lègò?* 'sell'; *déŋké ~ dèŋkè?* 'stand on tiptoe').

As in SM, there is no conditioning factor for the split of PMP *i* and *u* into high and mid-vowels. There is a tendency to height harmony of high and mid-vowels in adjacent syllables, and often (but far from always) a mid-vowel is found immediately following an *a,* e.g. *tòlòŋ, nènè?, lèhèr, pò?òn/pu?un/puhun* (see 3.1.2.1 - 3.1.2.4) and *taón* 'year' (< *tahun,* 3.4.1.2), *aós ~ aus* 'thirst(y)' (< *haus,* 3.8.2). But in many cases there is no conditioning factor, as is shown in the cognate sets in 3.1.2.1 - 3.1.2.4.

DISCUSSION

From the above description it appears that BH and IBN have a mutually agreeing pair of high vowels to which each of the other isolects have different correspondences. Since the high vowels, diphthongs, and mid-vowels in the other isolects all reflect PMP *i* and *u,* they must be the result of a split.

The dipthongisation of final-syllable high vowels in MIN and SWY can easily be explained as the result of a split which was conditioned by the following consonant (or, in MIN, by the historical antecedent of this consonant). It did not take place (or took place to a much smaller extent) in some other MIN and Middle Malay isolects (cf. the MIN of Tapan, Tamsin Medan 1980:70-75; and Semende (Middle Malay), Yuslizal Saleh et al. 1979:22). This phenomenon (which was realised in a different way for MIN and SWY) must be distinguished from the split of PMP *i* and *u* into high and mid-vowels in SM and JKT

---

[84]These four cases are: (1) *cəpò* 'without money', and *cəpó?* 'easy to break or tear'; (2) *ŋulòn* 'go west' (< JV), and *ŋulón* 'weep, cry a lot'; (3) *pòl* 'place to store s.th. in great number' (< ENG 'pool'?), and *pól* 'full' (< DU 'vol'); (4) *bòt,* in *səpatu bòt* (< ENG) 'boot', and *bót* 'bread' (< DU 'brood'?).

(and, to a lesser extent, in MIN and SWY). For the latter no comprehensive conditioning factor was found.

The question now is whether the split into high and mid-vowels took place before the divergence of the isolects, whereupon the secondary vowels merged again to *i* and *u* in BH and IBN, or whether PM still maintained the PMP high vowels, and the split occurred after the isolects began to diverge. In the latter case two developments are possible: the split may have taken place in each isolect separately, and BH and IBN did not follow, or BH and IBN may have been the first isolects to branch off from PM while maintaining the latter's high vowels, whereas the other isolects continued to have a common development, during which the split of PM *i* and *u* in high and mid-vowels took place. The answer to this question must be that PM, like PMP, had only a pair of high vowels, which were maintained in BH and IBN, and that the vowel split took place separately in each of the other isolects. This analysis is based on two considerations.

Firstly, the split into high and mid-vowels took place in each isolect in a different way. In SM and JKT mid-vowels occur very frequently in final as well as in penultimate syllables, whereas in SWY their existence is restricted to penultimate *o*. In MIN, lexemes with mid-vowels (from PMP *i* and *u*) are not frequent and must be loanwords from SM. These differences make it probable that vowel lowering was not part of the common history of these isolects. The lexical distribution of high and mid-vowels is also different for each isolect, as is shown in the following table.

TABLE 4: LEXICAL DISTRIBUTION OF HIGH AND MID-VOWELS

SM	MIN	SWY	JKT	meaning
oraŋ	*uraŋ*	*oxaŋ/uxaŋ*	*òraŋ*	human being
(BM) *ubat,*				
(BI) *obat*	*ubeʔ*	*obat*	*òbat*	medicine, drug
jatuh	*jatuᵊh*	*jatuᵊ(h)*	*jatò*	fall (v)
toloŋ/tuluŋ	*toloŋ*	*tuluŋ*	*tòlòŋ*	help, assistance
təlor/təlur	*taluᵊ*	*təlux*	*təlòr*	egg
lompat/lumpat	*lompeʔ/lumpeʔ*	*lumpat*	*lòmpat*	jump
putih	*putiᵊh*	*potiᵊ(h)*	*puti*	white
nene/k	*niniᵊ/ʔ*	---	*nènè/ʔ*	grandparent
tiŋgi	*tiŋgi,teŋgi*	*tiŋgi*	*tiŋgi*	high
pohon	*puhun*	*puun*	*pòʔòn/puʔun*	tree
			puhun	tree

Secondly, Prentice and Hakim Usman (1978:134) pointed out that BRU and KCI only reflect high vowels for PMP penultimate *i* and *u*.[85] Besides these two, Collins (1986b:184) found three other Malayic isolects with corresponding high vowels only, viz. Ulu Trengganu, Urak Lawoi', and Bacan. He goes on to say that it would not be unthinkable if PM already had a split into high and mid-vowels, and that an isolect or a group of contiguous isolects underwent a subsequent merger of these high and mid-vowels. But he finds it doubtful that this process could have taken place independently in five isolects which

---

[85]BRU, as a matter of fact, also retains PMP *i* and *u* in final syllables. KCI vowels in the final syllable underwent changes which altered the original PMP high vowels considerably, but no split into high and mid-vowels was involved.

were not in close contact with each other. To these five isolects can be added BH[86] and Bahasa Semende,[87] which maintain the PMP high vowels in the final as well as in the penultimate syllable, and IBN, which on the phonemic level only reflects *i* and *u* for PMP *i* and *u*. Furthermore, the OM of the inscriptions of Srivijaya, which date from the seventh century AD and are the oldest record of any form of Malay, only has four vowel phonemes viz. *a*, *ə*,[88] *i*, and *u*, the use of mid-vowels being restricted to SKT loanwords. But in the OM inscription of Kedu (Java) which dates from the ninth century, the inherited word *sa/popo* '(first degree relative in collateral line)' is found instead of an expected +*sə/pupu* (Teeuw 1959:146). Collins (1986b:186) concludes from this that the mid-vowels emerged in coastal Malay isolects spoken in areas where Malay was not autochthonous (particularly in Java) and in an era (ninth century AD) when borrowing and spread of linguistic features began to play an important role in coastal Malay. (Quoting Nothofer (1975), he mentions that the split of PMP *i* and *u* into high and mid-vowels is also seen in JV and Madurese.) In other words, there are isolects other than BH that did not undergo (or partly underwent) the split of PMP *i* and *u* into high and mid-vowels, and there are some indications that the origin of the split must be sought in areas where Malayic was not autochthonous. As a conclusion, I reconstruct only a series of high vowels (*i* and *u*) along with PM *a* and *ə*. I reconstruct PM penultimate *i* on the basis of BH, IBN, SWY *i*, SM MIN *i/e*, JKT *i/é/è*, and PM penultimate *u* on the basis of BH, IBN *u*, SM, MIN, SWY *u/o*, JKT *u/ó/ò*. I reconstruct PM final-syllable *i* on the basis of BH, IBN *i*, SM *i/e*, MIN *i/iʔ/e*, SWY *i/iʔ*, JKT *i/é/è*, and PM final-syllable *u* on the basis of BH, IBN *u*, SM *u/o*, MIN *u/uʔ/uy/o*, SWY *u/uʔ*, JKT *u/ó/ò*.

3.1.2.1 PM PENULTIMATE *i* > SM, MIN *i/e*, JKT *i/é/è*, OTHER ISOLECTS *i*

Examples:

hituŋ 'count (v)'; SM, BH *hituŋ*, MIN *etoŋ*, (Van der Toorn 1891) *hituʔŋ*, o.i. *ituŋ*;
lihər 'neck' (3.1.1.5IC);
niniʔ 'grandparent, ancestor'; SM *nene/k*, MIN *niniʔ/ʔ*, *ñiñiʔ/ʔ*, BH *nini*, IBN *iniʔ*, JKT *nènè/ʔ* (cf. 3.4.2 for -*ʔ*); MIN has also *nene/ŋ* in *si-nene/ŋ uraŋ* 'a well-to-do person, s.o. of rank' (cf. 5.4 for SM -*/k*, MIN -*/ʔ* / -*/ŋ*, JKT -*/ʔ*, and for the loss of *n-* in IBN);
sisip 'insert'; SM *sisip* 'insert', MIN *sisiʔ* 'add' (but cf. 4.4b);
bini 'wife'; MIN *biñi* (cf. 3.6.1.2 for *ñ*), o.i. *bini*;
kirim 'send (s.th.)'; SM, BH, JKT *kirim*, MIN *kirin*, SWY *kirim*, *kixim*, IBN *kirum* (cf. 4.4a).

3.1.2.2 PM PENULTIMATE *u* > BH, IBN *u*, JKT *u/ó/ò*, OTHER ISOLECTS *u/o*

Examples:

bulan 'moon; month'; a.i. *bulan*;

---

[86]Bahasa Banjar Bukit also retains PMP *i* and *u*, but this isolect is too closely related to BH (Abdurachman Ismail et al. 1979:7) to be considered as separate evidence.

[87]It is in fact indicative of the relative lateness of the SWY changes, that of the two other Middle Malay isolects for which there are data available (viz. BSM and Semende), neither diphthongises *i* and *u* before -ʔ and -*h*, nor does Semende have a split of penultimate *u* into *u* and *o*.

[88]A schwa as such is not represented in these inscriptions, but it can be inferred from the non-occurrence of a vowel symbol, or from the doubling of the following consonant grapheme. In a number of cases it is not distinguished from short *a* (Vikør 1988:71).

*ubat 'medicine, drug'; SM, SWY obat, MIN ube?, BH, IBN ubat, JKT òbat (< P(W)MP
*ubaj 'id.'; Dempwolff reconstructed *ubat and Zorc (1971) reconstructed Proto Philippine
*ubad, but cf. SUN ubar, Balinese ubad 'id.' which points to PMP *-j);
*huban 'grey or white (hair)'; SM (h)uban, MIN, JKT uban, BH huban, SWY oban 'id.',
IBN b/uban 'white-haired', ban/an 'prematurely white-haired' (with apocope: cf. 3.11c);
*buaŋ 'throw away'; IBN buay (cf. 3.2.3 for -y), o.i. buaŋ;
*putih 'white'; MIN puti²h, SWY poti²(h), JKT puti, o.i. putih;
*tuluŋ 'help, assistance'; SM toloŋ, tuluŋ, MIN toloŋ, o.i. tuluŋ.[89]

3.1.2.3 PM FINAL-SYLLABLE *i > BH, IBN i, SM i/e, MIN i/i²/e, SWY i/i², JKT i/é/è

Examples:

*habis 'all, entirely; used up, done with, finished off'; SWY, JKT abis 'done with, all used
up, finished off', MIN abih 'finished off, used up, completed; all' (Van der Toorn 1891),
IBN abis 1. 'all, the whole of, entirely' 2. 'used up, finished' (cf. 5.7 lemma 191);
*tiŋgi 'high'; MIN tiŋgi, teŋgi, IBN tiŋgi², o.i. tiŋgi;
*bəli 'buy'; SM, SWY, IBN, JKT bəli, MIN bali;
*bənih 'seed (for planting)'; SM, IBN bənih, MIN bani²h, bañi²h, SWY bəni²(h) 'id.', BH
banih 'rice plant';
*tarik 'pull'; MIN tari²? 'take (away), pick', IBN tarik 'pull taut', SWY tari²?, o.i. tarik
'pull';
*paŋgil 'call, summon, invite'; SM, SWY, IBN, JKT paŋgil, MIN paŋgi²;
*bibir 'lip, rim'; MIN bibi², SWY bibix, o.i. bibir;
*jahit 'sew, stitch'; SM, BH jahit, MIN jai?, o.i. jait;
*lagi? 'again, later on'; SM, BH, SWY lagi, MIN lagi, lage, JKT lagi, lagé 'again', IBN lagi?
'later on';
*pəlir 'testicle'; SM, IBN pəlir, BH palir, SWY pəlix, JKT pəlèr;
*aŋin 'wind (n)'; SM, MIN, BH, SWY, JKT aŋin;
*kəniŋ 'eyebrow'; SM kəniŋ, MIN kani²ŋ, kañi²ŋ, (3.6.1.2), BH kaniŋ 'eyebrow', o.i.
kəniŋ 'forehead' (the PM reconstruction for 'forehead' is *dahi, cf 3.4.2.6).

3.1.2.4 PM FINAL-SYLLABLE *u > BH, IBN u, SM u/o, MIN u/u²/uy/o, SWY u/u², JKT u/ó/ò

Examples:

*bəlum 'not yet'; SM bəlum, MIN (b)alun, BH balum, SWY (bə)lum, JKT bəlòm, bəlòn;
*bubur '(generic for) gruels'; SWY bubux, MIN bubu², o.i. bubur;
*bunuh 'kill'; MIN bunu²h, SWY bunu²(h), JKT bunu, o.i. bunuh;
*əmbun 'dew'; SM, SWY, JKT əmbun, MIN, BH ambun 'dew', IBN ŋ-əmbun 'expose to
the dew';
*habu 'ash, dust'; SM (h)abu, BH habu, o.i. abu;
*jatuh 'fall (v)'; SM jatuh, MIN jatu²h, JKT jatò 'fall', BH jatuh 'become', SWY jatu²(h) 'lie
at anchor', IBN jatuh 1. 'fail (in business)' 2. 'besmirched, degraded (name)';

---

[89] According to Wolff (1976:367) toloŋ is a Chinese loan, cf. CHI tō lôn 'patronise, help a man on'.

*kəluh 'sigh, complain'; SM kəluh, MIN kaluᵊh, JKT kəlu 'id.', SWY kaluᵊ(h) 'gasp for breath, recover breath';

*mulut 'lips, mouth'; SM, SWY, JKT mulut, MIN muluyʔ, IBN mulut 'lips, mouth';

*puhun 'stem; origin, basis; ask, beg (forgiveness)';⁹⁰ SM pohon, SWY puun, JKT puhun, pòʔòn, puʔun 'tree', SM mohon, JKT muhun 'ask, beg for', MIN puhun 1. 'origin, reason; the west' 2. 'beg, request', BH puhun 1. 'tree' 2. 'host, owner of the house', IBN puun 'origin, basis, start, stem';

*pukul 'strike, hit, knock'; MIN pukuᵊ, o.i. pukul;

*sambuŋ 'extend; extension piece'; MIN sambuᵊŋ, o.i. sambuŗ;

*təgur 'address; rebuke (v)'; SM təgur, MIN taguᵊ, BH tagur, JKT təgòr 'greet, address', IBN təgur 'chide, rebuke (usually children)';

*tərus 'right through, in a direct line through'; MIN taruyh, BH tarus, SWY təxus, o.i. tərus;

*tuŋkuʔ 'hearthstones, tripod'; SWY toŋku, IBN tuŋkuʔ, SM, MIN, BH tuŋku '(three) hearthstones, used for cooking';

*turut 'follow, obey'; SM, BH, JKT turut, MIN turuyʔ, SWY tuxut 'be in accordance with, follow', IBN turut 'offer no resistance';

*tutup 'close up, cover; lid, covering piece'; MIN tutuyʔ, o.i. tutup.

3.1.2.5 THE ORIGIN OF (SM) MID-VOWELS THROUGH CONTRACTION OF *a + *i/*u OR *i/*u + *a

In SM (and possibly in other isolects too) some mid-vowels originated from contraction of earlier high vowels with *a. There are seven examples of this:

(1) PMP *(ma-)iRaq 'red' > SM, MIN m/erah, SWY m/ira(h), IBN mirah, JKT m/èrè;

(2) PMP *Rahut(-an) 'whittle, split wood' > SM rot/an,⁹¹ MIN, JKT ròt/an, IBN rut/an 'rattan';

(3) PMP *(maR-)uliq 'be able to' > SM b/oleh, MIN b/uliᵊh, SWY b/oliᵊ(h), JKT b/ólé, b/òlè 'be allowed, obtain', IBN b-ulih 'get, obtain'; cf. also SM (məm)pər-oleh, MIN (mam)pa-uliᵊh, BH ba-ulih, SWY x/oliə(h) (x must be the result of a backformation from PM *(mb)Ar-ulih, see below), IBN ŋ-ulih 'get, obtain', and SM oleh/oleh 'present brought back from a journey'; (it is, however, also possible that o (as with e) originated through lowering instead of contraction);

(4) PMP *ba-isa-n 'tie between parents-in-law of a married couple' > SM besan, JKT bèsan 'id.', MIN bisan 'men whose wives are sisters', SWY bisan 'term of address to parents of child-in-law', IBN isan 'relatives of child-in-law' (3.5.1 UIC);

(5) PMP *saqup 'help, assist with work' (Blust 1980a) + *an > SM, MIN sop/an, JKT sòp/an 'showing respect (through courtesy, modesty or timidity)', BH sup/an 'shy, ashamed';

---

⁹⁰The same range of meanings is found with JV wit, cf. wit 'tree; origin, stem', wiwit 'begin, beginning', amit 'sorry! (in passing)', wét/an 'west'; see also 4.5 for the relation between *puhun and *ampun.
⁹¹According to Dempwolff, rot/an is originally JV (see Dempwolff 1938: *raut).

(6) BH ba/hira[92] 'defecate' vs SM *b/era/k,* JKT *b/èra/k* (*-/k* unexplained cf. 3.4.2.4 UIC), IBN *b/ira?* 'defecate'; cf. also BRU *baria/k* (with metathesis) 'defecate';

(7) BH *kam/ari/an* 'evening' vs SM *kəm/are/n, kəm/ari/n, kəlm/ari/n, kalam/are/n,* JKT *kəm/arè/n* 'yesterday' (cf. also 5.1.3).

Assuming that lowering of high vowels took place in each isolect independently, this contraction must have happened after the split of the six isolects. But it is more likely that the correspondences with mid-vowels in the non-SM isolects are SM loans. A reflex of PMP *(ma-)iRaq* is found in each isolect (except BH), but it is the usual term for 'red' in only two of them (SM and JKT): BH and SWY have *abaŋ,* IBN has *mansaw,* and MIN has *sirah;* also, SWY *m/ira(h)* has an apical trill, which indicates borrowing (see 3.7). SM, MIN *rot/an,* JKT *ròt/an,* IBN *rut/an* replaced an earlier PMP *quey,* which is still reflected in SWY and IBN *ui* 'rattan'. It is not likely that the same semantic shift took place separately in SM, MIN, IBN and JKT. SM *b/oleh* seems to be the result of contraction of *A and *u after the loss of *r on the morpheme boundary in *(mb)Ar-ulih.* But the other isolects (with the possible exception of JKT *b/ólé /b/òlè,* if this is not a loanword) do not reflect such a development: only SWY has a mid-vowel, but it also has a doublet that still reflects the *r of the PM prefix *(mb)Ar-* (cf. 6.1.1). JKT *b/èra/k* has the same unexplained final *k* as SM *b/era/k,* which points to borrowing from SM. In the above cases, most of the correspondences with mid-vowels are found in JKT, which borrowed heavily from SM.[93]

On the basis of these sets I reconstruct:

ma-irah 'red' (cf. also 5.7 lemma 149);
(mb)Ar-ulih 'get, obtain';
ba/isa(a)n 'tie between parents-in-law of a married couple';
sahup-an 'respectful';
ba/hira? 'defecate';
*kə-lə(hø)əm *hari* 'evening' (3.1.1.5; cf. also 5.1.3).

On the basis of SM *rot/an* no reconstruction is made; however, *raut* 'scrape with a knife (bamboo)' can be reconstructed on the basis of MIN *rauy?,* SWY *raut,* o.i. *raut* (cf. PMP *raut* 'id.').

3.1.3 VOWEL CHANGES IN ANTEPENULTIMATE SYLLABLES

In Malayic isolects the antepenultimate vowel is usually a schwa unless it is followed by *h* or by a vowel (as in SM *buaya* 'crocodile', *piutaŋ* 'debt', *biuku* 'turtle sp.', *kuilu* 'rabbit' (< POR)). This does not apply to MIN and BH, where antepenultimate high vowels often occur (and where *ə > a). The schwa in the other isolects must be the result of antepenultimate neutralisation. MIN and BH only partly avoided this phenomenon: PM *i and *u were usually retained, but in some cases there is disagreement between MIN, BH, and PMP (see below). MIN and BH may also have doublets with an antepenultimate

---

[92] Written 'babira' in Abdul Jebar.
[93] Waruno Mahdi (pers.comm.) explains the vowel contraction in the last syllable of SM *kəm/are/n* (etc.), JKT *kəm/arè/n* as a Bazaar Malay development. Only later on *kəm/are/n* became accepted in SM.

(neutralised) *a*. BH sometimes shows a high vowel which must have originated through assimilation with the following consonant. There are several reasons for regarding the MIN and BH antepenultimate high vowels as retentions, and for assuming that subsequent neutralisation of these vowels took place in the other isolects after they split from PM. Firstly, MIN and BH were not in close contact, and it is unlikely that their agreement is due to influence of one of these isolects on the other. Furthermore, the other isolects maintain antepenultimate high vowels in specific environments viz. before a vowel (usually *a*) or (in SM) *h*. IBN has ø corresponding to these high vowels.

Finally, the MIN and BH antepenultimate high vowels (usually) agree with high vowels in PMP or PMJ. Antepenultimate neutralisation is in agreement with the phonotactic tendency to disyllabicity in Austronesian languages (and in Malayic isolects in particular). This tendency is shown by the great majority of disyllabic lexemes in Austronesian languages and in lexical reconstructions for PMP, and by processes of apocope, syncope, and antepenultimate neutralisation in originally trisyllabic (or tetrasyllabic) lexemes.

The factors involving antepenultimate vowels are categorised as follows:

– those involving high vowels before consonants other than *h* (treated in 3.1.3.1),
– those involving high vowels before *h* or a vowel (see 3.1.3.2),
– those involving other vowels (indeterminate *a* or *ə*, which is represented as *A; see 3.1.3.3).

3.1.3.1 PM *i, *u > MIN, BH i, u, OTHER ISOLECTS ə │ C _ CV(C)V(C)

Examples:

*kuliliŋ 'go or turn around'; MIN kuliliᵊŋ, BH kuliliŋ, o.i. kəliliŋ;
*sumaŋət 'spirit (of a living being), soul, inspiration'; SM səmaŋat, MIN sumaŋeʔ, BH sumaŋat 'spirit, inspiration', SWY səmaŋay (-y unexplained), IBN səməŋat (penultimate ə unexplained; Mualang has səmaŋat) 'spirit, soul', JKT sumaŋet (a SUN loan?), səmaŋat (a SM loan?) 'inspiration' (cf. PMP *sumaŋed);
*surambiʔ 'eaves of a house'; SM sərambi, MIN, BH surambi 'front verandah', IBN serambiʔ 'a shed at the back of a house';
*b/in/antu 'child-in-law'; SM, IBN mənantu, MIN minantu, binantu, BH minantu, SWY nantu (with loss of antepenultimate syllable) (cf. PMP *b/in/antu 'son-in-law'; see 3.6.1.6 for *b > m);
*biruaŋ 'bear'; SM, IBN bəruaŋ, MIN, BH biruaŋ, SWY bəxuaŋ, JKT biruaŋ (< SUN?) (but cf. PMP *be(rR)uaŋ, Prentice (1974:44);
*tiŋadah 'look upwards'; SM təŋadah, MIN, BH tiŋadah, JKT təŋadè;
*tiŋaluŋ 'civet cat'; SM təŋaluŋ, MIN tiŋaluᵊŋ, BH tiŋaluŋ;
*tilañjaŋ 'naked'; SM, SWY təlañjaŋ, MIN tilañjaŋ, talañjaŋ, BH tilañjaŋ, IBN təlañjay (see 3.2.3 for IBN -ay);
*subəraŋ 'opposite bank or side'; SM, SWY JKT səbəraŋ, MIN, BH subaraŋ, IBN səbəray (see 3.2.3 for IBN -ay).

N.B. Sometimes antepenultimate neutralisation (to *a*) has also affected MIN or BH high vowels; as a result MIN can have an antepenultimate high vowel where BH has *a* or vice versa. In such cases I will reconstruct a PM high vowel, unless evidence from other isolects or other AN languages provides a more convincing explanation. For example:

*kulambu 'mosquito net'; SM, SWY, JKT kəlambu, MIN kulambu, kalambu, BH kalambu;
*tulad/an 'example, sample'; SM tuladan, təladan, MIN tuladan, BH taladan, SWY, IBN, JKT təladan;
*pilanduk 'mousedeer'; SM pəlanduk, MIN palandu³ʔ, BH pilanduk, IBN pəlanduk.

In a number of cases BH has a u which does not reflect PMP *u. This u possibly originated through colouring by a following (labial) consonant, for example, gumalan 'gamelan' (< JV), and jumbat/an (also jambat/an) 'balustraded gangway', SM jəmbat/an (also jambat/an in Wilkinson 1959), MIN jambat/an, JKT jəmbat/an 'id.'. According to Wilkinson, SM jambat/an, jəmbat/an comes from +jambat 'hold in the hand' + +an. Compare also BH kubaya 'k.o. dress', and SM kəbaya, MIN kabayo 'id.', which comes from AR qaba:ya (plural of qabaʔ 'k.o. tunic') according to Von de Wall. In one case the same assimilation seems also to have been at work in MIN, viz. SM səmbahyaŋ, MIN sumbayaŋ, sambayaŋ, BH sumbahyaŋ, sambahyaŋ, IBN səmbiaŋ, JKT səmbayaŋ 'prayer; pray', SWY səmba(h)yaŋ [sic] 'obligatory ritual prayer'. For SM jambat/an etc. and SM səmbah/yaŋ etc., I reconstruct:

*jambat-an 'balustraded gangway, jetty, bridge with handrail';
*səmbah-*hiaŋ 'pray to (the) god(s)'.

If both MIN and BH have antepenultimate a where PMP/PMJ show a high vowel, I will reconstruct *(A,i) or *(A,u), as it is not clear whether antepenultimate neutralisation took place on the PM level or later on, when MIN and BH were separate isolects; e.g.

*t(A,i)ŋgiliŋ 'anteater'; SM, SWY, JKT təŋgiliŋ, MIN taŋgili³ŋ, BH taŋgiliŋ, IBN təŋiliŋ (3.11); cf. PMJ *tiŋgiliŋ 'anteater'.

UNEXPLAINED IRREGULAR CORRESPONDENCE

All isolects (except IBN, which does not have a reflex) maintain antepenultimate i in SM, BH, SWY, JKT binataŋ, MIN binataŋ, minataŋ (3.6.1.6). No reconstruction is made for this correspondence set, which is most likely not inherited from PM.[94]

3.1.3.2 PM *i, *u > IBN ø, OTHER ISOLECTS i, u ∣ C _ (h)VCV(C)

Examples:

*biawak 'monitor lizard'; SM, BH biawak, MIN, SWY biawaʔ, IBN bayak (with possible metathesis of +w with a (non-phonemic) +[y]- glide, and subsequent loss of +w), JKT biñawak (nasalisation of +y unexplained);
*buhaya 'crocodile'; SM buaya, MIN, SWY buayo, BH buhaya (Fudiat Suryadikara et al. 1981:128), IBN baya, JKT buayè;

---

[94]PMP *bi(nN)a(ŋ)(Ct)aŋ 'beast, animal' (Blust 1970:119) is problematic. Jack Prentice (pers.comm.) drew my attention to the fact that many Austronesian languages lack a general term for 'animal'. As far as they have such a term, it is not amenable to the reconstruction of a PAN or PMP etymon. Brandes' (1884:175) connection between SM binataŋ, Ibanag bataŋ 'trap-net' and SUN pamataŋ 's.o. who hunts deer on horseback' is not entirely convincing.

*kuali 'cooking pot'; SM, MIN, JKT kuali, BH kuantan,[95] IBN kali;
*muhara(?) 'river mouth'; SM muara, MIN muaro, BH muhara, SWY muaro, JKT muarè;
*si-apa 'who'; SM, BH siapa, MIN, SWY siapo, IBN sapa, JKT siapè (cf. 5.5.2.4);

N.B. Other examples of *V > IBN ø | C _ (h)VCV(C):

pin/an 'plaything' < +pəin+an < +pain+an < +main 'play' + +an (6.5.5);
tik/i? 'climb, go up' < +təik+i? < +taik+i? < *naik + *i (3.6.1.2);
d/ulu? 'preceding, before, first' < *di *hulu(?) (3.1.3.3IC); cf. also di + udi, which becomes
d-udi 'following, after, behind', and di + atas, which becomes d-atas 'on, above, over';
k/ili? '(in the direction) downstream' < +kə+ +ili? < *kə + *hilir (3.7.4IC);
patu 'orphaned, without relations', cf. SM, MIN, BH, JKT piatu;
pasa 'fast (n)', cf. SM, BH puasa, MIN, SWY puaso, JKT puasè (< SKT).

3.1.3.3 PM ANTEPENULTIMATE *A FOR INDETERMINATE *a OR *ə

In a number of cases it is difficult to determine whether PM had antepenultimate *a or *ə, since in MIN and BH *a and *ə both became a, and in the other isolects all antepenultimate vowels as a rule became ə. The solution depends on a more general question, that is, whether PM had more than one non-high vowel in this position, and if not, which of *a or *ə occurred. This question has not yet been dealt with for PMP, and many scholars silently take it for granted that on this level only antepenultimate *a occurred (along with *i and *u).

Consequently, evidence for PM antepenultimate *a and/or *ə must be sought within the Malayic group, and this evidence does not yield a satisfactory solution. Some indication that PMP *a (with or without preceding laryngeal) became (pre-)PM *ə is given by the fact that such an *a was lost in all isolects, cf.

PMP *habaRat 'north-west monsoon' > SM barat 'west' (5.2.1);
PMP *q(ai)teluR 'egg' > *təlur (5.7 lemma 98);
PMP *qanibuŋ 'k.o. palm tree' > SM, SWY, IBN nibuŋ, MIN nibuʔəŋ (cf. PM *nibuŋ, 3.6.1.2);
PMP *hapejiq 'smarting, stinging pain' > SM pədih, MIN padiʔəh, BH padih, SWY pədiʔ(h), IBN pədis, bə-pədi? (with unexplained -?);
PMP *qanilaw 'k.o. tree: Grewia spp.' (Blust 1984a) > SM nilaw 'id.', SWY nilaw 'k.o. wood'.

---

[95] Kuantan must be a (originally high speech?) style variant. In a number of cases some (or all) isolects show a variant form that was created through replacement of an original final -CV sequence by another ending. This new ending is usually -/ntan, but -/ntaŋ and -/ntiŋ also occur. For example:
BH cuntan 'steal' < +curi (a North Indian loanword, cf. 3.4.1;
BH kuantan 'cooking pot' < *kuali (3.1.3.2);
Perak, Kedah antan 'pestle' < *halu (3.9.1);
BH kamintiŋ 'candlenut', SM kəmiri, kəmbiri, JKT kəmiri;
SM, MIN, BH, IBN santan, SWY santən (a JV loan?), JKT santən 'coconut milk', also SM sari 'flower; flower-like; pollen, quintessence' (< Old Javanese), JKT sari 'stamen';
SWY pantan 'restrain oneself from s.th. prohibited (because harmful)', o.i. pantaŋ 'taboo, thing not done, prohibition due to custom or superstition', and SM pəmali, MIN (rare) pamali 'id.'.
Sometimes the original form is no longer found in the isolects, but it can be traced through comparison with other languages, cf.
SM, MIN, IBN (Richards 1981; not usual), JKT jantan 'male, masculine', and SUN jalu, JV jalu (< SM?), Karo dalu/na 'masculine', Toba, Dairi dalu, JV dalo/n 'boar' yielding PWMP *Zalu 'male' (cf. Proto South-Sulawesi *jalu 'mate (animals)', Mills 1975 vol.2).
Brandes (1884:88) credits Van der Tuuk for explaining SM añjiŋ as a high speech variant of PMP *asu 'dog'. But this is doubtful, since the difference between *asu and añjiŋ does not fit in well with the formal variant pattern outlined above.

Most probably, loss of PMP *a took place after an intermediate stage in which it was weakened to ⁺ə.

Another indication that PM had an antepenultimate *ə is the cognate set SM (h)əmpədu, MIN ampadu, BH hampadu, SWY, IBN əmpədu 'gall bladder'. SM (2.1.3) and JKT (2.6.3) show a strong avoidance of initial schwa followed by a stop, and in SWY (2.4.3) and IBN (2.5.3) this sequence does not occur at all. But schwa followed by a cluster of a homorganic nasal + stop occurs regularly in these isolects, and it seems that inherited lexemes which in PMP had initial (laryngeal +) schwa + stop acquired a homorganic nasal preceding this stop. This is seen in MIN ampe?, BH ampat, other isolects əmpat 'four' (3.4.1.1), which came from PMP *hepat (7.2.5). The nasal in this set must be reconstructed for PM, because MIN and BH have it as well. So, the same avoidance of sequences of initial (laryngeal +) schwa + stop seems to have occurred in PM, and this possibly explains the nasal which is found in the cognate set SM (h)əmpədu etc., including its MIN and BH members. In that case *həmpədu must be reconstructed, but lack of other analogical cases makes this reconstruction premature.   Since it is not possible to determine with certainty whether antepenultimate *ə and/or *a occurred, I will reconstruct PM *A for the correspondence MIN, BH a, other isolects ə in antepenultimate position.

Examples:

*bAlaŋa? 'earthenware vessel'; SM bəlaŋa, MIN balaŋo, BH balaŋa, SWY bəlaŋo, IBN bəlaŋa?, JKT bəlaŋè;
*hAmpədu 'gall bladder' (see above);
*jArami? 'rice stalk, straw'; SM, SWY, JKT jərami, MIN jarami, IBN jərami?;
*tAliŋa(?) 'ear'; SM təliŋa, MIN taliŋo, BH taliŋa, SWY təliŋo.

In a number of cases SM has an antepenultimate a adjacent to h. This a does not necessarily reflect PM *a, since PM *ə also occurred adjacent to *h (cf. 3.1.1.5). But since SM ə adjacent to h is avoided (except in the sequence (h)əN- + stop (cf. 2.1.3), PM *ə may have become SM a in this environment. I will reconstruct *A for this SM a (/MIN, BH a, SWY ø(/a), IBN ø).

Examples:

*hArimaw 'tiger'; SM harimaw, rimaw, BH harimaw, MIN arimaw 'tiger', SWY ximaw (with loss of *ə (if *A was a schwa), or (if *A was an *a) with neutralisation of *a to schwa and subsequent loss, cf. also 2.4.3) 'wild cat'; IBN has rimaw 'tiger', but since tigers are not found in Borneo this is probably borrowed (from SM?);
*hAlu-an 'bows, forepart of a vessel'; SM, BH halu/an, MIN, SWY alu/an (with unexpected retention of a-; < SM?), IBN lu/an (3.11c);
*hAli'lipan 'centipede'; SM halipan (with haplology of second syllable), (li)lipan, MIN (a)li'lipan, BH halilipan, SWY (li)lipan, IBN ləlipan, ñəlipan;[96]
*bAharu 'new'; SM, MIN baharu, baru, IBN, JKT baru.

IRREGULAR CORRESPONDENCES

In the cognate set SM dahulu, dulu, MIN daulu, dulu, BH dahulu, SWY, JKT dulu, IBN dulu? 'before, first, ahead', SM, BH a preceding h corresponds to MIN a/ø, other isolects ø. From seemingly analogical cases like *bAharu and *hArimaw it would follow that ⁺dAhulu? be reconstructed. But it is more likely that SM dahulu etc. developed from *di *hulu(?).

---

[96]cf. Bruggeman's wordlist of IBN as spoken in Sarawak's Third Division (in an appendix to Scott 1956).

This assumption is supported by evidence from outside the Malayic group, e.g. Achehnese *dilèë*, Proto Chamic **di hleu,* Toba *jolo* 'before, first'. Given this assumption, the irregular development PM antepenultimate **i* > SM, MIN, BH *a* has to be accounted for. This is done for SM by the following rule order, and the presumption that **di* and **hulu(?)* were still independent lexemes at the time that rule I applied.

TABLE 5: THE DEVELOPMENT FROM PM **di*hulu(?)* TO SM *dahulu*

I   PM **h* > $\emptyset$ |#(C)V$_1$ _ V$_2$... (where V$_1 \neq$ V$_2$)

e.g.   **muhara(?)* > +*muara*
        *buhaya > +*buaya*
but: **di *hulu(?)* > +*di hulu(?)*

II   PM **V* > +*ə* | C_CVCV(C)

e.g.   +*baharu* > +*bəharu*
        +*di* + +*hulu(?)* > +*dihulu(?)* > +*dəhulu(?)*

III   +*ə* (adjacent to +*h*) > *a*

e.g.   +*bəharu* > SM *baharu*
        +*dəhulu(?)* > SM *dahulu*

Finally, MIN *daulu, dulu,* and BH *dahulu* must be borrowed from SM, and I reconstruct:

*di *hulu(?) 'before, first, ahead'.

3.2 THE PM DIPHTHONGS

In the inherited vocabulary of SM, MIN, BH, SWY, and IBN, there are two diphthongs which only occur lexeme-finally, and which mutually agree in these isolects: *-ay* and *-aw*. These diphthongs correspond to JKT *-é* and *-ò* respectively. Other diphthongs (in BH, IBN or JKT) are not inherited.[97] I reconstruct PM **-ay* on the basis of JKT *-é*, other isolects *-ay*, and PM **-aw* on the basis of JKT *-ò*, other isolects *-aw*.

3.2.1 PM **-ay* > JKT *-é*, OTHER ISOLECTS *-ay*

Examples:

*suŋay 'river'; SM, MIN, IBN *suŋay* 'river', SWY *suŋay* 'small river';
*baŋkay 'corpse'; JKT *baŋké*, o.i. *baŋkay*;
*tapay 'fermented rice, yeast'; SM, BH, SWY *tapay,* IBN *tapay*,[98] JKT *tapé* 'fermented rice', MIN *tapay* 'fermented corn';
*lantay 'floor of bamboo strips'; SM, MIN, BH *lantay* 'floor', SWY *lantay* 'floor of an elevated house', IBN *lantay* 'strips of bamboo forming a k.o. deck';
*hintay 'wait for; spy upon'; SM *(h)intay*, MIN *intay*, BH *hintay*, JKT *intey*[99] 'spy upon, watch', IBN *intay* 'wait for, watch for an opportunity';

---

[97]JKT diphthongs (viz. *oy, ay, ey,* and *aw,* see 2.6.1) do not occur in the inherited vocabulary, nor do BH and IBN *-uy* (which do not correspond with each other).
[98]Richards gives *tapay* and *tapey*; Scott gives only *tapé*; the forms *tapé* and *tapey* must be borrowed (from SAR, see Richards 1981).
[99]The phonetic realisation of this (unexplained) diphthong is unclear.

*rantay 'chain'; JKT ranté, SWY rantay (cf. 3.7 for r-), o.i. rantay.

3.2.2 PM *-aw > JKT -ò, OTHER ISOLECTS -aw

Examples:

*rantaw 'coastland, inlet; reach of a river; foreign country'; JKT ranté, MIN rantaw 'coastland, foreign country', SM rantaw, SWY rantaw (cf. 3.7 for r-) 'inlet, creek; reach of a river; foreign country', IBN rantaw 'reach of a river';
*hijaw 'green, unripe'; SM, BH hijaw, JKT ijò 'id.', o.i. ijaw 'green';
*kAr(ə)baw 'buffalo'; SM kərbaw, MIN kabaw, SWY kəbaw, IBN kərəbaw, kərəbo, JKT kərəbò (cf. 3.7.5);
*panaw 'white spots on the skin'; JKT panò? (3.4.2c), o.i. panaw;
*andaw 'day; daylight'; BH ma-landaw 'get up late, lie abed', and IBN apay andaw 'father of the day' (name of a star) (cf. 5.1.2).

IRREGULAR CORRESPONDENCES

JKT has -ó in kasó 'rafter'; other isolects have kasaw 'rafter', and I reconstruct:

*kasaw 'rafter'.

3.2.3 UNEXPLAINED IRREGULAR CORRESPONDENCES: IBN DIPHTHONGS

In a number of IBN lexemes final a(C) is replaced by -ay or -aw. There are also a few cases where expected +-i(C) or +-u(C) has been replaced by -aw. In general, -ay is found instead of expected +-a, +-a?, +-ar, +-an, +-aŋ, +-al or +-ah, and -aw is found instead of expected +-ap, +-am, +-as or +-i(C).

Examples:

-ay:
tuay 'old, mature'; SM tu(h)a, MIN, SWY tuo, BH tuha, JKT tuè;
umay 'farm(land)'; SM (h)uma, BH huma, SWY umo 'farm(land)';
bəsay 'big'; SM, JKT bəsar, BH basar (possibly a loan: gañal is more usual), SWY bəsa? (3.7.4 IC) 'big';
əŋgay < *əŋgən (3.1.1, 3.1.1.5);
jalay 'way, road'; o.i. jalan 'way, road';
pulay 'return (v)'; BH pulaŋ 'again', o.i. pulaŋ 'go home';
datay 'come, arrive; report (v)'; JKT datəŋ, o.i. datəŋ 'come';
gajay, gajah 'elephant'; JKT gajè, SWY gaja(h), o.i. gajah (< SKT);
uday, udah 'already, after, in the past'; SM, MIN sudah, udah, BH sudah, SWY udo, sudo, JKT udè (cf. 3.8.1 IC(1)) (< SKT çuddha- 'cleared, pure' and 'acquitted, complete', Gonda 1973:565);

-aw:
tasaw 'cut undergrowth'; SM, BH sasap 'hoe up weeds', MIN saso? 'cleared rice field';
diaw < *diəm (3.1.1, 3.1.1.5);
uraw < *urəm (3.1.1, 3.1.1.5);
bəraw < *bəras (3.1.1.3);

bəkaw 'trace, mark; showing trace of'; SM *bəkas*, MIN *bakeh*, BH *bakas* 'trace, mark; former', JKT *bəkas* 'used things, leftovers';

kibaw 'wave'; SM, BH *kibas* 'shaking vigorously', JKT *kibas* 'shake s.th. to get the dust off'; ·

maw, mas 'gold'; SM, SWY *(ə)mas*, MIN *ameh*, BH *amas*, JKT *əmas*;

tisaw, tisi 'edge, rim' (cf. 3.8.1 N.B. for *t-*); SM, MIN *sisi* 'id.';

ataw, ati 'liver; centre of the senses, mind'; SM, BH *hati*, MIN, SWY *ati*, JKT *ati?* 'id.';

mipaw, mipis (*m-* unexplained) 'thin'; SM *tipis, nipis*, MIN *tipih*, BH *nipis*, SWY, JKT *tipis* 'thin'.

Semantically differentiated doublets of the form *-aw/-ay* vs *-aC* also occur:

kətaw 'reaping', and *kətam* 'carpenter's plane' (< SM?); SM *kətam*, BH *katam* 'reaping; carpenter's plane', MIN *katam*, SWY *kətam* 'carpenter's plane';

silaw 'dusk', and *silap* 'conjuring' (< SM? cf. SM *silap* 'conjuring');

ləñaw 'lost', and *ləñap* 'die (in songs)'; SM *ləñap*, MIN *lañó?* 'vanish, disappear';

kitay 'we (incl.)', and *kita?* 'you (pl.)' < SAR? cf. SAR *kita?* 'you (pl.)', and *kita* 'we (incl.)' (Collins 1987:84), cf. also BRU *kita?* 'you (polite)'; SM, BH *kita*, SWY *kito* 'we (incl.)', MIN *kito* 'we (incl.); you (sg., to people one does not know)', JKT *kitè* 'I; we'.

On the basis of the above sets I reconstruct:

*bəsar 'big, large';	**sasAp* 'hoe up weeds';
*tuha(?) 'old (of people)';	**bəkas* 'trace, mark';
*huma(?) 'farm(land)' (3.4.2.5)	**kibas* 'wave (v)';
*jalan 'way, road';	**sisi* 'edge, rim';
*pulaŋ 'go back (home)';	**nipis* 'thin' (5.7 lemma 156);
*datəŋ 'come';	**əmas* 'gold';
*ləñap 'disappear, vanish';	**hati* 'liver, centre of senses';
	*kətəm 'reap, plane';
	*kita? 'we (incl.)' (3.4.2b, 3.4.2.4).

N.B. (1) In one case there are three variants, cf. *tiga, tigay, tigaw*, and SM, BH *tiga*, MIN, SWY *tigo*, JKT *tigè* 'three' (see also 5.3.1, 5.3.2). (2) In one case diphthongisation of IBN *+-ur* has occurred, cf. *jəmbuy* (also *jəmbi*) 'expose to the sun' (see 3.6.1.5 for the excrescent *b* in IBN).

3.3 SEMIVOWELS

As a rule, semivowels do not occur initially in the inherited vocabulary of the isolects at hand. The only exceptions to this are a few cases where *y-* originated through the loss of syllabicity of PM **i*:

SM *yaŋ* (2.1.3);
SM *yaŋ*, also given as *iaŋ, hiaŋ*, and *hiyaŋ* (Wilkinson 1959) 'divinity' (3.9.1);
SM *yaitu*, BH *yaitu* (2.1.3);
SM *yu*, also given as *iu* and *hiu*, JKT *yu* 'shark' (3.9.1).

In medial position *y* occurs in the environment *(a,u) _ (a,u)*. It agrees in all isolects, with the exception of MIN lexemes ending in *ø < *-r*, and a few other MIN and SWY isolated cases (where **y* merged with following **a* and became syllabic, i.e. MIN *i?* and SWY *i*). In other environments *-y-* sporadically occurs in the isolects, but it is a non-phonemic glide

when it follows a front vowel or when, in JKT, it occurs adjacent to schwa in a few lexemes which are probably all loanwords (2.6.3, including fn. 62).

Medial *w* is found in the environment *(a,i) _ (a,i)* in all isolects; in SWY it is also found between *a* and *o* (< *-a*; in MIN *-a* also became *-o,* but in this environment preceding **w* was lost). In some cases IBN has a corresponding *b* which reflects PM **b* (see 3.5, 3.5.1). Medial *w* is well attested between *a*'s, but it is hard to find a strong and convincing cognate set for *w* in the environments *a _ i* and *i _ a*. *Kawin* 'marry' is found in all isolects except IBN, but it must be a Persian loan.[100] Five other lexemes are each found in only three isolects:

(1) SM, MIN *jawijawi* 'k.o. tree', SWY *jawijawi* 'k.o. wood';
(2) SM, BH, JKT *giwaŋ* 'earstud' is a loan from SUN (Wilkinson 1959);
(3) SM *lewat*, BH *liwat*, JKT *lèwat* 'past, after; pass';
(4) SM, BH *kiwa*, IBN *kibaʔ* 'left-hand', SWY *kiwo* 'odd (number)'; here IBN reflects **b*: see 3.5, 3.5.1;
(5) SM *sewa*, SWY *siwa* (*-a* unexplained: a SM loan?), IBN *sua* (with contraction of *+i* and *+w*) 'rent (v)', JKT *səwè* (*ə* unexplained: should probably be *è*) 'hire, rent, engage for payment' < SKT *seva:* (Gonda 1973:95).

(1) and (3) are regular sets which I do not recognise as loanwords; however, they form too small a basis for the reconstruction of PM **w* between vowels other than **a*. (2) and (5) are not inherited. In (4), BH, SWY *w* came from **b*.

Dempwolff (1938) reconstructed the following PMP **-awi-* and **-iwa-* sequences with reflexes in the Malayic isolects:

PMP **kawil* 'fish hook'; SM, SWY, IBN *kail*, MIN *kaiʔ*;
PMP **kawit* 'hook'; SM, BH *gait* 'hook, catch on to', *kait* 'hook on to; crook-shaped; crook', MIN *kaiʔ*, SWY, IBN *kait* 'hook', JKT *gaèt* 'hook, catch on to' (cf. 3.4.1.4 UIC);
PMP **lawi* 'tail feather'; SM *lawi/lawi*;
PMP **Riwaŋ* 'lose one's balance'; SM *rewaŋ* 'yaw (of a ship)', MIN *rewaŋ* 'without direction, not knowing where to go';
PMP **tiwas* 'calamity'; SM *tewas* 'defeated', BH *tiwas* 'one's own fault'.

Blust (1970:120) reconstructed PMP **cawi* 'k.o. bird', with only a reflex in SM, *cawi/cawi* 'drongo'.

It appears that Malayic reflexes of PMP lexemes with **-awi-* and **-iwa-* sequences are well attested only for PMP **kawil* and **kawit*, and these reflexes have lost PMP **w*. PMP **w* must have been lost in this environment. SM *rewaŋ*, *tewas*, *lawi/lawi* and *cawi/cawi* are probably loanwords, although it is not clear where they came from.

I reconstruct PM **y* on the basis of *y* in the environments *a _ u, u _ a, a _ a,* and *u _ u* in all isolects in lexemes with a final consonant other than **r*; I reconstruct PM **y* on the basis of MIN *ø*, other isolects *-y-* in lexemes with final **r*.

I reconstruct PM **w* on the basis of *w* in the environment *a _ aC* in all isolects, and on the basis of MIN *ø*, other isolects *w* in the environment *a _ o#* in MIN, SWY and *a _ a#* in the other isolects. In final position semivowels are interpreted as second components of diphthongs, see 3.2, 3.2.1, 3.2.2.

---

[100]According to Von de Wall (cf. Persian *kawin, kabin* 'dower', in Steingass 1930). In SM a (Malaysian and Classical) spelling variant 'kahwin' occurs, which is phonotactically aberrant.

3.3.1 PM *y > MIN ∅, OTHER ISOLECTS y | _*ar; ELSEWHERE PM *y > ALL ISOLECTS y

Examples:

*bayar 'pay'; MIN baiᵊ, SWY baix (see below), o.i. bayar;

*layar 'sail'; MIN laiᵊ, SM, BH, IBN, JKT layar 'sail', SWY layar 'sail', but also SWY bə/layax (with fricative x, see 3.7) 'fastening for roofing material at the back of a roof';

*kAlimayar 'millipede'; SM kələmayar, MIN kalimaiᵊ, BH kalimañar 'id.' IBN əmbayar (with reduced antepenultimate syllables) 'centipede';

*layaŋ 'soar, be borne through the air'; a.i. layaŋ;

*hayam 'domesticated animal, pet animal, plaything'; SM (h)ayam, BH hayam, MIN, SWY, JKT ayam 'chicken, hen', IBN ayam 'plaything, toy, pet', uduk ayam 'pet dog'; (< PMP *qayam 1. 'domesticated'; 2. 'play'; cf. also KD pa-hayam-an 'livestock');

*dayuŋ 'oar'; MIN pan-dayuᵊŋ, o.i. dayuŋ;

*layuʔ 'withered, faded'; SM, BH, SWY, JKT layu, IBN layuʔ;

*guyaŋ 'rock, sway'; SM, MIN, goyaŋ, BH guyaŋ 'id.', IBN guyaŋ, JKT gòyaŋ, 'shake';

*puyu 'k.o. fish'; SM, MIN puyupuyu, IBN puyu;

*kayuʔ 'tree, wood'; IBN kayuʔ 'id.', JKT kayuʔ, o.i. kayu 'wood'.

IRREGULAR CORRESPONDENCES

In a few isolated cases, MIN exhibits a merger of *y with following *a (comparable to the regular merger in lexemes with ∅ < *r, see above):

MIN lampuiᵊŋ 'a ginger (used medicinally)', SM ləmpoyaŋ, ləmpuyaŋ, JKT ləmpuyaŋ 'id.', SWY ləmpuyaŋ 'k.o. shrub';

MIN saiʔ 'cut into slices'; SM, BH, IBN sayat 'id.';

In two other cases merger of *y with following *a (> i) is shown in SWY:

SWY baix < *bayar, and

SWY bais 'k.o. large nibung palm'; SM bayas, MIN bayeh, IBN (Richards 1981) bayas 'k.o. wild palm tree'.

Reconstructions made on the basis of the above cognate sets:

*lAmpuyaŋ 'ginger plant';
*sayat 'cut into slices';
*bayas 'k.o. palm tree'.

3.3.2 PM *w > MIN ∅, OTHER ISOLECTS w | *a _ a(ʔ)#;
      PM *w > ALL ISOLECTS w | *a _ aC (≠*ʔ)

Examples:

*sawaʔ 'python'; SM, BH sawa, IBN sawaʔ 'python', MIN sao 'snake living in swampy areas';

*lawaʔ/*laba(ʔ) 'spider'; SM labalaba, MIN lawah (-h unexplained), IBN əmpəlawaʔ, JKT labèlabè 'spider', BH gandaŋ lawa 'old spider web';

*ñawa 'soul, life; breath'; SM, BH ñawa, MIN ñao, SWY ñawo, JKT ñawè 'soul, life', IBN ñawa (Richards 1981) 'soul, life; breath'; cf. also BAC ba-ñawa 'breathe', SM məñawa 'breathe heavily, as in sleep' (PMP *ñawa 'id.', cf. Dempwolff 1938, and Blust 1978a:43);

*awak 'body; trunk of body; self'; SM awak 'id.', MIN awa? 'body; (formative for) personal pronoun(s)', BH, JKT awak 'body', SWY 'body; (pronoun for the 1st pers.)'; IBN awak 'space, gap, vacancy' may be a cognate, but its meaning is rather deviant.

3.4 THE PM VOICELESS STOPS

3.4.1 PM VOICELESS STOPS IN NON-FINAL POSITION

In non-final position the following voiceless stops are found in the Malayic isolects: p, t, c, k, and ?. They all occur intervocalically; p, t, c, and k also occur initially and post-nasally; p, t, c, and k agree in all isolects, and on the basis of their correspondences I reconstruct PM *p, *t, *c, and *k.

Of the PM voiceless stops, *c is more weakly attested than the others. Some well-established Malayic correspondence sets containing c are borrowed (e.g. SM curietc. 'steal' and SM cium etc. 'kiss', which are borrowed from a North Indian language, cf. Adelaar 1988:62). According to Zorc (1983:12-13), SM c is a secondarily developed reflex of PMP *s or *t, strengthened by an accent pattern favouring the last syllable. This may be correct, but the development in question may have taken place before the PM stage. At any rate, all isolects have c in at least some correspondence sets which are apparently not borrowed, and although the elimination of PM *c would certainly result in some more regular PM phonotactic and morphological patterns, this is not enough evidence for its absence in PM.

In a few cases, JKT has an intervocalic ? in the inherited vocabulary: it occurs between like vowels and is not in contrast with JKT h (cf. 2.6.1).

3.4.1.1 PM NON-FINAL *p > ALL ISOLECTS p

Examples:

*pisaŋ 'banana'; a.i. pisaŋ;
*pərut 'stomach, belly; intestines'; SM, JKT pərut, MIN paruy?, BH parut 'stomach, belly', SWY pəxut 'intestines', IBN pərut 'stomach, belly; intestines' (cf. 5.7 lemma 16);
*puluh 'ten'; MIN puluᵊh, SWY puluᵊ(h), JKT pulu, o.i. puluh;
*lipət 'fold (v,n)'; MIN lipe?, JKT lipət, o.i. lipat;
*kapur 'chalk, lime'; SM, BH kapur, MIN kapuᵊ, SWY kapux 'id.', IBN kapur, kapu? (3.7.4 IC) 'lime';
*tipu 'cheat, deceive'; a.i. tipu;
*əmpat 'four', cf. 3.1.3.3;
*rumput 'weed'; MIN rumpuy?, SWY xumput, o.i. rumput;
*impi 'dream (v,n)'; SM, JKT impi, m/impi, o.i. m/impi.

3.4.1.2 PM NON-FINAL *t > ALL ISOLECTS t

Examples:

*tahun 'year'; SM, BH tahun, JKT taòn, o.i. taun;
*tahi? 'excrement'; SM, BH tahi, SWY, JKT tai, IBN tai?;
*tulis 'write'; MIN tulih, o.i. tulis;
*turun 'go down, descend'; SWY tuxun, o.i. turun;

*hatəp 'roof, roofing thatch'; SM (h)atap, MIN ato², BH hatap, SWY, IBN atap, JKT atəp;
*batu 'stone'; a.i. batu;
*datəŋ 'come', cf. 3.2.3;
*mata 'eye'; SM, BH, IBN mata, MIN, SWY mato, JKT matè;
*gantuŋ 'hang'; JKT gantuŋ/an 'place to hang out s.th.; tools for hanging out s.th.', MIN gantuᵊŋ, o.i. gantuŋ 'hang';
*hantu 'ghost, demon'; SM, BH hantu, MIN, SWY, IBN antu.

UNEXPLAINED IRREGULAR CORRESPONDENCES

(1) IBN lost *t in the following cases:

anti²'wait for; until'; SM, MIN nanti, tanti, SWY tanti, (rare) nanti, JKT nanti;
aruh 'place (v)'; SM, BH taruh, MIN taruᵊh, SWY taruᵊ(h), JKT tarò 'place, put or keep in a place of safety; bet' (related to SM pəŋaruh 'influence'?);

It is not clear whether PM had an initial *n or *t for SM nanti etc.; I reconstruct for the above cognate sets:

*(nt)anti² 'wait for; until';
*taruh 'place (v); keep in a place of safety'.

(2) IBN has gañjaŋ 'naked' along with təlañjay 'naked' < *tilañjaŋ (3.1.3.1).

3.4.1.3 PM *c > ALL ISOLECTS c

Examples:

*cu(ŋ)kup 'enough, complete'; MIN cukuy², SWY cuŋkup, o.i. cukup;
*pəcah 'broken'; SM, IBN pəcah, MIN, BH pacah, SWY pəca(h), JKT pəcè;
*cəpat 'quick, agile'; SM, SWY, IBN cəpat, MIN cape², BH capat, JKT cəpət;
*bənci² 'hate; hatred'; SM, JKT bənci, MIN bañci 'hate', SWY bənci 'hatred', IBN bənci² 'abhor, shun';
*pucuk 'sprout, shoot; end part'; IBN pucuk 'top of a tree; part above the unbranched trunk', MIN pucuᵊ², o.i. pucuk 'sprout, shoot';
*kañciŋ 'button, bolt'; SM, BH, SWY kañciŋ, IBN kañcin (probably a SAR loan, cf. 3.6.3.3); the non-occurrence of a MIN cognate is probably due to the avoidance of a homonymic clash with kañciᵊŋ 'urinate'.

N.B. In IBN regressive dissimilation to t- took place if two syllables began with *c, e.g.

*cacat 'failure, spot, stain, defect'; IBN tacat 'incomplete (of a set)', MIN cace², o.i. cacat 'failure, stain, spot, defect';
*cucuk 'stab, pierce, prick'; SM, BH cucuk, MIN, SWY cucuᵊ², IBN tucuk, JKT còcòk;
*ciñcin 'finger-ring'; IBN tiñcin, SM, MIN, SWY, JKT ciñcin.

IRREGULAR CORRESPONDENCES

IBN lost *c- (which would have become ⁺t-) in ucu² 'grandchild', cf. SM, BH, JKT cucu, MIN cucu, cucuᵊ/ŋ, SWY cucu/ŋ; see 5.4 for this loss, and for MIN, SWY -/ŋ. For this set I reconstruct:

*cucu² 'grandchild'.

3.4.1.4 PM NON-FINAL *k > ALL ISOLECTS k

Examples:

*kutu 'head louse'; JKT kutuʔ (3.4.2c), o.i. kutu;
*k/anan 'right (hand)'; a.i. k/anan;
*kulit 'skin, bark'; SM, SWY, IBN, JKT kulit 'id.', MIN kuliʔ 'bark, peel';
*sakit 'ill, sick'; MIN sakiʔ, o.i. sakit;
*ikan 'fish'; BH n.c., o.i. ikan;
*bukit 'mountain, hill'; SM, BH, JKT bukit 'hill', MIN bukiʔ, SWY bukit 'hill, mountain',
IBN bukit 'mountain';
*aŋkat 'raise, lift, move'; MIN (rare) aŋkeʔ, o.i. aŋkat;
*laŋkah 'pace, step'; SWY laŋka(h), JKT laŋkè, o.i. laŋkah;
*buŋkuk 'bent'; MIN, SWY buŋkuᵊʔ, JKT bòŋkòk, o.i. buŋkuk.

UNEXPLAINED IRREGULAR CORRESPONDENCES

In six correspondence sets the isolects disagree in reflecting k or g:[101]

(1) SM, JKT kəmbar, MIN, BH kambar, SWY gəmbax, IBN gəmbar 'twin';
(2) SM, JKT gali, MIN, SWY, IBN kali 'dig';
(3) SM gosok, JKT gòsòk, MIN kusuᵊʔ, gusuᵊʔ, gosoʔ, BH kusuk, gusuk, SWY kosuᵊʔ,
IBN kusuk 'rub';
(4) SM kunday 'short queue', BH gunday 'hair', IBN gunday 'long hair at back of head',
JKT kòndé 'hair knot'; (BSM gunday 'lock of hair on crown of head');
(5) SM, BH, JKT kumpul, SWY kumpul, kumpuᵊʔ 'together, gathered', MIN kumpuᵊʔ
'heap, collection', IBN gumpul 'gather, pick' (cf. 3.7.2 UIC);
(6) SM, BH gait 'hook, catch on to', kait 'hook on to ; crook-shaped; crook', MIN kaiʔ,
SWY kait 'hook', IBN kait 1. 'crippled, deformed' 2. 'hook, catch', JKT gaèt 'hook, catch
on to', òraŋ gaèt 'thief', cf. PMP *kawit 'hook'.

In case (1), it is striking that two isolects, which can hardly have had any influence on one another, agree in having g-. But the four other isolects have k-, and the corresponding PMP form is *kembar.

In case (2) PMP has also a corresponding *k- (PMP *kali).

For (3) there is a corresponding PMP *gusuk (on the basis of Toba, JV, SM gosok 'rub', Ngaju maŋgosok 'rub', kusok 'rubbed', Malagasy kusuka 'rub'). But the *g of this reconstruction is questionable. Firstly, its Toba and Ngaju reflexes are loans. Toba o developed either from Proto Batak *e (a schwa, which came from PMP *e), or from Proto Batak *-ow (< PMP *-ew and *-aw, Adelaar 1981:12, 18; Adelaar 1988:68 fn.7). It does not reflect PMP *u, and hence Toba gosok cannot have developed from PMP *gusuk. Hardeland (1859) labelled Ngaju maŋgosok as a loan from Banjarese, which was apparently overlooked by Dempwolff. Secondly, the k- in Malagasy may reflect PMP *ng as well as *nk. Thirdly, Ngaju kusok, and Toba husuk (with a related meaning 'shake') may have developed from a PMP *kusuk. Moreover, JV has also kosok 'rub' along with gosok, and in Old Javanese only kusuk occurs. With the elimination of the Toba and Ngaju evidence, and in view of the indecisive role of the Malagasy and JV evidence, there is good reason to

---

[101]There is also SWY kampaŋ in anaʔ kampaŋ 'child of a prostitute', and Salako kampàkŋ 'commit adultery' (Adelaar, unpublished fieldnotes): these forms may be related to SM gampaŋ 'of little account; easy; light'; cf. also SM anak gampaŋ 'illegitimate (literally 'easy') child'.

substitute PMP *kusuk* for Dempwolff's *gusuk* on the basis of Toba *husuk* 'shake', SM *gosok* (dialectally also with *k-*), JV *gosok/kosok* (Old Javanese *kusuk*) 'rub', Ngaju *kusok* 'rubbed', and Malagasy *kusuka* 'rub'.

In (4) the correspondences are ultimately borrowed from TAM (cf. Van Ronkel 1902:112).

In (5) there is only one form with *g-*, and finally, in (6) there is an ancestral form PMP *kawit* 'hook'.

I do not know what caused the sporadic change of PMP *k* > *g* in the isolects. The agreement between SWY *gəmbax* and IBN *gəmbar* is probably coincidental, and the agreement between SM, BH *gait* and JKT *gaèt* is probably due to borrowing. I reconstruct:

kali 'dig';
kəmbar 'twins';
kait 'hook(n); hook, catch on to';
kusuk 'rub'.

3.4.2 PM VOICELESS STOPS IN FINAL POSITION

In final position the Malayic isolects have *p, t, k*, and *ʔ*. *-p* and *-t* agree in all isolects except MIN, where all final stops merged in *-ʔ* after raising or diphthongising some of the preceding vowels (cf. 3.1.1 - 3.1.2). *-k* agrees in SM, BH, IBN and JKT (it is realised as a glottal stop in SM and IBN). This SM, BH, IBN, JKT *-k* corresponds to *-ʔ* in SWY and MIN. Apart from *-ʔ*, SWY also has a *-k*, and conversely, IBN and JKT also have a *-ʔ*, each with a different origin. (IBN *-k* and *-ʔ* are both realised as [ʔ]: cf. 2.5.1 and fn. 48). In summary, we have the following regular correspondence sets:

MIN *-ʔ*, o.i. *-p*;
MIN *-ʔ*, o.i. *-t*;
MIN, SWY *-ʔ*, o.i. *-k* (realised as [ʔ] in SM and IBN).

And furthermore we have SWY *-k*, IBN *-ʔ* and JKT *-ʔ* which do not seem to have regular correspondences in the other isolects, and which need some further discussion.

(a) SWY

SWY *-k* is innovative. SWY regularly has *-ʔ* corresponding to *-k* (MIN *-ʔ*) in the other isolects, and in only three cases does it have *-k* corresponding to *-k* (MIN *-ʔ*) in the other isolects:

(1) *balik* 'wrong, upside down'; (cf. SM, BH *bali/k*, MIN *baliᵊ/ʔ*, JKT *bòlak/balik* < *baliʔ* 'reverse, go back', see below 3.4.2.4 UIC);
(2) *luk* 'curve in the blade of a knife'; cf. SM *luk* 'id.';
(3) *tasik* '(mythological) great lake around the axis of the world'; cf. SM *tasik* 'lake, sea', MIN (in texts) *tasiᵊʔ* 'lake', IBN *tasik* 'sea'.

These exceptions must be loanwords. They may be borrowed from SM: the meaning of *balik* matches with the usual meaning of SM *balik*. *Luk* is also found in SM, and in both SM and SWY, *luk* is irregular in its monosyllabic shape (cf. 4.2). *Tasik* is only found in traditional stories (the 'andayanday'), whereas it occurs as a usual form for 'sea' or 'lake' in SM and IBN. Furthermore, in Helfrich 22 lexemes with *-k* also contain the non-inherited apical trill *(r)*, whereas inherited *x* in combination with *-k* is found only in one case, where it

is, moreover, in free variation with *r*: *kəruk, kəxuk* 'cut or scratch out the meat of a coconut; crust in a pan' (SWY *r* is a loan phoneme, cf. 3.7). This fact strongly supports the idea that SWY -*k* is innovative, and hence that SWY -*ʔ* is the regular correspondence to MIN -*ʔ*, o.i. -*k*. A possible objection to a SM origin of *balik, luk* and *tasik* is that these lexemes should have had a -*ʔ* on account of the fact that SM realises final -*k* as [ʔ]. But then again, SWY speakers would most likely interpret this SM [ʔ] as [k] on account of the fact that the preceding vowels are monophthongs. (SWY, which has [-iᵊʔ], [-uᵊʔ], [-uk] and [-ik] sequences, does not have [-iʔ] or [-uʔ] sequences, 3.1.2.)

(b) IBN

The origin of IBN -*ʔ* is problematic. Zorc (1982:115) assumes that -*ʔ* is the result of a merger of PAN **S* (corresponding to -*s* in Formosan languages, and -*h*- (at morpheme boundaries) in Philippine languages), **H* (corresponding to -*h* in some Formosan languages, and to -*h*- at morpheme boundaries in some Philippine languages), and **ʔ* (corresponding to -*ʔ* in some Philippine and Formosan languages). The evidence for this theory is shown in the following correspondences (in my evaluation of Zorc's theory I restrict myself to criticism of his IBN evidence. In order to show that JKT -*ʔ* does not reflect any of the PAN laryngeals, evidence from this isolect is added to Zorc's list, see below).

Zorc's evidence for PAN **-S*:

PAN **CăliS* 'line, string' > IBN, JKT *tali*;
PAN **daqiS* 'forehead' > IBN *dai*;
PAN **ku(S)kuS* 'fingernail' > IBN *kuku* 'claw', JKT *kuku* 'fingernail';
PAN **pa:qaS*[102] 'thigh' > IBN *paah* (-*h* unexplained), JKT *pahè* (-*h*- unexplained);
PAN **tuqaS* 'old (people)' > IBN *tuay*, JKT *tuè*;
PAN **tĕbuS* 'sugarcane' > IBN *təbu*, JKT *təbuʔ*;
PHF **CĭŋaS* 'food particles caught between the teeth' > IBN *tiŋaʔ*.

To these reconstructions I add **Cu:meS* and **pa:RiS* (see below).[103]

Zorc's evidence for PAN **-H*:

PAN **qŭmaH* 'farm(land)' > IBN *umay*;
PHF **ba:RaH* 'live coals' > IBN *baraʔ*, JKT *barè*;
PAN **lĭmaH* 'five' > IBN *limaʔ*, JKT *limè*;
PAN **qi:SuH* 'shark' > IBN *iuʔ*, JKT *yu*;
PAN **qu:luH* 'head' > IBN *uluʔ* 'meaning [sic]';
PAN **baq(e)RuH* 'new' > IBN, JKT *baru*;
PHF **Ca:qiH* 'faeces' > IBN *taiʔ*, JKT *tai*;
PAN **ba:RuH* 'hibiscus' > IBN *baruʔ*;

---

[102]The **S* in **pa:qaS* is based only on intervocalic *s* in Bunun *p-in-asax* 'thigh'. This is assumed by Zorc (1982:119) to contain an infix -*in*- and to show metathesis of *s* and *x*. Other languages show no trace of **S*, and other Austronesianists have reconstructed PAN **paqa* (Dyen 1953:11; Ferrell 1969:225).
*Cu:maH should be reinterpreted as **Cu:meS* on account of Kavalan *tumes* 'louse' (Li 1982:489). PMP **pa:Rih* becomes PAN **pa:RiS* on account of Siraya *pagig* 'rayfish' (PAN **R, *S* > Siraya *g*, cf. *ougat* 'nerve' < PAN **uRaC, vigvig* 'lip' < PAN **biRbiR, vagiog* 'stormwind' < PAN **baRiuS, voukig* 'hair' < PAN **buSek* (with a regular metathesis, cf. Dahl 1976:33).
This affects the interpretation of Zorc's evidence in the following way: instead of two examples out of seven supporting the assumption that PAN **-S* > IBN -*ʔ*, there are three out of nine, and instead of 12 examples out of 18 supporting PAN **-H*> IBN -*ʔ*, there are now 11 examples out of 17.
In JKT this change alters the numbers as follows: instead of one out of seven examples reflecting PAN **S* > JKT -*ʔ*, there are now two examples out of eight, and instead of one out of eight examples reflecting PAN **-H*> JKT -*ʔ*, there now is no evidence at all for this change.
[103]See fn. 102.

PAN *CăquH 'know (how)' > IBN tau?, JKT tau 'id.', also tau? 'dunno!';

PAN *CunuH 'roast on fire' > IBN tunu;

PHF *dăkiH 'body dirt' > IBN daki?, JKT daki;

PHF *dăpaH 'sole (foot)' > IBN tapa? (t- unexplained) 'palm (hand); sole (foot)', JKT t/əl/apak 'id.';

PAN *kĭtaH 'we (incl.)' > IBN kitay 'id.', JKT kitè 'I; we';

PAN *si:kuH 'elbow' > IBN siku, JKT sikut (< JV);

PAN *tu:baH 'plant (n); fish poison' > IBN tubay;

PHF *ZaRa:miH 'rice stalk, straw' > IBN jərami?, JKT jərami;

PHF *Za:waH 'millet' > IBN jawa?;

(PAN *Cu:maH 'body louse' (> IBN tuma?, JKT tumè?) should be reinterpreted as *Cu:meS).[104]

Zorc's evidence for PAN *-?:

PHN *a:ku? 'admit, acknowledge' > IBN aku?, JKT aku;

PMP *qa:Ruhu? 'casuarina' > IBN ru?;

PHF *qila? 'like, note' > IBN ila? 'keep an eye on';

PHN *badi? 'knife' > IBN badi?;

PHF *b(al)aŋa? 'earthenware vessel' > IBN bəlaŋa?, JKT bəlaŋè;

PHN *bapa? 'father' > IBN bapa? 'father-in-law' (but cf. also apay 'father', cf. 4.3.1), JKT bapa? 'father';

PMP *bi:Ra? 'discharge, semen' > IBN bira? 'evacuate, discharge', JKT bèrak 'id.';

PMP *bŭka? 'open' > IBN buka?, JKT bukè;

PMP *Da:tu? 'chief' > IBN datu? 'nobleman, chief', JKT datu? 'ancestor';

PAN *kaka? 'elder sibling' > IBN aka?, JKT kaka/ŋ (cf. 5.4 N.B. for IBN loss of *k-, and JKT -/ŋ);

PHN *mama? 'male kin' > JKT mama/ŋ 'uncle' (cf. 5.4. N.B.);

PMP *muda? 'young, unripe' > IBN muda?, JKT mudè (Muhadjir 1981);

PHN *naŋka? 'Artocarpus sp.' > IBN naŋka?, JKT naŋkè;

PHF *Nasi? '(cooked) rice' > IBN asi? (with loss of initial +n), JKT nasi?;

PHN *ni:ni? '(form of address for female kin)' > IBN ini? (5.4 N.B.), JKT nini? 'grandmother';

PMP *pa:ku? 'nail' > IBN paku?, JKT paku;

PMP *pa:lu? 'strike, hit' > IBN palu?;

PMP *puki? 'vulva' > IBN puki?;

PHF *Sesi? 'meat, contents' > IBN isi? 'id.', JKT isi 'contents';

PMP *ta:bi? '(term for) greeting' > IBN tabi?, JKT tabé (both < SKT, see below);

PHN *ta:ma? 'hit the mark' > IBN tama? 'enter';

PHN *tĕku? 'bend' > IBN təku? 'bend', JKT təkuk (-k unexplained) 'fold (v)';

PHN *tĭku? 'bend' > IBN tiku?;

PMP *wăDa? '(existential)' > IBN n/aday 'there is not', JKT adè 'there is';

PHN *zĕra? 'warned by experience' > IBN jəra?.

Zorc's evidence for PAN *-ø:

PAN *măCaø 'eye' > IBN mata, JKT matè;

---

[104]See fn. 102.

PAN *su:suø 'breast' > IBN tusu? 'suck', tusu 'breast' (neither given by Zorc), JKT susu 'breast';

PAN *bătuø 'stone' > IBN batu 'stone', JKT batu in batu kòlar 'coral stone';

PAN *ku:Cuø 'head louse' > IBN kutu, JKT kutu?;

PAN *ăkuø 'I' > IBN aku;

PAN *bĕliø 'buy' > IBN, JKT bəli;

PAN *(k)ămiø 'we (excl.)' > IBN kami;

PAN *ĭøaø '(3rd pers. sg.)' > IBN ia, JKT iè.

Zorc's evidence for PMP *-h:

PHN *ănuh 'what(ever)?' > IBN anu?, JKT anu, anó 'something, someone, so-and-so';

PMP *bu:ŋah 'flower, fruit' > IBN buŋay, JKT buŋè;

PMP *bŭkuh 'joint, node' > IBN buku?, JKT buku;

PMP *dĕpah 'fathom' > IBN dəpa?, JKT dəpè?;

PHN *kĕnuh '(quotative particle)' > IBN kənu?;

PHN *lĕŋah 'Sesamum indicum' > IBN ləŋa?;

PHN *ŋĭluh 'set one's teeth on edge' > IBN ñilu? (ñ- unexplained), JKT ŋilu;

PHN *pakuh 'fern (Athyrium esculentum)' > IBN paku?, JKT paku;

PHN *peñuh 'turtle' > IBN pəñu?, JKT pəñu;

(PMP *pa:Rih 'rayfish' (> IBN pari?) should be reinterpreted as PAN *pa:RiS[105]);

PMP *săguh '(root crop) sago' > IBN sagu? 'balls of cooked sago', JKT sagu 'sago';

PHN *săwah '(snake) python' > IBN sawa?;

PHN *sĭDah 'they' > IBN sida?;

PMP *tŭlih 'earwax' > IBN tuli? 'having a suppurating ear', JKT tuli 'deaf';

PHN *zăzah 'carry (goods)' > IBN bə/jaja?.

Although Zorc's study of the PAN laryngeals is a very thoroughgoing one, and although his material, in many cases, shows a striking measure of agreement, I have some objections to his theory. Zorc considers IBN lexemes with a diphthongised last syllable to be evidence for the development of PAN *S, *H, and *? to IBN -?, which is unwarranted since there is no way of telling whether earlier (undiphthongised) forms of these lexemes had -? or not. The lexemes in question are: tuay, umay, kitay, tubay, buŋay, and n/aday; better evidence for *kĭtaH is IBN kita? 'you (plural)' (if this is not a SAR loan, cf. 3.2.3). Furthermore, IBN (and also Mualang) have a fossilised suffix -?, which marks transitivity (see 6.1.1). It is a reflex of the PAN locative focus marker, and it corresponds to -i in Philippine and Formosan languages (Wolff 1973:73, 77). It is not likely that the -? in IBN -i? reflects a PAN laryngeal, since no corresponding consonant is found in Formosan or Philippine languages. But more important criticism to Zorc's sound law is that there are too many unexplained IBN exceptions. To start with, his evidence for PAN *-S > IBN -? is too weak.[106] There are only three cases out of nine that support this correspondence, viz. tiŋa?, pari? and tuma?. In two cases the IBN reflexes do not provide evidence viz. tuay and paah (-h unexplained), and four reflexes provide counterevidence: təbu, tali, dai, and kuku. Evidence for PAN *-H > IBN -? is much stronger, although here too one is left with a considerable number of exceptions. Out of 17 examples, 11 (including kita? 'you (plural)') exemplify the change, viz. bara?, baru?, daki?, iu?, jawa?, jərami?, kita?, lima?, tai?, tapa?, and tau?; two do not provide evidence because of the diphthongisation of the last syllable (kitay not included), viz.

---

[105]See fn. 102.
[106]See fn. 102.

umay and *tubay*; one is based on an unconvincing reflex: according to Zorc, PAN *qu:luH* 'head; hilt; beginning; upper part of a river basin' became IBN *ulu?* 'meaning', but IBN *ulu* 'hilt; upper part (of a river basin); to lead, guide' is semantically more sound as a reflex; finally, three examples contradict the change viz. *baru* ('new'), *tunu,* and *siku.* More reconstructions with final *H* were provided by Tsuchida,[107] but Zorc doubts their validity, because Itbayaten (a Philippine language) and IBN cognates both reflect ø. This fact, and also the fact that Formosan languages often disagree (not only with Philippine languages and IBN but also with each other) in reflecting *H* also weighs against the argument for reconstructing this laryngeal.

PAN *? > IBN -? is well attested by a large number of cognate sets, although I do not agree with the use of the set PMP *wăDa?* > IBN *n/aday* as evidence because of the diphthongisation in IBN. I reject PMP *ta:bi?*, which is actually an erroneous reconstruction built on forms deriving from SKT *kṣantavya* 'to be forgiven and pardoned' (Gonda 1973: 640), and I also reject PHN *bapa?*. Zorc's evidence for PHN *bapa?* is Lampung, JV, SM, SUN *bapa?*, Madurese *əppa?* 'father', IBN *bapa?* 'father-in-law', Hanunoo *ba:pa?* 'uncle', Ilokano *ba:pa* 'parent or parent's siblings' and Sambal *ba:pa?*. But PHN *bapa?* is one of the exceptions to a phonotactic tendency which must have applied to PMP. The alleged reflexes of *bapa?* with initial *b-* presented by Zorc must be borrowed from SM, although I admit that I find Sambal *ba:pa?* somewhat puzzling in this respect. The above tendency to articulation-type harmony, and the history of SM *b/apa/k,* IBN *b/apa?* etc. (which derived from PM *apa(?)*), are discussed in 4.3 (including fns 130 and 131).

In cases where PAN *ø* is reconstructed, IBN agrees with -? in seven cases out of eight; for PAN *su:suø* 'breast' it has *tusu* 'breast' and *tusu?* 'suck' (note also *ŋ-aku?* 'confess' along with *aku* 'I'). The correspondence PMP *h* > IBN -? is also well attested (PMP *h* is the result of a merger of PAN *S* and *H*).

In the light of the above considerations it is hard to make out whether IBN supports Zorc's proposed sound law. His material supports PAN (/PHF/PHN/PMP) *? > IBN -?, but not PAN (/PHF) *S > IBN -?; it supports PAN (/PHF) *H > IBN -? in two thirds of the examples. Another interpretation of IBN -? is that it occurs in inherited vocabulary, and that, in an earlier stage of IBN, it was a non-phonemic glottal stop heard after any monophthongic final-vowel phoneme. This situation was later on altered by the monophthongisation of original final diphthongs, and by subsequent borrowing of lexemes with final vowel: lexemes with a final vowel or diphthong which developed from a PAN diphthong (e.g. *mati, padi, babi, buru,* and *ijaw, suŋay, baŋkay,* cf. 3.2.1, 3.2.2 and 7.1b), and loanwords, as a rule do not end with -?. The only loanwords ending in a -? are: *pala?* 'chief' ((< SM?) < SKT), *kiju?* 'cheese' ((< SM?) < POR), *cuka?* 'vinegar' ((< SM?) < SKT), *cabi?* 'chilli' ((< SM?) < SKT), *tabi?* '(term for) greeting' and *bapa?* (< SM; cf. also the more authentic IBN *apay* 'father' cf. 4.3.1). Finally, the apparent agreement in the occurrence of laryngeals between IBN on the one hand, and Formosan and Philippine languages on the other, may then be accounted for by the multitude of reconstructions with final laryngeals: if the majority of PAN (/PHN/PHF/PMP) reconstructions have a final laryngeal (as opposed to another consonant or ø), and if IBN -? was a phonetic glottal stop heard after each final-vowel phoneme in the history of IBN, then it is only self-evident that -? more often than not will agree with a PAN laryngeal. But if one follows the latter interpretation, one is left with a certain number of reconstructions ending in a final vowel which correspond with IBN

---

[107]cf. Tsuchida (1976:132-138), where he also distinguishes PAN *H_1* and *H_2*, and Zorc (1982:122).

lexemes ending in a final vowel, and these can hardly be interpreted as loanwords (e.g. PAN *măCaø > mata, PAN *bătuø > batu, PAN *ku:Cuø > kutu, etc.). There are also three reflexes of PAN proto-lexemes ending in a diphthong which have a -ʔ viz. bəriʔ 'give' (< PAN *beRey 'id.'), pandiʔ 'bathe' (< PAN *anduy 'id.'), and kayuʔ 'wood' (< PAN *kaSiw). Another point in favour of Zorc's assumption is that Tioman Malay also exhibits a final -ʔ, which seems to agree with the -ʔ in IBN (Collins 1985). There are exceptions, and information on Tioman Malay is rather limited, but the agreement is still significant.

TABLE 6: FINAL GLOTTAL STOPS IN TIOMAN MALAY AND IBN

Tioman Malay	Proto-language	IBN	meaning
[bukaʔ]	PMP *bŭkaʔ	bukaʔ	open
[ɛsɛʔ]	PHF *Sesiʔ	isiʔ	contents
[bʌɣɔʔ]	PAN *ba:RuH	baruʔ	hibiscus
[naŋkaʔ]	PHN *naŋkaʔ	naŋkaʔ	jackfruit
[sawaʔ]	PHN *săwah	sawaʔ	python
[pəñuʔ]	PHN *peñuh	pəñuʔ	tortoise
[nasiʔ]	PHF *Nasiʔ	asiʔ	cooked rice
[tali]	PAN *CăliS	tali	rope
[gutu]	PAN *ku:Cuø	kutu	louse
[matə^]	PAN *măCaø	mata	eye
[bʌtu]	PAN *bătuø	batu	stone

But cf. also Tioman bəɣi 'give', IBN bəriʔ.

A really satisfactory explanation for the origin of IBN -ʔ cannot be given at present. As long as I am not able to disprove Zorc's theory and to come up with a better explanation, I consider IBN -ʔ as the reflex of PAN (PHF/PHN), *S, *H, or *ʔ, or PMP *h, *ʔ. If IBN exhibits -ʔ, I reconstruct PM *-ʔ; if IBN has -ø, I reconstruct PM *ø; if IBN has no reflex, or a reflex with -ay, corresponding to SM, BH -a, MIN, SWY -o, JKT -è, I reconstruct PM *-(ʔ).

(c) JKT

JKT -ʔ is an innovation: no conditioning factor for its occurrence can be given, apart from the fact that some JKT sub-dialects show it more often and in more regular patterns than the sub-dialect of Mester.[108] It may be due to influence from SUN, where a non-phonemic [ʔ] is heard after all final vowels. Abdul Chaer gives quite a few variant forms from different sub-dialects. Especially in the case of lexemes reflecting PM *-a(ʔ), Abdul Chaer gives the Kebayoran as well as the Mester forms, e.g. apaʔ (Kebayoran), apè (Mester) 'what'; bawaʔ (Kebayoran) bawè (Mester) 'carry'. In Mester -ʔ may occur after any vowel, e.g.

asóʔ	take a rest	kutuʔ	body louse
cəntéʔ	k.o. plant	guruʔ	teacher
matiʔ	dead	cariʔ	look for
nasiʔ	cooked rice	jahèʔ	ginger
tumèʔ	head louse	pisòʔ	knife

Lexemes ending in é or ó often have a variant with a final èʔ or òʔ sequence respectively (cf. 3.1.2), e.g.

---

[108]From the chart in the Introduction to Abdul Chaer (1976:XVIII-XIX) it appears that in Kebayoran a glottal stop is heard after every final vowel (except ramè). The sub-dialect of Karet has a corresponding glottal stop, but it also lost ⁺-h (Kebayoran still has -h), giving rise to final vowels.

ja(h)é	~ *jahè?*	ginger
bégó	~ *bègò?*	stupid
légó	~ *lègò?*	sell

JKT *-?* differs from IBN *-?* in the following ways:

(1) it is found in loanwords (which have a final vowel in the lending language), e.g.

guru? 'teacher' (< SKT)
kunò? 'old, ancient' (< JV)
ləmari? 'cupboard' (< POR)
mani? 'semen' (< AR)

But in other loanwords it does not occur, e.g.

tèmpò 'time' (< POR)
waktu 'time' (< AR)

cf. also *kiju* 'cheese', *kəpalè* 'head', *cukè* 'vinegar', *cabé/cabè?* 'chilli' and *tabé* '(term of) greeting', the correspondences of which have *-?* in IBN (see above).

(2) it sometimes occurs after a final vowel which developed from a PAN diphthong (cf. 7.1b), e.g.

mati? 'dead' < PMP **matey*	*panò?* 'white spots on skin' < PMP **panaw*
ati? 'liver' < PMP **qatey*	*pisò?* 'knife' < PMP **pisaw*
kayu? 'wood' < PMP **kahiw*	*kəbò?* 'buffalo' < PMP **karebaw*

but cf. also:

padi 'rice plant' < PMP **pajey*
buru 'chase, hunt' < PMP **buRew*
rantò 'coastland, foreign country' < PMP **rantaw*
dami/damé 'quiet, peaceful' < PMP **damay* 'peace'

(3) it often has a variant with final *ø*, in which case the lower mid-vowels of the variant with *-?* usually change to higher mid-vowels (e.g. *bégó ~ bègò?, ja(h)é ~ jahè?*, see above).

In two cases it is found after a vowel that historically had a following *+h*, viz. *tuju?* 'seven', and *butu?* 'male sexual organ'.[109] Its occurrence does not correspond with the final laryngeals in PAN, as is shown in the JKT reflexes of the reconstructions given by Zorc (see above):[110]

– In only two cases does JKT *-?* correspond to PAN **S*, viz. **CebuS > təbu?* and **Cu:meS > tumè?*.

– In no case does it correspond to PAN **H*: in seven cases JKT has *ø* for PAN **H* (viz. *barè, baru, daki, jərami, limè, tai*, and *yu*), in one case it has *-k* (*təlapak)*, and in one case it has both *-?* and *ø* (viz. *tau* 'know' and *tau?* 'dunno').

– In four cases JKT *-?* corresponds to PAN **?* (viz. *bapa/?, datu/?, nasi?*, and *nini?*, but cf. also 5.4 N.B.); in six cases it has *-ø* for PAN **?* (*adè*, *ŋ-aku* ('confess'), *bukè, isi*,

---

[109] In modern Indonesian the meaning 'need' (*məmbutuhkan* 'need (v)') is prevailing (cf. JV *butuh* 'need, lack (v)'), but cf. SM *butuh*, BH, IBN *butuh*, SWY *butu?(h)* 'penis', MIN *butɨ?h* 'a good-for-nothing, a jerk'.
[110] See fn. 109.

naŋkè, paku, tabè); in one case it has *-ŋ* (*mama/ŋ*, a SUN loan? cf. 5.4.4), and in two cases it has *-k* (*bèra/k* and *təkuk*).

- In only one case does JKT have *-ʔ* corresponding to PMP **h* (viz. *dəpèʔ*); in eight cases it has ø (viz. *anu/anó, buku, buŋè, ŋilu, paku, pəñu, sagu,* and *tuli*).

- On the other hand, JKT *-ʔ* corresponds to PAN ø in two cases (viz. *kutuʔ, susuʔ*) and it has ø in four cases (viz. *batu, bəli, matè,* and *iè*).

Considering the above discussion and the sound correspondences between the isolects, I reconstruct PM final **k* on the basis of SM, BH, IBN, JKT *-k*, MIN, SWY *-ʔ*, and PM **ʔ* on the basis of IBN *-ʔ*, o.i. ø.

3.4.2.1 PM FINAL **p* > MIN *-ʔ*, OTHER ISOLECTS *-p*

Examples:

*hatəp 'roof, roofing thatch' (3.4.1.2);
*hi(ŋ)səp 'inhale, suck in'; SM *(h)isap*, MIN *isoʔ*, BH *hisap, isap*,[111] SWY *isap*, IBN *insap*, JKT *isəp*;
*lətup 'burst, bang'; SM, SWY *lətup*, BH *latup* 'burst, explode', MIN *latuyʔ* 'crack one's joints', IBN *lətup* 'a bang, crack';
*sisip 'add, insert' (3.1.2, 3.1.2.1 ; cf. also 4.4).

3.4.2.2 PM FINAL **t* > MIN *-ʔ*, OTHER ISOLECTS *-t*

Examples:

*bəlut 'crawling animal, eel'; SM, SWY *bəlut*, MIN *baluyʔ* 'id.', IBN *bəlut* 'worm';
*pulut 'sticky'; SM *pulut*, MIN *puluyʔ* 'id.', SWY, IBN *pulut* 'glutinous rice';
*səmpit 'narrow, tight'; SM, SWY, IBN, JKT *səmpit*, MIN *sampiʔ*;
*səsət 'have lost one's way'; SM, SWY *səsat*, MIN *saseʔ*, BH *sasat*, IBN *təsat* (cf. 3.8.1), JKT *səsət*.

UNEXPLAINED IRREGULAR CORRESPONDENCE

In three cases JKT has *-r* corresponding to *-t* in the other isolects. These cases are:

lalər 'fly (n)'; MIN *laleʔ*, o.i. *lalat*;
pusər 'navel'; SWY, IBN *pusat* 'id.', SM, BH *pusat*, MIN *puseʔ* 'centre, focus; navel';
ulər 'worm, grub, maggot'; SM *(h)ulat*, MIN *uleʔ*, BH *hulat*, SWY, IBN *ulat*.

Lalər, pusər, and *ulər* may be borrowed from JV (cf. JV *lalər, pusər, ulər,* with similar meanings). Another possibility is that JKT *-r* reflects PMP **-j*, in which case a separate PM reflex should be reconstructed (possibly **-d*?). But this is unlikely both because PMP **-j* is otherwise reflected as *-t*, and because nowhere else in the material is evidence found for the reconstruction of PM lexeme-final voiced stops.

Compare also other JKT reflexes for PMP lexemes with **-j*:

PMP **kunij* 'turmeric' > JKT *kuñit* 'turmeric' (cf. **kunit*, 3.6.1.2IC);
PMP **ubaj* 'medicine, drug' > JKT *òbat* 'id.' (3.1.2.2);

---

[111] Abdul Jebar gives both *hisap* and *isap*, without specifying whether they belong to BH or to BK.

In one case JKT has a doublet with -*d*:

PMP **qañud* 'drift, float' > JKT *añut, añud,* and SM *(h)añut,* MIN *añuy?,* BH *hañut,* SWY, IBN *añut* 'id.'.

I consider JKT *lalər, pusər,* and *ulər* loanwords from JV, and I reconstruct:

*lalət 'fly (n)';
*pusət 'navel';
*hulət 'worm, grub, maggot';
*hañut 'drift, float'.

3.4.2.3 PM FINAL **k* > MIN, SWY -*?*, OTHER ISOLECTS -*k*

Examples:

*luø(uə)k/*tA/luk 'bay, inlet; corner' (cf. 3.10); SM, IBN, JKT *tə/luk,* MIN *ta/luᵊ?,* BH *luuk, ta/luk* 'deep place in a river', SWY *tə/luᵊ?* 'bay, inlet';
*miñak 'oil'; MIN, SWY *miña?,* o.i. *miñak;*
*tətək 'cut up'; SM, IBN *tətak,* BH *tatak,* SWY *tata?* 'id.', MIN *tata?* 'carve, slash, delimit s.th.' (cf. PMP **tektek* 'chop off');
*tasik 'sea'; SM *tasik* 'lake, sea', MIN (in texts) *tasiᵊ?* 'lake', IBN *tasik* 'sea';
*bilik 'wickerwork of bamboo used for making partitions in houses; such a partition; compartment, apartment'; SM *bilik,* MIN *biliᵊ?* 'room, compartment', SWY *biliᵊ?* 'bedroom', IBN *bilik* 'room, especially of a longhouse', JKT *bilik* 'wickerwork of bamboo used as a screen or inner wall'.

IRREGULAR CORRESPONDENCES PM **-k* > -*t*

In a few cases IBN has -*t* corresponding to MIN, SWY -*?*, o.i. -*k*. In all these cases the final-syllable vowel is *i*, e.g.

cərədit, cərədik 'bright' < **cVr(ə)dik* (3.7.5);
itit 'duck'; SM, BH *itik,* MIN, SWY *itiᵊ?;*
gundit 'concubine'; SM, BH, JKT *gundik,* MIN *gundiᵊ?;*
rubit, rubik 'torn, frayed'; SM *robek* (< JV according to Klinkert), JKT *ròbèk;*
tarit, tarik 'pull, make taut' < **tarik* 'pull' (3.1.2.3).

These lexemes must be loanwords from SAR, where **-k* as a rule became *t* after *i* (Collins 1987:Appendix 4).[112] Other lexemes which are less likely to be borrowed show final *k,* like *tasik, bilik, tarik* (3.1.2.3), *p/udik* (3.5.2), etc.

---

[112]SAR palatalises final nasals which are preceded by front vowels (Collins 1987:35). Collins gives two examples of palatalisation of final velar stops preceded by front vowels: *bait* 'good' (SM *baik*), and *nait* 'go up, ascend' (SM *nait*). In two other examples it is not shown, viz. *péndék* 'short' (SM *pendek*) and *tasék* 'lake' (SM *tasik*). According to Collins (pers.comm.) these lexemes are probably not inherited. Blust (who had fieldwork experience with SAR) regards IBN lexemes with palatalised final stops as loans from SAR (Blust pers.comm.).
N.B. The occurrence of Tagalog *ma-ba?it* 'good' is possible evidence that SAR *bait* is not the result of palatalisation, but has developed from a PWMP form with final **t*. On the other hand, it is also possible that the Tagalog root *ba?it* is a loan from a Malayic isolect from North Borneo.

For the above cognate sets I reconstruct:

gundik 'concubine';
itik 'duck';
rubik 'torn, frayed'.

UNEXPLAINED IRREGULAR CORRESPONDENCES PM *-k* > IBN (phonemic) -*ʔ*

(1) IBN -*ʔ* corresponds to SWY -*ʔ*, o.i. -*k*, in :

badiʔ 'dagger'; SM, BH *badik*, MIN *badiᵊʔ* (Van den Toorn *badiʔ*), SWY *badiəʔ*;
curaʔ 'colour-scheme, pattern of cloth'; SM *corak*, MIN *curaʔ, coraʔ*, SWY *curaʔ*;
baiʔ, in baiʔ...pin... 'whether...or...'; SM *baik in baik...maupun...* 'whether...or..., both...and...';
gəmuʔ 'fat, stout'; SM, JKT *gəmuk*, SWY *gəmuᵊʔ*;
saraʔ 'part, separate'; SM, BH *sarak*, SWY *saxaʔ* 'separated, divorced'.

In the case of SM *badik* etc., it seems that the IBN correspondence is more authentic than those of the other isolects: cf. Zorc's PHN *badiʔ* 'knife', and I reconstruct PM *badiʔ* 'dagger'. But for the other cognate sets I reconstruct:

curak 'colour-scheme, pattern of cloth';
gəmuk 'fat, stout';
sarak 'part, separate'.

(2) In one case BH and SWY have -*t* corresponding to IBN -*ʔ*, o.i. -*k*:

BH *garut*, SWY *gaxut/garut*, IBN *garuʔ* 'scratch'; SM, JKT *garuk* 'id.', MIN *garuᵊʔ* 'rough, hoarse'. But there is also SM *garu* 'scraping with a blunt point', *garut* 'scraping against one another (of two surfaces in contact)', MIN *gayuʔ* 'scratch', BH *garu* 'scratch'.

The various endings of these lexemes reflect their sound-symbolic value. The variants probably already existed in PM, and slightly different meanings were associated with each of them in each isolect. I reconstruct:

*garuk/*garut/*garuʔ* 'scratching'.

3.4.2.4 PM FINAL *ʔ* > IBN -*ʔ*, OTHER ISOLECTS ∅

Examples (taken from Zorc's evidence in 3.4.2b):

akuʔ 'confess'; BH n.c., IBN *akuʔ*, o.i. *aku*;
anuʔ 'something, someone, so-and-so' (3.6.1.2);
baruʔ 'hibiscus'; SM, MIN, BH, SWY *baru*, IBN *baruʔ*;
bAlaŋaʔ 'earthenware vessel' (cf. 3.1.3.3);
bukuʔ 'joint, node'; IBN *bukuʔ*, o.i. *buku*;
butaʔ 'blind'; SM *buta*, MIN, SWY *buto*, IBN *butaʔ*, JKT *butè*;
dakiʔ 'body dirt'; IBN *dakiʔ*, o.i. *daki*;
datuʔ 'head of a clan'; SM *datu/k*, BH *datu*, MIN *datuᵊ/ʔ* '(title used for a *paŋulu*); grandfather', JKT *datu/ʔ* 'grandfather', SWY *datuᵊ/ʔ* 'head of a *pasar* (a settlement of foreign people)', IBN *datuʔ* 'nobleman, chief' (cf. 5.4 N.B. for SM -/*k*, MIN, SWY, JKT -/*ʔ*);
dəpaʔ 'fathom'; SM *dəpa*, MIN *dapo*, BH *dapa*, SWY *dəpo*, IBN *dəpaʔ*, JKT *dəpè*;
hilaʔ 'keep an eye on'; IBN *ilaʔ* (no cognates in the other isolects, but cf. PHF *qilaʔ*);
isiʔ 'meat, contents'; IBN *isiʔ* 'id.', o.i. *isi* (cf. 5.7 lemma 103);

*jaja? 'hawk, peddle'; SM, BH jaja 'id.', SWY jajo 'gadabout', IBN (Zorc) bə-jaja? 'carry (goods)', (Richards 1981) daja?, jaja? 'hawk, peddle';

*jArami? 'rice stalk, straw'; SM, SWY, JKT jərami, MIN jarami, IBN jərami?;

*jawa? 'millet'; SM jawa, SWY jawo, IBN jawa?;

*jəra? 'warned by experience'; SM jəra, MIN jaro, BH jara, SWY jəxo, IBN jəra?;

*kaka? 'older sibling'; SM kaka/k, BH kaka, SWY kaka/?, IBN aka?, JKT kaka/ŋ 'id.', MIN kako, kaka/? 'older sister' (cf. 5.4 N.B. for SM -/k, MIN, SWY -/?, JKT -/ŋ, and loss of *k in IBN);

*kənu? '(quotative particle)'; IBN kənu? (cf. PHN *kenuh);[113]

*kita? 'we (incl.)' (cf. 3.2.3);

*ləŋa? 'sesame plant; its oil'; SM ləŋa, BH laŋa, IBN ləŋa?, JKT ləñè;

*m/uda? 'young, unripe'; IBN muda? 'id.', SM muda, MIN, SWY mudo 'young, unripe, light (of colour)' (cf. also 6.8);

*naŋka? 'jackfruit'; SM, BH naŋka, SWY naŋko, IBN naŋka?, JKT naŋkè;

*nasi? 'cooked rice'; IBN asi? (loss of *n- unexplained[114]), o.i. nasi;

*ŋilu? 'on edge, of teeth'; IBN ñilu?, SWY ŋilu, ñilu, o.i. ŋilu (IBN, SWY ñ unexplained: possibly palatalisation due to following i?);

*paku? (1) 'nail'; IBN paku?, o.i. paku;

*paku? (2) 'fern'; SM, MIN, SWY paku, IBN paku?;

*palu? 'hit, beat'; SM, MIN, SWY palu, IBN palu?;

*pari? 'ray (k.o. fish)'; SM, MIN pari, IBN pari?;

*pəñu? 'turtle'; SM, JKT pəñu, IBN pəñu? (MIN pañu 'seabird preying on carrion' is probably not a cognate);

*puki? 'vulva'; SM, BH puki, IBN puki?;

*əru 'casuarina'; SM (ə)ru, IBN ru?; also BRU aru (Prentice pers.comm.);

*sagu? 'sago'; SM sago, sagu, BH, SWY sagu 'sago pith', MIN, JKT sagu 'doughy substance in plants', IBN sagu? 'balls of cooked sago';

*sawa? 'python' (3.3.2);

*sida? 'they'; IBN sida? 'they', BH sida 'you (polite)' (Fudiat Suryadikara et al. 1981:196), cf. also KCI sidə, Bukit Malay sida[115] 'you (sg., polite)';

*təku? 'bend (v)'; IBN təku? 'id.', SM təku 'bent', JKT təkuk (-k unexplained) 'fold (v)';

*tiku? 'bend (v)'; IBN tiku? 'id.', SM tiku 'a curve, bend';

*tiŋa? 'food particles caught between the teeth'; IBN tiŋa? (cf. PHF *CĭŋaS, no cognates in the o.i.);

*tuli? 'ear troubles'; IBN tuli? 'having a suppurating ear', SM, BH, SWY, JKT tuli 'deaf';

*tuma? 'body louse'; SM, BH tuma, MIN, SWY tumo, IBN tuma?, JKT tumè?.

UNEXPLAINED IRREGULAR CORRESPONDENCES SM -k, MIN -?, BH -k, SWY -?, JKT -k/-? corresponding to IBN -?

In the following cases one or more isolects have -k (MIN, SWY -?) corresponding to IBN ?, o.i. ø:

---

[113] cf. SM kono/n, Tagalog kunú?, Malagasy hono 'it is said', JV kono 'the thing discussed' (Dempwolff 1938: "das Besagte"), Sa'a 'unu 'say, think' < PMP *kunu 'id.'.

[114] Prentice (pers.comm.) drew my attention to Timugon Murut inasi 'beer made of rice or root crop' and Kadazan nasi 'id.' (where Timugon Murut in- and Kadazan n- are regular allomorphs of the infix -in- with vowel-initial roots). He suggests that SM, MIN, BH, SWY and JKT n- in nasi may be the remnant of a PM affix *in-/*-in- (through weakening and loss of initial vowel). But as yet the Malayic isolects do not provide evidence for the reconstruction of such a PM affix.

[115] For Banjar Bukit Malay see Abdurachman Ismail et al. 1979:100.

(1) SM, BH *bali/k* 'reverse, reciprocate', MIN *baliᵊ/ʔ* 'rear-side, contrary, reverse; go back', SWY *baliᵊ/ʔ* 'go back, do for a second time, restore', JKT *bòlak/balik* 'vice versa, both ways'; cf. also SM, JKT *kəmbali* 'back (adv)', *uaŋ* -- 'change (n)', MIN *kumbali* 'back (adv)';

(2) SM *bera/k*, BH *bahira*, IBN *biraʔ*, JKT *bèra/k* 'defecate' (cf. 3.1.2.5);

(3) SM *buka/k, buka*, MIN, *buka/ʔ, buko*, BH *buka*, SWY *buka/ʔ*, IBN *bukaʔ*, JKT *bukè* 'open', SWY *buko* 'large';

(4) SWY *pala/ʔ*, SM *kəpala*, MIN *kapalo*, BH *kapala*, IBN *palaʔ*, JKT *kəpalè* 'head';

(5) SM *pinta/k, pinta*, BH *pinta* 'a request', MIN *pinta/ʔ, pinto*, IBN *pintaʔ*, JKT *pintè* 'ask, beg, request';

(6) SM *tabi/k*, MIN, SWY *tabiᵊ/ʔ*, IBN *tabiʔ*, JKT *tabé* '(term for) greeting';

(7) SM *tapak, t/əl/apak*, MIN *tapaʔ*, BH *t/al/apak*, SWY *tapaʔ, t/əl/apak*, IBN *tapaʔ*, JKT *t/əl/apak* 'sole, palm', furthermore BH *tapak* 'hit with palm of hand', and JKT *tapak* 'footprint'; Zorc considers IBN *tapaʔ* a reflex of PHF *dăpaH* (see above);

(8) SM *ti/da/k*, MIN *in/daʔ*, BH *ka/da*, IBN *ən/daʔ*, JKT *n/da/ʔ* 'no, not'.

SM *tabi/k* etc. (< SKT) and *ti/da/k* etc. are often used in isolation, one being a term for greeting, and the other a negation. Final *-k/-ʔ* are often used as vocatives. Vocatives, greetings, and negations have in common that they can form a sentence in themselves, and that they are used in isolation. A possible explanation of the final *-k/-ʔ* in *tabi/k* etc., *ti/da/k* etc., and kinship terms is that it was originally a syntactic marker for words occurring in isolation (cf. 5.4 N.B. for *-k/-ʔ* in kinship terms). For the cases 1-5 and 7 no (tentative) explanation is available. In all these cases IBN has -ʔ; in the other isolects no regular pattern can be found in the occurrence of *-k/-ʔ*. SM *kəpala* etc. is borrowed (from SKT). IBN *tapaʔ* as a reflex of PHN *dapaH* is problematic, because the initial *t* in IBN (and in the other isolects) does not match with PHF *d-*. Furthermore, IBN *tapaʔ* does not match with the other isolects, which all have *-k/-ʔ* corresponding to IBN -ʔ. On the basis of the above correspondences I make the following reconstructions:

*bali*ʔ 'reverse; go back';
*buka*ʔ 'open';
*pinta*ʔ/*pintak* 'a request; request (v)';
tapa(ʔ,k) 'sole, palm';
*da*ʔ 'no, not'.

3.4.2.5 PM FINAL *(ʔ) > IBN A DOUBLET, OR NO REFLEX, OTHER ISOLECTS ∅

In case of doubt PM *-(ʔ)* is reconstructed, i.e. when IBN has a doublet, when dipthongisation applies to the IBN correspondence, or when there is no IBN correspondence. Examples:

ada(ʔ) '(not) exist, there is (not)'; SM, BH *ada*, MIN (usually in texts) *ado*, JKT *adè* 'exist, there is', IBN *n/aday* '(there is) no, not';
buŋa(ʔ) 'flower, blossom'; SM *buŋa*, MIN, SWY *buŋo*, IBN *buŋay* (cf. 3.2.3), JKT *buŋè*;

*hulu(?) 'head; upper part; upriver; handle'; SM *(h)ulu*, BH *hulu*, MIN *ulu* 'hilt; begin, first one; upstream area', SWY *ulu* 'begin, first one; upstream area', IBN *ulu?* 'meaning', but *ulu* 'handle, upper part; upriver';

*huma(?) 'farm(land)' (3.2.3);

*mama(?) 'uncle (mother's brother)'; SM *mama/k* 'uncle/aunt on mother's side', MIN *mama/?* 'uncle on mother's side', SWY *mama/?*, JKT *mama/?* 'uncle' (cf. 5.4 N.B. for SM /k, MIN, SWY, JKT -/?);

*susu(?) 'breast'; IBN *tusu*, JKT *susu?*, o.i. *susu*, but also IBN *tusu?* 'suck';

*tuba(?) '(plant providing) fish poison'; SM, BH *tuba*, MIN, SWY *tubo*;

*tuha(?) 'old (of people)' (3.2.3).

UNEXPLAINED IRREGULAR CORRESPONDENCE

IBN has -*h* corresponding to -ø in the other isolects in: IBN *paah* 'thigh'; SM, BH *paha*, MIN *pao*, JKT *pahè* (-*h*- unexplained), for which set I reconstruct:

*paha(?) 'thigh'.

3.4.2.6 PM *-Vø > ALL ISOLECTS - Vø

In other cases given by Zorc ø is reconstructed (i.e. if IBN has a final vowel).

Examples:

*aku 'I'; SM, BH, SWY, IBN *aku*;
*dahi 'forehead'; SM, BH *dahi*, MIN *dai*, IBN (poetic) *dai*;
*ia '(3rd pers. sg.)'; SM, IBN *ia*, MIN *io*,[116] BH *hia*[117] (*h*- unexplained), JKT *iè* (cf. PAN *iøaø*);
*kami 'we (excl.)'; JKT n.c., o.i. *kami*;
*kuku 'claw, nail'; IBN *kuku* 'claw', JKT *kuku?*, o.i. *kuku* 'claw, nail';
*siku 'elbow, angle'; (JKT *sikut* < JV), o.i. *siku*;
*tali 'rope, string'; a.i. *tali*;[118]
*təbu 'sugarcane'; SM, SWY, IBN *təbu*, MIN *tabu*, JKT *təbu?*;
*tunu 'roast'; SM, SWY, IBN *tunu*.

3.5 THE PM VOICED STOPS

The voiced stops that are found in the isolects are *b*, *d*, *j*, and *g*. They occur initially, intervocalically, and post-nasally (see 3.6.2). They agree in all isolects in all positions, except for IBN *b* in the environment *a _ a* which, as a rule, corresponds to *w* in the other isolects. On the basis of these correspondences PM *b*, *d*, *j*, and *g* are reconstructed. The correspondence IBN *b*, o.i. *w* in the environment l *a _ a* is reconstructed as PM *b*: a change of PM *b* to SM, MIN, BH, SWY, JKT *w* (with a decrease of closure) is phonetically more plausible than a change of PM *w* to IBN *b*; furthermore, IBN *b* in this environment is a retention from PMP (Blust 1981:459).

---

[116]See Van der Toorn (where it is written 'jo' *(yo))*.
[117]Written 'hiya' in Abdul Jebar.
[118]Written 'tapi' in Abdul Jebar.

3.5.1 PM *b* > IBN *b*, OTHER ISOLECTS *w* | *a* _ *a*; ELSEWHERE *b* > ALL ISOLECTS *b*

Examples:

bulət'round'; SM, BH, SWY *bulat*, MIN *buleʔ*, JKT *bulət*;

bəriʔ'give'; SM, JKT *bəri*, MIN, BH *bari* 'give', IBN *bəriʔ* 's.th. given';

bintaŋ 'star'; a.i. *bintaŋ*;

bubu 'bamboo basket trap for fish'; SM, BH, IBN, JKT *bubu*;

ləbih 'more, surplus'; SM, IBN *ləbih*, MIN *labiᵊh*, BH *labih*, SWY *ləbiᵊ(h)*, JKT *ləbi*;

təbu 'sugarcane' (3.4.2.6);

t/um/buh 'grow'; SM, BH, IBN *tumbuh*, SWY *tumbuᵊ(h)*, JKT *tumbu*;

tambah 'increase'; SWY *tamba(h)*, JKT *tambè*, o.i. *tambah*;

gambar'picture, image'; MIN *gamba*, SWY *gambax*, o.i. *gambar*;

ləmbut'soft, weak'; SM, SWY, IBN *ləmbut*, MIN *lambuyʔ*, BH *lambut*;

kaban 'companion, follower; herd, group'; SM, JKT *kawan*, BH *kawal* (-*l* unexplained)[119] 'friend, supporter, companion', MIN *kawan* 'companion, slave', SWY *kawan* 'follower, comrade; shoal, herd, group', IBN *kaban* '(human) relation; group, company, shoal, herd';

taban 'carry off, take prisoner'; SWY *tawan/an* 'prisoner of war', IBN *taban* 'carry off, carry with', o.i. *tawan* 'take prisoner, detain';

laban 'adversary, rival; fight, oppose'; SM, MIN, BH, SWY *lawan* 'id.', IBN *laban* 1. 'opposed to, against'; 2. 'because, on account of, by', JKT *lawan* 1. 'adversary, enemy'; 2. 's.o. who joins in eating rice';

*bah/*babah* '(position) under, below'; SM, MIN, BH *bawah*, SWY *bawa(h)*, IBN *bah* (Richards 1981), JKT *bawè*; cf. also SAR *bah* (Collins 1987:83), KD *ka-babah* (Dunselman 1949:62) 'under, below'.

UNEXPLAINED IRREGULAR CORRESPONDENCES

In a few cases IBN does not reflect *b-* :

antuʔ 'come to s.o.'s assistance'; SM, MIN, BH, JKT *bantu*;
aris 'division, boundary' (along with *baris* 'row, line', *garis* 'line, boundary', and *taris* 'line'); SWY *baxis*, o.i. *baris* 'straight line';
əŋkaruŋ 'grass lizard'; SM *bəŋkaruŋ*, MIN *biŋkaruᵊŋ*, BH *biŋkaruŋ/an*; (Wilkinson 1959 also Kedah Malay *məŋkaruŋ*, *cəŋkaruŋ*);
isan 'relatives of child-in-law' < *ba/isa(a)n* (3.1.2, 3.1.2.5);
isik 'whisper'; MIN, SWY *bisiᵊʔ*, o.i. *bisik*;
ukay 'no, not; is not' (along with *bukay* 'other'); SM, MIN, SWY *bukan*, MIN *bukan, ukan*, JKT *bukan* (a loan? see below).

A possible explanation for this loss is backformation, that is to say, *b* was reinterpreted in IBN as a prefix *b(ə)-*, which yielded the current lexemes. Another explanation is that IBN reflects the original state, and that the other isolects contain a petrified prefix *b-*. This may be the case with *isik/bisik*, since this lexeme often has the intransitive verbal prefix (cf.IBN *b-isik*, SM *bər-bisik*, etc.). But it would not apply to SM *bəŋkaruŋ* or SM *baris* etc., these being nouns that refer to something concrete. SM *bukan* etc. has cognates in many languages outside the Malayic group. These cognates all reflect PMP *ə in the final syllable, cf. PMP *buken* (Blust 1970) and *beken* (Blust 1980). In view of these proto-forms I interpret JKT *bukan* as a SM loan. On the basis of the above cognate sets I reconstruct:

---

[119] An orthographic error? or a form corresponding with SM *kawal* 'guard, escort' (< TAM, Van Ronkel 1902:111)?

bantu? 'help, assist';
baris '(straight) line';
bəŋkaruŋ 'grass lizard';
(b)isik 'whisper (v)';
bukən 'no, not; other' (cf. OM *bukan* 'other', Çoedès 1930:39-40, 78).

3.5.2 PM *d > d

Examples:

dada 'breast, chest'; SM, BH, IBN *dada*, MIN, SWY *dado*, JKT *dadè*;
duri? 'thorn'; SWY *duxi*, IBN *duri?*, o.i. *duri*;
hiduŋ 'nose'; SM, BH *hiduŋ*, MIN *iduᵊŋ*, o.i. *iduŋ*;
m/andi? 'bathe'; IBN *pandi?* (with *p-* through backformation), o.i. *mandi*;
pindah 'move, change place'; SWY *pinda(h)*, JKT *pindè*, o.i. *pindah*;
tadi? 'just now; previously'; IBN *tadi?*, JKT *tadi, tadé*, o.i. *tadi*;
tanda 'sign, mark'; MIN, SWY *tando*, JKT *tandè*, o.i. *tanda*;
m/udi/k 'go upstream, go back against the current'; SM, BH *m/udi/k*, MIN, SWY *m/udiᵊ/?*, IBN *p/udi/k* (backformation) 'go upstream', JKT *m/udi/k* 'go south; go back to one's native village' (see 5.2); for *m/-* and *-/k* see 6.8).

UNEXPLAINED IRREGULAR CORRESPONDENCE

SWY has *g-*, IBN *ø-*, corresponding to o.i. *d-* in SWY *gaxi*, IBN *ari*, o.i. *dari* 'from'; for this I reconstruct:

dari 'from'.

3.5.3 PM *j > ALL ISOLECTS *j*

Examples:

jahət 'evil-hearted, bad' (3.1.1.5 IC);
jalan 'way, road' (3.2.3);
jual 'sell'; a.i. *jual*;
tajəm 'pointed, sharp'; JKT *tajəm*, o.i. *tajam*;
tujuh 'seven'; MIN *tujuəh*, SWY *tujuᵊ(h)*, JKT *tuju?* (3.4.2c), o.i. *tujuh*;
iñjəm 'borrow'; SM, SWY *p/iñjam*, BH *iñjam*, IBN *iñjaw*, JKT *p/iñjəm*;
tuñjuk 'show (v); index finger'; SM *t/əl/uñjuk*, SWY *tuñjuᵊ?*, JKT *t/əl/uñjuk* 'index finger', MIN *tuñjuᵊ?*, BH, IBN *tuñjuk* 'point out, indicate; finger', SM *tuñjuk* 'point out, indicate'.

N.B. In IBN regressive dissimilation to *d* took place if two subsequent syllables began with *j*, e.g.:

jañji 'promise, agreement'; IBN *dañji* 'agreement', JKT *jaŋji* (3.6.2), o.i. *jañji* 'promise, agreement, term';
jijir 'row, line'; SM *jejer* 'orderly line', IBN *nijir (dijir)* 'put in a row'.

3.5.4 PM *g > ALL ISOLECTS *g*

Examples:

*dagiŋ 'meat, flesh'; *MIN dagiᵊŋ*, (IBN *dagin* from SAR, see 3.6.3.3 IC) o.i. *dagiŋ*,
*gigi 'tooth'; a.i. *gigi*;
*gəliʔ 'eerie feeling'; SM, SWY, JKT *gəli*, MIN *gali*, IBN *gəliʔ*;
*saŋgul 'bun of women's hair'; SM, BH, SWY, IBN *saŋgul*, MIN *saŋguᵊ*;
*taŋgaʔ 'house ladder, staircase'; SM, BH *taŋga*, MIN, SWY *taŋgo*, IBN *taŋgaʔ*, JKT *taŋgè*.

3.6 THE PM NASALS

3.6.1 INITIAL AND INTERVOCALIC NASALS

In initial and intervocalic position four nasals agree in all isolects: *m, n, n,* and *ŋ;* on the basis of their correspondences PM *m, *n, *ñ,* and *ŋ* are reconstructed.

3.6.1.1 PM INITIAL AND INTERVOCALIC *m > ALL ISOLECTS m

Examples:

*ma/buk 'intoxicated, mad, excited'; MIN, SWY *mabuᵊʔ*, JKT *mabòk*, o.i. *mabuk*;
*ma/kan 'eat'; IBN *makay* (cf. 3.2.3), o.i. *makan*;
*ma-lə(hø)əm 'night'; JKT *maləm*, o.i. *malam* (cf. 5.1.4);
*kami 'we (excl.)' (3.4.2.6);
*ləmah 'soft, weak'; SM *ləmah*, MIN, BH *lamah*, SWY *ləma(h)*, IBN (only in songs) *ləmah*, JKT *ləmè*;
*rumah 'house' (3.1.1.3).

3.6.1.2 PM INITIAL AND INTERVOCALIC *n > ALL ISOLECTS n

Examples:

*naik 'go up, ascend'; SM, BH *naik*, SWY *naiᵊʔ*, IBN *tik/iʔ* (3.1.3.2 N.B.), JKT *naèk*;
*nanah 'pus'; SWY *nana(h)*, JKT *nanè*, o.i. *nanah*;
*nibuŋ 'k.o. palm'; SM, SWY, IBN *nibuŋ*, MIN *nibuᵊŋ*;
*anuʔ 'something; someone, so-and-so'; IBN *anuʔ*, JKT *anu, anó*, o.i. *anu*;
*kənaʔ 'hit, affected; right, suitable'; SM *kəna*, MIN *kanay* (-*y* unexplained), BH *kana*, SWY *kəno*, IBN *kənaʔ*, JKT *kənè*;
*panas 'solar heat'; MIN *paneh*, o.i. *panas*.

IRREGULAR CORRESPONDENCES palatalisation of +*n* preceding +*i*

MIN usually has a doublet with a palatal nasal in lexemes with *n* preceding *i* or *iᵊ*, e.g.

bañiᵊh, baniᵊh '(plant-)seed' < *bənih (3.1.2.3);
ñiñiᵊ/ʔ, niniᵊ/ʔ 'grandmother' < *niniʔ (3.1.2.1);
kañiᵊŋ, kaniᵊŋ '(eye-)brow' < *kəniŋ (3.1.2.3);
barañi, barani 'daring; dare, venture'; SM, IBN, JKT *bərani*, SWY *bəxani*.

In one case there are variants with a palatalised and a non-palatalised nasal not directly preceding *i/iᵊ*:

ñamiᵊh, namiᵊh 'almost, all but' (no correspondences).

In *ja/ño, ja/no* 'says she/he' and *ti/ño, ti/no*[120] 'all right, let it be' no palatalisation is involved: these doublets reflect two variants of the MIN possessive suffix (-*ño* and -*no,* which themselves are reflexes of PM *-*ña* (see 5.5.1.3)). In a few other cases SWY has *n* before *i* corresponding to doublets with *n* and *ñ* in the other isolects:

niux 'coconut'; MIN *ñiuᵊ, niuᵊ*, SM, BH, IBN *ñiur*;
buni 1. 'sound' 2. 'hidden; hide'; MIN *buñi, buni* 1. 'sound' 2. 'not straight (in talking)', SM *buñi* 1. 'a sound' 2. 'hidden', BH *buñi* 'sound, voice', IBN, JKT *buñi* 'sound, noise'; (also SM, JKT *səmbuñi*, MIN *sambuñi, sambuni* 'hide; hidden');
suni 'quiet, empty, unoccupied, lonely'; MIN *suñi, suni*, SM, BH, IBN *suni.*

In one case MIN has doublets with *ñ* and *n* corresponding to a palatal nasal in the other isolects:

kuñiʔ, kuniʔ 'turmeric'; SM, SWY, IBN, JKT *kuñit*; cf. also BAC *kuñit* 'yellow'.

For all the above cases I reconstruct an alveolar nasal. Palatalisation of alveolar nasals before a high front vowel is much more probable than the change from *ñ* to *n* in this position. Moreover, Dempwolff (1938) reconstructed *n* in PMP lexemes on the basis of SM *bənih, nene/k, kəniŋ, bərani, ñiur, buñi* (2. 'hidden; hide'), and *kuñit*. He reconstructed PMP *buñi* 'announcement' on the basis of SM *buñi* 'sound' and two other reflexes which however are not valid: Tagalog *bunyi* 'celebrity' (a loan, as he indicated himself; probably from a Malay isolect; cf. also Wolff 1976:357) and Ngaju *nambuñi* 'announce confidentially' whose meaning is closer to *buñi* (2) than to *buñi* (1). Therefore it is very weak evidence for PMP *buñi*. The following reconstructions are made on the basis of the above cognate sets:

bArani 'daring; dare';
buni 's.th. that is not seen (s.th. hidden, or a sound)';
kunit 'turmeric; yellow' (cf. 5.7 lemma 150);
niur 'coconut';
suni 'quiet, deserted'.

UNEXPLAINED IRREGULAR CORRESPONDENCES

IBN does not reflect *n-* in:

asiʔ 'cooked rice' < *nasiʔ* (3.4.2.4).

For IBN *iniʔ* < *niniʔ*, see 5.4.2.

3.6.1.3 PM INITIAL AND INTERVOCALIC *ñ* > ALL ISOLECTS *ñ*

Examples:

ñaman 'pleasant, nice, comfortable'; SM, BH, JKT *ñaman*, IBN *ñamay* (cf. 3.2.3);
ñamuk 'mosquito'; MIN, SWY *ñamuᵊʔ*, o.i. *ñamuk*;
ñawa 'life, soul' (3.3.2);
añam 'weave, plait'; MIN *añam, ayam*, BH *ayam* (MIN, BH *y* unexplained),[121] o.i. *añam*;
kəñaŋ 'satisfied (hunger)'; MIN, BH *kañaŋ*, o.i. *kəñaŋ*;

---

[120]Short for *ka-andaʔ ati-ño-lah* 'as she/he pleases, whatever she/he likes' (literally 'the desire of her/his heart').
[121]Due to regressive dissimilation of nasality? cf. the substandard pronunciation [sampiyon] of DU *champignon* 'mushroom' (< French).

*taña?'ask'; SM taña, MIN, SWY taño, IBN taña?, JKT tañè.

3.6.1.4 PM INITIAL AND INTERVOCALIC *ŋ > ALL ISOLECTS ŋ

Examples:

*ŋaŋa(?) 'agape'; SM, BH ŋaŋa, MIN, SWY ŋaŋo, JKT ŋaŋè;
*ŋəri? 'afraid, panic-stricken'; SM, JKT ŋəri, MIN ŋari 'id.', IBN ŋəri? 'downhearted';
*ɲilu? 'on edge, of teeth' (3.4.2.4);
*laŋit 'sky'; MIN laŋi?, o.i. laŋit;
*haŋət '(non-solar) heat'; SM, BH haŋat, MIN aŋe?, SWY aŋat, JKT aŋət 'id.', IBN aŋat 'heat (also solar)';
*səŋət 'sting (of venomous insect)' (3.1.1.2).

3.6.1.5 EXCRESCENCE[122] OF VOICED STOPS IN IBN

IBN has developed a homorganic voiced stop after intervocalic nasals followed by a vowel + final r (i.e. in lexemes with an original +CVNVr-structure),[123] e.g.

bandir 'buttress of a tree'; SM, BH banir, MIN baniᵊ, SWY banix;
baŋgar 'rotten, rotting'; SM baŋar 'putrid', SWY baŋax 'rotten smell';
bəndar 'true, real'; SM bənar, MIN bana, BH banar, SWY bənax, JKT bənər;
dambar 'resin'; SM, BH damar, MIN dama, SWY damax;
iŋgar 'loud'; SM, BH iŋar, SWY iŋax;
jəmbuy, jəmbi 'expose to the sun' < *jəmur (3.2.3 N.B.2);
juŋgur 'snout'; SM juŋur, joŋor, MIN juŋuᵊ 'id.', SWY juŋux 'upperlip';
laŋgir 'tree bearing edible fruit; its bark (used for soap)'; SM laŋir 'k.o. shrub used for shampoo', SWY laŋix 'all that is used to wash one's hair';
səmbar 'a splint, fish (for strengthening posts etc.)'; SM səmar 'piece of wood lashed to a carrying-pole to strengthen it'.

A few lexemes have not undergone this development: timur 'east' (a loan, see 5.2.1), umur ((<SM?) < AR), amur 'dust, mud', jəŋur/jəŋur 'look with an expression of dislike', lumur 1. 'anoint, smear' 2. 'number' (lumur (2) < DU nummer?), pamur 'cloudy, of water', and raŋor (Bruggeman, see fn. 83 in this chapter) 'collide'. Excrescence of voiced stops also occurs in other isolects, but not on a regular basis (cf. 6.3.7 last N.B., and Adelaar 1988).

The following reconstructions are made on the basis of the above cognate sets:

*banir 'buttress of a tree';
*baŋar 'putrid';
*bənər 'true, real';
*damar 'resin';
*iŋar 'loud';
*juŋur 'snout';
*laŋir 'tree bearing edible fruit; its bark (used for soap)'.

---

[122]The term 'excrescence' is from Anttilla (1972:68), who assigns it to the origin of voiced stops in, for example, English finger, number, thunder, etc.
[123]This is the only environment in which excrescence can be observed. I did not find the same phenomenon in IBN lexemes of a CVNVl-structure.

3.6.1.6 HOMORGANIC NASAL SUBSTITUTION FOR STOPS IN ANTEPENULTIMATE SYLLABLES

In antepenultimate syllables stops are sometimes (and in IBN as a rule, cf. 3.11b) weakened to their homorganic nasals under the influence of a following nasal, e.g.

MIN *binataŋ, minataŋ*, SM, BH, SWY, JKT *binataŋ* 'animal' (3.1.3.1 UIC); **b/in/antu* 'child-in-law', MIN *binantu, minantu*, SM, IBN *mənantu*, BH *minantu*, SWY *nantu* (3.1.3.1), cf. PMP **binantu*.

This phenomenon is also shown in the correspondence set MIN *kamanakan* 'relative, kin, family, esp. uterine heirs under Minangkabau law' (Wilkinson 1959), BH *kamanakan*, JKT *kəpònakan* 'one's sibling's children', but *kəpònakan* is a loan (according to Wilkinson 1959), and no reconstruction is made.

3.6.2 PRECONSONANTAL NASALS

Preconsonantally, nasals occur before stops or *s*. As a rule, they are homorganic to a following stop. However, there are some exceptions in JKT, where heterorganic nasal + stop clusters are found; the nasal is usually a velar one. These clusters could occur in lexemes which originally (i.e. in pre-PM) had heterorganic clusters, e.g.

diŋdiŋ (only one variant; PMP **DiŋDiŋ*) 'wall';
buŋbun, bumbun (PMP **buŋbun*) 'cylindrical vessel made of a joint of bamboo';
juŋjuŋ, juñjuŋ (PMP **zuŋzuŋ*; SUN *juŋjuŋ*) 'carry on the head';
bunbunan, bumbunan (PMP **bunbun*) 'fontanelle'.

But in other cases they have developed from homorganic consonant clusters, as in some loanwords and inherited lexemes, e.g.

dəŋdè, dəndè (< SKT; SUN *dəŋda*) 'punishment';
təŋtu, təntu (< Old Javanese; SUN *taŋtu*) 'certain(ly)';
toŋton, tonton (< JV; SUN *toŋton*) 'watch (a movie, play)';
jaŋji (only one variant; PMP **zanzi*; SUN *jaŋji*, JV *jaŋji, jañji*) 'promise, agreement'.

As can be seen from the above examples, most of the forms with these clusters have variants with a cluster consisting of a homorganic nasal + stop. Furthermore, the same phenomenon is seen in SUN and, to a lesser extent, in JV. Nothofer (1975:99, 194) reconstructs PMJ **-ŋC-* on the basis of these languages. It is not inconceivable that these JKT heterorganic consonant clusters are a retention from a stage prior to PM (and hence should be reconstructed for PM), but for the time being I prefer to interpret them as an innovation. There are several reasons for this. Firstly, these JKT clusters are sometimes the result of a tendency to dissimilate nasals in homorganic consonant clusters, as can be seen in loanwords like *dəŋdè/dəndè, təŋtu/ təntu*, or in inherited lexemes like *jaŋji*. Secondly, if JKT heterorganic clusters were a retention, PMP reduplicated root morphemes like e.g. **gemgem* 'make a fist' and **DemDem* 'keep quiet' would have become JKT *+gəmgəm* and *+dəmdəm*. But *gəŋgəm* 'hold tight' and *dəndəm* 'long for; grudge' are found instead, and also, the heterorganic clusters almost invariably have a velar nasal, and rarely *m* or *n*. Thirdly, most of the lexemes that exhibit a heterorganic cluster have a variant with a homorganic one. The variants with a heterorganic cluster may be loans from neighbouring languages, or

otherwise, the tendency to velarise the nasal may be due to an areal feature in West Java. This feature would first have originated in SUN, and would gradually have spread in neighbouring languages. Finally, JKT is the only isolect which exhibits heterorganic nasal clusters: the other isolects unanimously show homorganisation of the nasal (except for -ŋs-). The ultimate decision as to whether these heterorganic clusters have to be interpreted as retentions or as innovations depends on more insight into the sociolinguistic situation in JKT.

On the basis of JKT *m, ŋ*, other isolects *m*, I reconstruct PM **m* before a homorganic stop (**b* or **p*). See 3.4.1.1 and 3.5.1 for examples. On the basis of JKT *n, ŋ*, o.i. *n*, I reconstruct PM **n* before **d* and **t*. See 3.4.1.2 and 3.5.2 for examples. On the basis of JKT *ñ, ŋ*, o.i. *ñ*, I reconstruct PM **ñ* before a homorganic stop (**j* or **c*). See 3.4.1.3 and 3.5.3 for examples. On the basis of *ŋ* in all isolects, I reconstruct PM **ŋ* before a homorganic stop (**g* or **k*). See 3.4.1.4 and 3.5.4 for examples. *s* is usually preceded by *ŋ* in SM, BH, SWY and JKT, and by *n* in MIN and IBN.[124] Since these clusters probably reflect a PAN cluster **-ŋs-*,[125] I reconstruct PM **-ŋs-* on the basis of this correspondence. See 3.8.1 for examples.

UNEXPLAINED IRREGULAR CORRESPONDENCES

In the following cases the isolects disagree in showing prenasalisation:

SWY *cuŋkup*, o.i. *cukup* 'enough, complete' < **cu(ŋ)kup* (3.4.1.3);
SM *muntah*, MIN, IBN *mutah*, SWY *muta(h)*, JKT *muntè* 'vomit (v)';
SM *otak*, BH *utak*, SWY *otaʔ*, IBN *untak*, JKT *òtak* (*a* unexplained; 3.1.1A) cf. PMP **u(n)tek* 'id.';
IBN *səmpaʔ* (-ʔ unexplained, cf. 3.9.3UIC), SM *səpah*, MIN, BH *sapah*, SWY *səpa(h)* 'betel cud';
BH, IBN *undaŋ*, o.i. *udaŋ* 'prawn, crayfish';
SM *məntah*, MIN *matah*, BH *mantah*,[126] SWY *mata(h)*, IBN *mataʔ* (3.9.3UIC), JKT *məntè* (in Wilkinson 1959 also Kedah, Negeri Sembilan *matah*) 'uncooked, unripe'.

These cognate sets yield the following reconstructions:

*m/u(n)tah 'vomit (v)';
*undaŋ 'prawn, crayfish';
*u(n)tək 'brain';
*sə(m)pah 'betel cud';
*mə̆ntah/*m/atah 'raw, unripe'.

3.6.3 FINAL NASALS

In final position *m, n* and *ŋ* occur in all isolects. None of the isolects has final *ñ; n* and *ŋ* agree in all isolects, and PM final **n* and **ŋ* are reconstructed. Final *m* agrees in all isolects but MIN, where lexemes only have *-m* when it is preceded by *a*; otherwise they have a corresponding *-n*. I reconstruct PM final **m* on the basis of all isolects *-m*, and furthermore on the basis of MIN *n*, other isolects *m*, if the preceding vowel is *i, e, u* or *o*.

---

[124] According to Moussay (1981:23) *n* is an apico-dental. (Moussay (1981:25) describes MIN *s* as an alveolar, and so does Asmah (1964:1) for IBN *s*.) It is also noteworthy that Moussay (1981:32) has -ŋs- clusters corresponding to -ns- clusters in Van der Toorn and Thaib, cf. *baŋso* 'people' (Moussay) vs *banso* 'id.' (Van der Toorn, Thaib). Moussay's description is based on the MIN isolect of Padang.
[125] cf. Blust (1970:111 + nn.9-10) and Dahl (1976:99-100).
[126] In Abdul Jebar 'mantan' is found, which is probably a misprint.

3.6.3.1 PM FINAL *m > MIN n, OTHER ISOLECTS m | (i,e,u,o) _#; ELSEWHERE *-m > OTHER ISOLECTS m

Examples:

*bəlum 'not yet'; SM bəlum, MIN balun, alun, BH balum, SWY bəlum, lum, JKT bəlòm, bəlòn;
*kirim 'send' (3.1.2, 3.1.2.1; cf. also 4.4);
*ənəm 'six'; SM, SWY, IBN ənam, MIN, BH anam, JKT ənəm;
*tanəm 'plant (v)'; JKT tanəm, SM, MIN, BH, IBN tanam;
*hitəm 'black, dark'; SM hitam, MIN, SWY itam, JKT itəm.

3.6.3.2 PM FINAL *n > ALL ISOLECTS n

Examples:

*bulan 'moon, month' (3.1.2.2);
*simpən 'keep, store, save'; JKT simpən, SM, MIN, BH, IBN simpan;
*aŋin 'wind' (3.1.2.3);
*əmbun 'dew';
*puhun 'stem, origin, basis; beg, ask (forgiveness)' (3.1.2.4).

3.6.3.3 PM FINAL *ŋ > ALL ISOLECTS ŋ

Examples:

*linduŋ 'shaded, protected, covered'; SM, BH, SWY, JKT linduŋ 'id.', MIN lindu³ŋ 'shadow, protection against the sun'; 'shaded, protected, covered', IBN linduŋ 'shaded, screened, covered';
*ulaŋ 'repeat'; SM, BH, SWY ulaŋ 'repeatedly; repeat', MIN, IBN ulaŋ 'repeat';
*dindiŋ 'wall, partition'; MIN dindi³ŋ, JKT diŋdiŋ, o.i. dindiŋ.

IRREGULAR CORRESPONDENCES

In a number of cases IBN exhibits -n corresponding to -ŋ in other isolects. In all these cases IBN -n is preceded by i. This points to borrowing from SAR, where original final nasals preceded by i became alveolar (Collins 1987:35; cf. also 3.4.2.3IC).

Examples:

dagin 'meat, flesh' < *dagiŋ (3.5.4);
kañcin 'button, bolt, lock' < *kañciŋ (3.4.1.3);
pusin 'turn around'; SM, BH pusiŋ, MIN pusi³ŋ 'id.', JKT pusiŋ 'be dizzy, have a headache';
antin 'earring'; SM antiŋ, JKT antiŋ/antiŋ 'pendent, hanging down and swinging'; also SM, BH, SWY, JKT antiŋ/antiŋ, MIN anti³ŋ/anti³ŋ 'earring';
kambin 'goat'; SM, SWY, JKT kambiŋ, MIN kambi³ŋ;
guntin, guntiŋ 'scissors'; MIN gunti³ŋ, o.i. guntiŋ;
dacin 'steelyard'; SM, BH, SWY daciŋ, MIN daci³ŋ, JKT dacin (-n unexplained);
sətukin 'stocking', and bəlakin 'tar' (from 'blacking') are from English.

IBN usually has -*iŋ* corresponding to PM *-*iŋ* (in lexemes which are not suspect of borrowing), cf. *dindiŋ* < *dindiŋ*, *kəniŋ* < *kəniŋ* (3.1.2.3), and *kəliliŋ* < *kuliliŋ* (3.1.3.1). On the other hand, *dacin, kambin, sətukin,* and *bəlakin* are obviously loanwords, *dacin* being from Chinese, *sətukin* and *bəlakin* from English, and *kambin* referring to a domestic animal among coastal Malays, and not among the Ibans (Richards 1981). These loanwords must have come into the language through SAR. Compare also doublets *pirin* 'saucer' and *piriŋ* 1. 'an offer' 2. 'distribute', the former of which has the same meaning as SM, BH, SWY, JKT *piriŋ*, MIN *piriᵊŋ* (and probably also the same meaning as a supposed SAR *pirin?*). *Piriŋ* maintained the final velar, but has a quite divergent meaning. *Pirin* must be borrowed from SAR, and *piriŋ* with its divergent meaning must be the regular cognate of SM *piriŋ*.

Reconstructions:

antiŋ 'pendent, hanging down and swinging';
guntiŋ 'shears, scissors';
piriŋ 's.th. to offer with, saucer';
pusiŋ 'turn around'.

3.7 THE PM LIQUIDS

Three liquids are found: all isolects have *l* and *r*, and SWY also has *x* (2.4.1). *l,* which occurs in inherited lexemes in initial, intervocalic, and final position, agrees in all isolects in initial and intervocalic position. In final position it agrees in all isolects except MIN, which has a corresponding ø. But if a corresponding MIN form is suffixed with -*i* or -*an,* a morphophonological -*l*- reappears (2.2.2 B). PM initial and intervocalic *l* is reconstructed on the basis of *l* in all isolects, and PM final *l* is reconstructed on the basis of SM, BH, SWY, IBN, JKT -*l,* MIN -ø (or -*l*- at morpheme boundaries). Initial and intervocalic *r* agrees in all isolects but SWY, which has a corresponding *r* and *x*. Final *r* agrees in all isolects but MIN and SWY. MIN has a corresponding ø, and SWY a corresponding *x* and *r*. If a corresponding MIN form is suffixed with -*i* or -*an,* a morphophonemic -*r*- reappears (2.2.2 B). As to SWY *x* and *r,* the question arises whether they are reflexes of proto-phonemes that have merged everywhere else in the isolects, or whether their distinction is an innovation. In the latter case they may be the result of a split, or one of them is a loan phoneme. For PMP, a velar *R* and an apical trill *r* were reconstructed. If these proto-phonemes really existed (which recent research has been making increasingly doubtful), they must have merged in the Malayic isolects. At any rate, SWY *x* and *r* do not reflect the distinction between *R* and *r*: it turns out that *x* is inherited, and that *r* is a loan phoneme.

Compare the following list which contains all PMP lexical items with *R* and *r* from the 200-item basic wordlist (Blust forthcoming) that have a SWY reflex (where the SWY reflex differs semantically from PMP, the meaning of this reflex is added next to it):

	PMP	SWY	meaning
2.	*ka-wiRi	kiri	left side
23.	*DaRaq/*DaReq	daxa(h)	blood
25.	*liqeR	liax	neck
41.	*kaRat	kaxat gəxman	bite (v)
		'gnash one's teeth'	
44.	*deŋeR	dəŋar/dəŋax	hear
48.	*(ma-)tiDuR	tidur	sleep (v)
52.	*DiRi	dixi	stand (v)
		'(2nd pers. sg., reflexive pronoun)'	
54.	*(ma-)Ruanay	moanay	man, male
61.	*Rumaq	rumah	house
64.	*tutur	tutux	say (speak, utter)
68.	*ZaRum	jaxum	needle
78.	*taRaq	taxa(h)	cut, hack
87.	*baReq	baxa(h)	swell (an abscess)
88.	*peReq	pəxa(h)	squeeze
98.	*q(a,i)teluR	təlux	egg
105.	*ikuR	ikuə?	tail
106.	*ulaR	ulax	snake
112.	*buRuk	buruə? 'rotten (egg)'	rotten (in general)
		buxuə? 'rotten, gone, in ruins, worn out'	
115.	*akaR/*wakaR	akax 'root, creeper'	root
	*uRat	uxat 'tendon, nerve'	root
122.	*wahiR	ayiə?	water
123.	*aliR/*aluR	alur/an 'gully, fairway' (a loan)	flow (v)
131.	*Rabun	xabun 'not seeing very well'	cloud (not a raincloud)
142.	*beReqat	bəxat	heavy
149.	*ma-iRaq	mira(h) (a loan? cf. 3.1.2(d) N.B.2)	red
153.	*(ma-)Raya	rayo	big
166.	*bener	bənax	true, correct
168.	*waRi	axi	day

It appears that for the 25 proto-lexemes with *R, SWY has 16 reflexes with x, four with r, three with ø, and two with variants with x and r. For the two PMP proto-lexemes with *r SWY has only reflexes with x. In other words, it seems that most inherited SWY lexemes have x, and the SWY distinction between x and r does not reflect the distinction of PMP *R and *r.[127] This has already been pointed out by Prentice and Hakim Usman in their discussion of KCI h and r: the lexical distribution of KCI h and r does not match that of PMP *R and *r, nor does the lexical distribution of SWY x and r (Prentice and Hakim Usman 1978:131-132). From this and on the following grounds it must be concluded that the SWY distinction is not inherited:

---

[127]An additional difficulty with the comparison of SWY x/r and KCI h/r with PMP *R/*r is that many reconstructions with *r are actually based on borrowings (Wolff 1974), and although in some cases this is quite evident (cf. many of Dempwolff's reconstructions on the basis of SM and Ngaju, Dyen 1956), it is not so in many other cases. A thorough study of borrowing among Austronesian languages would give a better insight into the PMP *R/*r distinction.

(a) More than 135 known loanwords in Helfrich have *r,* whereas only four of the SWY lexemes known to me as loanwords have *x*[128] viz. *pikix* 'think', *kəxtas* 'paper', *kabax* 'news' (all from AR), and *sətəxo* 'silk' (< SKT). Also, in many cases there are minimal pairs with *x* and *r* with related meanings, and of which the member with *r* agrees in meaning with a SM correspondence, e.g.

xapat 'often, over and over' and *rapat,* SM *rapat* 'closely packed, fitting closely';
xəba(h) 'hanging down, as a full rice ear' and *rəba(h),* SM *rəbah* 'fall down (of heavy things)';
xampas 'cut the grass very short (with a particular k.o. knife)' and *rampas,* SM *rampas* 'rob, take with force';
xañjaw 'construction for catching tigers' and *rañjaw,* SM *rañjaw* 'caltrop'.

Furthermore, *x* and *r* are often not distinguished (Helfrich 1904:99), and in Helfrich's wordlist there are many variants (e.g. *dəŋar/dəŋax, baxaŋ/baraŋ,* etc.).[129]

(b) Three SWY affixes have an *x* (viz. *bəx-, pəx-* and *təx-*), whereas none has *r.*

(c) PMP **-D* is reflected as SWY *-x* (MIN *-ø/(-r-),* o.i. *-r*), e.g. PMP **bayaD* 'pay' > SWY *baix,* and **wakaD* 'root' > SWY *akax* 'root, creeper' (see 7.1e). For a change from a retroflex apical stop to a velar or uvular fricative one would expect the intermediate stage of an apical trill (i.e. PMP **-D* > *+r* > SWY *x*); if SWY *r* (which also occurs in final position) were inherited, one wonders why it did not change into *x,* whereas PMP **-D* did.

(d) In contradistinction to *x, r* often co-occurs with (non-inherited) *-k* (cf.4.3.2).

As to the origin of SWY *r,* it is interesting to compare it to the apical trill in KCI and Lampung. As in SWY, in these isolects *r* occurs mainly in loanwords, whereas KCI *h* and Lampung *x* belong to the inherited phonemes (Prentice & Hakim Usman 1978:129-132; Walker 1976:3-4). Walker says about *r* in Lampung as spoken in Way Lima:

> The trill r occurs in unassimilated loanwords. Most of these have an alternate in which r is replaced by x. Speakers differ as to how much r is used in the Indonesian words which are freely used in Lampung contexts. Some speakers reproduce the Indonesian r; others automatically substitute x in almost every instance...

It is very likely that the south Sumatran languages originally had a fricative (which became KCI *h/-ø* by subsequent decrease in closure) and that the apical trill was introduced through borrowing from SM and/or other isolects. In this process some inherited SWY and KCI forms may also have been replaced by SM forms containing *r.* This resulted in the confusing picture found in present-day SWY and KCI, where a few frequently used loanwords have acquired the inherited reflex (e.g. SWY *pikix* 'think'; KCI *kahay?* 'friendly, intimate' (< AR *qari:b*)), and where SWY *x* and KCI *h* may have been obligatorily or optionally replaced by an apical trill in some lexemes (e.g. SWY *dəŋax/dəŋar* 'hear'; *karam* 'sink'; KCI *kamahò/kamarò* 'drought'; *jarèwŋ* 'needle'). In this way are also explained cases like *xurut/rurut* 'pull, jerk (v)', where *x* and *r* occur simultaneously within one lexeme.

---

[128]There is also *axaŋ* 'charcoal' which must be borrowed because it does not reflect *d* for PMP **j* in PMP **gajeŋ,* but as all other Malayic isolects show a corresponding *r* instead of *d, axaŋ* must be a very early loan.
[129]Moreover, Helfrich himself is not always consistent about writing *x* and *r,* cf. (p.79) *kirim* and (p.212) *kixim* 'send (s.th.)'.

TABLE 7: THE DISTRIBUTION OF SWY *r*, SWY *x*, KCI *r*, AND KCI *h*

SWY	KCI	(language of origin)	meaning
x : r			
gəx/man	garmeŋ	PMP *geraqam	molar tooth
kuxus	kurawh	PMP *kurus	thin (people)
jəxami	jaramoy	PMP *ZaRami	straw, stubble
təxus	tərawh	PMP *terus	straight, through
jaxum	jarèwŋ	PMP *ZaRum	needle
r : h			
kənduri	kanuhay	Persian	religious feast
kərip (relative)	kahayʔ	AR	friendly, intimate
mura(h)	muhah	JV	cheap
karam	kahaŋ	PMP *kaRem	sink (v)
r/x : h			
cəray/cəxay	cəhè		separate (v)
kirim/kixim	kihayn	PMP *kirim	send
x : h			
kəxiŋ	kəhayn	PMP *keriŋ	dry
pikix	pikè	< AR	think
sətəxo	sutò	< SKT	silk
uxaŋ	uhaŋ	PMP *uRaŋ	human being
axi	ahay	PMP *waRi	day
r : r			
timur	timur	PMP *hatimuR	east (probably < SM, cf. 5.2.1)
guriŋ	guroyn		fry
buruŋ	burewŋ		bird
piraʔ	piraʔ		silver
suaro	suarò	< SKT	voice

Dempwolff gives PMP *buruŋ* 'bird' and *pirak* 'silver', but these are doubtful reconstructions (cf. also Adelaar 1989).

The fact that *x* and *r* are often interchanged makes one refrain from drawing a sharp line between inherited lexemes and loanwords along the distinction of *x* and *r*. Although, generally speaking, lexemes containing *r* are likely to be loanwords, they should not a priori be considered as such and be discarded as a basis for lexical reconstruction. The occurrence of SWY *x* and *r*, and their casual interchangeability should rather be seen as the effect of a sound shift at work (in this case from a velar fricative to an apical trill).

In initial and intervocalic position SWY *x/r* agrees with *r* in the other isolects, and on the basis of this correspondence I reconstruct PM initial and intervocalic *r*. Lexeme finally, SWY *x/r* agrees with MIN *-ø/* (-*r*- at morpheme-boundaries), other isolects *r*, and on the basis of this correspondence I reconstruct PM *-r*.

PM *r* was a (velar or uvular) fricative. Collins (1986b:181-183) observes that PMP *R* (which was probably a velar fricative) and *r* have merged and are reflected as a fricative in the peninsular Malayic isolects; isolects outside the peninsula mostly have an apical trill.

According to Collins it is most probable that the Malayic isolects outside the peninsula fronted an original velar fricative to an apical trill, whereas the peninsular Malayic isolects remained more faithful to the original pronunciation. This fits in well with the fact that SWY *x* is the regular reflex of PMP **R* and **r*. The remark by Collins about the majority of the non-peninsular isolects having an apical trill as a reflex of PMP **R/*r* is a little overstated. In fact almost all local Malayic isolects of Sumatra have a fricative. Fokker (1895:27-28), who made a study of the SM phonology,[130] recorded a "guttural r" for the area where he did his research (West Borneo). Collins himself recorded a velar fricative for SAR, and he regards this sound as a retention from 'Proto Malay' (Collins 1987:41-42). Fokker (1895:27-28) furthermore mentioned (quoting Klinkert) a guttural *r* in the Riau archipelago, and (referring to Helfrich) the *x* and *r* in Middle Malay.[131] A uvular *r* is reported for Deli Malay (spoken in and around Medan, Sumatra) by Roolvink (1953:6).

In Van der Toorn and Moussay's descriptions, MIN has an apical *r*, but Tamsin Medan (1980:89-90, 153-155) showed that 16 out of the 25 regions into which he divided the MIN area have a velar fricative, against nine with an apical trill. The regions where apical *r* occurs are coastal areas, and the extreme north. The MIN isolect of Koto Gadang, which formed the basis of Van der Toorn's study, is spoken in one of the areas that have an apical trill (the area of Koto Tuo in Tamsin Medan's dialectography).

So, apparently coastal Malayic isolects of western and northern Borneo and most Malayic isolects of Sumatra also have a fricative as the regular reflex of PMP **R* and **r*. (BH has an corresponding apical trill, and has a fricative in some sound-symbolic lexemes: Durdje Durasid pers.comm.). That Bahasa Indonesia has an apical trill instead finds its explanation in the fact that for quite a long period it has been a lingua franca in large areas outside its native domain (especially Java). Most of its recent developments have been taking place in these areas. Speakers of BI usually have another language as their mother tongue, and the most important of these, JV and SUN, have an apical trill. Many eastern Indonesian Malayic isolects also have an apical trill. With the exception of BAC, they also are lingue franche in areas where they are not native.

SWY *x/r*, SM, JKT *r* is also found as first component of a consonant cluster, as is seen in Table 8. The examples in this table show that lexemes with a preconsonantal *r* or *x* (hereafter pre-C-R) sometimes have a variant with a vowel (usually schwa) breaking the consonant cluster. It is also seen that, corresponding to a pre-C-R in SM, SWY, and JKT, MIN has ø in inherited lexemes, and -ra- in loanwords. In loanwords, BH has a corresponding -ra- or (in *kartas* and *sarbu*) a pre-C-R, and IBN has -rə-. According to Richards (1981:XV), penultimate schwa in IBN trisyllables like *tərəbay* and *kərəja* (< SM < SKT) is not pronounced, and is written as a spelling convention. Due to the small number of examples, it is difficult to determine what correspondence BH and IBN have in inherited lexemes. In the light of some inherited lexemes like BH *tajun, kabat,* IBN *cəlaŋ, kədil,* and moreover of

---

[130]Although the title of his study *(Malay phonetics)* does not suggest this the use of the term 'phonetics' here is in agreement with the nineteenth century terminology.

[131]He even asserted that he had heard this sound in JKT: "I have heard this sound in West Borneo, at Singapore and Batavia [now Jakarta]. Of course at Batavia, the pronunciation of Sundanese and Javanese people cannot be taken into account. Europeans born in that place very seldom give the guttural its particular value, generally substituting for it a strong palatal" (Fokker 1895:27).

TABLE 8: THE DISTRIBUTION OF r PRECEDING A CONSONANT

SM	MIN	BH	SWY	IBN	JKT	meaning
bərkas	--	--	bəxkas	bərəkas	--	bundle
bərkat	barakat	barakat	bərkat	bərəkat	bərkè	blessing (< AR)
bərnas, bərənas	--	barunas	bəxnas	--	--	rice ears
bərsih, bərisi, bərisih	barasiᵊh	barasih	--	bərəsi	bərsi, bərəsi	clean
bərsin	basin	--	bəxsin	bərəsin	--	sneeze (v)
bərtih	batiᵊh	--	bəxtiᵊ(h)	--	--	roasted rice
cərdik	cadiᵊʔ	--	cərədiʔ	cərədit (<SAR)	cərdik	clever, bright
cərlaŋ	--	--	cəxlaŋ, cərlaŋ	cəlaŋ	--	wide open (eyes)
cərmin, (Klinkert) cərəmin	camin	caramin	cərmin	cərəmin	--	mirror
jərnih, jənih	janiᵊh	jaranih	jəxəniᵊ(h)	--	--	pure
kərəbat, kəbat, kərbat	kabeʔ	kabat	kəbat	kəbat	--	bind, lash together
kərbaw	kabaw	--	kəbaw	kərəbo, kərəbaw	kəbòʔ, kərəbó	buffalo
kərdil, kərədil	kadi, kadih	--	--	kərdil	kərdil	stunted
kərdut, kədut, kərut, kərudut	--	--	--	kədut	kərut, kədut	crease, wrinkle
kərtas	karateh	kartas	kəxtas	kərətas	--	paper (< AR)
pərcik, pərcit	(pacaʔ)	puracit	--	pərañcit, pərəñcit	pərcik	squirt (v)
sərbu	--	sarbu	--	--	sərbu	attack, invade
tərban/tərəban	tabaŋ	tarabaŋ	təxbaŋ	tərəbay	tərbaŋ	fly (v)
tərbit	tabiʔ	--	təxbit	tərəbit (Richards)	tərbit	emerge
tərjaŋ/tərajaŋ	--	tirajaŋ, tarajaŋ	--	tərajaŋ	tərjaŋ, tərəjaŋ	kick (v)
tərjun	tajun	tajun	tərjun	tərəjun	--	leap down
bər-	ba-	ba-	bəx-	bə(r)-	bər-	(verbal affix)
pər-	pa-	pa-	pəx-	pə(r)-	pər-	(verbal affix)
tər-	ta-	ta-	təx-	tə(r)-	tər-	(verbal affix)

BH and IBN reflexes of *(mb)Ar-, *pAr-, and *tAr- (cf. 6.1 - 6.2) I assume that BH and IBN have lost a pre-C-R. But for some other lexemes a different explanation is required. Most loanwords found in BH as well as in IBN must be borrowed from SM, since SM is the most probable source for lexical borrowing common to BH and IBN, and as a lingua franca it has been very influential in other languages. BH forms like *barunas, tirajaŋ, puracit*, and IBN forms like *tərajaŋ* and *pəraňcit*, have (penultimate or antepenultimate) vowels which do not occur in SM. It is possible that these vowels are due to secondary developments, but as yet there is no evidence for this. For the forms containing them (and perhaps also for BH *tarabaŋ*, IBN *tərəbay*, which has a cognate in all other isolects) the hypothesis of a trisyllabic PM reconstruction with a *-rV*-sequence is more appropriate.

The pre-C-R in SM, SWY, and JKT may then be explained by syncope of the following vowel as a result of the tendency to disyllabicity of lexemes (4.5). This tendency, together with a counteractive tendency to avoid heterorganic consonant clusters, is probably the reason for the occurrence of variants, one with a pre-C-R, and the other with ə breaking the consonant cluster (e.g. SM *bərnas/bərənas*; SWY *səxkit/səxəkit* 'wooden spoon'). Another reason for assuming that pre-C-R's originated through syncope of the following vowel is that (in SM, SWY, JKT) they are always preceded by ə. This ə may be the result of antepenultimate neutralisation in an originally trisyllabic form. As yet I do not know what the situation in PM was, and for the time being I reconstruct:

(a) *-r(ə)- if BH and IBN have a corresponding ø;
(b) *-rV- if BH and/or IBN have a following vowel (other than IBN ə).

I also reconstruct *-rV- if SM, SWY, or JKT have a variant form with a vowel other than ə between the pre-C-R and the following consonant (cf. SM *kərdut/kərudut*). In both cases MIN has a corresponding ø. As for the antepenultimate vowel, I reconstruct *i/*u on the basis of BH *i/u*, (MIN *a*), other isolects ə, and *A on the basis of BH (and MIN) *a*, other isolects ə. If no BH cognate is available, I reconstruct *V.

3.7.1 PM NON-FINAL *l > ALL ISOLECTS l

Examples:

*lama? 'long (time), old (things)'; SM *lama*, MIN, SWY *lamo*, IBN *lama?*, JKT *lamè*;
*lima? 'five'; SM, BH *lima*, MIN, SWY *limo*, IBN *lima?*, JKT *limè*;
*laki 'husband'; BH, JKT *laki* 'husband; male', o.i. *laki* 'husband';
*baləs 'reply, reciprocate; revenge (v)'; MIN *baleh*, JKT *baləs*, o.i. *balas*;
*malu 'shy, ashamed'; a.i. *malu*;
*tulaŋ 'bone'; a.i. *tulaŋ*.

3.7.2 PM FINAL *l > MIN ø (AT MORPHEME-BOUNDARIES -l-), OTHER ISOLECTS l

Examples:

*gatəl 'itch, itchy; sensual'; JKT *gatəl*, MIN *gata*, o.i. *gatal*;
*jəŋkal 'span of thumb and middle finger'; SM *jəŋkal*, MIN *jaŋka*, BH *jaŋkal*;
*təbəl 'thick'; JKT *təbəl*, MIN *taba*, BH *tabal*, o.i. *təbal*.

UNEXPLAINED IRREGULAR CORRESPONDENCES

(1) In a few cases SWY has -ʔ corresponding to -l (MIN ø) in the other isolects. SM and MIN sometimes agree with this reflex in having -k and -ʔ respectively, or in having doublets with -k/-ʔ and -l (MIN ø). In the one case where IBN has a cognate, it ends with -ʔ. The cases with SWY -ʔ are:

ambiᵊʔ 'take, get, fetch'; IBN ambiʔ, SM ambil, ambik, MIN ambiᵊ, ambeʔ, BH, JKT ambil (cf. also IBN sambiʔ 'get in exchange', MIN s/ambiᵊ, o.i. s/ambil 'simultaneously, along with');
kəciᵊʔ 'small'; SM kəcil, kəcik, MIN kaciᵊʔ, BH kacil, JKT kəcil, kəcit (-t unexplained);
kumpuᵊʔ (also kumpul) 'together, gathered'; SM, BH, JKT kumpul 'id.', MIN kumpuᵊʔ 'heap, collection', IBN gumpul 'gather, pick' (3.4.1.4; cf. also SM kəlompok 'group'; SWY kəlumpuk must be a loan, see 3.4.2a).

I do not have an explanation for these correspondences. Judging by the low frequency of inherited lexemes with final -il(/-iᵊ) in the isolects, one may assume that this ending is not in favour in SWY, (pre-) MIN and IBN. The SWY and IBN reflexes should also be seen in the light of the change *-r > -ʔ in these isolects (see next paragraph). On the basis of the above sets I make the following reconstructions:

*ambil/*ambik 'take, get, fetch';
*kəcil/*kəcik 'small';
*kumpul/*kumpuk 'together; gather'.

(2) In one case IBN has r- corresponding to l- in the other isolects:

rumbaʔ 'race, contest'; SM lomba, lumba, MIN lomba/n (< +lumbaʔ + +an: see 4.5), BH lumba, JKT lòmbè.

Besides lagiʔ 1. 'later on' 2. 'more' IBN has also agiʔ 'more' < *lagiʔ (3.1.2, 3.1.2.3). On the basis of SM lomba etc. I reconstruct:

*lumbaʔ 'race, contest'.

3.7.3 PM NON-FINAL *r > SWY x/r, OTHER ISOLECTS r

Examples:

*rusaʔ 'sambhur deer'; SM rusa, MIN ruso, SWY xuso, IBN rusaʔ, JKT rusè;
*rusuk 'side, flank'; SM, IBN, JKT rusuk, MIN rusuᵊʔ, SWY xusuᵊʔ, rusuᵊʔ;
*raup 'scoop with both hands'; SWY xaup, raup, o.i. raup 'id.', MIN (sa-)rauyʔ 'a handful';
*rəbaʔ 'disorderly mass of dry wood'; SM, BH rəba, MIN rabo, SWY xəbo, IBN rəbaʔ;
*rinduʔ 'like, long for'; SM, MIN, JKT rindu, SWY xindu, rindu 'long for', IBN rinduʔ 'like, love, be gladdened';
*bərat 'heavy, important'; SM, IBN, JKT bərat, MIN bareʔ, BH barat, SWY bəxat;
*buruŋ 1. 'bird'; 2. 'omen';[132] IBN buruŋ 'id.', MIN (rare) buruᵊŋ (5.7(97)), o.i. buruŋ 'bird';

---

[132] In Achehnese buroŋ means 'spirit of a woman who died in childbirth; tormentor of women in childbed'. It is not improbable that buruŋ etc. with the meaning of 'bird' is borrowed: most AN languages have a reflex of PAN *manuk for 'bird', and the r in many languages is not the regular correspondence to SM r. Compare Ngaju buroŋ, Malagasy vorona, SWY buruŋ, KCI burewŋ, whereas SM r : Ngaju h : Malagasy ø/z : SWY x : KCI h; the SWY and KCI reflexes suggest that buruŋ etc. (for 'bird') is also innovative within the Malayic

*kuraŋ 'less, short of, lacking'; SWY kuraŋ, (rare) kuxaŋ, o.i. kuraŋ;
*darah 'blood'; SWY daxa(h), JKT darè, o.i. darah.

UNEXPLAINED IRREGULAR CORRESPONDENCE

IBN reflects ø- for *r- in:

asay 'feel; feeling'; SM, BH rasa, MIN raso, SWY xaso, raso, JKT rasè 'feeling, flavour; feel, taste' (< SKT).[133]

3.7.4 PM FINAL *r > MIN ø (AT MORPHEME BOUNDARIES -r-), SWY x/r, OTHER
        ISOLECTS r

Examples:

*liur 'saliva'; SM, BH, IBN liur, MIN liuᵊ, SWY liux;
*tampar 'slap with the hand'; MIN tampa, BH tampar 'hit with the fist', SWY tampax, o.i. tampar 'slap with the hand';
*haur 'k.o. bamboo'; SM (h)aur, BH haur, MIN auᵊ, SWY aux, IBN aur.

IRREGULAR CORRESPONDENCES

In the following cases IBN and/or SWY have -ʔ corresponding to a final -r in the other isolects, e.g.

SWY ayiᵊʔ, IBN aiʔ 'water'; SM air, ayər (cf. 2.1.3), BH air, MIN aiᵊ, JKT aèr;[134]
SWY bəsaʔ 'big' < *bəsar (3.2.3);
SWY buntaʔ 'round'; SM buntar, MIN bunta;
IBN buntuʔ 'dead, bloated (of fish)'; SM buntur 'oversatiated, completely "full"' (Blust 1980 a);
IBN gəlaʔ 'nickname'; SM, JKT gəlar, MIN gala, BH galar 'title, surname';
SWY ikuᵊʔ, IBN ikuʔ 'tail'; SM ekor, MIN ikuᵊ;
IBN iliʔ 'downstream'; SM (h)ilir, BH hilir, MIN iliᵊ, SWY ilix 'id.', JKT ilir 'north';
IBN kapuʔ, kapur 'lime, chalk' < *kapur (3.4.1.1);
IBN salaʔ in bau salaʔ 'smell of roasting flesh'; SM səlar 'branding', MIN sala, saŋla 'broiling, cooking at an open fire' (cf. 4.6);
IBN təluʔ 'egg' < *təlur (5.7 lemma 98);
SWY tiduᵊʔ 'sleep (v)'; SM, JKT tidur 'id.', MIN tiduᵊ 'lie down; sleep' (IBN tinduk 'sleep (v)' is a loanword from the Tamanic languages, Adelaar in press a)

These correspondences may point to a shift of +-r to ʔ in SWY and IBN. On the basis thereof the following reconstructions are made:

*air 'water';
*buntar 'round';
*gəlar 'title, surname';

group. A semantic shift from 'omen' to 'bird' may first have happened in the Malayic isolects (cf. IBN, where both meanings still apply), whereupon buruŋ with the meaning 'bird' was borrowed in non-Malayic isolects.

[133] Although it is conceivable that the split of PM into the different Malayic isolects postdated the introduction of the first SKT loan-words, I prefer not to conjecture on this possibility, and not to make a reconstruction on the basis of correspondence sets that were ultimately borrowed from SKT.

[134] Abdul Chaer's 'aer' (without diacritic) is probably a misprint.

*ikur 'tail';
*hilir 'flow down; downstream';
*saŋ(ə)lar 'fry without oil' (cf. 4.6);
*tidur 'sleep'.

UNEXPLAINED IRREGULAR CORRESPONDENCE

IBN reflects -s for *-r in:

buŋas 'virgin, first (fruits)' (cf. also 3.6.1.5); SM, JKT buŋar, MIN buŋa.

On the basis of SM buŋar etc. I reconstruct:

*buŋar 'first fruits'.[135]

3.7.5 PM *r > SM, JKT r, SWY r, x, BH, IBN Vr, MIN ø l _*(V)C

Examples (see above for correspondence sets):

*bVr(ə)kas 'bundle' *kAr(ə)bat 'fasten together'
*bArisih 'clean' *kAr(ə)baw 'buffalo'
*bArunas 'rapidly growing (of rice ears, etc.)' *kVr(ə)dil 'stunted'
*bVr(ə)sin 'sneeze' *kVrudut 'crease, wrinkle'
*bVr(ə)tih 'roasted rice' *pura(ñ)ci(kt) 'squirt (v)'
*cVr(ə)dik 'clever, bright' *tAr(ə)baŋ 'fly (v)'
*cVr(ə)laŋ 'wide open (eyes)' *tVr(ə)bit 'emerge'
*cAr(ə)min 'mirror' *tirajaŋ 'kick (v)'
*jAr(ə)nih 'pure, transparent' *tVr(ə)jun 'leap down'

3.8 THE PM SIBILANT

The sibilant s occurs and agrees initially, intervocalically, and post-nasally in all isolects, and in these positions PM *s is reconstructed. The nasal preceding s in clusters is ŋ in SM, BH, SWY, and JKT; it is n in MIN and IBN. In final position s occurs in all isolects but MIN, where a corresponding -h is found. On the basis of SM, BH, SWY, IBN, JKT -s, and MIN -h, PM *-s is reconstructed. Most vowels directly preceding PM *-s underwent a change in MIN: *a, *ə > e (3.1.1, 3.1.1.3, 3.1.1.5), and *u > uy (3.1.2, 3.1.2.4).

3.8.1 PM NON-FINAL *s > ALL ISOLECTS s

Examples:

*saraŋ 'nest'; IBN saraŋ 'container', and saraŋ mañi 'bee's nest', SWY saxaŋ, o.i. saraŋ 'nest';
*sayap 'wing' (3.1.1.3);
*sumpah 'oath, ordeal; imprecation'; JKT sumpè, SWY sumpa(h), o.i. sumpah;
*(b)isik 'whisper' (3.5.1 UIC);
*tasik 'sea' (3.4.2.3);

---

[135]But cf. Proto South-Sulawesi *buŋas 'first fruits' (Mills 1975 vol.2). .

*isi? 'contents, flesh' (3.4.2.4);

*buŋsu 'youngest born'; SM, SWY buŋsu, boŋsu, MIN bunsu, bonsu, BH buŋsu, IBN bunsu;

*laŋsAt 'k.o. tree with sour fruits (Lansium domesticum)'; SM, SWY, BH laŋsat (in Wilkinson 1959 and Klinkert also SM lansat), MIN lanse?, IBN lənsat (ə unexplained); (JKT laŋsəp is probably not a cognate (< JV laŋsəb 'id.' ?);

*laŋsuŋ 'direct(ly), straight'; SM, BH, JKT laŋsuŋ 'id.', MIN lansuᵖŋ 'finished, ended'.

N.B. In IBN regressive dissimilation to t took place if two syllables began with *s, e.g.

*suŋsaŋ 'upside down, against the grain'; SM soŋsaŋ, MIN sonsaŋ, IBN tunsaŋ, o.i. suŋsaŋ;

*səsət 'have lost one's way' (IBN təsat, 3.4.2.2);

*sələsay 'settled'; SM sələsay, MIN salasay, SWY (sə-)ləsay (apparently reanalysed as a derived form with prefix sə-), IBN tələsay, JKT sələsé;

*sisik 'fish scale'; SM, BH, JKT sisik, MIN, SWY sisiᵖ?, IBN tisik.

UNEXPLAINED IRREGULAR CORRESPONDENCES

(1) Sporadic loss of JKT s-. In JKT s- is often omitted in some frequently occurring lexemes. In only one case (ajè) is the variant with ⁺s- lost altogether; in some cases the meanings have diverged. For example:

ampé 'until', and sampé 'complete, arrived, reached'; o.i. sampay 'arrived, reached; until'; atu, satu 'one' < +sa-batu, cf. 5.3.2);

ayè, sayè '(1st pers. sg.)', SM saya, sahaya, MIN, SWY sayo, 'servant, slave; (1st pers. sg.)' < SKT;

amè 'with, and; by (actor preposition); to(wards)', cf. SM, BH, IBN sama, MIN, SWY samo, JKT samè 'together, same, equal', < SKT;

ajè 'only', SM saja, sahaja, SWY sajo 'only; on purpose', MIN sajo, IBN saja?, aja? 'only, simply'; also SM səŋaja, MIN saŋajo, siŋajo, BH saŋaja 'on purpose' < SKT sahaja 'innate, by nature' (Gonda 1973:390);

udè, sudè 'already', but this lexeme has cognates with and without s-, cf. SM sudah, MIN sudah, udah, SWY udo, sudo, and IBN uday, udah (< SKT, cf. 3.2.3).

(2) In one case JKT has s- corresponding to MIN ø-, SM s- or d-, o.i. d-:

SM, JKT sampiŋ 'side, flank, border', and SM, BH, SWY dampiŋ 'close, near(by)', MIN ampiᵖŋ (Wilkinson 1959 hampiŋ) 'id.'.

It is likely that SM, JKT s-, and SM, BH, SWY d- reflect earlier clitics, and I tentatively reconstruct:

*(h)ampiŋ 'close (to), near(by)'.[136]

(3) Sporadic change of *s to BH h-. In initial position BH sometimes has h- for s- in the other isolects, or it has variants with both s- and h-, e.g.

BH hiŋal (cf. 3.1.1.2 UIC for i), SM, JKT səŋal, IBN səŋak (-k unexplained) 'out of breath, breathing with difficulty';

BH hual, SM soal, MIN sua, JKT sòal 'problem, affair, question' (< AR).

---

[136]cf. d/- in d/ari < *ari (3.5.2 UIC), da/hulu, d/ulu < *di *hulu? (3.1.3.3), and IBN alam, o.i. dalam/daləm < *(d-)aləm (3.1.1, 3.1.1.5).

cf. also *sagan/hagan* 'for', *sidin/hidin* '(3rd pers. sg. honorific)' (Abdul Jebar p.18). In view of the number of unexplained correspondences in *saŋal/hiŋal* etc., no PM reconstruction is made.

3.8.2 PM FINAL *s > MIN -h, OTHER ISOLECTS -s

Examples:

atas 'on, above; upper part'; MIN *ateh*, o.i. *atas*;
haus 'thirst(y); worn out'; SM *(h)aus*, BH *haus*, MIN *auyh* 'thirst(y); worn out, eroded', SWY *aus* 'thirst', IBN *aus*, JKT *aus*, *aós* 'thirst(y)';
ruas 'internode'; MIN *rueh*, SWY *xuas*, o.i. *ruas*.

UNEXPLAINED IRREGULAR CORRESPONDENCES: sporadic change of *-s to -h.

In a number of cases some isolects reflect -h instead of -s (in these cases, the last-syllable vowels in the corresponding MIN forms indicate that the following -h came from +-h rather than from +-s), e.g.

SM *paras*, *parah*, MIN, BH *parah* 'squeeze, press', SWY *paxa(h)* 'squeeze, press, milk', IBN *parah*, JKT *paras*, *parè* 'press, squeeze';
SM *raih*, MIN *raiᵊh* 'draw towards oneself, scoop in, annex', BH *raih* 'ask a young person to marry one's child', SWY *ma-raiᵊ(h)*, *ma-xaiᵊ(h)*, IBN *rais*, *raih* 'cock a gun, take off the safety catch';
SM *ta/lah*, MIN *lah*, SWY *la(h)* 'already', BH *laas* 'finished, used up' (cf. 3.10 and 4.5; cf. also JV *talas* 'finished, used up', and SUN *laas* 'lose colour, become tasteless; solve, cease to exist');
SM, MIN, BH *tarah*, SWY *taxa(h)*, IBN (Richards 1981) *taras*, *tarah* 'plane with an adze, shape, do some rough-hewing';
SM *ubas*, IBN *ubah*, 'k.o.tree used for firewood'.

I do not have an explanation for this change; palatalisation of final *-as-sequences and/or the change from *-s to -h is seen as a regular change in many peninsular Malay isolects, and in many (Malayic and other) isolects in Sumatra (including Achehnese, MIN, KCI).[137] On phonetic grounds it is likely that in the above cases final *s must be reconstructed. But for SM *paras* etc. there are PMP forms available with *-s and with *-q (PMP *q > PM *h), and in this case I reconstruct PM doublets.

On the basis of the above sets I reconstruct for PM:

laøas 'finished, used up' (see also 3.10);
*paras/*parah* 'squeeze, press';(cf. PMP *peReqes*, PMP *peRaq*);
rais 'draw towards oneself, scoop in';
tarAs 'plane with an adze, shape, do some rough-hewing'.

---

[137]The same variation can in some cases be observed between Malayic and JV or Toba, cf. PM *habis* 'all, entirely; used up, done with, finished off' vs JV *k/abèh* 'all' (cf. 3.1.2.3 and 5.7 lemma 191); SM *lamah* 'soft, weak, slack' vs JV *lamas* 'supple, flexible; weak; elegant'; SM *luruh* 'dropping, being shed, especially of leaves and fruit' vs Toba *rurus* 'fall off (leaves), fall out (hair)'.

3.9 THE PM GLOTTAL SPIRANT

The glottal spirant *h* occurs in all isolects; in SWY *h* occurs in final position in Helfrich, but is left out in the more recent publication of Aliana et al. (1979) Compare Table 9:

TABLE 9: THE DISTRIBUTION OF *h* IN THE MALAYIC ISOLECTS

	SM	MIN	BH	SWY	IBN	JKT
initially	*(h)*	-	*h*	-	-	-
between like vowels	*h*	*h*	*h*	-	-	*h*
between unlike vowels	*(h)*	-	*h*	-	-	-
finally	*h*	*h*	*h*	*(h)*	*h*	-

This table shows that *h* is most common in final position: SM, MIN, BH, and IBN show it, while SWY reflects it in older sources. JKT lost *+-h* in the sub-dialect of Mester (see Chapter 1), but from the comparative wordlist of JKT sub-dialects in the introduction to Abdul Chaer (pp.XVIII - XIX) it appears that in Kebayoran there is a final *-h* in (at least) those cases where SM, MIN, SWY (Helfrich) and IBN have a corresponding *-h*.[138] On the basis of SM, MIN, BH, IBN *-h*, SWY *-(h)*, and JKT *-ø* (Kebayoran *-h*), PM final **h* is reconstructed. Table 9 also shows that *h* between like vowels occurs in SM, MIN, BH, and JKT (in JKT a corresponding *-ʔ-* is sometimes found, which is not in phonemic contrast with *h*, cf. 2.6). In SM, MIN, and JKT, however, two like vowels never occur adjacently, whereas in BH, they do. In this environment SM, MIN, and JKT *h* does not necessarily reflect a proto-phoneme, and its presence may be motivated by a phonotactic constraint. Finally, it appears that initially and intervocalically between unlike vowels, *h* only appears in SM and BH; in SM it is, moreover, almost always lost or optional in these environments, even when retained in the spelling. Since

(1) the occurrence of BH *h* between like vowels is not conditioned by a phonotactic constraint (see also 3.10),

(2) BH *h* in all positions is a stable phoneme which as a rule (and contrary to SM) never alternates with *ø*, and finally,

(3) BH *h* reflects PMP **q*, whereas BH *ø* between like vowels does not (see 3.10),

I will use BH as a test language for PM non-final **h*. If BH has *h*, PM **h* is reconstructed, but if BH has *ø* (in Abdul Jebar's orthography sometimes an apostrophe, 2.3.1) in correspondence with SM, MIN, and JKT *h* between like vowels, PM **ø* is reconstructed (see 3.10). If no BH cognate is available, **h* is reconstructed initially and between unlike vowels when a corresponding *h* is attested in SM, and **(h)* is reconstructed when SM, MIN, JKT have *h* between like vowels.

3.9.1 PM NON-FINAL **h* > BH *h*, SM *h*, *ø*, OTHER ISOLECTS *ø* (EXCEPT BETWEEN LIKE VOWELS)

Examples:

*halu 'pounder, pestle'; SM, MIN, IBN *alu*, BH *halu*, JKT *aluʔ*;

---

[138] Kebayoran in some cases also has *-h* corresponding to *ø* in the other isolects, e.g. *sayah* 'I', *sǝpèdah* 'bicycle', *diah* '(3rd pers. sg.)', *apah* 'what?'. But as a rule it has a non-phonemic glottal stop after final vowels. Karet agrees with Kebayoran in so far as it shows a final glottal stop wherever Kebayoran does, and it has *ø* wherever Kebayoran has *-h* (Abdul Chaer 1976:XIX).

*hulət 'worm, maggot'; SM (h)ulat, MIN uleʔ, BH hulat, SWY, IBN ulat (cf. PMP *qulej; JKT ulər must be a loan, cf. 3.4.2.2UIC);

*hirup 'lap up, absorb'; SM (h)irup, MIN iruyʔ, BH hirup, SWY, JKT irup 'id.', IBN irup 'drink (v)';

*hiuʔ 'shark'; SM (h)iu, yu (3.3), MIN, SWY iu, BH hiu, IBN iuʔ, JKT yu (3.3);

*hiaŋ 'divinity'; SM (h)iaŋ,[139] MIN iaŋ 'id.', IBN (Richards 1981) yaŋ 'tutelary spirit appearing in dreams or visions'; cf. also *səmbah-*hiaŋ 'pray to (the) god(s)' (3.1.3.1 N.B.);

*hAlu-an 'bows, forepart of a vessel' (3.1.3.3);

*sahut 'answer (v, n)'; SM, BH, SWY[140] sahut, MIN sauyʔ, IBN, JKT saut;

*tihaŋ 'mast, post, pillar'; BH tihaŋ, o.i. tiaŋ;

*tuha(ʔ) 'old, mature (of people)' (3.2.3).

3.9.2 PM *h > SWY, IBN ø, OTHER ISOLECTS h (BETWEEN LIKE VOWELS)

(cf. also 3.1.1.5 IC for the assimilation of *ə to a)

Examples:

*dAhək 'phlegm, mucus'; MIN dahaʔ, SWY dahaʔ,[141] IBN daak, o.i. dahak (cf. 3.1.1.5 IC);

*tahən 'strong, able to endure; keep, detain'; SWY, IBN taan, o.i. tahan;

*pahət 'chisel'; SM, BH, JKT pahat, MIN paeʔ, SWY, IBN paat;

*gahəm 'molar tooth' (3.1.1.5IC);

(see also *puhun (3.1.2, 3.1.2.4) and *jahət (3.1.1.5IC)).

3.9.3 PM FINAL *h > SWY (-h), JKT ø, OTHER ISOLECTS -h

Examples:

*labuh 'fall, hang down'; SM, BH labuh, MIN labuᵊh 'let down, lower by means of a strong rope or cable', SWY labuᵊ(h) 1. 'be anchored' 2. 'fall (rain)' 3. 'start, begin' 4. mə-labuᵊ(h)-ka(n) hukuman 'pass sentence', JKT pəlabuan (also SM pə-labuh-an, MIN pa-labuᵊh-an) 'anchorage, harbour';

*pilih 'choose'; MIN piliᵊh, SWY piliᵊ(h), JKT pili, o.i. pilih;

*susah 'difficult, troublesome'; SM susah, SWY susa(h), JKT susè 'id.', MIN (rare) susah 'trouble, worry, burden', BH susah 'poor', IBN tusah 'troubled, sad' (< PMP *suqsaq; Dempwolff reconstructed *susaq, but cf. Karo suhsah 'difficult, troublesome').

UNEXPLAINED IRREGULAR CORRESPONDENCES

In the following cognate sets one or more isolects do not reflect *-h:

(a) SM, with -k:

SM bəŋik 'catch in the breath', and sakit bəŋik 'fowl tuberculosis', JKT bəŋèk 'bronchial asthma'; IBN bəŋih 'have a dry cough';

---

[139] SM also has orthographic variants 'yang' and 'hiyang'.
[140] In some instances Helfrich's orthography seems to be inconsistent, especially when he uses one lexical entry for both SWY and BSM (in BSM h some imes occurs intervocalically).
[141] See fn. 140.

(b) SWY, with ø:

SWY *sudo, udo* 'already, after, in the past' (ultimately < SKT, cf. 3.2.3);

(c) SWY, with -ʔ:

SWY *limpaʔ* 'flow over'; SM, MIN *limpah*;

(d) SWY and IBN, with -ʔ and -k respectively:

SWY *panduᵊʔ* 'burn up brushwood', IBN *tuguŋ panduk* 'a heap of sticks etc., for burning'; SM *pandu, panduh* 'burn up brushwood';

(e) IBN, with -ʔ:

IBN *basuʔ* 'wash'; SM *basuh*, MIN *basuᵊh*, SWY *basuᵊ(h)*, JKT *basu* (ultimately borrowed from JV);[142]

IBN *bəpədiʔ, pəpədiʔ* 'smart (v,n)'; SM *pədih*, MIN *padiᵊh*, BH *padiʾh*, SWY *pədiᵊ(h)*;

IBN *kəmiʔ* 'urinate'; SM *kəmih*, BH *kamih*, SWY *kəmiᵊ(h)*, JKT *kəmih* (-h unexplained; probably a loan, cf. Table 4 (31)).

IBN *ləlaʔ* 'fall asleep'; SM *ləlah*, JKT *ləlè* 'tired, exhausted', BH *ba-lalah* 'walk without direction' (a cognate?), SWY *ləla(h)* 'pant from exhaustion';

IBN *ŋə-luaʔ* 'spit out'; SM, MIN *luah*;

IBN *mataʔ* 'uncooked, unripe' < *m/əntah/*m/atah* (3.6.2UIC);

IBN *səmpaʔ* 'what is rejected after chewing betel' < *sə(m)pah* 'betel cud (3.6.2 UIC);

IBN *ayaʔ* 'uncle'; SM *ayah* 'father (polite)', MIN *ayah* (rare, according to Van der Toorn) 'father', BH *ayah*, JKT *ayè* 'father'.

In one case SM has -h/-ø where other isolects have -ø:

SM *contoh, conto* (with a variant *cinto*; these variants are ultimately from CHI, cf. Klinkert), BH *cuntu*, IBN *cunto* (a loan, cf. 3.1.2b), JKT *cóntó* 'model, example, pattern'.

In one case IBN has -h where SM has -ø:

SM *dura* 'anxiety, disquiet', IBN *durah/durah* 'make a lot of noise, out of fear of s.th.'.

On the basis of these correspondences I reconstruct:

bəŋi(hk) 'have difficulty in breathing, cough'   *limpa(hk)* 'flow over'
pədih 'smart (v,n)'   *luah* 'spit out'
dura(h) 'disquiet, anxiety'   *ayah* 'father'

---

[142]Dempwolff reconstructed PMP *basuq* 'wash' with support from JV *wasóh*, SM *basuh* 'id.', and Fijian *savu-i, savuya* 'clean the canoe'. In view of its similarity in meaning and form with PMP *baseq* 'wet', the formally deviant Fijian reflex, and the fact that JV *-óh* may reflect PMP *-eq* as well as *-uq* (see below), *basuq* must be discarded as a PAN/PMP reconstruction. Dyen (1965:295-296) gives additional support from Atayal dialects (*mahuq* and *mahuʔ*), Paiwan dialects (*v-in-tuʔ* and *v-in-atuq*), and Pazeh (*ba-batsuʔ*), all meaning 'wash (clothes)'. But Atayal and Pazeh *u* may reflect PAN *e as well as *u, whereas Paiwan *u* < PAN *u. So only Paiwan provides evidence for PAN *basuq. Dyen obtained his information from Ogawa and Asai's *The myths and traditions of the Formosan native tribes* (Taihoku 1935), a source which is not available to me. Ferrell (1982), however, gives Paiwan *v/n/ateq* (from *vateq*), and I suppose the last-syllable *u*'s in *v-in-tuʔ* and *v-in-atuq* are due to misprint or wrong perception in Ogawa and Asai.

I believe that SM *basuh* etc. is borrowed from JV, and that JV *wasóh* is a cognate of SM *basah* 'wet' etc. Nothofer (1975:77 + n.48) points out that PMP *-eq# regularly becomes JV *-oh*. He considers JV *basah* 'dissolve, decay' as a loan. However, for some reason he does not continue this line of reasoning, and he does not reinterpret JV *wasóh* as a reflex of PAN/PMP *baseq. He reconstructs PMJ *Bassəh* as well as *Bassuh* (p.185). (His evidence from Old Javanese viz. *wasəh* 'washed' versus *aŋ-wasuh* 'wash' are two variants of the same word, as is shown by the large number of variants with *u* and *ə* in the Old Javanese dictionaries.

As to the correspondence JV *w-* : SM *b-*, the change from semivowel to homorganic stop in SM borrowed lexicon is not unusual, cf. JV *wolanda* 'Netherlands; Dutch' > SM *bəlanda*; DU *wiŋkəl* 'shop' > SM *beŋkel* 'workshop'; DU *wortəl* 'carrot' > SM *bortol*; DU *yüni* 'June' > SM *juni*; DU *yənevər* 'Dutch gin' > SM *jənewər*; DU *yas* 'coat' > SM *jas*; ENG *yard* > SM *yar, jar*; ENG *waistcoat* > SM *beskat*.

*k/əm/ih 'urinate'   *pandu(hk) 'burn up brushwood'
*ləlah 'exhausted, tired'

N.B. In the above reconstructions PM *-h agrees with PMP *-q. For SM ayah etc. I reconstruct *ayah[143] because of PWMP *ayaq (Blust 1979). PWMP had a vocative suffix *-q; the proto-phoneme *q in final position is generally reflected as h in the Malayic isolects (see 7.1.i). Blust (1979:234-235) assumes that at the time PWMP *-q became *-h (and in this way would have lost its force as a vocative marker), SM resisted this sound change in many vocatives. It is possible that IBN did so too, and even retained the glottal stop in aya? whereas SM, MIN, BH did not (they reflect PWMP *-q in ayah, which may be due to the fact that it became a polite term, and therefore would not be used as a vocative any more).

3.10 PM (INTERVOCALIC) *ø

BH provides the strongest evidence for PM *h (3.9); it has retained *h in all positions and its occurrence is never optional. Moreover, its occurrence between like vowels is not determined by a phonotactic rule, and here BH distinguishes h and ø.[144] For two reasons this distinction is relevant for the reconstruction of PM. Firstly, the other isolects either allow only h (i.e. SM, MIN, and JKT) or ø (i.e. SWY and (phonetically as a long vowel) IBN), or have contracted the adjacent vowels after affixation (4.5). Secondly, the BH h : ø distinction seems to reflect the distinction between PAN/PMP *q and PAN *S (/PMP *h) (7.1.i).

I reconstruct PM *ø on the basis of ø between like vowels in BH (corresponding to SM, MIN, JKT h, SWY, IBN ø). The function of *ø is to indicate that in its place there was no PM phoneme, although it is likely that there was a (non-phonemic) glottal stop, as is the case in BH (2.3.1).

Only a few examples reflect PM *ø:

*baøah 'flood'; SM bah (with vowel contraction),[145] BH, IBN baah;
*laøas 'finished, used up' (3.8.2 UIC);
*tuøət 'knee' (3.1.1.5IC);
*luø(uə)k 'bay, inlet; corner' (3.4.2.3);
*puø(uə)t 'vagina'; SM, IBN am/put, MIN am/puy? 'copulate', BH puut 'vagina' (according to Klinkert, SM amput also occurs with meaning 'vagina' in some literary texts; cf. also P(W)MP *pue(Ct) 'bottom, buttocks' (Blust 1970)).

N.B. From the little material available on KD (all taken from Dunselman (1949, 1950) it appears that this isolect also distinguishes between PM *h and PM *ø between like vowels, or between vowels one of which is *ə. Whereas PM intervocalic *h became KD h, PM *ø between like vowels (or between vowel and *ə) became glottal stop (written as an

---

[143]Cognates of SM ayah etc. are found in some other AN languages, and they refer to various kinship relations of parental generation (Blust 1980c:224-225). On the other hand, the southern Dravidian languages have the following correspondence set: TAM ayya: 'father, respectable man', Malayalam ayyan 'father, lord', Kannada ayya, aya 'father, grandfather, master, lord, teacher', Telugu ayya, aya 'father' (Burrow & Emeneau 1961:15 (entry no.163)). It would be interesting to know more about the possible relationship between these southern Dravidian lexemes and SM ayah etc.

[144]Abdul Jebar sometimes writes an apostrophe for a non-phonemic glottal stop in BH (2.3.1).

[145]The non-occurrence of -h- is probably due to a phonotactic constraint, as there are no inherited lexemes with a CVhVh- sequence in the Malayic isolects.

apostrophe) in one example. This example unfortunately does not have a cognate in BH. Compare:

KD *bu²uk* 'hair of the head' (Dunselman 1949:81); MIN *a/bu²²*, BH *buuk*; Wilkinson (1959) also gives BRU *buhuk*; cf. also Urak Lawoi' Malay (spoken in Southwest Thailand) *bo²*; cf. PAN *buSek*.

Compare also KD *tuha* (Dunselman 1949:62) < *tuha(²)* (3.9.1); KD *tahutn* (1949:148) < *tahun* (3.4.1.2); *ñahut* (1949:149) < *sahut* (3.9.1); *puhutn* (1949:64, 161) 'tree' and 'owner' < *puhun* (3.1.2.4). For *bu²uk* etc. I reconstruct:

buø(uə)k 'hair of the head'.

3.11 CHANGES IN IBN ANTEPENULTIMATE SYLLABLES AND IN ADJACENT CONSONANTS

(A) In IBN original homorganic nasal + stop clusters following an antepenultimate vowel were reduced to their homorganic nasal, e.g.

təŋal/i² 'do s.th. separately, independently', and *tuŋgal* 'single', cf. SM, BH *tuŋgal*, MIN *tuŋga* 'sole, unique, alone', SWY *sə-tuŋgal* 'mixed with each other';
təŋəlam 'sink', cf. SM, SWY *təŋgəlam*, MIN, BH *tiŋgalam*, JKT *təŋgələm*;
jəmatan 'jetty; a bridge' (< *jambat-an* 'bridge (with handrail), jetty', 3.1.3.1);
rəmutan 'the rambutan, *Nephelium lappaceum*', cf. SM, BH, JKT *rambutan*, MIN *rambutan*, *rambuy²tan*, SWY *rambutan*;
səmilan, also *səmbilan* 'nine' < *(ə)sa² ambil-an* (see 5.3.1, 5.3.2).

N.B. cf. also the following loanwords (which must have been introduced via SM or SAR):

jənila '(western type) window' < SM *jəndela* < POR *janela*;
mənira 'flag' < SM *bəndera* < POR *bandeira*;
səməlih 'slaughter ritually' (Richards 1981) < SM *səmbəlih*;[146]
səmako 'tobacco' (+t- > s unexplained) < SM *təmbakaw* < POR *tabaco*;
səñata 'weapon, armour' < SM *səñjata* < SKT).

The above correspondence sets yield the following PM reconstructions:

tuŋgal 'single, alone';
tiŋgələm 'sink';
rambut-an '*Nephelium lappaceum*'.

(B) Nasal assimilation of stops takes place between the initial consonant of the antepenultimate and penultimate syllables: when one of them is a nasal and the other a stop, the latter is replaced by its homorganic nasal, e.g.

nəŋəri 'town, city' < SM *nəgəri* < SKT;
mənua 'country, home country, district', cf. SM *bənua* 'continent', BH *banua* 'place of birth, village', BSM *bənua* 'island' (Helfrich); MIN *banu²* (instead of expected +banuo) remains unexplained).

N.B. (1) The same nasalisation process is seen in trisyllabic verbs, e.g.

ñəmilik 'live together' from *sə-bilik* 'the inmates of one bilik';

---

[146] According to Dempwolff (1937:45), SM *səmbəlih* derives from AR *b'ismi'lla:hi* 'in the Name of God'. (For a phonological and semantic justification of this etymology, see also Adelaar 1989:5-7).

ñəməray 'cross; swim across', cf. *sə-bəray* 'the opposite side'. These trisyllabic verbs are derived from compounds consisting of *sə-* 'one' (5.3.1) + a noun; the same nasalisation process is sometimes seen in Malaysian SM, cf. *məŋənəpikan, məŋəñampiŋkan* 'push aside, discard (as irrelevant)', (Indonesian SM *məŋətəpikan, məŋəsampiŋkan*), from *kə təpi* 'to the edge, side, border', and *kə sampiŋ* 'to the side').

(2) In one case an original homorganic nasal + voiceless stop became homorganic nasal + voiced stop:

kəmbu/an 'keep, possess', cf. SM *əmpu/ña* 1. 'owner' 2. 'own (v)' 3. 'possession', MIN *pu/ño* 'have, be the owner of', *nam/pu* 'the one who owns', *am/pu* 'owning, owner of', SWU *əm/pu-o* 'owner (of)', JKT *pu/ñè* 'belonging to'.

The above correspondence sets yield the following PM reconstructions:

*bAnua 'land, home country';
*(əm)pu 'owning, owner of, belonging to'.

(C) Original initial vowels or **hV-* were lost, e.g.

luan 'bows, forepart of a vessel' < **hAlu-an* (3.1.3.3);
rimaw 'tiger' (also SM *rimaw*) < **hArimaw* (3.1.3.3);
ban/an '(prematurely) white-haired' < **huban* (3.1.2.2) + **-an*; cf. IBN *uban* 'grey or white hair', *b/uban* 'white-haired';
laman, təŋah laman '(in) the open space before the house', cf. SM, BH *halaman*, MIN (Van der Toorn), JKT *alaman* 'front yard, page of a book', SWY *laman* 'front yard, village square';
lia? 'ginger', cf. SM *(h)alia* (also Urak Lawoï *liya*, SD *ahià?* 'id.' (with regular metathesis of liquid and *h*, and subsequent loss of *+l*, cf. Adelaar forthcoming);
piun 'opium' < SM *apiun* < AR *afyu:n* 1. 'opium' 2. '(name of a city in Turkey, where opium is cultivated)'.

N.B. (1) *Rian* 'durian, *Durio zibethinus*' probably originated through antepenultimate neutralisation, subsequent assimilation of *+d* to *+r*, and apocope: *+durian* > *+dərian* > *+r(ə)rian* > *rian*; cf. SM, MIN, BH *duri/an* (JKT *durèn* < JV); cf. also **duri?* 'thorn' (3.5.2).

(2) The same apocope is sometimes also seen in other isolects, e.g. SM *dəp/an* 'ahead, next' < **hadəp* (3.1.1.5) + **an*; SWY *laman, ximaw* (see above).

The above comparisons lead to the following PM reconstructions:

*halaman 'open space before the house, front yard';
*halia? 'ginger'.

3.12 SUMMARY OF CHAPTER 3

Comparison of the phonemes in the six isolects yields the following PM phoneme inventory: PM had four vowels, **a, *ə, *i,* and **u. *ə* did not occur lexeme finally (but it did occur in final syllables). It is not clear whether **a* and/or **ə* occurred in antepenultimate syllables. There were two diphthongs (**-ay* and **-aw*): both occurred lexeme finally only, and both are analysable as **a* + a semivowel. There were two semivowels: **y* occurred intervocalically adjacent to **a* or **u,* and lexeme finally following **a; *w* occurred intervocalically between **a*'s, and lexeme finally following **a.*

There was a series of voiceless stops (*p, *t, *c, *k, and *ʔ) and a series of voiced stops (*b, *d, *j, *g): *ʔ occurred in final position only, and its reconstruction is uncertain. Voiced stops and *c did not occur in final position.

A series of nasals (*m, *n, *ñ, *ŋ) occurred in initial and medial position, and, except for *ñ, also in final position. In medial position they occurred intervocalically and before homorganic stops.

There were two liquids (*r and *l), and two fricatives (*s and *h). *h between like vowels (or between vowels one of which was a schwa) must be distinguished from ∅.

For a summary of PM consonant clusters, see 4.2: for a phoneme chart, see 4.1.

CHAPTER 4

PROTO MALAYIC WORD STRUCTURE

4.1 THE PM PHONEME SYSTEM

The following proto-phonemes were reconstructed in the previous chapter:

VOWELS

	front	back
high	*i	*u
mid		*ə
low		*a

(diphthongs: *-ay, *-aw)

CONSONANTS

		labial	dental	alveolar	palatal	velar	glottal
stops	voiceless	*p	*t		*c	*k	*ʔ
	voiced	*b		*d	*j	*g	
nasals		*m		*n	*ñ	*ŋ	
fricatives				*s (4.3.2)			*h
liquids				*l		*r (3.7)	
semivowels		*w			*y		

4.2 THE PM CANONICAL SHAPE AND PHONOTACTIC CONSTRAINTS

Most PM lexemes are disyllabic; there are also monosyllabic, trisyllabic, and even tetrasyllabic reconstructions, but the great majority are disyllabic:

$$*[C\ V(N)][C\ V(N)]\ C\ V\ (N))C\ V\ C$$

The following phonotactic constraints apply to this canon:

(1) Final C is never a voiced stop (3.5) or a palatal (3.4.2).

(2) *ʔ occurs only as final C (3.4.1, 3.4.2).

(3) Any C can be ø, e.g. *aku, *mata, *daun, *tuøət, etc.

(4) Consonant clusters consist of a nasal + homorganic stop, or a velar nasal + *s; they occur only intervocalically (3.6.2).

(5) Semivowels only occur intervocalically: *w occurs only between *a's, and *y occurs between vowels other than *i or *ə (3.3, 3.3.1, 3.3.2); in final position semivowels are treated as part of a diphthong.

(6) *ə only occurs as V_2 if C_3 is a *ʔ, ø (2.6.3 and 3.1.1), or *h.

102

(7) Diphthongs only occur lexeme finally (i.e. as V_2, and if $C_3 = \emptyset$).

(8) Initial schwa, or schwa preceded by initial *h, is never followed by a stop.[147]

4.3 ARTICULATION-TYPE HARMONY

4.3.1 ARTICULATION-TYPE HARMONY APPLYING TO HOMORGANIC STOPS, NASALS AND SEMIVOWELS IN INITIAL AND MEDIAL POSITION

In an article on Malay consonant-harmony (Adelaar 1983), I assert that in an earlier stage of SM there must have been a tendency to articulation-type harmony which disallowed initial stops followed by medial homorganic nasals or homorganic stops differing in voice. For instance, lexemes with a structure like that of *papan, bibit, mamak, cucu, jajar, ñañi, kikir* etc. are usual, but lexemes with a structure like that of *bomoh, bepaŋ, paman, kaŋa, kugah*, etc. are not. The tendency did not apply to combinations of *t* with *d* or *n,* which may be due to the fact that, being a supradental, *t* is not homorganic to *d* and *n,* which are alveolars. There were nine unexplained exceptions to this constraint: *bapa* 'father', *bepaŋ* 'k.o. sweetmeat', *kugah* 'k.o. shrub', *bomo(h)* 'sorcerer', *pamah* 'low-lying (land)', *giŋin* 'k.o. herb', *kaŋa* 'k.o. tree', *kaŋar* 'bird of prey', and *danaw* 'lake'.

That the constraint also applied to combinations of homorganic stops or nasals with semivowels seems likely, but there are still a few unexplained counterexamples. This tendency to articulation-type harmony was probably already at work in PMP. Although Chrétien never formulated this tendency, it is borne out by his lexicostatistical study of the PMP morph (Chrétien 1965 and Adelaar 1983:63-65). Adelaar (1983) did not include other Malayic isolects, which are therefore inspected in this study in the hope of determining to what extent articulation-type harmony was operative in PM. I will do so by giving all disyllabic lexemes in each isolect with a cognate in at least one other isolect, which are exceptional to the consonant-harmony constraint. (A full list of exceptions in each isolect is given in Appendix I.) As in Adelaar (1983), I organise the lexemes according to the pattern to which they belong.

These patterns are defined as follows:

I   initial stop followed by homorganic stop differing in voice (e.g. *bapa*);

II   initial stop followed by prenasalised homorganic stop differing in voice (e.g. SM *bimpaw* 'handkerchief, towel', < CHI);

III   initial stop followed by homorganic nasal (e.g. SM *danaw, pamah*);

VI   initial nasal followed by (prenasalised) homorganic stop (e.g. SM *mabuk, mimpi*);

V   initial stop or nasal followed by homorganic semivowel (e.g. SM *bawaŋ*).

---

[147]This observation is based on the following considerations:
Initial schwa (or *hə-*) is never followed by a stop in SWY (2.4.3) and IBN (2.5.3), and it is only in a very few cases followed by a stop in SM (2.1.3) and JKT (2.6.3). On the other hand, initial schwa is regularly followed by a homorganic nasal + stop cluster (and sometimes by a liquid or nasal) in these isolects. PMP forms with a *əCVC(V)(C) pattern obtained a əNCVC(V)(C) pattern in SM, SWY, IBN and JKT, and a aNCVC(V)(C) pattern in MIN and BH. As all isolects have reflexes with a homorganic nasal, nasal accretion must already have taken place in PM, cf. PMP *hepat > PM *əmpat, PMP *e(N)bun > PM *əmbun, PMP *e(N)tut > PM *kəntut (cf. 7.2.5). An exception to this phonotactic rule is SM *əjan,* MIN *ajan,* BH *hajan* 'squeeze out by pressure' (which has a variant *rəjan,* see Chapter 2 fn.21). Dyen reconstructed PMP *eZen* 'squeeze' on the basis of SM *əjan* and Toba *odon,* JV *ədən* 'id.' (as a correction to *eden* in Dempwolff 1938, Dyen 1951:536-537). On account of BH *hajan* this reconstruction has to be readjusted to PMP *qeZen.*

The only Malayic cognate set belonging to Pattern I is SM *b/apa, b/apa/k, b/apa/ŋ*, MIN *apa/ʔ*, *b/apa/ʔ, pa/ʔ*, BH *b/apa*, SWY *b/apa/ʔ*, JKT *b/apa/ʔ* 'father', IBN *b/apaʔ* 'father-in-law' and *apay* 'father'. No cognate set belongs to Pattern II. The cognate sets belonging to Pattern III are (1) SM *pamah* 'low-lying (land)', SWY *pama(h)* 'swamp'; (2) a.i. (except JKT) *danaw* 'lake, pool'; (3) MIN *ja/ño*, BH *ja-ña* (and also KD *ja-ña*, Dunselman 1949:69) 'what she/he says, his/her words'; in BH (and KD) this is still a morphologically complex form (cf. BH *ja-ku* 'my words, what I say').[148]

The cognate sets belonging to Pattern IV are (1) SM, BH, JKT *m/ampus*, MIN *m/ampuyh* 'dead, wiped out (coarse)'; also SM *(h)ampus* 'id.', MIN (Van der Toorn) *ampuyh* 'wiped out'; (2) SM, MIN *m/umbaŋ* 'coconut in its earliest stage of growth'; (3) a.i. *m/impi* 'dream (v, n)'; also SM, JKT *impi* 'dream, hope for, fancy; vision, illusion'; (4) SM, BH, IBN *mabuk*, MIN, SWY *mabuʔ*, JKT *mabòk* 'intoxicated'.

The cognate sets belonging to Pattern V are (1) a.i. *bawaŋ* 'onion'; (2) SM, JKT *bawal* 'k.o. fish'; (3) SM *p/awaŋ*, MIN *p/awaŋ, pu/awaŋ* 'guide, shipmaster; expert in any art believed to need the use of magic'; this is originally a contracted form *(ʔpu + ʔawaŋ)*, as can still be seen in MIN *puawaŋ* (Adelaar 1983:61). From the 11 correspondence sets given above two must be discarded, because they are historically complex: one is MIN *ja/ño*, BH *ja/ña*, and the other is SM, MIN *p/awaŋ*, MIN *pu/awaŋ*.

So there are actually nine correspondence sets which are not subject to the consonant-harmony constraint, and which yield the following PM lexemes: SM *b/apa/k* etc., SM *danaw* etc., SM *pamah* etc., SM *m/ampus, hampus* etc., SM *m/umbaŋ* etc., SM *m/impi, impi* etc., SM *ma/buk* etc., SM *bawal* etc., and SM *bawaŋ* etc.

SM *b/apa/k* etc. developed from a proto-form without *b-*. This proto-form was probably PM **apa(ʔ)*, but it may also have been a TAM loanword.[149] *b/-* remains problematic. It is observed in some other kin terms and may be a fossilised prefix, but its exact nature is not clear.[150]

---

[148]cf. Tioman Malay *yə diə* '(s)he said'. These forms are possibly from *+ia* + a personal pronoun (*+d-ia*, cf. 5.5.1, 5.5.1.2), with desyllabification in Tioman Malay, MIN, BH, and KD, and subsequent change from *+y-* to *j-* in MIN, BH, and KD (analogous to the change of *+y-* to *j-* in loanwords, cf. fn. 125). Another possibility is that they are allegro forms for *+ujar* + a personal pronoun, cf. Selako *ja-ku, jar-e* 'I/she/he said'.

[149]According to Asmah Haji Omar (1975), the TAM terms *amma:* 'mother', *appa:* 'father', and *ayya:* 'father, respectable man' are the sources for SM *əma/k, b/apa/k* and *mama/k* respectively. She found the terms in Burrow and Emeneau (1961), which has furthermore *ma:ma:* 'mother's brother'. These four TAM kin terms have cognates throughout the Dravidian language family (Burrow & Emeneau 1961). PM has **(ə)ma(ʔ)*, **apa(ʔ), *ayah* and **mama(ʔ)* respectively (3.4.2.5; 3.9.3 UIC), and at least **(ə)ma(ʔ)* and **apa(ʔ)* have replaced PMP terms (viz. PMP **ina* 'mother' and PMP **ama* 'father'). These facts suggest that **(ə)ma(ʔ), *apa(ʔ), *ayah* and **mama(ʔ)* were borrowed from Dravidian languages. This either implies that TAM already had an influence on PM, or, if not, that the four kin terms in question cannot be reconstructed for PM on account of their being borrowed.

[150]Apart from IBN *apay*, MIN *apa/ʔ* and Madurese *əppaʔ* 'father' (3.4.2b) there are many other correspondences of SM *b/apa/k* without initial *b/-*, cf. Toba, Ngaju *apa* '(one's own) father', Soboyo *n-apa* 'grandmother', Kapuas *apa/ŋ* 'father', Mualang *apay* 'father' (Adelaar 1983:59 n.9), Malagasy *z/afy* (< *+i-ape*) 'descendants', SD *apàʔ* '(speaker's own) father' (Adelaar unpublished fieldnotes), SUN *apa, bapa* 'father'. The notion of 'one's own father' is prevalent in these correspondences, and it is opposed to a wider meaning of father (other people's father, father-in-law) in the correspondences with *b/-* or in other available terms, cf. IBN *bapaʔ* 'father-in-law', Toba *ama-ŋ* 'father, a term which can be used towards young people, and hence by a father to his son' (Van der Tuuk), Ngaju *bapa*, SD *bapàʔ* '(someone else's) father', Maranao *bapa/ʔ* 'uncle, father-in-law', Hanunoo *bapa/ʔ* 'uncle'. In Mualang and in several other languages (including some Moluccan ones) *bapa(k)* occurs with the meaning 'father', but it is felt as a borrowing.

An initial *b/-* as in SM *b/apa/k* etc. is also observed in a number of other kin terms, or in terms which are associated with this semantic category (such as domestic, maidservant, or female (of animals)). Compare: *b/uda/k* 'lad, lass' (also 'domestic, slave') and **m/udaʔ* 'young, unripe' (3.4.2.4), and furthermore SD *kamudàʔ* 'youngster', SM *maʔuda* 'mother's younger sister' MIN *udo* 'older brother' [sic];

The ancestral forms of SM *m/ampus, hampus* etc. and SM *m/impi, impi* etc. may have been morphologically complex at the PM level because they show reflexes with and without *m-*. But no isolect has a productive prefix *m-* (or infix *-um- (-əm-)* < PAN/PMP **m-/ *-um-,* Wolff 1973:73), so that it remains uncertain whether *m/-* in *m/impi, m/ampus, m/ampuyh* was still a living prefix or not in PM (cf. 6.8). In the same way, it is also possible that SM *mabuk* etc. and SM, MIN *mumbaŋ* were still morphologically complex in PM, but one is not sure (SM *ma/buk* etc. < PAN **ma-buSek,* Blust 1976:125 n.18). SM, JKT *bawal* may be a TAM loan (Wilkinson 1959). *Bawaŋ* 'onion, garlic' (all isolects) or, more generally, 'bulb' (Wilkinson 1959), is widespread all over Indonesia and in the Philippines (where it means 'garlic' as against *lasuna* 'onion' (< SKT), or against *sibuyas* (< Spanish *sèbollas* 'onions'). The different varieties of onion and garlic were introduced into Southeast Asia from India, China and Europe, and most of these varieties do not grow very well on the wet soil of Malaysia and the Archipelago (Burkill 1966:99-103). It is possible that SM *bawaŋ* developed from a PM **babaŋ* 'bulb', and was later on borrowed from one or more Malayic isolects into languages all over Indonesia and the Philippines (in the Philippines its specific meaning may be derived from SM *bawaŋ putih* 'garlic'). In this context it is worth mentioning that in European languages there is a similar confusion in the etymology of terms for 'garlic' and 'onion'.[151]

Dempwolff reconstructed PMP **bawaŋ* 'onion' on the basis of Tagalog *bawaŋ* 'garlic', Toba *baoaŋ,* JV, SM, Ngaju *bawaŋ* 'onion'. But Toba *baoaŋ* is not inherited (Adelaar 1981:13), and JV has also *brambaŋ* 'red onion'. Ngaju *bawaŋ* is ambivalent, and could be from **bawaŋ* as well as **babaŋ.* It is likely that the reflexes of Dempwolff's **bawaŋ* are loanwords from SM, and that SM *bawaŋ* developed from PM **babaŋ* (or **baŋ*). At present

---

b/esa/n 'the relationship of persons whose children have intermarried', and SD *isàtn* 'term of address or reference for siblings of those who are married to one's siblings' (one is in a *ba-isàtn* relationship with the siblings of those who are married to one's siblings); (the same absence of *b/-* and the same m   ning is observed in the corresponding form in several other isolects in western Borneo);
b/iras 'wife's sister's husband, husband's brother's wife', Tagalog *bílas* 'id.' (but probably borrowed from SM), and Manobo *izas* 'spouses of two people who are related to one another (congenerationally', Timugon Murut *ilas* 'husband's brother's wife'; probably related to SUN *ŋ-iras* 1.' do s.th. in combination with s.th. else, combine two functions' 2. (also *ŋa-hiras*) 'ask other people's assistance for work one cannot perform alone (e.g. construction work, harvesting)';
bə/t/ina 'female (of animals)' and PMP **ina* 'mother', **t-ina* '(id., as a term of reference)'
ba/bu 'maidservant' (in Old Javanese and in dialectal SUN it occurs as 'mother' or as 'servant'); possibly related to SM *ibu, bu,* JV *mbo?* 'mother' ?
Embaloh (a Tamanic language of West Kalimantan) *b/aki?* 'grandfather' (a Malayic borrowing) and PM **aki?* 'grandfather'.
In *b/apa/k, b/uda/k, bə/t/ina* (and *ba/bu?*), the variants with **b(a)-* seem to convey the notion of 'one who acts as -, or is considered as - [the one designated with the variant without **b(a)-*]'. Whether this is evidence for a fossilised prefix is a matter of further investigation.
[151]Borrowing of words for 'onion' has crossed the divisions into Romance, Germanic and Slavonic languages. Compare the terms that are related to 1. *oignon* (French) 2. *Zwiebel* (German) 3. *luk* (Russian), all meaning 'onion', in ten European languages (in the list below, (1), (2) and (3) should be read 'related to 1, 2, 3' respectively):
French: *oignon* 'onion' (1), *ail* 'garlic' (none of 1,2,3);
Spanish: *cebolla* 'onion' (2), *ajo* 'garlic' (none of 1,2,3);
Italian: *cipolla* 'onion' (2), *aglio* 'garlic' (none of 1,2,3);
English: *onion* (1), *garlic, leek* (3);
Swedish: *lök* 'onion' (3), *vitlök* 'garlic' (3);
Dutch: *ui,* (Flemish) *ajuin* (1), Eastern Dutch (vernacular) *siepel* (2) all meaning 'onion', and *knoflook,* (Flemish) *look* 'garlic' (3);
German: *Zwiebel* 'onion' (2), *Knoblauch* 'garlic' (3);
Polish: *cebula* 'onion' (2), *czosnek* 'garlic' (none of 1,2,3);
Bulgarian: *luk* 'onion' (3), *byal luk* 'garlic' (3), *česán* 'id.' (none of 1,2,3);
Russian: *luk* 'onion' (3), *česnok* 'garlic' (none of 1,2,3).
The etymologies of the terms related to 1, 2, 3, are (1) Latin *unio, unionis* 'unity, union, a k.o. large pearl, a single onion', (3) Old Teutonic **lauko-* 'onion' (*The Shorter Oxford English Dictionary* revised and edited by C.T. Onions), and (2) Latin *cepula,* a diminutive of *cepa* 'onion' (*Duden*). Note that Bulgarian *byal luk* and Swedish *vitlök* literally mean 'white onion', cf. SM *bawaŋputih* (also a common circumscription in other (non-Indo-European) languages). Note also that correspondences to *luk* (3) mean 'garlic' as well as 'onion'.

there is not sufficient evidence to discount each of the nine PM lexemes which do not conform to the articulation-type harmony of homorganic consonants, although it is clear that there was a tendency to such a consonant constraint in PM (or at least in a not-too-remote period preceding PM).

4.3.2 AN ARTICULATION-TYPE HARMONY APPLYING TO *s AND PALATALS

It is worthwhile extending the line of questioning followed above by investigating whether PM had a constraint on combinations of *s and palatals in initial and intervocalic position. If it had, it is probable that PM *s was a palatal. An original palatal value of *s is also suggested by the fact that nasalisation of initial s results in its replacement by ñ in all isolects. (On the other hand, PM *-ŋs- was reconstructed on the basis of a nasal + *s cluster in the isolects (cf. 3.6.2).

The data however do not allow a definite conclusion (see Appendix II). They show that:

(1) none of the isolects have inherited lexemes containing an initial palatal stop or nasal in combination with intervocalic s;

(2) IBN has only combinations of initial s with intervocalic y;

(3) combinations of initial s and intervocalic palatal stops or nasals are particularly rare in BH and SWY;

(4) the following correspondence sets exist:
SM səjak, səñjak, MIN saja?, JKT səjək (səjəg) 'since';
SM səjuk, MIN sajuᵊ? 'cold, fresh';
SM, JKT səñum, MIN sañun 'smile (v)';
SM, JKT səñap 'deserted, lonely', SWY səñap 'dizzy, stupefied, numb'; cf. also BK señap (səñap?) 'deserted, lonely'; (MIN saño? 'id.' is only in local use, Van der Toorn);
a.i. (except IBN) sayaŋ 'pining, longing, pitying; love, affection';
SM, SWY, IBN, JKT sayap, MIN sayo? 'wing' < *sayap (3.1.1.3);
SM, BH, IBN, JKT sayur 'vegetable(s)';
SM, BH, IBN sayat 'cut, slice off' < *sayat (3.3.1IC).

Of the above examples one set contains too few examples to yield a PM reconstruction (SM səjuk, MIN saju?). The sets SM səjak etc., səñap etc., and səñum etc. have each three or more reflexes, but none of them has a BH or IBN reflex. Moreover, the co-occurrence of forms like SM tiñjak 'tread on' and SM jəjak 'trample on' (labelled as JKT, and ultimately as a borrowing from JV or SUN, in Wilkinson 1959) strongly suggest that səjak was originally polymorphemic. But then again, if it ever was, it is not clear whether it was still in PM or only in a period prior to it .

Although apart from IBN each isolect has some lexemes with a combination of initial s and an intervocalic palatal stop or nasal, the only well-attested combination in all Malayic isolects is initial s with intervocalic y. It is very likely that PM contained lexemes with this combination. It is also likely that PM did not have a constraint on the combination initial *s and intervocalic palatal stop or nasal, unless an explanation is found for the occurrence of səjak etc., səjuk etc., and səñum etc. (such an explanation may be that these lexemes were morphologically complex, but as yet there is no sufficient evidence for this). On the other hand, no inherited lexemes with initial palatal consonant and intervocalic s are found in any of the isolects, and it is very probable that PM did not have lexemes containing this

combination. This information is not sufficient to make definite conclusions about the phonetic character of PM *s. Although there may have been a structural relationship between the ancestor of *s* on the one hand and the ancestors of *c, j,* and *ñ* on the other, it is not clear what this relationship was, or if it existed at the PM level or only prior to it. On the basis of *s* in all isolects (which is an alveolar) I assume that PM *s was also an alveolar.

Reconstructions:

sayaŋ 'pining, longing, pitying';
sayur 'vegetable(s)';
səñap 'deserted, lonely';
səjək 'since';
səñum 'smile (v)'.

4.4 A CONSTRAINT ON FINAL LABIALS PRECEDED BY *i

The evidence presented here is not sufficient to state the existence of such a rule as a fact, but it is rather striking that:

(a) Only one correspondence set was found for the reconstruction of PM *-im, viz. *kirim (3.1.2.1). But *kirim has an irregular IBN reflex *kirum,* whereas the usual IBN term for 'send' is *pait.*[152] SM and JKT do not have other inherited lexemes with *-im/-em* (JKT *-im/-ém/-èm*), and SWY has only *udim* 'faeces'. BH has *siim* 'quiet, desolate', and *pijim* 'close the eyes'. In IBN no final *-im* sequences occur, nor in MIN, where *-m became *n* after a high vowel (3.6.3, 3.6.3.1). (cf. also SM *kəlim* etc. in 3.1.1.2 UIC (2).)

(b) Reflexes of *-ip are infrequent in the isolects, although they are not as exceptional as reflexes of *-im. No full correspondence set yielding a sound PM reconstruction was found, and the only examples which reflect *-ip in more than one isolect are:

(1) SM, BH, SWY, JKT *kacip,* IBN *kacit* 'betel-nut scissors'; this is a JV loan according to Wilkinson (1959) and Van der Tuuk (Von de Wall 1880, II:475);

(2) SM, JKT *cicip* 'taste (v)', BH *cicip* 'examine';

(3) SM *lañcip, liñcip,* BH *liñcip,* JKT *lañcip* 'smooth and pointed';

(4) SM *(h)intip,* BH, JKT *intip* 'spy, lurk', which is given as a JV loan in Wilkinson (1959) and Klinkert;

(5) SM *sisip* 'insert', MIN *sisiʔ* 'add' (3.1.2.1);

(6) SM, JKT *sirip,* BH *sirit* 'fin';

(7) SM, JKT *kutip* 'extract (v)';

(8) SM, JKT *kədip, kəlip* 'flicker, blink'.

In Dempwolff (1938) there are three PMP reconstructions with *-ip that are reflected in one or more Malayic isolects:

(1) PMP *sisip 'insert', cf. SM *sisip,* MIN *sisiʔ*;

---

[152]*pait* 'send' is also found in BAC. Languages outside the Malayic group that have a cognate of *pait* for 'send' are Ngaju (ait), Kadazan ((popo-)ovit), the Chamic languages (Rhade *moit,* Jarai *mo-it*), and

(2) PMP *ketip* 'squeeze between the fingernails', which became SM *kətip* 'nip or bite, of small insects';

(3) PMP *quDip* 'live', which became SM *(h)idup*, MIN *iduyʔ*, BH *hidup*, SWY, IBN, JKT *idup* 'id.' (the vowels were metathesised, a feature in which the Malayic isolects distinguish themselves from most other AN languages).[153]

In summary, (1) there is only one possible reconstruction with a final *-im* sequence (and a weak one at that), (2) there are only eight poorly represented correspondence sets which reflect *-ip*, and (3) there is a metathesis of vowels in the only PMP reconstruction with *-ip* that is well represented in the isolects viz. PMP *quDip* > *hidup*. These arguments clearly point to a PM constraint on final labials preceded by *i*.

4.5 VOWEL CONTRACTION IN PM LEXEMES OF MORE THAN TWO SYLLABLES

Original trisyllables with laryngeals separating like vowels, or separating schwa and another vowel, were contracted to PM disyllabic lexemes. This process involved loss of the PAN/PMP laryngeal, and the development of a new PM vowel with the same colouring as the contracted vowels if these were like ones; PAN/PMP combinations of *a* or a high vowel with *ə* were reduced to PM *a* or a high vowel respectively, e.g.

PMP	*beReqat* 'heavy'	> PM	*bərat*
	tuqela(n,ŋ) 'bone'		*tulaŋ*
	qahelu 'pestle'		*halu*
	peReqes 'squeeze'		*pərəs*
	qijuhuŋ 'nose'		*hiduŋ*
PAN	*tinuʔun*[154] 'weave'	> PM	*tənun*
	b(in)aHi[155]/*bineHi* 'woman'		*bini* 'wife'
	bineSiq/*beneSiq*[156] '(planting) seed'		*bənih*

Contraction of vowels also happened in PAN reconstructions which have more than two syllables through affixation or doubling of the root morpheme, e.g.

PMP	*ma-kaʔen* 'eat'	> PM	*ma/kan*
	ka + *wanan* 'right(-side)'		
	> (pre-PM *ka* + *anan*)		*k/anan* (cf. 3.4.1.4 and 7.1C)
PAN	*ka* + *luSeq* 'tear (n)'		*kəluh*[157] 'sigh, complain'
PWMP	*(dD)ehuk(-(dD)ehuk)*[158] 'sit, dwell'		*duduk*

Achehnese *(peuʼét).* Thc Kadazan form suggests that these cognates originally reflect a causative form (cf. Kadazan *-ovit* 'bring, carry', *popo-ovit*'(cause to be brought, carried =) send'.

[153]According to Blust (1981:463), the only other non-Malayic languages with a reflex of PMP *quDip* with metathesised vowels are Rejang *(idup* 'living, alive') and SUN *(hirup* 'living, alive', and *hurip* 'revive; flourish, thrive (as vegetation)'). However, Balinese and Sasak *idup* 'live' also shows this metathesis.

[154]See Dyen 1965b:292.

[155]Blust (1982b) reconstructs *baHi(/*b-in-aHi),* but Nothofer (1984) reconstructs *beHi(/*b-in-eHi).*

[156]Blust originally reconstructed *bineSiq* (1982a:289), but in Blust (1984a) he gives a doublet *beneSiq* (on account of Manobo *beniʔ*).

[157]However, the contraction in SM *kəluh* etc. is possibly a post-PM development. If contraction had already taken place in PM, PAN *ka-luSeq* should have become PM *kaluh,* SM +*kaluh,* SWY +*kaluʔ(h),* JKT +*kalu.* But it is also possible that PM *kəluh* did occur, and was a regular reflex of PAN *ke-luSeq,* as it is uncertain whether PAN/PMP had *a* or schwa in antepenultimate syllables with non-high vowels.

[158]PWMP *(dD)ehuk* 'sit, dwell' is reconstructed here on the basis of SUN *diuk,* Ngaju *oŋ/dok,* Achehnese *duëk,* Proto Chamic *dōk* and PM *duduk* 'id.'.

In one PM lexeme contraction did not take place (although it did take place in the indiviual isolects), i.e.

PMP　　　*baqeRu 'new'　　　　　>　　*bAharu (3.1.3.3)

PAN/PMP trisyllables with differing vowels other than schwa which were adjacent, or were separated by a laryngeal, remained trisyllabic in PM, e.g.

PMP　　　*kaluaŋ 'bat'　　　　　>　　*kAluaŋ

　　　　　*buqaya 'crocodile'　　　　*buhaya

　　　　　*baŋkuaŋ 'k.o. plant'　　　*bAŋkuaŋ

　　　　　*be(r,R)uaŋ 'bear'　　　　*biruaŋ (3.1.3.1)

　　　　　*ka + *wiRi 'left(-side)'
　　　　　> pre-PM *ka + *iri　　　>　　*kA/iri (cf. 5.7 (2) and 7.1c)

Contraction of adjacent vowels in trisyllables (or tetrasyllables) has also been an ongoing process in the isolects after the PM stage. All isolects contracted adjacent vowels of original polymorphemic or reduplicated PM forms. In many cases the uncontracted and morphologically non-complex root morpheme still occurs. For example:

SM lu/tut, MIN lu/tuyʔ, SWY ən/tuat 'knee', and BH tuut (3.10);
BH am/pun 'apology, forgiveness; ownership, property', o.i. 'pardon, forgiveness', and SM mohon, MIN puhun, JKT muhun 'beg, request' (also SM pohon 'tree' etc., 3.1.2.4);
SM,JKT tə/luk, MIN ta/luyʔ, BH ta/luk, SWY tə/luᵊʔ 'bay, inlet', and BH luuk 'bay; cavity under water' (3.10);
SM, IBN am/put, MIN am/puyʔ 'copulate', and BH puut 'vagina' (3.10);
MIN a/buᵊʔ, IBN buuk 'hair of the head', cf. also KD bu'uk 'id.' (Dunselman 1949:81; 3.10).
SM, SWY, JKT bəban, MIN baban 'load', and SM, BH bawa, MIN bao, SWY bawo, IBN baiʔ 'bring, take, carry', which yields PM *ba(ʔ) 'bring, take, carry'; this proto-lexeme acquired the (now fossilised) suffix -iʔ in IBN (cf. 6.1.1), and was reduplicated in the o.i.; bəban, baban is the result of reduplication of *ba(ʔ) + suffixation of the nominal affix *-an, thus *ba(ʔ)+*ba(ʔ)+*an, which became +bəba(ʔ)an > SM, SWY, JKT bəban, MIN baban.

As there is no way of telling whether contraction in the above examples had already taken place in PM or took place in each isolect independently, I reconstruct a complex form with vowel contraction when both the complex form and the root morpheme are reflected in one and the same isolect. For example, *tAluk is reconstructed along with *luø(uə)k because both are reflected in BH, and *ampun is reconstructed along with *puhun because most isolects reflect both of them.

N.B. The following cognate set must have developed from *(əm)pu - *hiaŋ 'Lord God; ancestor' (cf. *(əm)pu (3.11B) and *hiaŋ (3.9.1)), with loss of *h and desyllabification of intervocalic *i to y:

SM mo/yaŋ 'great-grandparent', MIN pu/yaŋ (Van der Toorn) 'great-grandmother', SWY po/yaŋ 'great-grandfather, ancestors; patriarch', IBN akiʔ-pu/yaŋ 'great-grandfather', JKT nènè/k-mo/yaŋ (also SM nene/k-mo/yaŋ) 'ancestor'; also (in Wilkinson 1959): po/yaŋ (Sumatra) 'patriarch, old man', (MIN) 'shaman'.

4.6 Syncope of the Penultimate Vowel of Trisyllables

Blust (1982a) describes a historical development in SM phonology which involved the loss of penultimate schwa in original PAN trisyllables, and the subsequent assimilation or elimination of one of the components of the resulting consonant cluster ("cluster-reduction"). He (p.285) calls this development "shwa-syncope" For example:

PAN *qalesem 'sour, acid' > SM, BH, SWY (m-)asam 'sour, acid', MIN, IBN asam 'sour fruit', m/asam 'sour', JKT asəm 'sour fruit; sour';

PAN *patelaŋ 'spotted, striped' > SM pəlaŋ 'striped, banded';

PAN *qaŋeliC / *qaŋeSiC 'burning smell' > SM (h)aŋit 'foul-smelling (dirty linen, dung, armpits etc.)', MIN aŋiʔ, BH, SWY aŋit 'burning smell', IBN aŋit 'fragrance' (cf. Blust 1980b:37-38).

In two of Blust's examples a penultimate high vowel (instead of schwa) is lost:

PAN *paŋuDaN 'pandanus' > o.i. pandan;

PAN *qaNiCu 'ghost, spirit of the dead' > SM hantu 'generic for invisible spirits of evil that work in darkness or secrecy', MIN, SWY, IBN antu 'ghost, demon'.

A third, originally polymorphemic example may be added here:

PAN *C-um-ubuq 'grow' > *t/um/buh 'id.' (3.5.1; 5.6.4).

Comparison of SM with the other isolects shows that in most of Blust's examples schwa-syncope has taken place in pre-PM. But in two of his examples the isolects disagree in reflecting a trisyllable, or a disyllable with a consonant cluster, e.g.

PAN *timeRaq 'tin, lead' > SWY tima(h), JKT timè, o.i. timah; but MIN also has timarah, timbarah 'tin foil'; MIN timah is assumed to be a loan (from SM), and MIN timarah, timbarah are explained as reflections of a Proto Minangkabau-Malay +timrah in which a schwa (> MIN a) was inserted to break an unusual heterorganic consonant cluster. In timbarah the insertion of schwa was preceded by an excrescent b (or "obstruent-insertion", Blust 1982a:288 fn.8).

PAN *saŋelaR 'fry without oil' > PM *saŋ(ə)lar > SM səlar 'branding', MIN sala, saŋla 'broiling', IBN bau salaʔ 'smell of roasting flesh' (3.7.4IC).

A few other cognate sets can be added to these two:

PM *pura(ñ)ci(kt) 'squirt' > SM pərcik, pərcit, BH puracit, IBN pərañcit, pərəñcit, JKT pərcik (3.7, 3.7.5);

PM *tirajaŋ 'kick (v)' > SM tərjaŋ, tərajaŋ, BH tirajaŋ, IBN tərajaŋ, JKT tərjaŋ, tərəjaŋ (3.7, 3.7.5);

PM *bArunas 'full, of rice ears' > SM bərnas, bərənas, BH barunas, SWY bəxnas (3.7, 3.7.5);

MIN paraʔ, SWY pəlaʔ 'garden', cf. also OM parlak 'id.' (Aichele 1942-43:52).

From the reflexes in the Malayic isolects for PAN *timeRaq,*saŋelaR, and for PM *pura(ñ)ci(kt), *tirajaŋ and *bArunas, it appears that syncope was still an ongoing process in PM.

Although almost all of the examples given in Blust (1982) show the elimination of schwa, I prefer the term 'vowel-syncope' to 'shwa-syncope' because of PM *pura(ñ)ci(kt), *tirajaŋ and *bArunas in addition to PAN *paŋuDan and *qaNiCu.

On the basis of the above cognate sets I reconstruct for PM:

*asəm 'sour fruit', and
*m-asəm 'sour, acid';
*pəlaŋ 'striped, banded' (only in SM);
*haɲit 'smelling as if burned';
*pandan 'pandanus';
*hantu 'ghost, spirit of the dead';
*tim(ə)rah 'tin, lead';
*pVr(ə)lak 'garden' (cf. Toba porlak 'id.').

PROTO MALAYIC LEXICON

In this chapter an attempt is made to reconstruct parts of the PM basic lexicon pertaining to well-defined semantic fields. These fields are: times of the day (5.1); directional terms (5.2); numerals (5.3); basic kinship terms (5.4); pronouns (5.5); parts of the body (5.6).

This chapter also includes a modified Swadesh 200-item basic wordlist for PM (5.7) built on evidence of 200-item basic wordlists for the six isolects as well as on additional information. The list is meant to be a tool for measuring the relative position of other isolects vis-à-vis PM, and for counting the lexical retention rate of the six isolects.

5.1 TIMES OF THE DAY

Four major parts of the day are distinguished in the isolects: morning (5.1.1); noon, middle of the day (5.1.2); evening (5.1.3); night (5.1.4).

5.1.1 MORNING

Pagi is the regular term for 'morning' in all isolects except BH, which has *ba/isuk/an*. The root *isuk* is probably from JV (cf. JV *ésok* 'morning', from which probably also originated SM *b/esok* 'tomorrow'). As a PM term for 'morning' I will reconstruct *pagi*.

5.1.2 NOON, MIDDLE OF THE DAY

The term for 'noon, middle of the day' agrees in all isolects: SM (Malaysian) *təŋah hari*, MIN *taŋah ari*, BH *taŋah*[159] *hari*, SWY *təŋa(h) axi*, IBN *təŋah ari*, JKT *təŋari* (with sandhi); cf. also 'tamhahari' in Pigafetta's 1,521-word basic wordlist, which is the oldest western source for Malay (Pigafetta 1972:67). SM (Indonesian) and JKT *siaŋ* are sometimes used for 'middle of the day (c. 11 a.m. - 2 p.m.)', but in all isolects (see Aliana et al. p.86 for SWY) it means 'daylight (as opposed to the dark)', and (except in JKT) it also means 'clear, limpid, bright, clean' and (in derivations) 'weed, clear (woods, bushes, etc), clean'.

For PM 'noon' *təŋah *hari* is reconstructed; *siaŋ* is reconstructed with the meaning 'bright, clean; clean, weed (v)'. An interesting cognate set in this context is BH *ma-landaw* 'oversleep', IBN *s/andaw* in *gaway s/andaw ari* (loss of *+l-* unexplained) 'k.o. festival ending at midday', and (Richards) *andaw* 'rise late, lie abed (from laziness)', *apay andaw* 'father of the day' (name of a star) (< *andaw*, 3.2.2). BH *landaw* and IBN *andaw* seem to reflect PMP *qalejaw* 'day, daylight', although the sound correspondences are irregular.

---

[159]Durdje and Djantera 1978:45.

Either BH *l*- reflects PMP *-l-*, and PMP *qa*- was lost (cf. 3.1.3.3), which asks for an explanation of PMP *e* > IBN *a* and PMP *l* > IBN ø, or PMP *e* was lost through syncope, and BH *l*- remains unexplained. Moreover, it is not clear why BH and IBN have a cluster -*nd*-.

5.1.3 EVENING

SM and SWY have *pətaŋ*, MIN *pataŋ* for 'late afternoon, evening'; this meaning must also be secondary, since in IBN and JKT, as well as outside the Malayic group, *pataŋ, pətəŋ*, (JV *pətəŋ*) means 'dark, obscure'. BH has *kam/ari/an* 'evening, yesterday', which must be a contraction of **kə-lə(hø)əm* + **hari-an*. IBN has *ləm/ay* which is from **(kələ(hø)əm* + **hari*: the loss of **r* is unexplained, but cognates from other Ibanic isolects (in West Kalimantan) clearly show the development, cf. (in kabupaten Sanggau) Ketungau *mələyɛy*, Kerabat *mələyi* (both with metathesis), Mahap *mari*, Benawas *mayi*, Taman *bəmayi*, (in kabupaten Kapuas Hulu:) Suhait *gələmay* (Adelaar unpublished fieldnotes). SM (BI) *sore* and JKT *sòré* is a loan from JV.

Most likely, **pətəŋ* originally meant 'dark, obscure', and its reflexes were later on used metaphorically for 'evening' in SM, MIN, and SWY. This shift of meaning is analogous to that of SM, MIN, and JKT *siaŋ*. The PM term for 'evening' was presumably a compound **kə-lə(hø)əm *hari* which is still reflected in BH *kam/ari/an* and IBN *ləm/ay*; in other isolects this compound is still reflected in the term for 'yesterday', cf. SM *kəm/ari/n, kəm/are/n* (3.1.2.5), *kəlm/ari/n*, (Klinkert) *kəlam/ari, kalam/are/n*, SWY *kəm/axi*, JKT *kəm/arè/n* 'id.'; cf. also 'calamari' in Pigafetta's list (p.71), and BRU *kamai* 'yesterday' (in Wilkinson's (1959) presentation *kamai* has a schwa, but according to Prentice, schwa does not occur in BRU).

5.1.4 NIGHT

All isolects agree in having *malam* (JKT *maləm*) for 'night', which points to PM **maləm*. But it is likely that *malam* etc. is related to **kə-lə(hø)əm*, and I reconstruct a PM etymon **ma-lə(hø)əm*.[160]

5.2 DIRECTIONAL TERMS

No terms for cardinal points of the compass are reconstructable for PM. This is not surprising if one compares the isolects with other languages all over the world: terms for cardinal points as a rule originate as secondary semantic developments in lexemes primarily referring to (1) celestial bodies and rising/setting of the sun (as a consequence of this, midday/night, coming/going, etc.); (2) atmospheric features (warm/cold, names of winds); (3) other, more general directions ('up' vs 'down', 'upriver' vs 'downriver', 'right' vs 'left' etc.); and (4) environment-specific features (mountains, cliffs, rocky places, rough country etc.) (C.H. Brown 1982:5).

---

[160]BH *halam* 'yesterday, earlier' and Salako *àapm* 'morning', Mualang *lam* 'early morning' may be related.

In what follows here I first give the terms for cardinal points in the isolects, and show that they are not suitable for the reconstruction of PM cardinal points (5.2.1). Then I give other directional terms on the basis of which PM directional terms can be reconstructed (5.2.2).

5.2.1 CARDINAL POINTS

NORTH

All isolects except JKT have a SKT loan to denote 'north': SM, IBN (Richards), SWY (Aliana et al.) *utara,* MIN *utaro* (SWY *-a* and IBN pre-penultimate *u* point to very recent borrowing, undoubtedly via SM). JKT *ilir* 'north' developed from **hilir* 'downstream' (3.7.3 IC): the JKT term for 'south', *udik,* also means 'rural, provincial' and developed from **udi/k* 'upstream, upriver area', cf. also SM *udi/k* 'upstream area', BH *udi/k* 'countryside, place of origin', and BSM *udi/ʔ* 'upstream area'. The use of lexemes referring to the course of a river to denote 'north' and 'south' is related to the north-south orientation of rivers in the area around Jakarta (which is situated on Java's north coast).

SOUTH

SM *səlat/an* developed from a morphologically complex directional term referring to the Strait of Malacca (cf. SM *səlat* 'strait, narrows' < **səlat* 3.1.1.5). SWY, IBN *səlatan* and MIN *salatan* are very likely SM loans, since the coincidence of 'south' and 'straits' is only evident from the geographical position of the inhabitants of the Malay Peninsula (and, in theory, of the inhabitants of Banjarmasin). JKT *udik* < **udi/k,* see above.

EAST

Timur occurs in SM, BH, SWY, and IBN, and *timuᵊ* occurs in MIN; the usual term for 'east', however, is *ujuᵊŋ* in MIN and *mata ari idup* or *mata ari tumbuh* in IBN. IBN *timur* must be borrowed, because this lexeme did not undergo excrescence of a voiced stop after a nasal before *-r* (3.6.1.5). In SWY *timur* as well as *barat* 'west' have an apical trill and are probably borrowed (3.7), just as *utara* and *səlatan* are (see above; cf. also Semende (another Middle Malay isolect) *mataxi idup* 'east' and *mataxi mati* 'west'). It is likely that *timur/timuᵊ* in the other isolects is borrowed from SM. The meaning 'east' of *timur/timuᵊ* is innovative and developed from PMP **hatimuR* 'south-east monsoon'. With an innovative meaning 'east' in SM and BH, and no inherited correspondences in the other isolects, no PM etymon meaning 'east' can be reconstructed. PM **timur* (as a descendant of PMP **hatimuR* and still reflected in at least SM *timur*) must still have meant 'south-east monsoon'. JKT *wèt/an* < JV *wèt/an* 'east'.

WEST

Barat is found in SM, MIN, SWY, and IBN (Richards), but it is clearly a loan in MIN (where *+bareʔ* would be expected instead of *barat*). More usual terms for 'west' are *puhun* and *mato ari mati, mantari mati* (Van der Toorn) in MIN, and *mata ari turun* or *mata ari mati* in IBN. SWY *barat* is probably a loan (see above). JKT *kulòn* < JV *k/ulo/n* 'west'. SM

barat is a reflex of PMP **habaRat* 'north-west monsoon'. PM **barat* should be reconstructed with the same meaning as PMP **habaRat:* as it appears that none of the other isolects has an inherited reflex of PMP **habaRat,* the shift of meaning to 'west' can not be attributed to PM and must be considered as a SM innovation.

So it appears that in most isolects terms for cardinal points are borrowed or have originated through semantic extension in lexemes of a primary meaning closely associated with a cardinal point.

5.2.2 OTHER DIRECTIONAL TERMS

Important directional terms in AN languages are reflexes of PMP **lahud* 'towards the sea' and **Daya* 'towards the interior', and furthermore reflexes of PMP **habaRat* and **hatimuR* indicating the directions from which the seasonal winds blow. Since the meanings 'east' and 'west' are innovative and cannot be attributed to PM, the PM ancestor of SM *barat* and *timur* must have maintained the meaning 'north-west monsoon' and 'south-east monsoon' respectively (5.2.1). As to PMP **lahud* and **Daya, *lahud* changed its meaning to 'sea' in the isolects, whereas **Daya* only survived in compounds such as *barat-daya* 'south-west' and *oraŋ daya/k* 'Dayak (= inhabitant of the interior of Borneo)'. Nevertheless, the old directional meanings can still be reconstructed for PM. The terminology of the SM compass card becomes logically more consistent if we presume that the terms *utara* and *səlat/an* have replaced earlier terms +*laut* 'north' and +*daya* 'south' (Brandes 1884:102; Adelaar 1989:10). That this happened is still witnessed by the presence of the terms *timur-laut* 'north-east', *barat-laut* 'north-west' and *barat-daya* 'south-west': these are compound terms consisting of the closest directional terms in the original SM compass card system. The substitution must have taken place after the Malay power centre moved from South Sumatra to the west coast of the Malay Peninsula. In the Palembang area of South Sumatra (where Srivijaya was presumably located), the sea is in a northern direction, as opposed to the land, which is in a southern direction. On the other hand, in Malacca, which became the Malay political and cultural centre later on, the sea (and the Strait of Malacca) is to the south *(səlat/an),* and the interior is to the north *(utara).*

I reconstruct PM **daya?* 'towards the interior, up-country' on the basis of SM *barat-daya* 'south-west' SM *daya/k* 'Dayak', IBN *daya?* 'Dayak; up-country, inland'. I reconstruct PM **laut* 'towards the sea' on the basis of the directional terms SM *barat-laut* and *timur-laut.* Today, SM, BH, SWY, JKT *laut,* MIN *lauy?* means 'sea', and it replaced an earlier **tasik* 'sea' (5.7 lemma 124).

IBN *laut* means 'Malay'.[161]

5.3 NUMERALS

5.3.1 RECONSTRUCTION OF THE NUMERALS

---

[161]In IBN (and in many other Bornean languages) *laut* means 'Malay (i.e. coastal people)'. The SM term for 'south-east' is *təŋgara.* The origin of this term is unknown (it could be a TAM loan, where *ten-kara* means 'south bank', Menon [Govindankutty] pers.comm.).

TABLE 10: THE NUMERALS

	SM	MIN	BH	SWY	IBN	JKT	PM
1	s(u)atu, (ə)sa, sə-	cieʔ, sa-	asa, sa-	so, sə-	saʔ, sə-	(s)atu, sə-	*əsaʔ, *sA-
2	dua	duo	dua	duo	dua	duè	*dua(ʔ)
3	tiga	tigo	tiga	tigo	tiga	tigè	*təlu
4	əmpat	ampeʔ	ampat	əmpat	əmpat	əmpat	*əmpat
5	lima	limo	lima	limo	limaʔ	limè	*limaʔ
6	ənam	anam	anam	ənam	ənam	ənəm	*ənəm
7	tujuh	tujuᵊh	pitu	tujuᵊ(h)	tujuh	tujuʔ	*tujuh
8	(də)lapan	(sa)lapan	walu	dəlapan	(də)lapan	dəlapan .	*dua(ʔ) alap-an
9	səmbilan	sambilan	saŋa	səmbilan/ səlapan	səmilan	səmbilan	*əsaʔ ambil-an/ *(ə)saʔ alap-an
10	sə-puluh	sa-puluᵊh	sa-puluh	sə-puluᵊ(h)	sə-puluh	sə-pulu	*sA-puluh
11	sə-bəlas	sa-baleh	sa-walas	sə-bəlas	sə-bəlas	sə-bəlas	*sA-puluh əsaʔ
12	dua bəlas	duo baleh	dua walas	duo bəlas	dua bəlas	duè bəlas	*sA-puluh dua(ʔ)
13	tiga bəlas	tigo baleh	talu walas	tigo bəlas	tiga bəlas	tigè bəlas	*sA-puluh telu
20	dua puluh	duo puluᵊh	dua puluh	duo puluᵊ(h)	dua puluh	duè pulu	*dua(ʔ) puluh
21	dua puluh satu	duo puluᵊh cieʔ	sa-likur	duo puluᵊ(h) so†	dua puluh saʔ	duè pulu atu	*dua(ʔ) puluh əsaʔ
22	dua puluh dua	duo puluᵊh duo	dua likur	duo puluᵊ(h)	dua puluh dua	duè pulu duè	*dua(ʔ) puluh dua(ʔ)
23	dua puluh tiga	duo puluᵊh tigo	talu likur	duo puluᵊ(h) tigo†	dua puluh tiga	duè pulu tigè	*dua(ʔ) puluh təlu
30	tiga puluh	tigo puluᵊh	talu puluh	tigo puluᵊ(h)	tiga puluh	tigè pulu	*təlupuluh
100	sə-ratus	sa-ratuyh	sa-ratus	sə-ratus	sə-ratus	sə-ratus	*sA-ratus
200	dua ratus	duo ratuyh	dua ratus	duo ratus	dua ratus	duè ratus	*dua(ʔ) ratus
1,000	sə-ribu	sa-ribu	sa-ribu	sə-ribu	sə-ribu	sə-ribu	*sA-ribu
2,000	dua ribu	duo ribu	dua ribu	duo ribu	dua ribu	duè ribu	*dua(ʔ) ribu

† Helfrich gives *so likur* 'twenty-one', *duo likur* 'twenty-two', *tigo likur* twenty-three', etc. along with forms on the basis of *puluᵊ(h)*.

5.3.2 EXPLANATION OF THE RECONSTRUCTIONS

'1' – *əsa?, *sA-. All isolects agree in having at least a free and a cliticised form for 'one'. The cliticised form in all isolects reflects a proto-form *sA-. The free forms show more variety: SM *satu, suatu* agree with IBN *satu* and JKT *(s)atu* (3.8.1). *Satu* 'one' and *suatu* 'a (certain)' have undergone a semantic differentiation. Brandes (1884:162) derived SM *suatu* from ⁺*sa-* + ⁺*batu* 'one stone'. Gerth van Wijk (1889:198-199) supposed that SM *suatu* originated from JV *sə-watu* 'one stone' (rather than from a pre-SM ⁺*sa-* + ⁺*batu*). But supposing a JV origin is unnecessary on account of the fact that in the history of SM (and JKT), (1) *b underwent lenition between *a's, (2) PM antepenultimate *a was neutralised to schwa and subsequently lost, and (3) finally, following ⁺w was vocalised to u (Blust 1974a:134), and apparently also in JKT. Thus: ⁺*sa-batu* > ⁺*sawatu* > ⁺*səwatu* > ⁺*swatu* > ⁺*suatu* > SM *s(u)atu*, JKT *(s)atu*.[162] BH *asa* agrees with SWY *so* and IBN *sa?*, and also with SM *sa* '1', *əsa* 'One (in reference to "God's Unity")', which leads to a reconstruction *əsa? (in SWY and IBN initial ə never occurs before consonants other than nasals, 2.4.3 and 2.5.3). The origin of MIN *cie?* is obscure.

'2' – *dua(?). Final *(?) is reconstructed on account of the IBN derivations *pədua?/bədua?* 'divide' along with *dua*.

'3' – In spite of the agreement among the isolects in reflecting *tiga* etc., *təlu is reconstructed on account of OM, KD *talu*, BAC *tolu*. *təlu is not reconstructed with *-(?) because KD does not have a final glottal stop whereas it usually agrees with IBN in showing one. BH also has *talu* but this is considered a JV loan, see '8' below. In SM *təlu is still retained in *buah kəras təlu* 'a hard fruit with three pips', and *buah salak təlu* 'a salak fruit with three pips'.

'7' – BH *pitu* agrees with PMP *pitu, but it is considered as a JV loan (see '8' below). BAC also has *pitu*, which remains unexplained here.

'8' – The BH numeral system was heavily influenced by JV, e.g. *talu, pitu, walu, saŋa,* and the formant *likur*. Whereas *talu* and *pitu* could as well have developed from PMP *telu and *pitu, this is not the case with *walu* and *saŋa*. A Malayic reflex of PMP *walu would have lost its initial semivowel, and *saŋa* (which formally agrees with JV *saŋa*) does not agree in form with *siwa. On the other hand, BH *walu* cannot derive directly from JV *wolu*, but was probably borrowed from JV in a period when the latter still had *a* in the penultimate syllable of this lexeme (cf. Old Javanese *wwalu*). In any case, lexemes with initial *w* are not inherited in the Malayic isolects.

The form for 'eight' in the other isolects developed from an original compound which underwent contraction. This can still be seen from the archaic SM variant *du/alap/an* (Gerth van Wijk p.198; Wilkinson (1959) referring to the Classical Malay text *Busta:n us-Sala:ti:n*), from BAC *du/alap/aŋ*, from KD *du/alap/atn* (Dunselman 1950:322) and from 'duolappan' and *dualapan* in Pigafetta's Malay wordlist (Pigafetta 1972:72). *dua(?) alap-an originally meant 'two taken away (from ten)', and *əsa?-alap-an 'one taken away (from ten)', cf. Kedah Malay *alap* 'gather fruit by means of long pole to which a knife or hook is attached' (Wilkinson 1959) < PMP *alap 'take'. Blust (1981:467 n.6) gives the following explanation for MIN *s/alap/an* 'eight': ancestral forms of *s/əmbil/an* and *s/alap/an* were originally synonymous for 'nine', and MIN *s/alap/an* 'eight' is the result of a change of meaning after

---

[162]cf. MIN and Negeri Sembilan Malay *suaraŋ* 'joint property of husband and wife', < ⁺*sa* + ⁺*baraŋ* (*baraŋ* 'thing, wares, stuff, something'); cf. also JV *siji* 'one', which consists of *sə-* + *wiji* 'one seed'.

the loss of function of the original morphology of these forms. This must also be the case with the SWY doublet *s/əlap/an* 'eight'. Initial *l-* in IBN *l/əlap/an* is due to antepenultimate assimilation, cf. 3.11.

'9' – SM *s/əmbil/an* and its cognates must have developed from a compound form **əsa? ambil-an* ('one taken away (from ten)' = nine, cf. Klinkert). This compound form has replaced PMP **siwa,* and is sometimes circumscribed in SM as *kuraŋ əsa* or *kuraŋ satu* 'minus one'; such a circumscription is also used in larger numbers with 'nine' as the last component, e.g. *sə-ratus kuraŋ əsa* or *sə-ratus kuraŋ satu* 'ninety-nine', and *kuraŋ əsa əmpat puluh* 'thirty-nine' (Gerth van Wijk p.205). BH *saŋa* is a JV loan (see '8' above), and in IBN *s/əmil/an* the original **-mb-* was weakened to *-m-* (3.11A). BAC has *s/alap/aŋ* 'nine', hence **(ə)sa? alap-an* next to **(ə)sa?ambil-an.*

'11' – '19' – All isolects exhibit *bəlas* etc. for the formation of numbers between ten and twenty. Nevertheless I reconstruct **sA-puluh* with a following cardinal number for PM numerals between ten and twenty, on account of OM, BRU and BAC. In OM 'twelve' occurs in the following line of the Kedukan Bukit inscription (Palembang):

> daŋan jalān   saribu tlurātus sapulu dua vañak-ña
> with   infantry 1,000 300   ten   two quantity-their
> with an infantry of 1,312 soldiers in number (Çoedès 1930:34)

According to Prentice (pers.comm.) in BRU numbers between ten and twenty are formed with *sapulu* + cardinal number (although Ray gives them on the basis of *bəlas,* Ray 1913:62). In BAC one finds *sapulu sabua, sapulu dua, sapulu tolu* for 'eleven', 'twelve', 'thirteen' etc. (Stokhof 1980:97-98). These three isolects agree with PMP, where teens are formed with **sa-puluq* + a cardinal number.

'21' -'29' – BH (and optionally SWY) *likur* in compounds denoting numbers between twenty and thirty must be from JV: apart from BH and SWY, *likur* sometimes also occurs in SM, and, outside the Malayic group, in Maanyan Dayak, Ngaju, and Balinese. In Maanyan Dayak (and Ngaju?) *likor* is not inherited.[163]

5.4 BASIC KINSHIP TERMS

In this section those kinship terms are reconstructed which cover well-defined concepts in the isolects, and for which strong comparative evidence is available. No attempt is made to reconstruct the PM kinship system, that is, the set of relations that hold together in a structural whole the concepts that are denoted by PM kinship terms. The following concepts are investigated: ancestors (5.4.1); grandparents (5.4.2); parents (5.4.3); aunts, uncles, cousins, nephews, nieces (5.4.4); siblings (5.4.5); children, grandchildren, great-grandchildren (5.4.6); in-laws (5.4.7).

N.B. In many cases kinship terms underwent formal changes. SM kinship terms originally ending in a vowel (< **Vø,* or < **V?)* acquired a *-k* which is described as a fossilised vocative

---

[163]*Likor* as a formant of numbers between 20 and 30 in Barito languages is a JV loan (probably introduced via Banjarese). As has been pointed out by earlier scholars, JV *likur* had an original meaning 'behind' (< PMP **likud* 'back'), and formations like *ro-likur, təlu-likur* etc. literally meant 'two after (twenty), three after (twenty) etc., or two, three, etc. back again'. Ngaju *rikor, likor* must be borrowed, as this language has *likut* 'back, behind' as a regular reflex of PMP **likud.* Maanyan *likor* must also be borrowed, because in the inherited vocabulary of this language, PMP **-li-* became *-di-* (Dahl1951:54-55), and Maanyan lexemes containing *-li-* sequences must have been borrowed after this change had ceased to be operative.

suffix by Blust (1979). Blust reconstructed PWMP *-q as a vocative marker. He argues that after the change of PMP *q to h in JV and SM, the vocative forms in these languages would have lost much of the efficiency inherent in a final glottal stop in calling. For this reason a final glottal stop was reintroduced (which is phonemically -k, cf. 2.1.1). MIN, SWY, and JKT also have reflexes of this fossilised suffix (cf. MIN, SWY -ʔ, JKT -k and -ʔ). The same suffix is found in reflexes of *datuʔ (which is also a kinship term in some isolects, see below). It is possibly the same suffix as in negations and in tabi/k etc., in which case the suffix may be interpreted as a syntactic device used for words in isolation (including vocatives, negations, and greetings, cf. 3.4.2.4 UIC).

Examples:

*niniʔ 'grandparent' > SM nene/k, MIN niniᵊ/ʔ, ñiñiᵊ/ʔ, JKT nènè/ʔ vs BH nini, IBN iniʔ (3.1.2.1);
*kakaʔ 'older sibling' > SM kaka/k, MIN kaka/ʔ, kako, vs BH kaka, IBN akaʔ (3.4.2.4);
*datuʔ 'head of a clan' > SM datu/k[164] MIN, SWY datuᵊ/ʔ, JKT datu/ʔ vs BH datu, IBN datuʔ (3.4.2.4);
*apa(ʔ) 'father' > SM b/apa/k[165] MIN (b)apa/ʔ, SWY, JKT b/apa/ʔ vs BH b/apa, IBN apay 'id.', b/apaʔ 'father-in-law';
*mama(ʔ) 'maternal uncle' > SM mama/k, MIN, SWY mama/ʔ vs JKT mama/ŋ (see below) (3.4.2.5); (see below for *ma(ʔ) and *adi).

There are also other vocative suffixes: SM, MIN, SWY, and JKT exhibit a sporadic -ŋ, e.g. SM kaka/ŋ, bapa/ŋ, datu/ŋ, cucu/ŋ, indu/ŋ (especially in Old Malayo-Javanese literature, Wilkinson 1959; cf. also ana/ŋ 'child'); MIN cucuᵊ/ŋ, BH adi/ŋ, cucu/ŋ (3.4.1.3IC), SWY adi/ŋ, ibu/ŋ (see below), cucu/ŋ (3.4.1.3IC), JKT mama/ŋ (3.4.2.5), kaka/ŋ. Another suffix, SM -(a)nda, is honorific (cf. 5.5.1.3 N.B.). In IBN these fossilised suffixes do not occur, but in this isolect another formal change is observed: PM kinship terms of which the first and the second syllable are similar in structure lose their initial consonant, e.g. *cucuʔ > ucuʔ (3.4.1.3); *niniʔ > iniʔ (3.1.2.1); *kakaʔ > akaʔ (3.4.2.4); *cicit 'great-grandchild' > icit (see below).

5.4.1 ANCESTORS

Most isolects have a cognate of SM mo/yaŋ to denote 'ancestor' or 'great-grandparent', which yields *(əm)pu - *hiaŋ 'Lord God; ancestor' (4.5 N.B.). SM datu/k-nene/k 'ancestors' also occurs (Klinkert), but datu/k comes from *datuʔ 'head of a clan' (3.4.2.4) (< PMP *datu 'id.'); cf. 3.1.2.1 for nene/k.

5.4.2 GRANDPARENTS

For 'grandparent' *niniʔ is reconstructed (3.1.2.1). Reflexes of *datuʔ (3.4.2.4) meaning 'grandfather' also occur in SM, MIN, SWY (with unexplained -k in Aliana et al. p.86), and JKT, but except for JKT, in none of the isolects is this the primary meaning. *niniʔ was not distinguished for sex (and still is not in Malaysian SM): SM aki, IBN akiʔ

---

[164]In Jawi (Arabic) script these lexemes are written as follows (in my orthography I substitute ʔ for a hamza): 'bapa', 'bapaʔ', 'bapak', 'bapaŋ' (vocative form); 'maʔ' (but: 'əmak'); 'kakeʔ'; 'dik', 'adik'; 'datu', 'datuʔ'; 'nenek'; 'mamak'; 'induk'; 'əñciʔ' (cf. Klinkert).
[165]See fn. 164.

occur with the explicit meaning 'grandfather', yielding *aki? 'id.' (< PWMP *aki/*laki, Blust 1979:211). SM kake/k, JKT kakè/? 'grandfather' are from JV (Klinkert 1916; Wilkinson 1959).

5.4.3 PARENTS

For 'father' *apa(?) was reconstructed (4.3.1) as well as *ayah (3.9.3ICe). Reflexes of *ayah are not as common as those of *apa(?), and they are mainly used as a polite form, so it is likely that the common term for 'father' was *apa(?) in PM. Other lexemes for 'father' are BH abah, SWY ba?, and JKT babè. For 'mother' the isolects have: SM (ə)ma/k, ibu, indu/k, MIN (m)anday, ama/?, induᵊ/?, BH uma, SWY ma/?, ənduᵊ/?, IBN inday, JKT əña?, ibó/ibu. SM ibu (and hence also JKT ibó/ibu) used to be a polite form next to the more usual əma/k. Nowadays əma/k has been replaced by ibu, and it survives only in Malaysian SM. JKT əña? is a Chinese loan (Kähler 1966). I do not know the origin of MIN (m)anday, nor do I know the explanation of u- in BH uma.[166] On the basis of SM (ə)ma/k, MIN ama/?, BH uma, and SWY ma/?, *(ə)ma(?) 'mother' is reconstructed. Two other words for mother can be reconstructed. *indu? has reflexes in SM (indu/? 'dam'), MIN (indu/k 'dam'), and IBN (indu? 'woman; female'). The fact that its reflexes in SM and MIN do not refer to a mother of humans is not an impediment: a constant devaluation of old terms of address (from formal to informal and rude), and a constant creation of new ones, is observed in the isolects. *ina is reflected in IBN ina 'mother' (less usual than inday), and furthermore (with a semantic shift and fossilised affixes) in SM ina/ŋ 'duenna, governess of an unmarried girl of high rank', MIN, SWY ina/ŋ 'look after, nurse (children)', IBN ina/ŋ 'nurse (children), rear (animals)' (Wilkinson (1959) also gives JKT iña 'child's nurse', which is not found in Abdul Chaer), and in SM bə/t/ina, MIN ba/t/ino, SWY bə/t/ino, t/ino 'female (of animals)' (according to Aliana et al. t/ino means 'woman'). The more usual IBN term for 'mother', inday, may also be a reflex of *ina, although I do not have an explanation for its -d-.

5.4.4 AUNTS, UNCLES, COUSINS, NEPHEWS, NIECES

Apart from *mama(?) 'mother's brother' (3.4.2.5) no specific terms for parent's siblings, cousins, or sibling's children can be reconstructed. These concepts are denoted by different terms and/or circumscriptions. Compare the following list (which is not exhaustive):

(a) parent's siblings:

SM pa/k sawdara 'father's brother', bapa/k muda 'father's brother next in age', pa/k tua 'father's older brother', sawdara ma/k 'mother's brother', ma/k sawdara 'mother's sister', ma/k (kə)cik 'mother's younger sister', bibi 'id.' (< JV or Balinese, Wilkinson 1959), ua(k) 'father's older sister'; cf. also mama/k for which Wilkinson gives 'maternal uncle; (loosely) any uncle', and for which Klinkert gives 'aunt (in general)', and paman 'uncle (in general)' (< JV);

MIN (m)anday 'mother's sister' (also 'mother'), mama/k 'mother's brother', bapa/k, apa/k, pa/k 'father's brother' (also 'father');

SWY mərajo 'mother's brother', wa? 'mother's older sister', ibu/ŋ 'mother's younger sister', mama/? 'uncle (in general?)';

IBN ibu? 'aunt', aya? 'uncle';

---

[166]The u in BH uma may be due to rounding of *ə as a result of assimilation to a following labial nasal, cf. BH jumbatan, gumalan (3.1.3.1 N.B.).

JKT (in order of frequency, cf. Grijns 1980:208-209) *uaʔ, əñcaŋ, əndé, bapa/ʔ (gədé)* 'parent's older brother', *uaʔ, əñcaŋ, əmaʔ gədé, (əma/ʔ) əndé* 'parent's older sister', *mama/ŋ, əñci/ŋ, mama/ʔ, əñci/ʔ* 'parent's younger brother', *əñci/ŋ, əñcè/ʔ, bibi, əñci/ʔ* 'parent's younger sister'. *Mama/ʔ, mama/ŋ* also occur with the restricted meaning 'mother's older brother' in some parts of Jakarta (*mama/ŋ* could be a SUN loan, cf. SUN *mama/ŋ* '(term of address for parent's younger brother)'; Abdul Chaer also gives *óm* 'uncle' and *tantə*[167] 'aunt' (both from DU).

(b) cousins:

SM *sawdara sə/pupu*, MIN *dan/s/anak* (also 'siblings' and 'nephews / nieces'), BH *sa/pupu* (Durdje & Djantera 1978:42), SWY *moanay b/apa/ʔ* 'father's brother's son', *kəlaway b/apa/ʔ* 'father's brother's daughter (man speaking)', *moanay ənduᵊ/ʔ* 'mother's sister's son (woman speaking)', *kəlaway ənduᵊ/ʔ* 'mother's sister's daughter (man speaking)', *ibu/ŋ* 'mother's younger sister's child', *uaʔ* 'mother's older sister's child' (other terms for cousin relations are not given in Helfrich), IBN *pətuŋgal*, JKT (in order of frequency, cf. Grijns 1980:208-209) *misan/an, kəponakan, misan, ponakan* (these terms are also used for sibling's children, see below).

(c) sibling's children:

SM *anak sawdara, anak pənakan, anak s/anak, kəponakan* (< JV), *kəmanakan* (< MIN); MIN *dan/s/anak, s/anak* (specifically for 'sister's child (man speaking)'; *kamanakan*; SWY *anaʔ bəlay* 'sister's child (man speaking)', *ənduᵊ/ʔ bə/dəŋan s/anaʔ* 'brother's child (man speaking)' (no other terms were given by Helfrich); IBN *indu ʔ* 'niece', *akan* 'nephew'; JKT (in order of frequency, Grijns 1980:209): *kəponakan, ponakan, misanan*.

5.4.5 SIBLINGS

Reflexes of **kakaʔ* 'older sibling' agree in all isolects but MIN (3.4.2.4). In MIN *kako, kaka/ʔ* means 'older sister', and *ambo* is used for 'older brother'.[168] In some Malayic isolects (SM as spoken in Java (Wilkinson 1959), SAR (according to Blust pers.comm.)), a reflex **kakaʔ* (> SAR *kakaʔ*) refers to 'older sister', whereas *abaŋ* refers to 'older brother'. JKT also has *abaŋ* meaning 'older brother'. In MIN, BH, and SWY, no cognate of *abaŋ* is found. IBN *abaŋ* is a term of address to Malay men, and so is BRU *awaŋ* (Prentice pers.comm.). **adiʔ* 'younger sibling' is reconstructed on the basis of SM *adi/k*, MIN *adi, adiᵊ/ʔ*, BH, SWY *adi/ŋ*, IBN *adiʔ*, JKT *adé, adi/k*.

Cross-siblings are only found in SWY: *moanay* 'brother (woman speaking)', and *kəlaway* 'sister (man speaking)'. At least one of these terms has a PAN ancestor (*moanay* < PAN **maRuanay* 'id.', Blust 1980c:238-239). Cross-sibling terms are a general Austronesian phenomenon, and their occurrence in SWY could be a retention. In that case **mA(r)(w)anay* 'brother (woman speaking)' should be reconstructed. Although the concept must have existed, no reconstruction is made for SWY *kəlaway*. This form apparently has no cognates within the Malayic group, and outside the Malayic group it has a correspondence only in Rejang (LeBar 1972), which is spoken in an area directly bordering on the SWY area.

---

[167] Abdul Chaer has *tanté*, but the diacritic here is a misprint.
[168] According to De Josselin de Jong (1951:44-45); according to Van der Toorn *ambo* occurs only regionally. Moussay (1981:159) gives *uda/udo*.

5.4.6 CHILDREN, GRANDCHILDREN, GREAT-GRANDCHILDREN

For 'child, offspring' *anak was reconstructed (3.1.1.3), and for 'grandchild' *cucu?
(3.4.1.3IC). SM, BH cicit, MIN cici?,[169] and IBN icit 'great-grandchild' yields *cicit 'id.'.
JKT also has buyut 'id.', but in SM and BH buyut means 'ancestor (in the fourth and fifth
generation)' (no cognates in the other isolects).

5.4.7 IN-LAWS

The term for 'parent-in-law' is as follows in the isolects: SM məntua, MIN mantuo,
mintuo, BH mintuha, SWY ən/tuo, IBN (m)əntua, JKT mərtuè; JKT mərtuè must be from JV
(JV mərtua 'parent-in-law'), and PM *mintuha 'parent-in-law' is reconstructed.

The PM term for the tie between parents-in-law of a married couple was *ba/isa(a)n
(3.5.1UIC); BH has waraŋ (< JV waraŋ 'id.'). On the basis of reflexes in all isolects
*b/in/antu 'child-in-law' was reconstructed (3.1.3.1).

The general term for 'sibling-in-law' in all isolects is ipar (MIN ipa, SWY (Aliana et al.,
p.86) ipax, (Helfrich) pax buntiŋ '(term of address to a bride, used by older sister of
bridegroom)', hence *ipar 'sibling-in-law' (< PMP *hi(N)paR 'the other side; (those of the
other side =) sibling-in-law', Adelaar 1988:72).

N.B. SM biras refers to 'wife's sister's husband; husband's brother's wife'. It has an IBN
cognate which differs semantically: biras 'one who, unable to discharge a debt, was taken as
a servant by his creditor; the descendants of a biras in relation to the descendants of his
master'. The other isolects do not have cognates (in Tagalog bílas[170] is found, with the same
meaning as SM biras). But IBN has a term which is equivalent to SM biras, viz. ipar duay
'spouse of sibling-in-law'; cf. also IBN ipar səmbar 'sibling-in-law where brothers have
married sisters (?)' [sic].

5.5 PRONOUNS

5.5.1 PERSONAL PRONOUNS

	SM	MIN	BH	SWY	IBN	JKT
1ST PERSON						
singular	saya	den	ulun	aku	aku	guè
	sahaya	ambo	unda	sayo		kitè
	(d)aku	sayo	aku	awa?		anè
	beta	(awa?-)				(s)ayè
(clitic)	ku-/-ku		ku-/-ku	ku-/-ku		
dual					kami dua(y)	
plural excl.	kami	kami	kami	kami	kami	
incl.	kita	kito	kita	kito	kitay	

---

[169]According to Thaib; other sources define cici? further back (cf. Van der Toorn 'great-grandchild'; cf. also De
Josselin de Jong 1951:45).
[170]However, this is probably a Malayic loan, since other Philippine and Bornean cognates of SM biras
generally do not reflect SM b-, cf. among others Manobo izas 'spouses of two people who are related to one
another (co-generationally)', Timugon Murut ilas 'husband's brother's wife' (Prentice pers.comm.).

2ND PERSON						
singular	(əŋ)kaw (dikaw) tuan, tuanhamba mega, miga (kamu) (awak)	aŋ, kau datuə? tuan (awa?-)	(i)kam kau (sam)pian andika	kaba(n) kuti	nuan di?	(ə)lu énté disitu
(clitic)	kaw-/-mu		-mu, -kaw			
dual					kita?dua	
plural	kamu əŋkaw + səkalian	(same as singular +kalian)		kaba(n)	kita?	
(clitic)	-mu			-mu		

3RD PERSON						
singular	ia (dia) bəliaw	iño, yo baliaw	iña sidin, hidin	dio	ia	diè
(clitic)	-ña	-ño, -o, -no, -e	-ña	-ño,-o		-ñè
dual					si-duay	
plural	(d)ia məreka (itu) marika (itu)	(awa?-)	iña sidin, hidin bubuhanña		sida?	
(clitic)	-ña		-ña			

GENERAL REMARKS

(1) In SM, MIN, and SWY many personal pronouns are not marked for number, and in JKT number distinction was lost in the entire pronoun system. In JKT the plural of a subject is indicated by putting *padè* before the verb (Muhadjir 1981:41). IBN is the only isolect that makes a systematic distinction between singular and plural; it also has a dual series.

In many AN languages there is an ongoing process of replacement by plural personal pronouns of singular ones that were felt to be too familiar, too impolite and/or too direct. It is quite possible that already in PM distinctions like polite versus familiar played a crucial role. However, a singular/plural distinction is reconstructed for PM because it also apparently existed in PMP, and because the PMP plural personal pronouns are reflected in the isolects with retention of their plural meaning (viz. all isolects except JKT reflect PMP *(i)kita*, and PMP *(i)kami*, Classical Malay and SWY reflect PMP *(i)kamu*, and IBN reflects PMP *si-iDa*, all with retention of the plural notion).

(2) In Classical Malay there occurs a series of personal pronouns viz. *daku*, *dikaw*, and *dia*, which is a variant of the series *aku*, *əŋkaw* and *ia* (*dia* is still a very frequent variant of *ia*).

The members of this series are used as direct objects, as emphatic forms, and after the prepositions *akan* and *dəŋan* (in modern SM *ia* is restricted to subject position and to the written language, while *dia* is universal). The vowel difference in the pair *əŋkaw* and *dikaw* raises the question whether **əŋkau(ʔ)*, **kau(ʔ)*, or **ikau(ʔ)* must be reconstructed. Both **əŋkau(ʔ)* and **ikau(ʔ)* would result in SM *əŋkaw*, the latter after (1) antepenultimate neutralisation and (2) nasal insertion between initial *ə* and following stop (4.2 + fn.128). In each case the last vowel would have become a diphthong to yield a disyllabic form. But an original **əŋkau(ʔ)* cannot account for *dikaw*, nor for MIN *kau* 'you (to women)' or BH *kau* 'you' (Asfandi Adul) 1976:145-146); *dikaw* must have developed from **ikau(ʔ)* (with diphthongisation of **-au-*, and prefixation of **(d)i-*), or from **kau(ʔ)* (with prefixation of **di-*, and subsequent diphthongisation of **-au-*). But the reconstruction of **ikau(ʔ)* assumes (unexplained) loss of the initial vowel in MIN and BH (where *kau* occurs).

Most likely, **kau(ʔ)* was the ancestor of SM *əŋ/kaw*, *di/kaw*, MIN, BH *kau*, and SM *əŋ/kaw* acquired an epenthetic *əN-* after its disyllabicity was lost in fast speech.

5.5.1.1 FIRST PERSON

*aku '(1st pers. sg.)' was reconstructed on the basis of evidence from four isolects (3.4.2.6). Clitic forms of *aku* occur in three isolects (SM, BH, SWY), and **ku-/*-ku* are reconstructed. **ku-*[171] was an agentive pronoun precliticised to a verbal base, and **-ku* a postclitic used as an object (with a verb) or a possessive pronoun (with a noun).

SM *saya, sahaya*, MIN, SWY *sayo*, JKT *sayè, ayè*, MIN *ambo*, and BH *ulun*, originally mean 'slave, servant' (*saya* etc. < SKT). JKT *guè* is borrowed from Hokkien Chinese (Wilkinson 1959; Leo 1975:5), and· *anè* from Arabic. The original meaning of SM *awak*, MIN, SWY *awaʔ* is 'body' (cf. **awak*, 3.3.2); it is sometimes used as, or in combination with, a personal pronoun (in SM as a first or second person, in MIN as a first, second, or third person, in SWY as a first person).

SM *beta* is found in classical literature, and it is still used in Malayic isolects of eastern· Indonesia. Its original meaning is 'slave, servant'. According to Wilkinson it is borrowed from Hindi/Urdu, but its form suggests a connection with **kitaʔ*.[172]

The origin of BH *unda* is obscure.

PM **kami* '(1st pers.pl. excl.)' was reconstructed on the basis of *kami* in all isolects except JKT (3.4.2.6); **kitaʔ* '(1st pers.pl. incl.)' was reconstructed on the basis of evidence from all isolects (3.2.3 and 3.4.2.4).

5.5.1.2 SECOND PERSON

SM *əŋkaw*, MIN, BH *kau*, yield **kau(ʔ)* (cf. 5.5.1). Cliticised variants are found only in two isolects: SM *kaw-*, an actor prefix before patient-oriented verbs, and SM *-kaw*, BH *-kau* (or *-kaw?*), an object pronoun or a possessive pronoun.

---

[171]In contradistinction to **-ku*, no reflex of **ku-* is attested in the OM inscriptions, which may be an indication that **ku-* is a later (post-PM) development.
[172]Klinkert has a variant *bcita* along with *bcta*. Rémy Haaksma (1933:22) derived these forms from *⁺(a)ba* (> SM *awak* 'person; body, (one)self', and > Toba *iba* 'person; one(self); I') + *⁺ita*. I find a connection with *+ita*

Kamu '(2nd pers.pl)' is not attested in MIN, BH, IBN, or JKT, but it is found in BAC and in many languages outside the Malayic group (cf. PMP **(i)-kamu* '(id.)', Blust 1977a:11). Therefore PM **kamu(?)* is reconstructed. A postclitic variant (but unspecified for number) *-mu* occurs in SM, BH, and SWY (as an object or possessive pronoun), and **-mu(?)* '(2nd pers. pronominal postclitic (unspecified for number?) indicating object or possessor)' is reconstructed. (The use of SM *kamu* as a singular pronoun is a recent development).

The origins of SM *mega* (classical texts; related to *məreka, marika?*), MIN *aŋ* 'you (to men)', and IBN *nuan, di?* are unknown to me (cf. also Perak Malay *mika* '(2nd pers.sg.)'). MIN *aŋ* has a correspondence *haŋ* in Kedah and Perlis (cf. also Haŋ Tuah, the name of a Malay culture hero); IBN *nuan* is possibly related to SM *tuan*.

The origins of *ikam, kaba(n),* and *kuti* are obscure. BH *ikam* may be a cognate of *kamu,* but it could also be borrowed from one of the Dayak languages (Kayan also has *ikam* '(2nd pers. pl.)', Blust 1977b:100). *Kaba(n)* may somehow be derived from **kaban* 'companion, follower; herd, group' (3.5.1); *kuti* is also found in Lampung (the Wai Lima isolect of Lampung has *kuti* '(2nd pers.pl.)', Walker 1976:43).

Other terms were borrowed or originally had another meaning which became secondary, e.g. SM, MIN *tuan* 'lord, master; you', *tuan/hamba* 'your servant's lord, my lord, you', *əŋkaw səkalian,* (short) *kalian* 'you all, you together' (*səkalian* 'all, together'). BH *sampian (pian), andika* < JV (JV *sampé(y)an, əndika*). JKT *(ə)lu* < Hokkien Chinese (Wilkinson 1959; Leo 1975:6); *énté* < AR; *disitu* is a polite form originally meaning 'there'.

5.5.1.3 THIRD PERSON

For the third person SM and JKT have variant forms (SM *dia/ia,* JKT *diè/iè*) of which one member agrees with IBN *ia,* and the other with SWY *dio.* SWY *dio,* JKT *diè* may originally have had a distribution parallel to that of SM *dia;* it is also possible that they were borrowed from SM. The MIN and BH forms are problematic: they may have have developed from a morphologically complex pre-PM form **i-ni-ia,* or from a (post-?) PM combination **ia + *ña* with contraction of **ia* to *i* (or to *a* in MIN *año/ano;* pronouns are often subject to irregular shortening and contraction), cf. SM *diaña* '(3rd pers. sg./pl.)' (Gerth van Wijk p.221; Klinkert). **ia* was reconstructed for the third person singular (3.4.2.6), which is also in agreement with PMP **si-ia* (Blust 1977a:11).

A cliticised form of the third person is SM, BH *-ña,* MIN *-ño/-no/-o/-e;* in IBN the full form is used. SM, MIN, BH, SWY, and JKT reflect **-ña;* the PMP ancestral form of **-ña* is **ni-a* (Blust 1977a:10-11) in which a merger of the alveolar nasal with following **i* into a palatal nasal had not yet taken place. Some isolects have a third person plural pronoun. One of these, IBN *sida?,* reflects PMP **si-iDa* (Blust 1977a:11), and PM **sida?* is reconstructed.

The other pronouns are later developments: SM *məreka (itu)/ marika (itu)* does not occur in early Classical Malay texts, nor does it have correspondences in other isolects that are not felt to be originally SM. *Marika* is described as 'a people' in Howison (1801) and as 'people, persons; an armour-bearer, esquire' in Marsden (1812). It may be borrowed from Old Javanese, which has *marika,* an emphatic article which apparently also functioned as a

---

very likely, but the connection of *b-* with *⁺(a)ba* is too conjectural. *Beta* and *beita* may be analysed as a derivation from an (unidentified) element **bA + *ita.*

third person pronoun.[173] BH *bubuhañña* (with epenthetic *h*?) 'they; her/his group, family' is derived from *bubu* 'group, family' (Durdje & Djantera 1978:43; Abdul Jebar 1977).

N.B. OM has three other postcliticised personal pronouns. According to De Casparis -*ma:mu* is a second person possessive pronoun (De Casparis 1956:35), -*ta* is "probably a second person polite form" (p.21), and -*da* (/-*nda* after vowels) is "a honorific equivalent of -*ña*" (p.3). A semantic shift from first person plural (incl.) to second person was already shown in BRU *kita?* '(2nd pers. polite)', IBN and SAR *kita?* '(2nd pers. pl.)' < **kita?* (3.2.3). -*(n)da* is also found in SM, where it became a polite style marker suffixed to kinship terms in (archaic) formal language (especially in letter writing). Compare *anak-da* or *ana-nda, ibu-nda, bapa-nda, cucu-nda,* etc. As cliticised pronominal forms were already attested for PMP (Blust 1977a:11), it is likely that PM had them as well, and that -*ma:mu*, -*ta*, and -*(n)da* are reflexes of these clitics. But it is not clear whether -*ma:mu* already occurred in PM as a suffix in its present shape or whether it was a complex form which developed into a single suffix after PM. Nor is it clear if the semantic shifts shown in -*ta* and -*(n)da* took place before or after PM. I reconstruct PM **-(ma)mu(?)*, **-ta?* and **-(n)da?*, to which I tentatively attribute the meanings '(2nd pers. pl.)', '(1st pers. pl. incl.)', and '(3rd pers. pl.)' because of their formal agreement with **kamu(?)*, **kita?* and **sida?* respectively.

TABLE 11: THE RECONSTRUCTED PERSONAL PRONOUNS

	singular	(clitic)	plural	(clitic)
1st person	**aku*	**-ku, *ku-*	**kita?* (incl.) **kami* (excl.)	**-ta?*
2nd person	**kau(?)*	**-mu(?)*	**kamu(?)*	**-mu/* **-(ma)mu(?)*
3rd person	**ia*	**-na*	**sida?*	**-(n)da?*

5.5.2 OTHER PRONOUNS

	SM	MIN	BH	SWY	IBN	JKT
demonstrative						
(this)	*ini*	*iko/ko*	*ini/nih*	*ini*	*tu?*	*ini/ni*
(that)	*itu*	*itu/tu*	*ituh/tuh*	*itu*	*ña?*	*itu/tu/tó/(ònò)*
(yonder)	*nu(n)*				*ñin*	*(ònò)*
locative						
(here)	*-sini*	*siko*	*sia/sini*	*sini*	*ditu?*	*sini/mari*
(there)	*-situ*	*situ*	*situ*	*situ*	*dia?*	*situ*
(yonder)	*-sana*	*sinan*	*sana*	*sano*	*diin*	*sònò*
relative	*yaŋ*	*na(N)*	*naŋ*	*ño*	*ti/kə*	*yaŋ/ñaŋ/ñəŋ*

---

[173]Zoctmulder gives for Old Javanese *marika* 'emphatic article, including *ika*', and for *ika* 'demonstrative pronoun: that, those (of what is at some distance from the speaker); often: he, she, it (equivalent to a personal pronoun of the third person)'.

interrogative						
(what)	apa	apo, a	apa	(tu)apo	nama	apè
(which)	mana	mano/ma	mana	mano	ni	manè
(where)	-mana	-mano/-ma	mana	-mano	-ni	-manè
(who)	siapa	siapo/sia	siapa	siapo	sapa	siapè
(when)	bila/kapan	bilo	bila	kəbilo	kəmaya	kapan
(how much, how many)	bər/apa	bar/a	sa'apa	bəxapo	bər/apa məsak	bər/apè
indefinite	anu	anu	anu	anu	anu'	anu/anó

5.5.2.1 DEMONSTRATIVE PRONOUNS

Most isolects have a bipartite series of demonstratives, but it is likely that PM had a tripartite series: this would be in analogy with the locative pronouns (5.5.2.2) and with the IBN demonstratives. Whether JKT had three demonstratives is uncertain: the sources are not clear as to whether *ònò* is a variant of *itu/tu/tó*, or is in series with it (and with *ini/ni*). OM has *inan* 'that' (Çoedès 1930:66). This agrees with the MIN locative *s/inan*, and points to **inan* 'yonder' along with **(i)ni(?)* 'this' and **(i)tu(?)* 'that'. Another possibility is that the last part of the reflexes for 'yonder' in SM, BH, and SWY (*di-s/ana, s/ana*, and *s/ano* respectively) reflect PM **ana(?)* 'that (yonder)'. (JKT *ònò* must be borrowed because of the irregular vowel correspondences).

More dialect material is required to determine the shape of this third demonstrative pronoun, and for the time being I reconstruct **(i)ni(?)* 'this', **(i)tu(?)* 'that' and **(i)na(n)* or **(a)na(?)* 'that (over there)'. The final **(n)* in **(i)na(n)* was possibly a suffix, which would unify this form with **(a)na(?)*. The origin of MIN *iko (ko)* is unclear (in JV *ika* occurs, but this means 'that (yonder)' and is probably not a source). Nor do I have an explanation for the IBN demonstratives (note their agreement with Tagalog, which has *itó* referring to things close to the speaker).

5.5.2.2 LOCATIVE PRONOUNS

In all isolects, locative pronouns are derived from demonstratives. In IBN this happened through prefixation of *di-*, and in the other isolects through prefixation of *s-*; in SM locative pronouns are usually precliticised by a locative preposition (viz. *di* 'in, at', *kə* 'to(wards)', and *dari* 'from'). This leads to the following reconstructions:

*(?)-(i)ni(?) 'here';
*(?)-(i)tu(?) 'there';
*(?)-(i)na(n) / *(?)-(a)na(?) 'yonder'.

5.5.2.3 RELATIVE CLAUSE MARKER

The relative clause markers disagree in most isolects: SM and JKT have *yaŋ*, which must have developed from ^{+i}a + $^{+}$-*ŋ* (2.1.3, 3.3, 6.1.2); JKT *ñaŋ* and *ñeŋ* are probably variants of *yaŋ*.

MIN *naN* has a final nasal that varies according to the initial phoneme of the following word.[174] *naN* may be related to *s/inan*, as its shape suggests; but it could also be related to BH *naŋ*. The origin of BH *naŋ* is also uncertain.

SWY *ño* is probably a secondary development of the third person possessive pronoun. Finally, the origin of IBN *ti* and *kə* are obscure; *ti* must be a contracted form, since in MUA *tay* (written 'tai') is found. No attempt at a reconstruction is made (the lack of comparative agreement makes it quite likely that PM did not have a relative clause marker).

5.5.2.4 INTERROGATIVE PRONOUNS

Along with full interrogative pronouns, MIN has a series of short forms (*a, ma, bar/a* etc., written 'á', 'má', 'bará' etc. in Van der Toorn). **apa* 'what (interrogative)' was reconstructed on the basis of reflexes in all isolects (3.1.1.4; IBN has *apa* (< SM or SAR?) besides *nama*). **mana(?)* 'which' is reconstructed on the basis of regular correspondences in all isolects but IBN. It should, however, be pointed out that MIN *á* could also be a contraction of an earlier +*aha*. This would correspond with SUN *aha* and KD *ahe* 'what?' and would eventually lead to a PM form **aha*.[175]

In all isolects, the interrogative pronoun referring to place consists of a locative preposition prefixed optionally in SM to the interrogative pronoun meaning 'which', hence **-mana(?)* 'where'.

*si-apa 'who' has been reconstructed on the basis of evidence from all isolects (3.1.3.2). Alongside **si-apa *sai* 'id.' is also reconstructed on the basis of KD *saé*; it is not found elsewhere in the Malayic group, but cf. PMP **sa(y)i* 'id.'.

SM *bila* etc. < SKT *vela* 'time, moment'; (Indonesian) SM, JKT *kapan* < JV *kapan* 'when'; I do not know the origin of IBN *kəmaya*. No reconstruction is made for this meaning.

SM, *bər/apa*, MIN *bar/á*, SWY *bəxapo*, JKT *bər/apè* 'how much, how many' derive from PM **bara?* '(a marker of uncertainty or indefiniteness of object or number)' + **apa*. **bara?* and **apa* did not yet form a compound in PM (Adelaar in press b). IBN *bər/apa* is borrowed from SM and occurs along with an apparently more original *məsak*. **bara?* seems to be reflected in IBN *bara?* 'not serious, frivolous'.

5.5.2.5 INDEFINITE PRONOUN

*anu? 'something; someone, so-and-so' has been reconstructed (3.4.2, 3.4.2.4, 3.6.1.2).

---

[174]Generally speaking *naN* becomes *naŋ* before initial velars, *nam* before initial labials (including *m-?*), *nan* before initial *t, d, c, j,* and *s,* and *ná* before liquids and vowels (Van der Toorn 1899:39-40), but the examples (Van der Toorn, p.40) are not quite in agreement with this.

[175]The final vowel in KD *ahe* replaced an earlier -+*a*. The substitution of -*c(?)* for +-*a*(C) is seen in a number of other KD (and sometimes SD) forms, cf. KD *garc?* (SD *garàpm*) 'salt' < +*garapm* < SM *garam; nape?* (SD *napàtn*) 'not yet' < +*napatn* < **hadəp-an*; SD *sape* 'who?' < **si-apa*. Seeing that KD and SD changed original diphthongs to monophthongs, this seems to be basically the same phenomenon as the substitution of diphthongs for final +-*a*(C) syllables in IBN (3.2.3).

Compare also SM *bara/ŋ* 1. 'thing, object' 2. 'marker of indefiniteness or uncertainty' 3. 'marker of possibility or hope' (as in *bara/ŋ-bara/ŋ* 'may it happen that', *bara/ŋ/kali* 'perhaps, maybe'). *Bara/ŋ* consists of reflexes of PM **bara?* (5.5.2.4) and PM **-ŋ*, which was still a ligature after quantifiers and after pronouns introducing a relative clause. *Bara/ŋ* is, however, a post-PM derivation (Adelaar in press b).

5.5.2.6 SYNOPSIS OF THE RECONSTRUCTED PRONOUNS (OTHER THAN PERSONAL PRONOUNS)

demonstrative pronouns: **(i)ni(?)* 'this',
     *(i)tu(?) 'that',
     (i)na(n)/(a)na(?) 'that (yonder)';

locative pronouns: **(?)-(i)ni(?)* 'here',
     *(?)-(i)tu(?) 'there',
     *(?)-(i)na(n), *(?)-(a)na(?) 'yonder';

interrogative pronouns: **apa* 'what' (? **aha* 'id.'), **mana(?)* 'which',
     *-mana(?) 'where',
     *si-apa, *sai 'who';

indefinite pronoun: **anu?* 'something; someone, so-and-so'.

5.6 PARTS OF THE BODY

5.6.1 INTERNAL BODY PARTS

The following terms for internal body parts have been reconstructed:

*hati 'liver' (3.2.3);
*pərut 'belly, stomach; intestines' (3.4.1.1);
*u(n)tək 'brain' (3.6.2UIC);
*tulaŋ 'bone' (3.7.1);
*darah 'blood' (3.7.3);
*dagiŋ 'meat, flesh' (3.5.4), and **isi?* 'meat, contents' (3.4.2.4) (see also 5.7 lemma 103);
*hAmpədu 'gall bladder' (3.1.3.3).

Other PM terms for internal body parts that can be reconstructed are:

*jantuŋ 'heart'; MIN *jantuᵊŋ*, o.i. *jantuŋ*;
*kura? 'spleen'; SM, BH *kura*, MIN *kuro*, SWY *kuxo*, IBN *kura?*, JKT *kurè*;
*tian 'uterus'; SM, MIN, BH *tian*;
*urat 'vein, sinew'; MIN *ure?*, SWY *uxat, urat*, o.i. *urat*.

5.6.2 EXTERNAL BODY PARTS: GENERAL

Two reconstructions belong to this category:

*kulit 'skin, bark' (3.4.1.4);
*bulu 'body hair, fur, feather', which is based on *bulu* 'id.' in all isolects.

5.6.3 EXTERNAL BODY PARTS: THE HEAD

Parts of the head for which PM terms have already been reconstructed are:

*bibir 'lip; rim, edge' (3.1.2.3)
*buø(uə)k 'hair of head' (3.10);
*dahi 'forehead' (3.4.2.6);
*gahəm 'molar tooth' (3.9.2);
*gigi 'tooth' (3.5.4);
*hiduŋ 'nose' (3.5.2);

*hulu(?) 'head; upper part of river; hilt' (3.4.2.5);
*lihər 'neck' (3.1.2, 3.1.2.1);
*mata 'eye' (3.4.1.2);
*mulut 'lips, mouth' (3.1.2.4);
*tAliŋa(?) 'ear' (3.1.3.3).

Other reconstructions:

*rambut 'hair of the head'; SM, BH, JKT rambut, MIN rambuy?; *rambut and *buø(uə)k
both have the same meaning: *buø(uə)k is a regular reflex of PAN *buSek, but a reflex of
*rambut is found more often within the Malayic group;
*dilah 'tongue'; SM, MIN lidah, SWY lida(h), IBN dilah, JKT lidè (lidah, lida(h) and lidè
have undergone metathesis, cf. PMP *dilaq);
*pipi(?) 'cheek'; SM, MIN, BH pipi, JKT pipi? (no SWY word for 'cheek' was found; IBN
has kuyu?);
*dagu? 'chin'; SM, BH, SWY dagu, MIN daguə? (-? unexplained), IBN dagu?.

For other concepts no strong cognate sets are available, e.g.

SM rahaŋ, IBN raaŋ 'jaw' (no correspondences in other isolects);
BH muha, IBN mua 'face'[176] (SM has muka, MIN muko, JKT mukè < SKT; SWY has dai
'face' < *dahi 'forehead' (3.4.2.6));
SM, BH bahu, MIN, IBN bau 'shoulder' < SKT;
SM ubun/ubun, MIN bubun/bubun, BH bumbun/an, IBN bubun/aji, ubun/aji, JKT
bumbun/an, bunbun/an (3.6.2) 'fontanelle'; the ancestral form is uncertain, and a doublet is
reconstructed: *bu(m)bun/*ubun 'fontanelle'.

5.6.4 EXTERNAL BODY PARTS: THE TRUNK

Terms for parts of the trunk that have already been reconstructed are:

*awak 'body' (3.3.2);
*dada 'breast, chest' (3.5.2);
*pəlir 'testicle' (3.1.2.3);
*puki? 'vulva' (3.4.2.4);

*pusət 'navel, centre' (3.4.2.2UIC);
*puø(uə)t 'vagina' (3.10);
*rusuk 'side, flank' (3.7.3);
*susu(?) 'breast' (3.4.2.5).

Other terms that can be reconstructed are:

*bAlakaŋ 'back'; SM, SWY, IBN bəlakaŋ, BH balakaŋ (see also 5.7 lemma 13);
*bulu 'hair (on the skin)' (5.6.2);
*butuh 'penis'; SM, BH, IBN butuh, SWY butuə(h) 'penis', MIN butuəh 'a good-for-
nothing, a jerk';
*puŋguŋ 'lower part of the back'; SM puŋguŋ 'buttocks', (Indonesian SM) 'back', MIN
puŋguəŋ 'back', SWY puŋguŋ 'buttocks', IBN puŋguŋ 'loins, waist'.

---

[176]Muha, mua may also derive from SKT mukha, through BH and IBN having maintained h (which became
IBN ø) instead of k (as found in the other isolects). Another possibility is that muha and mua (and, for that
matter, Balinese mua 'id.') are inherited, and are cognates of Proto Oceanic *muqa 'front, to precede' (Milke
1968:158), which would yield PMP *muqa 'front, to precede'.

Another reconstruction for 'body' is:

*tubuh; SM, BH, IBN *tubuh*, MIN *tubuᵓh*, SWY *tubuᵓ(h)*, JKT *tubu* (*tubuh and *t/um/buh 'grow' are related, the former reflecting PAN *Cubuq '(shoot? growing body?)' and the latter reflecting the PAN verbal derivation *C-um-ubuq 'grow', cf. 4.6).

A term for 'armpit' was only found in SM *(kətiak)* and BH *(katiak)*.

5.6.5 EXTERNAL BODY PARTS: THE LIMBS

Terms for parts of the limbs that have already been reconstructed are:

*buku? 'joint, node' (3.4.2.4); *siku 'elbow' (3.4.2.6);
*kuku 'claw, nail' (3.4.2.6); *tuøət 'knee' (3.10);
*paha(?) 'thigh' (3.4.2.5UIC); *tuñjuk 'index finger' (3.5.3).
*ruas 'internode' (3.8.2);

Other reconstructions are:

*bətis 'part of leg between knee and ankle'; MIN *batih*, SM, SWY, IBN *bətis* 'id.' (Indonesian SM 'calf'), BH *batis* 'leg', JKT *bətis* 'calf of leg';
*buah *bətis 'calf of leg'; SM, IBN (Richards 1981) *buah bətis*, MIN *buah batih*;
*tulaŋ *bətis 'shin'; SM, IBN (Richards 1981) *tulaŋ bətis*, MIN *tulaŋ batih*;
*tumit 'heel'; SM, BH, SWY, IBN *tumit*, MIN *tumi?*;
*jari 'finger'; SM, MIN, JKT *jari*, SWY *jaxi, jari* 'id.', IBN *jari* 'hand' (see also 5.7 lemma 1);
*kaki 'leg, foot'; SM, MIN, IBN, JKT *kaki* (see also 5.7 lemma 4);
*kAliŋkiŋ 'little finger'; SM, JKT *kəliŋkiŋ*, MIN *kaliŋkiᵓŋ* 'id.', IBN *kələŋkiŋ* (penultimate ə unexplained);
*ləŋan 'arm (from shoulder to wrist); sleeve'; SM, SWY, IBN *ləŋan*, MIN *laŋan*, cf. PMP *leŋen;
*taŋan 'hand'; IBN *taŋan* (probably a loan, *jari* is usual, see above), o.i. *taŋan*.

5.6.6 BODY PARTS OF ANIMALS

In this category the following terms have already been reconstructed:

*ikur 'tail' (3.7.4 IC, 5.7 lemma 105);
*juŋur 'snout' (3.6.1.5);
*sayap 'wing' (3.1.1.3, 5.7 lemma 100);
*səŋət 'sting of an insect' (3.1.1.2);
*sisik 'fish scale' (3.8.1N.B.).

5.7 A 200-ITEM BASIC WORDLIST FOR PROTO MALAYIC

This list is Hudson's variant of the Swadesh 200-item basic wordlist as modified by Blust in his forthcoming study on Austronesian lexicostatistics (Hudson 1967; Blust forthcoming). Its purpose is to provide a tool for measuring the relative affinity of other isolects, languages, or language groups with Proto Malayic. Through comparison of closely related isolects it is

possible to eliminate a number of lexical innovations which are not common to all these isolects, and in this way to make a basic wordlist for the proto-language with more retentions than corresponding wordlists for each of the isolects. Of course such a list requires much critical insight and a great knowledge of as many dialectal variations as possible. Moreover, the result will always be hypothetical in more than one respect: (1) it consists of proto-lexemes, and (2) the relative frequency of each proto-lexeme is also hypothetical. Therefore the list presented here should be seen as a first attempt. Hopefully other scholars will amplify and improve it, or develop an alternative and more reliable means for testing the relative position of PM and the Malayic group within the Austronesian language family.

The list is built on comparative evidence from basic wordlists for each of the isolects forming the basis of this study. These wordlists were collected as follows:

— For SM I made one myself, and took my own knowledge as a guide;

— For MIN, BH and JKT, lists filled out by native speakers were used; these lists were collected by Blust, who will use them in his lexicostatistical study (Blust forthcoming);

— For IBN I made a list on the basis of Richards (1981) and, to a very limited extent, on the basis of the English-Iban phrase book (Borneo Literature Bureau 1967). The use of Richards yielded reliable results, because each entry includes synonyms (if any) with the most frequently occurring one in capitals;

— For SWY part of the list could be drawn from the vocabulary included in Aliana et al. (1979), and for the other part I used my own judgement in collecting lexical items from Helfrich. Where I was not certain or did not find a convincing item, I left a blank in the wordlist.

The PM 200-item basic wordlist is built up in the following way:

(a) Loanwords are not considered as evidence. In the list, they are shown between brackets. The source language of a loanword may be indicated either in the list or in the explanation which follows the list (see 5.7.1). The source language is given in the explanation when the identification of a particular lexeme as loanword requires more support, or when this is convenient (as in the case of 'think' (21) where ultimately the whole set derives from AR *fikr*).

(b) When three or more isolects have corresponding lexemes, this yields a reconstruction in the wordlist, unless these lexemes are loanwords or unless the evidence militates against a reconstruction in any other way. Such cases will be discussed in the explanation.

(c) If in a particular set a lexeme does not have correspondences but there is sufficient evidence from outside the six isolects to suppose that it had a PM ancestor belonging to the basic vocabulary, this ancestor is (also) given in the list, e.g. 'other' (190): SM, MIN, BH, SWY *lain,* JKT *laèn,* but IBN *bukay; *bukən* is given, because IBN is assumed to retain (with some alternations) the original PM form for this concept. In two cases a reconstruction without reflexes in the lists for each isolect is given: *hulu(?)* (5.7.1 lemma 24) and *təlu* (5.7.1 lemma 199).

No.	meaning	SM	MIN	BH	SWY	IBN	JKT	PM
1.	hand	taŋan	taŋan	taŋan	taŋan	jari	taŋan	*taŋan
2.	left side	kiri	kida	kiwa	kido	kiba?	kiri	*kA-iri/*kiba?
3.	right side	kanan	suo?	kanan	kanan	kanan	kanan	*k/anan
4.	leg/foot	kaki	kaki, tuŋkay	batis	kətiŋ	kaki	kaki	*kaki
5.	walk/go	bər-jalan	ba-jalan,pai	ba-jalan,tulak	bə-jalan	bə-jalay, mansaŋ	jalan	*((mb)Ar-)jalan
6.	road, path	jalan	jalan	jalan	jalan	jalay	jalan	*jalan
7.	come	dataŋ	dataŋ	dataŋ	dataŋ	datay	dataŋ	*datəŋ
8.	turn (v)	belok	belo?	biluk	--	biluk	pəŋkòl	*biluk
9.	swim	bə-rənaŋ	ba-ranaŋ	ba-kuñuŋ	bə-dənaŋ	səməray	bə-rənaŋ	(*(mb)A-rənaŋ)
10.	dirty	kotor	kumuᵊh	rigat, cumuh	kama(h), kumuᵊ(h)	kamah	jəmbər <JV	(*kamah/*kumuh)
11.	dust	dəbu	abu	dabu	dəbu	dəbu, abu, abuk, apuk	abu	*dəbu
12.	skin	kulit	kuli?	kulit	bawa?, kulit	kulit	kulit	*kulit
13.	back	bəlakaŋ	puŋguᵊŋ	balakaŋ	bəlakaŋ	bəlakaŋ	bòkòŋ <SUN	*bAlakaŋ
14.	belly	pərut (bəsar)	paruy?	parut	pəxut bəsa?	pərut	pərut	*pərut
15.	bone	tulaŋ	tulaŋ	tulaŋ	tulaŋ	tulaŋ	tulaŋ	*tulaŋ
16.	guts	isi pərut, pərut muda	tamunsu, isi paruy?	ucus < JV	pəxut mudo	pərut	ucus < JV	*pərut
17.	liver	hati	ati	hati	ati	ataw, ati	ati	*hati
18.	breast	susu	susu	susu	susu	tusu	tété	*susu(?)
19.	shoulder	bahu, pundak	bau	bahu	bakix	bau	pundak	--
20.	know (things)	tahu	tau, abeh	tahu	tau	təmu	tau?	*tahu(?)
21.	think	bər-pikir	piki?	ba-pikir	pikix, pikir	pikir	mikir	(all < AR)
22.	be afraid	takut	takuy?	takut-an	takut	takut	takut	*takut
23.	blood	darah	darah	darah	daxa(h)	darah	darè	*darah
24.	head	kəpala	kapalo	kapala	pala?	pala?	kəpalè	*hulu(?)
25.	neck	leher	li(h)iᵊ	gulu	liax	rəkuŋ	lèhèr	*lihər
26.	hair of head	rambut	abuᵊ?	rambut	gumba?, rumba?	buuk	rambut	*buø(uə)k, *rambut
27.	nose	hiduŋ	iduᵊŋ	hiduŋ	idu	iduŋ	iduŋ	*hiduŋ
28.	breathe	bər-napas< AR	aŋo?	hinak	--	ñawa	napas < AR	*ñawa
29.	sniff, smell	cium	ciun, idu	cium	cium	siun	cium	--
30.	mouth	mulut	muñcuᵊŋ	muntuŋ	mulut	mulut, ñawa	mulut	*mulut
31.	tooth	gigi	gigi, gigih	gigi	gigi	gigi, ŋəli?	gigi	*gigi
32.	tongue	lidah	lidah	ilat	lida(h)	dilah	lidè	*dilah
33.	laugh (v)	tər-tawa	gala?	ta-tawa	tə-tawo	kə/tawa?	kə-tawè	*tawa?
34.	cry (v)	taŋis	taŋih	taŋis	taŋis	sabak	taŋis	*taŋis
35.	vomit (v)	muntah	mutah	muak	muta(h)	mutah	muntè	*m/u(n)tah
36.	spit	ludah	liuᵊ, ludah	ludah	--	luaʔ, ludah	ludè	*ludah
37.	eat	makan	makan	makan	makan	əmpaʔ, makay	makan	*ma/kan
38.	chew	kuñah, mamah	kuñah, mamah	mamah	kuña(h), mama(h)	gulum, kuñah	kuñè	*kuñah, *mamah
39.	cook (v)	masak, tanak	masa?, tana?	masak, tanak	tana?	sumay, panduk	masak	*m/asak, *tanak
40.	drink (v)	minum	minun	kinum	minum	irup	minum	*inum
41.	bite (v)	gigit	gigi?	igut	gigit	gigit, kətup	gigit	*gigit

No.	meaning	SM	MIN	BH	SWY	IBN	JKT	PM
42.	suck	(h)isap	iso?	hisap	isap	insap, tusu?	isəp	*hi(ŋ)səp
43.	ear	təliŋa	taliŋo	taliŋa	təliŋo	pəndiŋ	kupiŋ < JV	*tAliŋa(?)
44.	hear	dəŋar	daŋa	daŋar	dəŋax	diŋa	dəŋər	*dəŋər
45.	eye	mata	mato	mata	mato	mata	matè	*mata
46.	see	lihat	lie?	lihat	ina?	pəda?	liat	*lihat
47.	yawn (v)	kuap	kuo?	kuap	uap	ŋuap	uap	*uap
48.	sleep (v)	tidur	tidu?, lalo?	guriŋ	tiduə?	tinduk	mòlòr < SUN	*tidur
49.	lie down	bər-bariŋ	ba-gole?	ba-rabah, ba-huñur	--	gali?	rəbè	*--
50.	dream (v)	bər-mimpi	ba-mimpi, ba-rasian	mimpi	mimpi	mimpi	impi	*m/impi/*impi
51.	sit	duduk	dudu??	duduk	dudu??	duduk	duduk	*duduk
52.	stand	bər-diri	taga?	ba-diri	bə-təga?	bə-diri	diri	*diri
53.	human being	oraŋ	uraŋ	uraŋ	oxaŋ, uxaŋ, jəmo	uraŋ	òraŋ	(*uraŋ)
54.	man, male	lakilaki	lakilaki	lalakian	laki	laki	ləlaki	*laki(-laki)
55.	woman, female	pərəmpuan	padusi	babinian	tino	indu?	pərəmpuan	--
56.	child	anak	ana?	anak	ana?	anak	anak	*anak
57.	husband	laki, suami < SKT	laki	suami < SKT	laki	laki	laki	*laki
58.	wife	bini, istri < SKT	bini	bini	bini	bini	bini	*bini
59.	mother	ibu, (ə)ma?	mande, ama?	uma, induŋ	əndu??, ma?	inday	əña?	*(ə)ma(?), *indu?, *ina
60.	father	bapak	apak	abah, bapa	ba?	apay	babè	*apa(?)
61.	house	rumah	rumah	rumah	xuma(h)	rumah	rumè	*rumah
62.	thatch, roof	atap	ato?	hatap	atap	atap	atəp	*hatəp
63.	name	nama < SKT	namo < SKT	ŋaran < JV	namo < SKT	nama < SKT	namè < SKT	--
64.	say	bər-kata < SKT	ece?	ujar, tutur	bə-kato < SKT	səbut, bə-jaku?	òmòŋ	*tutur
65.	rope	tali	tali	tali	tali	tali	tali	*tali
66.	tie up, fasten	ikat	ike?, kabe?	ikat, jarat	ikat, kəbat	ikat, kəbat	ikət	*ikət
67.	sew	jahit	jai?	jahit	jait	jait	jait	*jahit
68.	needle	jarum	jarun, piñjay	jarum	jaxum	jarum	jarum	*jarum
69.	hunt (v)	buru	buru	buru, andup	buru	giga?, asu	buru	*buru
70.	shoot (arrow)	tembak, panah	temba?, panah	timbak	timba?, pana(h)	panah	tèmbak, panè	*timbak, *panah
71.	stab	tikam	tikam	suduk, tikam	tusu??, tikam	tusuk	tikəm	*tikəm, *tusuk
72.	hit, slap	pukul	toko?, laka?	tukul	palu, pukul	palu?, pukul	pukul	*pukul, *palu?
73.	steal	curi	cilo?, mali?ŋ	cuntan	maliŋ	curi	curi	*maliŋ
74.	kill	bunuh	bunu?h	bunuh	bunu?(h)	bunuh	bunu	*bunuh
75.	die, be dead	mati	mati	mati	mati	mati	mati	*m/ati
76.	live	hidup	iduy?	hidup	idup	idup	idup	*hidup
77.	scratch (an itch)	garuk	gauy?	garu, garuk	gaxut, garut	garu?	garuk	*garut, *garuk, *garu?
78.	cut, hack	tətak, tarah	tata?, tarah	tatak, tarah	təta?, taxa(h)	tətak, tarah	pòtòŋ	*tətək, tarAs
79.	stick (wood)	kayu	kayu	kayu	kayu	kayu?	kayu?	*kayu?
80.	split (VTR)	bəlah	balah	balah	bəla(h)	bəlah	bəlè	*bəlah
81.	sharp	tajam	tajam	tajam	tajam	tajam	tajəm	*tajəm
82.	dull, blunt	tumpul	--	landap	tumpul	tumpul	tumpul	*tumpul
83.	work (in field)	bə-kərja	ba-karajo	ba-huma	--	b-umay	kərjè	*(mb)Ar-huma(?)
84.	plant (v)	tanam	tanam	tanam	--	tanam	tanəm	*tanəm
85.	choose	pilih	pili?h	pilih	pili?(h)	pilih	pili	*pilih
86.	grow (VTR)	tumbuh	tumbu?h	tumbuh	tumbu?(h)	tumbuh	tumbuh	*t/um/buh

No.	meaning	SM	MIN	BH	SWY	IBN	JKT	PM
87.	swell (abscess)	*bəŋkak*	*baŋkaʔ*	*baŋkak*	*bəŋkaʔ*	*bəŋkak*	*bəŋkak*	**bəŋkak*
88.	squeeze	*pəras*	*rameh*	*paras, parah*	*pəxa(h)*	*pəruk, pərah*	*(pəras)*	**pərəs, *pərah*
89.	hold (in fist)	*pəgaŋ*	*paciʔ*	*piŋkut, pagaŋ*	*pəgaŋ*	*pəgay*	*pəgaŋ*	**pəgaŋ*
90.	dig	*gali*	*kali*	*tabuk*	*kali*	*kali*	*gali*	**kali*
91.	buy	*bəli*	*bali*	*tukar*	*bəli*	*bəli*	*bəli*	**bəli*
92.	open, uncover	*buka*	*bukaʔ*	*buka*	*bukaʔ*	*bukaʔ*	*bukè, bukaʔ*	**bukaʔ*
93.	pound, beat	*tumbuk*	*tumbuˀʔ*	*tumbuk*	--	*tutuk*	*tumbuk*	**tumbuk, *tutuk*
94.	throw	*lempar*	*ampoʔ, bae*	*tawak*	*limpar*	*tikaw*	*lèmpar*	--
95.	fall (v)	*jatuh*	*jatuˀh*	*gugur*	--	*labuh*	*jatò*	**jatuh, *labuh*
96.	dog	*añjiŋ*	*añjiˀŋ*	*hadupan*	*añjiŋ*	*ukuy, asuʔ*	*añjiŋ*	**asuʔ*
97.	bird	*buruŋ*	*buruˀŋ, uŋgeh*	*buruŋ*	*buruŋ*	*buruŋ*	*buruŋ*	**buruŋ*
98.	egg	*təlor*	*taluˀ*	*hintalu, hantalu (< JV)*	*təlux*	*təluʔ*	*təlòr*	**təlur*
99.	feather	*bulu*	*bulu*	*bulu*	*bulu*	*bulu*	*bulu*	**bulu*
100.	wing	*sayap*	*sayoʔ*	*halar*	*sayap*	*sayap*	*sayap*	**sayap*
101.	fly (v)	*tərban*	*tabaŋ*	*tarabaŋ*	*təxban*	*tərəban*	*tərban*	**tÁr(ə)baŋ*
102.	rat	*tikus*	*mañciʔ*	*tikus*	*tikus*	*cit, tikus*	*tikus*	**tikus*
103.	meat, flesh	*dagiŋ*	*dagiˀŋ*	*dagiŋ*	*dagiŋ*	*isiʔ, dagin < SAR*	*dagiŋ*	**dagiŋ, *isiʔ*
104.	fat, grease	*ləmak*	*lamaʔ*	*lamak*	*gəmuʔʔ*	*ləmak*	*gəmuk*	**ləmək*
105.	tail	*ekor*	*ikuʔ*	*buntut, ikuŋ*	*ikuˀʔ*	*ikuʔ*	*buntut*	**ikur*
106.	snake	*ular*	*ula*	*ular*	*ulax*	*ular*	*ulər*	**ulər*
107.	worm	*(h)ulat, caciŋ*	*uleʔ, caciˀŋ*	*hulat, caciŋ*	*ulat*	*ulat, bəlut*	*caciŋ*	**hulət, *caciŋ*
108.	louse	*kutu*	*kutu*	*kutu*	*kutu*	*kutu*	*kutu*	**kutu*
109.	mosquito	*ñamuk*	*raŋiʔ, ñamuˀʔ*	*ñamuk*	*ñamuˀʔ*	*ñamuk*	*ñamuk*	**ñamuk*
110.	spider	*labalaba*	*lawah*	*sabaŋkaŋ*	*ləlawa [sic]*	*əmpəlawaʔ*	*labèlabè*	**lawaʔ, *laba(ʔ)*
111.	fish	*ikan*	*ikan, lauʔʔ*	*iwak < JV*	*ikan*	*ikan*	*ikan*	**ikan*
112.	rotten	*busuk*	*busuˀʔ*	*buruk*	*busuˀʔ*	*busuk, buruk*	*busuk*	**busuk*
113.	branch	*dahan*	*daan, capaŋ*	*dahan, cakaŋ*	*daan*	*daan*	*cabaŋ*	**dahan*
114.	leaf	*daun*	*daun*	*duanb*	*daun*	*daun*	*daòn*	**daun*
115.	root	*akar*	*aka*	*akar*	*akax*	*urat, akar*	*akar*	**akar*
116.	flower	*buŋa*	*buŋo*	*kambaŋ < JV*	*buŋo*	*buŋay*	*kəmbaŋ < JV*	**buŋa(ʔ)*
117.	fruit	*buah*	*buah*	*buah*	*bua(h)*	*buah*	*buè*	**buah*
118.	grass	*rumput*	*rumpuyʔ*	*rumput*	*xumput*	*rumput*	*rumput*	**rumput*
119.	earth, soil	*tanah*	*tanah*	*tanah*	*tana(h)*	*tanah*	*tanah*	**tanah*
120.	stone	*batu*	*batu*	*batu*	*batu*	*batu*	*batu*	**batu*
121.	sand	*pasir*	*pasiˀ*	*pasir, karaŋan*	*pasir, buŋin*	*pasir*	*pasir*	**pasir*
122.	(fresh) water	*air*	*aiˀ*	*bañu < JV*	*aiˀʔ*	*aiʔ*	*aèr*	**air*
123.	flow	*alir*	*iliˀ*	*alir*	--	--	*alir*	**alir*
124.	sea, ocean	*laut*	*lauyʔ*	*laut*	*laut*	*tasik*	*laót*	**tasik*
125.	salt	*garam*	*garam*	*uyah < JV*	*gaxam, garam*	*garam*	*garəm*	**sira*
126.	lake	*danaw, tasik*	*danaw*	*danaw*	*danaw*	*danaw, ləpuŋ*	*əmpaŋ*	**danaw*
127.	woods, forest	*hutan*	*utan, rimbo*	*hutan*	*ximbo, utan*	*utan, kampuŋ*	*utan*	**hutan, *rimbaʔ*
128.	sky	*laŋit*	*laŋiʔ*	*laŋit*	*laŋit*	*laŋit*	*laŋit*	**laŋit*
129.	moon	*bulan*	*bulan*	*bulan*	*bulan*	*bulan*	*bulan*	**bulan*
130.	star	*bintaŋ*	*bintaŋ*	*bintaŋ*	*bintaŋ*	*bintaŋ*	*bintaŋ*	**bintaŋ*
131.	cloud	*awan*	*awan*	*awan*	*aban, awan*	*rəmaŋ*	*awan*	**a(bw)an*
132.	fog	*kabut*	*kabuyʔ*	*kabut*	*kabut*	--	*kabut*	**kabut*

No.	meaning	SM	MIN	BH	SWY	IBN	JKT	PM
133.	rain	*hujan*	ujan	hujan	ujan	ujan	ujan	*hujan
134.	thunder	*guntur, guruh*	guruᵊh	guntur	guruᵊ(h)	guntur	gəlèdèk < SUN	*guntur, *guruh
135.	lightning	*kilat*	kileʔ	kilat	kilat	ɲitar,kilat	kilat	*kilat
136.	wind	*aŋin*	aŋin	aŋin	aŋin	ribut	aŋin	*aŋin
137.	blow	*bər-tiup*	bar-ambuyh	ba-tiup	(tiəp)	tiup	tiup	*t/iup
138.	warm	*panas*	paneh	panas	panas	panas	gərè, panas	*panas
139.	cold	*diŋin*	diŋin	diŋin	diŋin	cəlap	diŋin	*diŋin
140.	dry	*kəriŋ*	kariᵊŋ	kariŋ	kəxiŋ	raŋkay, raŋkaŋ	kəriŋ	*kəriŋ
141.	wet	*basah*	basah	basah, jandaw	basa(h)	basah	basè	*basah
142.	heavy	*bərat*	bareʔ	barat	bəxat	bərat	bərat	*bərat
143.	fire	*api*	api	api	api	api	api, baraʔ, barè	*api
144.	burn (VTR)	*bakar*	baka	banam	tunu	tunu	bakar	*tunu, *bakar
145.	smoke	*asap*	asoʔ	asap	asap	asap	asəp	*asəp
146.	ash	*abu*	abu	habu	abu	abu, abuk, abus	abu	*habu
147.	black	*hitam*	itam	hiraŋ < JV	itam	cəlum	itəm	*hitəm
148.	white	*putih*	putiᵊh	putih	potiᵊ(h)	burak	puti	*putih
149.	red	*merah*	sirah	habaŋ < JV	abaŋ < JV	mansaw	mèrè	*(ma-)irah
150.	yellow	*kuniŋ*	kuniᵊŋ	kuniŋ	kuniŋ	kuniŋ	kuniŋ	*kunit
151.	green	*hijaw*	ijaw	hijaw	ijaw	ijaw	ijó	*hijaw
152.	small	*kəcil*	kəteʔ	halus	kəciᵊʔ	mit	kəcil	*kəcil, *kətik
153.	big	*bəsar*	gadaŋ	ganal	bəsaʔ	bəsay, raya	gədé < JV	*bəsar, *raya
154.	short	*pendek*	pandaʔ, siŋkeʔ	pindik, handap	pandaʔ	pandak, pandaw	pèndèk	*pandak, *pindik
155.	long	*pañjaŋ*	pañjaŋ	pañjaŋ	pañjaŋ	pañjay	pañjaŋ	*pañjaŋ
156.	thin	*tipis, nipis*	mipih, nipih	nipis	tipis, nipis	mipaw, mipih, nipih, mipis, pipis	tipis	*nipis, *m/ipis
157.	thick	*təbal*	tabal	kanda	təbal	təbal	təbəl	*təbəl
158.	narrow	*səmpit*	sampiʔ	kipit, sampit	səmpit	səmpit	səmpit	*səmpit
159.	wide	*lebar*	laweh	libar	libax	luas	lèbar	*libar
160.	ill, sick	*sakit*	sakiʔ, padiᵊh	sakit, gariŋ	sakit	sakit	sakit	*sakit
161.	ashamed	*malu*	malu	supan	malu	malu	malu	*malu
162.	old	*tua*	gaeʔ	tuha	--	tuay	tuè	*tuha(ʔ)
163.	new	*baru, baharu*	baru	hañar	əmpay	baru	baru	*bAharu
164.	good	*baik*	eloʔ, baiᵊʔ	baik	iluᵊʔ	manah	baèk	*baik
165.	bad, evil	*jahat*	jaeʔ	jahat	jaat	jaiʔ	jahat	*jahət
166.	true	*bənar*	bana, batuᵊ	banar, bujur	bənax	amat	bənər	*bənər
167.	night	*malam*	malam	malam	malam	malam	maləm	*ma-lə(hø)əm
168.	day	*hari*	ari	hari	axi	ari	ari	*hari
169.	year	*tahun*	taun	tahun	taun	taun	taòn	*tahun
170.	when?	*kapan < JV, bila(mana)*	bilo	bila	--	kəmaya	kapan < JV	--
171.	hide	*bər-səmbuñi*	mandoʔ	sambuñi	buni	pəlam	səmbuñi	*buni
172.	climb	*naik*	naiᵊʔ	naik	naiᵊʔ	tik/iʔ	naèk	*naik
173.	at	*di*	di	di	di	di	di	*di
174.	inside	*(di)dalam*	(di)dalam	(di)dalam	(di)dalam	(di)alam	daləm	*(d-)aləm
175.	on (top of)	*(di)atas*	(di)ateh	(di)atas	(di)atas	(d-)atas	(di)atas	*atas
176.	below	*(di)bawah*	(di)bawah, baruᵊh	(di)bawah	(di)bawa(h)	(di)bah, baruh	(di)kòlòŋ	*bah, *babah

No.	meaning	SM	MIN	BH	SWY	IBN	JKT	PM
177.	this	ini	iko, ko	ini, ni(h)	ini	tuʔ	ini	*(i)ni(ʔ)
178.	that	itu	itu, tu	tu, itu(h)	itu	ñaʔ, ñin	ònò	*(i)tu(ʔ), *(i)na(n), *(a)na(ʔ)
179.	near	dəkat	dakeʔ	parak	dampiŋ	dampih, dampiŋ, dampir	dəkət	*dəkət, *(h)ampiŋ
180.	far	jauh	jauᵊh	jauh	jauᵊ(h)	jauh	jau	*jauh
181.	where?	di mana	di ma, di mano	di mana	(di)mano	dini	(di)manè	*-mana(ʔ)
182.	I	saya, aku	den, ambo	aku	aku	aku	guè < CHI	*aku
183.	you (sg.)	əŋkaw	aŋ, kau	ikam	kaba, kuti	diʔ, nuan	lu < CHI	*kau(ʔ)
184.	(s)he	ia, dia	iño	iña	dio	ia	diè	*ia
185.	we	kita, kami	kito, kami	kita, kami	kito, kami	kitay, kami	kitè	*kitaʔ, *kami
186.	you (pl.)	kamu	(aŋ, kau, kalian)	(bubuhan) ikam	kamu, kaba, kuti	kitaʔ	lu	*kamu(ʔ)
187.	they	məreka	iño	bubuhan-ña	dio	sidaʔ	diè	*sidaʔ
188.	what?	apa	apo, a	apa	tuapo, apo	nama	apè	*apa
189.	who?	siapa	siapo, sia	siapa	sapo	sapa	siapè	*si-apa, *sai
190.	other	lain	lain	lain	lain	bukay	laèn	*bukən
191.	all	səmua	sadoño	samuñaan, samua	səgaloyo, gəgaloyo	abis, magaŋ	səmuè	*habis
192.	and/with	dan	jo, dan	lawan < JV	ŋan	əŋgaw	amè < SKT	--
193.	if	kalaw	kalaw, jiko	lamun, jika	amun, jəkalaw	əntiʔ, tiʔ	kaló	--
194.	how?	bagay-mana	ba	kaya apa < JV	--	kati	bagi-manè	--
195.	no, not	tidak	indaʔ	kada	nido	naday, əndaʔ	əŋgaʔ	*-daʔ
196.	count (v)	hituŋ	etoŋ	hituŋ	ituŋ, bilaŋ	təsaʔ, ituŋ	ituŋ	*hituŋ
197.	one	satu, suatu	cieʔ	asa	so	saʔ	atu	*əsaʔ
198.	two	dua	duo	dua	duo	dua	duè	*dua(ʔ)
199.	three	tiga	tigo	talu, tiga	tigo	tiga	tigè	*təlu
200.	four	əmpat	ampeʔ	ampat	əmpat	əmapt	əmpat	*əmpat

5.7.1 EXPLANATION

2. SM, JKT *k/iri*, BAC *ka/iri* < PMP **ka-wiRi*; BH *kiwa*, IBN *kiba?* < PMP **kiwa* (cf. 7.3.7 N.B. (d) for IBN unexpected -*b*-); MIN *kida*, SWY *kido* (loss of -*l* unexplained) have a JV correspondence *kidal*, and a SM one (SM *kidal* 'left-handed', which is used on Java; a loan?).

4. Cognates of BH *batis* refer to the part between knee and ankle (cf. **bətis*, 5.6.5).

5. **(mb)Ar-* : see 6.1.1.

9. **(mb)A-rənaŋ* is not attested outside the Malayic group. Inside the group, it is found with an irregular sound correspondence in SWY, and it does not occur in BH or IBN. It may be a reflex of PMP **laŋuy* 'swim' which, in many languages, has a reflex with initial *n* (and has therefore been reconstructed with a doublet **naŋuy*, cf. Blust forthcoming). PMP **(ln)aŋuy* would have become **laŋi* or **naŋi* in PM, and, when **(mb)Ar-* was prefixed, it may have lost through backformation its final *i*, after the latter was reinterpreted as an out-of-place transitive-marking suffix. Thus: PMP **(ln)aŋuy* > PM **(mb)Ar-* + **(ln)aŋi* > *+bər-naŋi* > *+bər(ə)naŋ-i* > SM *bərənaŋ*. IBN *səməray* is a backformation on the basis of *səbəray* 'across, on the other side' (which becomes *ñəməray* after prefixation of N-, 3.11b).

11. BH *habu*, o.i. *abu* 'ash, dust' < **habu* (3.1.2.4).

13. In modern Indonesian, *puŋguŋ* has taken the meaning 'middle part of the back', but in modern Malaysian and in Classical Malay *bəlakaŋ* refers to 'back' and *puŋguŋ* to 'buttocks' (Howison 1801; Marsden 1812; Klinkert 1916) .

16. No separate reconstructions are made for 'belly' and 'intestines', cf. 3.4.1.1; cf. also SM *pərut muda*, MIN *paruy?* mudo, SWY *pəxut mudo* 'intestines', and SM *pərut bəsar*, SWY *pəxut bəsa?* and MIN *paruy? kapuy?* 'belly, stomach'.

18. cf. 3.4.2.5.

19. *Bahu* etc. < SKT; *pundak* etc. < JV; *bakix* < ?

21. *Pikir* etc. < AR.

24. SM *kəpala* etc. are borrowed from SKT. SWY, IBN *pala?* (-*?* unexplained) (with loss of the same syllable) must derive from the same source. All isolects have a reflex of PM **hulu(?)* 'head; head of a river, upriver; hilt' (3.4.2.5). Presumably **hulu(?)* was the original term for 'head', and it retained mainly a metaphorical meaning after *kəpala, pala?* etc. came into use (in SM *hulu* is also still used for the head of a royal person).

26. Although *rambut* etc. occurs more often within the Malayic group, **buø(uə)k* is a regular reflex of PMP **buh(ue)k* (cf. also BAC *buok*, KD *bu'uk*).

28. SM, JKT *napas* < AR. IBN *ñawa* < **ñawa* 'soul, life; breath' (3.3.2).

29. SM *cium* etc. is a North Indian borrowing (3.4.1).

33. **tawa?* is reconstructed without prefix indicating unintentionality, because the isolects do not agree in the use of such a prefix.

38. No further differentiation of meaning between **kuñah* and **mamah* can be given on the basis of the (often contrary) descriptions of the reflexes in the isolects.

39. **m/asak* 'cooked, done, ripe'; **tanək* 'cook rice'.

SM, MIN, BH *m/asak* 'cooked, ripe', and *tanak* 'cook rice'. SWY *tana?* 'cook rice' (for *masa?* Helfrich only gives 'ripe, done'). IBN has *sumay* 'cook, boil' (2nd Division of Sarawak; cf. SD *ñuman* 'id.' Ina Anak Kalom & Hudson 1970:290) and *panduk* (used elsewhere); IBN does not distinguish between rice cooking and other cooking. Nor does JKT: *masak* also applies to rice, and no reflex of **tanak* is given. The final **ə* in **tanak* is based on evidence from outside the Malayic group (cf. PMP **tanek* 'cook, prepare food').

40. The SM, MIN, SWY and JKT fossilised prefix *m/-* is not reflected in BH, which has *k/-* instead. BH *k/inum* may be borrowed from one of the neighbouring Barito languages (cf. Dusun Lawang *kinum* 'drink (v)' in Hudson 1967). Other possibilities are that PM had a variant form **k/inum* which occurred along with **m/inum*, and which was retained in BH, or that PM still had a living affix **-um- / *(u)m-* which is still reflected in SM, SWY and JKT *m/inum*, MIN *m/inun* (cf. 6.8).

44. IBN *diŋa* is possibly a reflex of PMP **diŋa* 'hear' (Blust 1984a), but it could also be borrowed from another (non-Malayic) isolect from Borneo. (Correspondences are found in the north-east Barito isolects, cf. Lawangan *diŋa?*, Tabojan *ŋɨdɨŋɨ* Hudson 1967).

47. According to Blust, SM (MIN, BH) *k-* originated through backformation: **maN-uap >* *⁺ma-Nuap >* SM, MIN, BH *k/uap*. If so, SWY, JKT *ŋ-uap* still reflects the original PM shape, cf. PM **uap <* PMP **huab* 'id.' (Blust 1970:144). Another interpretation (Prentice, pers.comm.) is that SM, MIN, BH *k-* is a fossilised prefix **kA-* (denoting unintentionality).

48. See 3.7.4 IC for **tidur*, it is also attested in JKT. IBN *tinduk* is not a cognate (3.7.4 IC). BH *guriŋ*: cf. KD *guriŋ* 'id.' (but without expected *-kŋ*, which is the regular KD reflex of PM **-ŋ*).

49) BH *ba-rabah*, JKT *rəbè* have cognates with different meaning, cf. SM, IBN *rəbah*, MIN *rabah* 'fall down', and SWY *rəba(h)* 'id.' and *xəba(h)* 'bent down (like a full rice ear)' (cf. 3.7a), hence **rəbah* 'lie down; falling or hanging down'.

50. SM has also *impi* 'dream (v)' and *impi-an* 'dream, fantasy, utopia' (Iskandar 1970; cf. also 4.3.1).

53. Although the general meaning of the reflex of **uraŋ* is 'human being' in all isolects, this etymon should be glossed 'outsider'. This is shown by (1) some specific meanings of its reflexes, cf. SM *nəgəri oraŋ* 'abroad', *baraŋ oraŋ* 'someone else's belongings'; (2) evidence from Malayic Dayak isolects, where reflexes of **uraŋ* still specifically refer to 'strangers, people other than one's relatives or village members' (cf. also Urak Lawoi' *urak* 'man, person who is not Urak Lawoi'); (3) a derivation of PM **uraŋ* from PMP **(t-)uRaŋ* 'outsider' (Adelaar in press b). SWY *jəmo*, SM *jəlma* 'assumption of human form' < SKT *janma-* 'birth, (state of) existence' (Gonda 1973:121).

55. SM and JKT agree, but it is very unlikely that *pərəmpuan* with its four syllables and polymorphemic structure already occurred in PM; JKT *pərəmpuan* is probably a loan from SM. The etymology of SM *pərəmpuan* is uncertain: Von de Wall derives it from *ampu* 's.th. that is supported, sustained, etc.; possession, belonging(s)' (*pər/əmpu/an*, *pər/ampu/an* 'woman, s.o. who has to be supported'); Klinkert derives it from *ampu* 'loincloth for women' (*pər/əmpu/an*, *pər/ampu/an* 'one who wears an *ampu*'). BH *ba/bini/an* possibly originated by analogy with *la/laki/an* (on the basis of *bini*), but it could also be a reflex of PAN **ba+b-in-aHi* (Blust 1982b:19) with **-an*. SWY *t/ino* < PM **ina* 'mother' (5.4.3).

59. cf. 5.4.3

60. The origin of SM *baba*, JKT *babè* is unclear; it is often associated with the Chinese in Malaysia and Indonesia.

63. *Nama* etc. < SKT; BH *ŋaran* < JV (JV *ŋ/aran* < PMP **ajan*). No reconstruction for 'name' can be made; but there is **gəlar* 'title, surname' (3.7.2 IC; cf. Proto Batak **gəlar* 'id.' on the basis of Toba *goar*, Simalungun *gor/an* (with contraction of *+o* + *+a*), Dairi *gərar*, Karo *gəlar* 'name (in general)', Adelaar 1981:passim).

64. cf. also SM *tutur*, MIN *tutuᵊ* 'talk, speak', SWY *tutux* 'address s.o. by his/her surname'.

66. MIN *kabeʔ*, SWY *kəbat* seem to be equivalent in meaning to *ikeʔ*, *ikat*, but in SM and IBN *kəbat* means 'wrap, bind' as against *ikat* 'tie, bind'. On the basis of SM, IBN *kəbat*, BH *kabat* 'wrap, bind', MIN *kabeʔ*, SWY *kəbat* 'wrap, tie, bind', I reconstruct **kəbAt* 'wrap, bind'.

73. BH *cuntan* must be a style variant of *+curi* (cf. fn. 82). JKT *curi* does not occur in Abdul Chaer, who gives *còlòŋ*. *Curi* etc. is a North Indian loanword (3.4.1), and its occurrence is restricted to Malayic isolects; *maliŋ* is widespread outside this group, cf. SM, BH, JV, Ngaju *maliŋ* 'thief by night', Malagasy *madina* 'be careful' (< PMP **maliŋ*).

78. **tətək* 'hack, carve, cut'; cf. also SM *tətak* 'id.', MIN *tataʔ* 'delimit, fence off; carve, cut', cf. PMP **tektek* 'hack, hew, cut'. **tarAs* 'plane with an adze; shape, do some rough-hewing' (3.8.2 UIC). SM *potoŋ*, (MIN *potoŋ* a loan? cf. 3.1.2 N.B. (2)), BH *putuŋ*, JKT *pòtòŋ* 'cut, slice' < **putuŋ*.

83. SM *kərja* (*mə-ŋərja-kan*), MIN *karajo*, JKT *kərjè* < SKT. BH *ba-huma* and IBN *b-umay* derive from **huma(ʔ)* and refer to working a field.

86. See 5.6.4 for the relation between **t/um/buh* and **tubuh*.

88. JKT *pəras* must be an Indonesianism in the informant's speech (Mrs Montolalu, see Blust forthcoming). Abdul Chaer gives *pərəs* and *pərè* (3.8.2 UIC).

91. According to Abdul Jebar (1977:2) and Fudiat Suryadikara et al. (1981:22) BH *hurup* is usual, and *tukar* is BK; cf. BAC *tukaʔ* 'buy', MIN *tuka* '(ex)change, swap; buy'; cognates in the other isolects mean '(ex)change, swap', viz. SM, IBN *tukar*, *tukax*, JKT *tukər*, *tukar*, hence **tukər* '(ex)change, swap'.

93. IBN *tutuk*: cf. PMP **tuktuk* 'id.', and SWY *tutuᵊʔ* *tampi* 'plant dry rice field plot by plot'.

95. **labuh* is reconstructed along with **jatuh*, as it is well attested outside the Malayic group.

96. IBN *asuʔ* reflects PMP **asu* 'dog'; cf. also SAR *asu(ʔ)* (Collins 1987:81), and SM *gigi asu* 'canine teeth'. SM *añjiŋ* etc. agrees in form with SD *eñekŋ* 'domestic pig', which suggests a proto-form **añjiŋ* 'domestic animal'. Outside the Malayic group, *añjiŋ* only has a correspondence in SUN (SUN *añjiŋ* 'id.').

97. According to Wilkinson MIN *uŋgeh* was preferred to *buruᵊŋ* after the latter acquired the connotation of 'penis'.

100. BH *halar* must be a loan from JV: the form maintained *h-* which was lost in modern JV (but still present in older forms of JV, including Old Javanese), cf. also BH *habaŋ* 'red' < JV *abaŋ*, BH *hañar* 'new' < JV *añar*, BH *hiraŋ* 'black' < JV *irəŋ*. SM, BH *harap*, MIN *aroʔ*, SWY, IBN *arap*, JKT *arəp* 'hope (v)' must also have been borrowed from JV at a stage

where this language still exhibited *h-* for PMP **q-*. Compare Old Javanese *harəp* 1. 'the front, fore part' 2. '(to stand before), to wish, desire, be on the point of (about to)': *harəp* reflects PMP **qadep,* which in PM developed into **hadəp* '(be) in front of, before' (3.1.1.5). (From this interpretation it follows that I reject PMP **qarep* 'like (v)'.

102. MIN *mañciʔ,* IBN *cit* derive from an onomatopoeia for the sound which a mouse/rat makes.

103. IBN *dagin* < SAR (3.6.3.3 IC). IBN *isiʔ* has many cognates outside the Malayic group, whereas *dagiŋ* meaning 'meat, flesh' is only found in JV (in the Batak isolects it is also found with the meaning 'body', cf. Toba, Karo *dagiŋ*).

104. **ləmək* 'fat, grease (n)' < PMP **lemek. *gəmuk* 'fat, stout': 3.4.2.3 UIC (1).

105. BH and JKT *buntut* may be loans from JV (JV *buntut* 'tail'). SM *buntut* 'butt, posterior, fag-end; stern (of a ship)', MIN *buntuyʔ* 'female sex-organ', SWY *buntut* 'behind of animals or people', IBN *buntut* 'end, inmost part' yield PM **buntut* 'butt, posterior, end'.

107. cf. SWY *caciŋ* 'intestinal worm'; SWY *ulat* 'worm, maggot, insect' (3.4.2.2UIC); IBN *bəlut* < **bəlut* 'crawling animal, eel' (3.4.2.2). **hulət* is a generic for worms, maggots of the caterpillar type, and **caciŋ* refers to snakelike types of worms (cf. Wilkinson (1959) for SM *ulat* and *caciŋ*).

109. According to Thaib and Van der Toorn, MIN *ñamuəʔ* is 'mosquito' and *raŋiʔ* is 'gnat', which conforms to the meanings of cognates in other isolects; cf. SM, JKT *rəŋit,* BH *raŋit,* SWY *xəŋit, rəŋit* 'minute fly'; PM **rəŋit* 'id.'.

111. MIN *lauəʔ* and SM, BH, IBN *lauk* 'fish (or meat) as a side-dish to rice', hence **lauk* 'id.', and **ikan* 'fish (in general)'.

112. SM *buruk,* MIN *buruəʔ,* SWY *buxuəʔ* 'rotten, worn out', BH *buruk* 'rotten, putrid; old, worn out', IBN *buruk* 'rotten, septic', PM **buruk* 'worn out, gone bad, in decay'. SM, IBN (Richards 1981) *busuk,* MIN, SWY *busuəʔ* 'rotten, stinking, putrid', JKT *busuk* 'rotten, gone bad, stinking; bad, evil, worn out, ugly', PM **busuk* 'stinking, rotten'. SM, MIN, and SWY make a clear distinction between *buruk* etc. 'rotten, in decay' and *busuk* etc. 'rotten, stinking', whereas in BH and JKT both meanings have merged in one lexeme.

113. cf. PMP **daqan* 'branch'. JKT *cabaŋ*: cf. SM *cawaŋ, cabaŋ,* IBN *cabaŋ* 'branch, prong, bifurcation', SWY *cabaŋ* 'forked branch; anything with the shape of a forked branch', yielding **cabaŋ* 'forked branch'.

123. MIN *aliə* 'slippery (of a path)'. *lliə* < **hilir* 'flow down; downstream' (3.7.3 IC).

124. SM *tasik* originally meant 'sea'; the meaning 'lake' is a recent development; **laut* 'towards the sea' (5.2.2).

125. Pigafetta gives 'garan sira' for 'salt' (Pigafetta 1972:66; '-n' is a result of Pigafetta's Italian perception of *-m* in *garam*). Bausani's assumption that 'garan' and 'sira' are synonyms is wrong (Pigafetta 1972:78): both words form a compound which must literally have meant 'grain of salt', and therefore PM **garəm* 'grain' (3.1.1B and 3.1.1.5) and PM **sira* 'salt' (< PMP **qasiRa* 'id.') are reconstructed. In Sasak *garəm* retained the meaning 'grain', and some Malayic isolects (among others BM) still have a reflex of **sira* (cf. *sira(h)* in Wilkinson 1959).

126. See lemma 124 for SM *tasik*.

127. MIN *utan* occurs only in written language (Van der Toorn). In SM, BH, SWY and JKT, *hutan/utan* refers to 'wood(s)', and *rimba* etc. to '(virgin) forest'; IBN *rimba?* 'forest cut but not yet burnt'; PM **rimba?* 'virgin forest'.

144. SM, SWY *bɘnam*, MIN *banam* 'immerse in water or mud', PM **bɘnɘm* 'id.'; BH *banam* must be borrowed from JV (cf. JV *bɘnɘm* 'roast under hot ash; lay in ashes, burn down'). SM, SWY, IBN *tunu* 'burn up' < **tunu*.

148. IBN *putih* 'white' also occurs.

149. SM *merah*, JKT *mèrè*, and MIN *sirah* possibly derive from the same proto-form *(*irah?)*; but a segmentation of *sirah* requires an explanation of its initial constituent *(*s(i)-)*, which is not available. BH *habaŋ* and SWY *abaŋ* are probably borrowed from JV (cf. lemma 100 for the *h-* in *habaŋ* 'wing').

150. SM, SWY, IBN, JKT *kuñit*, MIN *kuñi?*, *kuni?* 'curcuma', BAC *kuñit* 'yellow' < **kunit* 'turmeric; yellow' (3.6.1.2IC). It is likely that BAC *kuñit* is inherited and that *kuniŋ* etc. is a loan: judging from its phonological shape, *kuniŋ* etc. must be a loan from Karo (PMP **-j* > *-ŋ* in Northern Batak isolects, Adelaar 1981:13-14).[177]

152. Correspondences of BH *halus* have a different meaning: SM *(h)alus*, MIN *aluyh*, JKT *alus* 'refined, fine, delicate', SWY *alus* 'fine, tender, decent', IBN *alus* 'smooth, fine, small'. *Halus* may be borrowed from JV (JV, Old Javanese *alus* 'refined, fine, delicate'; Old Javanese *alus* is still analysable as a derivation from *lus*, a noun). On the other hand, if *halus* etc. < JV/Old Javanese *alus*, then (SM), BH *h-* remains unexplained. SWY *kɘci??* underwent palatalisation of **t* due to following **i*. IBN *mit* has a cognate in BRU *damit*; both may be related to SM *(dɘ)dɘmit* 'ghost ("orang halus")' (Prentice pers.comm.).

153. MIN *gadaŋ*: cf. also KCI *gɘdéŋ* (Prentice & Hakim Usman 1978:124). In Sumatra and Negeri Sembilan *gɘdaŋ* is a general term for 'big', whereas *bɘsar* means 'magnate', cf. *oraŋ bɘsar* 'id.' (Wilkinson 1959). BAC has *ra:* which must be a cognate of SM *raya* (Collins 1986a:142), and hence of IBN *raya* (and of KD *aya?*, with unexplained loss of **r-*). In contrast to *bɘsar*, *raya* has many cognates outside the Malayic group (cf. PMP **Raya* 'id.'). Compare also SM *raya*, MIN *rayo*, JKT *rayè* 'great', SWY *rayo* 'big, strong (of build); (excessive marker)'.

154. In this doublet sound-symbolic vowel variation is involved. As other scholars pointed out before, in a number of Malayic lexemes, front vowels are associated with smallness, whereas back vowels are associated with moderate size, and *a* with very large size.

156. SM (Indonesian), SWY, JKT *t-* originated through backformation.

159. *Laweh* has a SM cognate *lawas* 'spacious, broad' (BH *lawas* 'long (time)' must be from JV, cf. JV *lawas* 'id.'). *Luas* has the following cognates: SM, JKT *luas* 'clear, open, unobstructed (of an open field)', BH *luas* 'wide', IBN *luas* 'open space; spacious, cleared'; PM **luas* 'clear, spacious, unobstructed (of an open field)'.

161. BH *supan* has the following correspondences: SM, MIN, JKT *sop/an* 'showing respect (through courtesy, modesty or timidity)'. As suggested by Blust (1980a:142), *sop/an* is possibly a contraction of **saup + *-an*, cf. IBN *saup* 'help, assistance'. But contrary to Blust, I consider *sop/an* as a Malayic, and not a JV, form (cf. 3.1.2.5): it occurs neither in

---

[177]A Karo origin for *kuniŋ* 'yellow' was suggested earlier by Aichclc.

Old Javanese, nor in old dictionaries of JV (*sopan* does not occur in Gericke and Roorda (1901), Jansz (1913), or Pigeaud (1938); it is found in Horne (1974)).

163. BH *hañar* < JV (cf. JV *añar* 'id.'). BH maintained initial and intervocalic *h* in JV loanwords, whereas this *h* was lost in modern standard JV.

164. SM *elok*, MIN *eloʔ*, SWY *iluʔ?* 'beautiful, pretty; good, right'; JKT *éloʔ* 'beautiful, pretty', PM **iluk* 'beautiful, pretty, nice'.

170. SM *bilamana* < *bila* + *mana* (5.5.2.4); SM *bila* etc. < SKT *vela* 'time, moment' (Klinkert 1916). SM (Indonesian), JKT *kapan* < JV *k/apa/n* 'when'.

171. The first syllable in SM, JKT *sambuñi*, BH *sambuñi* remains unexplained.

176. JKT *kòlòŋ* literally means 'space under s.th. (usually a Malay house)', cf. JV, SUN, SM *koloŋ* 'id.'. MIN *baruᵃh* originally meant 'land below', as SM *baruh* still does; hence **baruh* 'land below'.

179. cf. also BAC *dampiŋ* 'near, adjacent' (SM, BH *dampiŋ*, MIN *dampiᵃŋ* cf. 3.8.1 UIC (2)).

188. IBN *nama* is related to SM *nama* 'name' and it is ultimately derived from SKT (cf. also Urak Lawoi' *nama* 'what'). The same semantic change also occurs in other Austronesian languages (cf. also Tok Pisin *wanem* 'what' < English 'what name').

190. **bukən* is given on the following grounds:

(a) Reflexes meaning 'other' are still found in IBN and in OM inscriptions (Çoedès 1930:39-40, 78).

(b) Outside the Malayic group, reflexes meaning 'other' occur in Borneo, Philippines (Blust 1980a:119; Prentice 1974:58), and in Chamic languages.[178] But the Malayic isolects differ from other languages (with the exception of the Chamic ones) in that they have penultimate *u* (in other languages PMP **e* is reflected).

(c) BAC has *ma-lain* 'other' and KD, SD have *lain* 'id.', but these must be loans because of their *-n* (PM **-n* > BAC *-ŋ*, cf. *diŋiŋ* 'cold' < **diŋin*, *ŋaŋ* 'with' (with apocope of first syllable) < **dəŋən*, and *tahuŋ* 'year' < **tahun*; PM **-n* > KD, SD *-tn*, cf. **tahun* 'year' > KD, SD *tahutn*, and **bukən* 'other' > KD *bukatn*, SD *bukàtn*). In other isolects there is no such device for differentiating between inherited and borrowed lexemes.

(d) Some applications of SM *bukan* like *bukan/bukan* for 'nonsense; nonsensical' are understood better if one assumes 'other' as the original meaning; also, *bukan* is not used as a complete negation (except in a one-word reply), but as a contrastive negative requiring or implying an alternative predicate, e.g.

> *Bukan ayah-na   mə/lain/kan Ahmad yaŋ   məmaŋgil-ña.*
> not   father-his but   Ahmad who call-him
> It was not his father, but Ahmad, who was calling him.

The same semantic shift as between IBN *bukay*, OM *bukan* 'other' and other isolects *bukan* 'not' must have taken place between SM *lai/n* etc. 'other' and SUN *lain* 'not' (= SM *bukan*). See Adelaar (1988:71) for the origin of *lai/n*.

---

[178]Lee reconstructed Proto Chamic **tukOn* 'other' on the basis of Roglai *tuket* and Rhade *mkan*, but this should be reinterpreted as Proto Chamic **bukOn* on account of Cham *bukan, bikan, bakan* (Aymonier & Cabaton 1906:334), Jarai *pŏkɔ̆n*, and on account of the initial labial in Rhade.

191. SM *səmua,* BH *samua, samuñaan,* JKT *səmuè* < SKT. SWY *səgaloyo, gəgaloyo* (as well as SM *səgala,* MIN *sagalo,* BH *sagala,* JKT *səgalè* 'id.', IBN *səgala* 'round, whole' are borrowed from SKT via TAM (< TAM *sagala* < SKT *sakala* 'complete, entire, all', Gonda 1973:162). MIN *sadoño* < +*sa-ado-ño.* IBN *abis* and MIN (Van der Toorn) *abih* have, along with the meaning 'finished, used up' (which they share with their cognates in other isolects) also the meaning 'all, complete',[179] hence the reconstruction **habis* 'all, complete; used up, finished'. Other cognates are Proto Chamic **ʔabih* 'all' (Lee 1966), Achehnese *(h)abéh* 'finished, completed, used up; complete(ly), the whole of', Malagasy *avy* 'all' and Old Javanese *hawis* 'finished, completely gone, nothing left' (modern JV *wis* 'already; enough'). Dempwolff (1938) reconstructed PMP **abiq* 'all' on the basis of JV *k/abeh* and Malagasy *avy* 'all', but with the above evidence a PMP form **qabis* is more justified.

192. cf. also JKT *dəŋən* 'with'. MIN *jo* may be an allegro form of *juo* 'also; only'. The origin of IBN *əŋgaw* is unclear (possibly < **dəŋən,* with unexplained loss of **d,* and with (equally unexplained) excrescent *g*). SWY *ŋan* must be short for *dəŋan* 'with'. The same origin was suggested by Aichele (1942-43:42 footnote 3) for SM *dan,* which would have developed from *dəŋan* (originally 'companion'), in the same way as JV *lan* 'and' from *lawan* 'partner, adversary; with'. The different local varieties of SAR have *ŋan* or *dəŋan* for 'and' (Collins 1987:84). Collins (unpublished fieldnotes) supposes that BAC *ŋaŋ* 'and' derives from **dəŋən.* Another origin of SM, MIN *dan* may be **dua(ʔ)-(a)n,* cf. KD *dua* 'and'; such an etymology would also have a parallel in JV, viz. JV *ro* 'two' and *ka/ro* 'with, and'.

193. *(jikalaw), kalaw, kalo* < *jika* (< SKT) + AR *law* 'if'; *jika* is still found as *jaka* in old manuscripts and as *joko* in MIN (Wilkinson 1959).

194. SM *bagai/mana,* JKT *bagi/manè* < +*bagay* 'kind, variety, species' (< TAM) + +*mana* 'which, what' (5.5.2.4, 5.5.2.6). MIN *baa* < +*ba* + *a* 'what' (5.5.2.4). IBN *kati* < ?.

195. The common element in this correspondence set (minus JKT) is the last syllable *-da/-do* (to which SM *-k,* MIN *-ʔ* is attached, see 5.4 N.B.). I do not have an explanation for the different first syllables (but cf. Deli Malay (East Sumatra) *tei,* Urak Lawoi' *tët* and, outside the Malayic family, Malagasy *tsy,* all with the meaning 'no(t)' (Adelaar 1989:42 n.40).

196) *Bilaŋ* 'count (v)' is found in all isolects; in SM, MIN, and JKT it also means 'say' (a very common semantic development); on the basis of this set PM **bilaŋ* is reconstructed.

---

[179]This range of meanings is also found in other languages, cf. German *alles* 'everything' and *alle* 'all (adjective)' versus *alle* 'used up, finished'.

CHAPTER 6

PROTO MALAYIC AFFIXES

In this chapter some of the affixes that occur in the Malayic isolects are treated. Only the affixation of certain word classes, namely that of verbs and nouns, is discussed; reduplication is not treated.

Ongoing research has yet to produce a watertight set of criteria for word-class membership in SM, and this topic needs much further investigation in each of the isolects. I restrict myself here to a syntactical definition of verbs and nouns in SM. The implicit application of these definitions (and that of precategorials, see below) to other isolects has proven to be satisfactory for the aims pursued in this chapter. In SM, verbs are defined as lexical entities which can be modified by the negation marker *tidak*.

Verbs are divided into two main categories:

(a) transitive verbs (VTR), which can occur in object-oriented constructions and which govern an object (O);

(b) intransitive verbs (VI), which do not occur in object-oriented constructions. Intransitive verbs are further subdivided into dynamic intransitive verbs (VDI) which refer to an action, a process, or a change of state, and stative intransitive verbs (VSI) which refer to a state of affairs.

Syntactic differences between VSIs and VDIs include the following:

(1) VSIs cannot be modified by the aspect markers *təlah, akan,* and *sədaŋ* (denoting perfect, non-commenced, and ongoing action respectively; other aspect markers such as *sudah* (completed action) occur with both VSIs and VDIs).

(2) Unlike VSIs, VDIs cannot modify nouns directly, but must be placed within a relative clause (cf. *anak nakal, anak yaŋ nakal* 'naughty child', but **anak tidur* vs *anak yaŋ tidur* 'sleeping child').

(3) VDIs cannot function as complements of *(məñ)jadi* 'become' (to be distinguished from *jadi* 'succeed, manage').

Nouns are lexical entities that can be subject or object of a sentence, that can occur following a preposition, and that can be modified by other nouns, VSIs, demonstratives, numerals, and relative clauses. Verbs sometimes also occur in positions reserved for nouns: when they do, they are considered as nominalised verbals that refer to the process as such.[180] Lexemes in the Malayic isolects usually belong to a certain word class (or more than one word class). But they do not always do so: many occur only as the base of a derivation, in which case only the derivative has word-class membership. They may also occur unaffixed, but only as

---

[180]The concepts VSI and VDI, as well as the definitions of VSIs, VDIs, VTRs and nouns, and the descriptions of some of the SM affixes presented here are from Prentice (1987, n.d.). The idea of distinguishing between

an element of a compound. These lexemes are called precategorials.[181] Since precategorials are usually the base of verbal and nominal derivations, they will be included here. Precategorial bases will be indicated by a preceding hyphen, and their meanings will be given between brackets.

Other verb classes are not distinguished. Apart from VTRs and VDIs, Muhadjir (1981:13), in his treatment of word classes in JKT, also distinguishes "semi-transitives" along with adjectives, intransitive verbs, and transitive verbs (which are comparable to VSIs, VDIs, and VTRs respectively in this study). Semi-transitives have co-constituents, but these are complements, not objects: since the semi-transitives have no object-oriented form these complements can only be made subject by the use of transitive (derived) verb forms. The co-constituents are often introduced with *amè* (also translatable as 'with, towards'). Reciprocal verbs all belong to the category of semi-transitives because they usually also have co-constituents. Muhadjir's "semi-transitives" could be distinguished on the same terms for SM as for JKT: SM also has a class of verbs which can govern a complement which, however, cannot function as subject vis-à-vis these verbs (e.g. *suka* 'like (s.th. or s.o.)', *tahu* 'know (s.th.)', *bəlajar* 'study, learn (s.th.)', *bərbicara* 'speak (a language)'; these verbs do not require a preposition). Rather than adding another subclass of verbs to the ones already distinguished here, I prefer to consider Muhadjir's semi-transitives as a subclass of VDIs that have the faculty of taking a complement (which may or may not be introduced by a preposition), which is in correlation with their semantics.

As far as PM is concerned, I do not make any attempt to reconstruct the word-class system although I assume that comparable word classes must have existed. The precategorials in particular have an uncertain status in PM.

I confine myself here mainly to the study of living affixes, although I give a short discussion of the evidence from some fossilised affixes in 6.8.[182] There are verbal and nominal affixes. Verbal affixes are derivational or inflectional (the latter are focus-markers); nominal affixes are derivational. Cliticised reflexes of **saʔ* (> MIN, BH *sa*, o.i. *sə-*) are treated in 5.3.2; reflexes of the clitic **-ña* (> SM, BH *-ña*, MIN *-ño*, *-no*, *-o*, *-e*, SWY *-ño*, *-o*, JKT *-ñè*) are treated in 5.5.1.3. Transitivising affixes (SM *-i*, *-kan*, BH *-i*, *-akan*, SWY *-i*, *-ka(n)/-kə*, IBN *-ka*, JKT *-in*; SM *(məm)pər-*, MIN *(mam)pa-*, BH *(ta)pa-*, IBN *p(ə)-*) and reflexes of PM **(mb)Ar-* (> SM *bər-*, MIN, BH *ba-*, o.i. *bə-*) have already been treated in Adelaar (1984a); a short account of this treatment is given below in 6.1.1, 6.1.2.

Verbs can be either simple or derived. For instance, transitive verbs are either transitive sui generis or secondarily derived from nouns or other verbs. Derived transitive verbs are subject to some of the same morphological processes as primary verbal bases (the derivative

---

VDIs and VSIs by testing their valency with the verb *məñjadi* 'become' is from Hein Steinhauer. Prentice and Steinhauer's criteria are tentative, and have not yet been tested systematically.

[181]The definition of precategorials differs from one author to the other. For instance, Muhadjir's definition includes root morphemes with multiple-class membership, components of (partial or entire) reduplication that do not occur in other environments (e.g. *kurè*, *grak* and *còdòt* would be precategorial roots in the words *kurèkurè* 'turtle', *grakgrikgruk* 'sound like one who has a cough or a cold', and *cəcòdòt* 'bat'), lexemes denoting a unit of money such as *pérak* 'rupiah', and even loanwords deriving from prepositional phrases in the lending language (cf. *alakadar* 'what is available (adv.)' < AR ᶜalā qadr 'commensurate with, corresponding to, according to'). I prefer to narrow down the definition of precategorials to roots that do not occur in isolation, that is, roots which only occur in derivations and in compounds. I see no need to make a further segmentation of loanwords which derive from morphologically complex forms, compounds or prepositional phrases in the lending language, at least not if they are fully lexicalised in the borrowing language, as is the case with AR ᶜalā qadr > SM, JKT *alakadar*.

[182]Other fossilised affixes are SM *-ər-*, and *-əl-*, and their correspondences. Although admittedly they also deserve attention in the study of Proto Malayic, I will not treat them here.

affixes are usually lost in these processes). So there is an ordering of derivational rules, and verbal derivation occurs prior to other derivations. This rule ordering is implicit in the description of the Malayic affixes. For example, in SM, *bəsar* 'big' is a VSI, which becomes nominalised through circumfixation of *kə-_-an₁* (*kəbəsaran* 'largeness, greatness'); *air* 'water' is a noun. *Bəsar* and *air* become VTRs through suffixation of *-kan* and *-i* respectively, and as such they can become abstract nouns, i.e. nouns referring to the act or event as such in the same way as *kirim* 'send' (a VTR), viz. through affixation of *pəN-_ -an* (6.7.3). Thus:

adjective	verb	nominal derivative
	kirim 'send (O)'	*pəŋiriman* 'consignment'
bəsar 'big'	*bəsarkan* 'enlarge (O)'	*pəmbəsaran* 'enlargement'
air 'water'	*airi* 'irrigate (O)'	*pəŋairan* 'irrigation'

In what follows here derivations like *bəsarkan* and *airi* will be considered as VTRs, and *pəmbəsaran* and *pəŋairan* as nouns derived from these VTRs (the fact that they ultimately derive from the VSI *bəsar* and the noun *air* is not relevant for the *pəN- -an* derivation). There are many shortcomings in the following presentation of the morphology of the six isolects. These are mainly due to the lack of information for some isolects, to the variety of methods used for describing data in the individual isolects, and to the inadequacy of some of these descriptions. The handicaps were felt most in the BH and SWY material, but they also play a role in the treatment of the other isolects.

6.1 EARLIER RECONSTRUCTIONS

In Adelaar (1984) I discussed the transitivising affixes and the intransitive verbal prefixes *bər-/ba-/bə-* in the Malayic isolects. Ras (1970:439-443) and Collins (1986b:190-193) had already given attention to SM *-kan*, and Roolvink (1965) had studied the interrelationship of SM *bər-* and *(məm)pər-* in Classical Malay texts. In Adelaar (in press b) I give evidence for a PM linker *ŋ*.

6.1.1 THE PM INTRANSITIVE AND TRANSITIVE VERB MARKERS

Ras argued that SM *-kan* is a relatively recent suffix compared to *-i* (and *-an*) on account of differences in their morphophonemic behaviour in some (Northern) peninsular isolects. Also, in these isolects *-kan* (Kelantan *-kɛ*) is productive, whereas *-i* is unproductive and survived only as a fossil.

Collins compared SM *-kan* to Banjarese *-akan* and BAC *-akaŋ* and *akaŋ*. He concluded that a preposition **akan* has to be reconstructed for this correspondence, and that Banjarese and BAC reflect intermediate stages of the development from this preposition into a suffix as in SM.

In Adelaar (1984) I check the findings of both Ras and Collins against evidence from the other isolects. I agree with Ras that SM *-kan* etc. is younger than SM *-i* etc., and with Collins that a preposition has to be reconstructed as the PM ancestor of SM *-kan*. However, this preposition should be **akAn* instead of **akan* on account of data from languages outside the Malayic group and the fact that PM retained schwa in last syllables (cf. PAN **akən*,

TABLE 13: THE AFFIXES COMPARED IN CHAPTER 6

		SM	MIN	BH	SWY	IBN	JKT
(verbal) prefix denoting unintentionality	6.2	*tər-*	*ta-*	*ta-*	*təx-*	*tə-*	*tə-*
focus-marking prefixes	6.3	*məN(1)-*	*maN(1)-*	*maN(1)-*	*(mə)N(1)-*	*N-*	*N(1)-, əN(1)-, ŋə(1)-*
	6.3	*ø-*	*ø-*	*ø-*	*ø-*	*ø-*	*ø-*
	6.3	*di-*	*di-*	*di-*	*di-*	*di-*	*di-*
intransitive verbal prefix	6.3	*məN(2)-*	*maN(2)-*	*maN(2)-*	*(mə)N(2)-*	*əN-*	*N(2)-, əN(2)-, ŋə(2)-*
subjunctive suffix†	6.4.1	--	--	--	--	--	--
transitive affix††	6.4.2	--	--	--	--	--	--
(verbal and nominal) suffix denoting plurality of subject, diffuse action, reciprocity	6.5	*-an(1)*	*-an(1)*	*-an(1)*	*-an(1)*	*-an(1)*	*-an(1)*
nominalising suffix referring to goal or place of an action	6.5	*-an(2)*	*-an(2)*	*-an(2)*	*-an(2)*	*-an(2)*	*-an(2)*
circumfix denoting unintentionality	6.6.1	*kə- -an(1)*	*ka- -an(1)*	*ka- -an(1)*	*kə- -an(1)*	--	*kə-, kə- -an(1)*
nominalising circumfix referring to a quality, process, event	6.6.2	*kə- -an(2)*	*ka- -an(2)*	*ka- -an(2)*	*kə- -an(2)*	--	*kə- -an(2)*
(nominal) agent- and instrument-prefixes	6.7.1	*pəN-*	*paN-*	*paN-*	*pəN-*	*pəN-*	*pəN-*
	6.7.2	*pər-*	*pa-*	*pa-*	*(pər-)*	--	--
circumfixes forming abstract nouns and nouns referring to place (or actor, goal, instrument)	6.7.3	*pəN- -an*	*paN- -an*	*paN- -an*	*pəN- -an*	--	*pəN- -an*
	6.7.4	*pər- -an*	*pa- -an*	*pa- -an*	*pəx- -an*	--	*(pər- -an)*
fossilised affixes	6.8	*m/-, -/əm/-*	*m/-, -/um/-*	*m/-*	*m/-, -/əm/-*	*m/-*	*m/-, -/əm/-*

†only reflected in KD *(-a?)*, SD *(-à?)*, and Old Malay *(-a)*

††only reflected in KD, SD and Old Malay *(maka-)*

Pawley & Reid 1976:59). The reasons I give for reconstructing a PM preposition are that in SM a preposition *akan* still exists and, preceded by a VSI or VDI, is often interchangeable with a derived VTR with *-kan* (especially in Classical Malay texts). BH *-akan* still shows the full form of the original preposition (before the loss of **a-*). Moreover, BH *-akan* can also be suffixed to a verb that already has a suffix *-i*, whereas the opposite (suffixation of *-i* to a verb that already has *-akan*) does not occur. Other indications for not reconstructing a suffix is that there is no formal agreement among the correspondences of *-kan,* cf. SM, MIN *-kan,* BH *-akan,* SWY *-ka(n)/-kə,* IBN *-ka,* JKT *-in.* Moreover, IBN *-ka* is a clitic that still occurs independently in some sources (e.g. Scott 1956), whereas the older suffixes are lost in IBN (it is not improbable that SWY *-ka(n)/-kə* are also cliticised forms of a preposition *kə* 'to(wards); in order to'). JKT *-in,* which combines the functions of *-kan* etc. and *-i* in the other isolects, is formally identical with *-in* in Balinese (from which it borrowed heavily); it is probably a loan morpheme from Balinese. Finally, there are sometimes formally different correspondences between closely related (sub)dialects: most regional forms of MIN have *-an* instead of *-kan,* and KD, which is very close to IBN, has *-atn* or *-an* corresponding to IBN *-ka.*

I reconstruct **-i* on the basis of SM, MIN, BH, SWY *-i,* and IBN (fossilised) *-/i?.*[183] **-i* was a locative-oriented transitivising suffix; when added to transitive verbal bases, it could also add the notion of multiple action or plurality of object or subject.

Roolvink found that in Classical Malay texts *bər-* was originally transitive, and that it corresponded to *pər-* in passive constructions. Before the end of the eighteenth century *bər-* and *pər-* became disconnected, the former now being an intransitive verbal marker, and the latter acquiring an active counterpart in *məmpər-* (= *məN-* + *pər-*). In the light of comparison with the other isolects (and with Kedah, Pattani and Jakun Malay) I agree with this opinion in-so-far as I reconstruct **pAr-* as a prefix forming transitive verbs, and **(mb)Ar-* as one forming intransitive verbs, and assume that there was a paradigmatic relation between **pAr-* and **(mb)Ar-* . **pAr-* occurred affixed to adjectives (but not exclusively so) to form causatives, and affixed to nouns it conveyed the meaning 'treat or use O as a (noun), turn O into (noun)'. **(mb)Ar-* occurred with intransitive verbs and nouns. With nouns it must have meant 'possess, contain, wear, use, produce, acquire (noun)' or, if the noun referred to a profession or mutual relationship, 'assume the quality of (noun)'. The initial consonant in **(mb)Ar-* is uncertain because of the correspondence of OM *mar-* with evidence from outside the Malayic group (cf. Toba *mar-,* Old Javanese *mər-,* Tagalog *mag-*), which points to **mAr-* (against evidence in all contemporary Malayic isolects pointing to **bAr-*).

Traces of a paradigmatic relationship between **pAr-* and **(mb)Ar-* are to be found in the isolects:

(a) In the formation of deverbal nouns, SM *bər-* is replaced by *pər-,* and MIN *ba-* is replaced by *pa-,* e.g. SM *bərjañji* 'promise, make an agreement', and *pərjañjian* 'promise, agreement, testament', MIN *bajalan* 'walk, go', and *pajalanan* 'trip, journey; distance covered (during a journey)'; *pər-/pa-* also applies when a deverbal noun is formed on the basis of a verb with *(məm)pər-/(mam)pa-,* e.g. SM *(məm)pərkəcil* 'make smaller' and

---

[183]Despite the IBN reflex I do not reconstruct a final glottal stop for this proto-suffix, since Formosan cognates of *-i / -i?* do not exhibit a corresponding *-s, -h* or *-?,* nor do Philippine cognates have a corresponding *-h* or *-?* (Adelaar 1984a:419). Timugon Murut has a referent-focus suffix *-i?* (used in phrases with atemporal aspect; Prentice 1971). A final glottal stop after vowels is shown in many Bornean languages, and it is quite possible that this is an areal feature (which, as an alternative explanation for IBN *?,* affected IBN as well).

pərkəcilan 'minimising, reducing in size', MIN *(mam)padamaykan* 'reconcile, pacify' and *padamayan* 'reconciliation; place of reconciliation'.

(b) The unproductive BH prefix *tapa-* expresses unexpected or involuntary action. It is probable that it originally consisted of ⁺*ta-* + ⁺*pa-*, and that the second part of this sequence was an alternant of ⁺*ba-* in this particular environment, cf. *bahurup* 'exchange'[184] and *tapahurup* 'exchanged (by accident, erroneously)'; *bahual* 'be quarrelling' and *tapahual* 'get into a fight (unasked for)'.

(c) IBN *bə-* in combination with *-ka* yields transitive verbs; when these verbs are object-oriented (which includes verbs in the imperative mood) *bə-* alternates with *pə-*, e.g. *bəjalay* 'walk, go, move' and *bəjalayka* 'make O go, move O', *pəjalayka!* 'move it!', *dipəjalayka* 'be moved'.

6.1.2 A PM LINKER

The PM (or ligature) **ŋ* was apparently used between quantifiers and following nouns, and after pronouns introducing relative clauses: these are at least the positions in which SM maintained a fossilised remnant of this linker, cf. SM *bara/ŋ* (5.5.2.5) and SM *ya/ŋ* (2.1.3, 3.3, 5.5.2.3). PM **ŋ* apparently developed from PMP **ŋ*, but its use became very restricted. Reflexes of PMP **ŋ* are found in a good many other Austronesian languages, where they often also occur in numeral compounds (between digits and higher order numerals). In some languages they are used as a linker between the constituents of almost any kind of noun phrase, as in Tagalog. That PM **ŋ* was not yet fossilised can be deduced from the fact that *ya/ŋ* and *bara/ŋ* cannot be reconstructed for PM (Adelaar in press b).

6.2 SM *tər-* AND ITS CORRESPONDENCES

6.2.1 SM

With VDIs and VTRs, *tər-* denotes an 'accidental' state, process or action; the term accidental is used to cover such concepts as involuntary, unmotivated, agentless, sudden, and unexpected action (or state resulting therefrom). Prefixed to VDIs and precategorials *tər-* forms active verbs, and prefixed to VTRs it forms verbs that are active or passive, depending on the context in which they occur. The transitivising suffixes are almost always elided. For example:

tidur 'sleep' *tərtidur* 'fall asleep';
-pəkik > *məməkik* 'scream' *tərpəkik* 'scream involuntarily';
honnat 'respect (n,v)' > *hormati* 'respect, pay homage to' *tərhormat* 'respected, esteemed';
buka 'open' *tərbuka* 'open(ed); opened (by mistake)';
dapat 'acquire, get' > *dapati* 'come across, catch, find' *tərdapat* 'occurring'.

Compare also the constructions used in the following examples:

(1)　　*tərbawa*　　*aŋin*
　　　　tər- + carry　wind
　　　　taken by the wind

---

[184]cf. Abdul Jebar; *bahurup* also means 'buy', cf. 5.7 lemma 91.

(2) *Saya tərbawa   hujuŋ atap itu.*
     I   *tər-* + carry end   roof that
     I took by accident a straw from the roof with me.

(3) *Tərgali -lah ia kəpada tajaw.*
     tər- + dig -*lah* he on   vase
     While digging, he hit upon a vase.
     (-*lah* is an emphatic particle, also used to mark a fronted predicate.)

(4) *Ia təsərdawasərdawa səpərti tagar   buñi-ña.*
     he *tər-* + burp(ing)   like   thunder sound-its
     He burped with the power of thunder.

N.B. In Indonesian SM *tər-* + a transitive verbal base is only used as an object-oriented form, but in Malaysian SM and in Classical Malay texts this construction is also used with agent-orientation, as in the second and third sentences given above.

On the basis of VTRs *tər-* can also denote a potential action, and consequently an impossible action if the resulting form is preceded by a negation. In Gerth van Wijk's examples the resulting form is object oriented if the base is a VTR, e.g.

(5) *Kaki-ña tidak təraŋkat.*
     leg-her not *tər-* + lift
     She couldn't lift her leg.

(6) *Hati-ña   tiada tərtahan.*
     liver-his not *tər-* + restrain
     His heart[185] could not be restrained.

(7) *baraŋ yaŋ   tərbawa   oleh sə-oraŋ*
     good which *tər-* + carry by   a/one-person
     goods that can be carried by one person

(8) *tiada tərkatakan*
     not *tər-* + say
     unspeakable, indescribable/not to be said, ineffable (of blasphemies etc.)

(9) *Papan itu səmua-ña tərpakay.*
     plank that all-its   *tər-* + use
     Those planks can all be used.

Prefixed to VSIs, *tər-* denotes a superlative degree, e.g.

baik 'good'	*tərbaik* 'very good, best'
kəcil 'small'	*tərkəcil* 'very small, smallest'
baru 'new'	*tərbaru* 'very new, newest'

6.2.2 MIN

MIN *ta-* prefixed to VDIs and VTRs is equivalent to SM *tər-*. It denotes that the subject either commits the act involuntarily, or is affected by the act. The will or motivation to

---

[185] *Hati* is translated as 'heart' in this context, as the liver is considered the centre of emotions in Malay culture.

perform the act is beyond the subject. Forms prefixed with *ta-* are irrespective of agent/object-orientation or tense (*-i* and *-kan* are not elided when *ta-* is prefixed), e.g.

(10)   *Uraŋ   tu   tajua.*
          person   that   *ta-* + sell
          That person is selling/has sold (by accident) (or That person is sold/was sold).

(11)   *Kudo tu   tajua.*
          horse that *ta-* + sell
          That horse is sold/was sold.

(12)   *Pupuyʔ lah sudah   tapabueʔ.*
          flute   *lah* already *ta-* + make
          The flute has been made.

(13)   *Indaʔ tanantikan   do   di   den.*
          not   *ta-* + wait.for particle by   me
          I cannot wait for you.

(14)   *Lah   bañaʔ tajuᵊʔ tajalani,   bañaʔ lah   karaŋ talieʔi.*
          (particle) many bay   *ta-* + go by many particle rock *ta-* + see O
          Many bays were passed, many rocks were seen.

N.B. *Lah* preceding a verb is a perfect tense marker; in other positions it is an emphatic particle.

Prefixed to a noun, *ta-* means 'be affected by / affect (noun) (irrespective of one's own will); fall, lapse into (noun); succeed in acquiring (noun)', e.g.

(15)   *Takudokudo   samiᵊŋ pañcarian-ño.*
          ta- + horse + reduplication all   earning-his
          Everything he earns is lost on horses.

(16)   *Sa-bulan   ño yo   bakuli,   lah   takabaw   di ño.*
          one-month he indeed work.as.a.coolie particle *ta* + buffalo by him
          He's been a coolie for a month, and he's already got hold of a buffalo.

(17)   *Tatulaŋ   sakiʔ-ño di   ambo tu.*
          ta- + bone pain-its on/by me   that
          It hurts me very much (right to the bone I feel the pain of it).

(18)   *Barulah   tabatu   ambo kali, ambo antikan.*
          as.soon.as *ta-* + stone I   dig I   stop (with O)
          As soon as I hit upon a stone, I stopped digging.

Prefixed to VSIs *ta-* denotes an excessive degree, e.g.

         siŋkeʔ 'short'   *tasiŋkeʔ* 'too short'
         dalam 'deep'   *tadalam* 'too deep'
         kuniᵊŋ 'yellow'   *takuniᵊŋ* 'too yellow'
         labiᵊh 'more'   *talabiᵊh* 'too much'

N.B. According to Van der Toorn *ta-* can also mean that the subject acquires, or is affected by, the quality of the VSI, irrespective of the subject's own will. It is more probable that forms with this meaning are derived from inchoative VDIs which in turn were derived from VSIs. For example:

(19)   *Tasirah   muko-ño.*
         ta- + red face-her
         She turned red (all of a sudden) (cf. *mañirah* 'turn red').

(20)   *Taputiᵊh   gigih-ño sadaŋ ño gala?.*
         ta- + white tooth-his while he laugh
         When he laughs the white of his teeth appears.

6.2.3 BH

Like SM *tər-*, BH *ta-* denotes unintentionality when prefixed to VDIs and VTRs; when prefixed to VTRs[186] it also denotes feasibility. In most of Asfandi's examples the subject of *ta-* forms is in the object role, but there are some exceptions (*-i* and *-akan* are maintained when *ta-* is prefixed), e.g.

lihati 'see (O)'	*talihati* 'see (O) by coincidence'
tulak 'go'	*tatulak* 'go after all' ('jadi pergi', Asfandi 1976:27)
padahakan 'tell, report'	*tapadahakan* 'told, reported'
padahi 'advise'	*tapadahi* 'be advised, receive advice'

Prefixed to VSIs *ta-* denotes a comparative degree, e.g.

baik 'good'	*tabaik* 'better'
kurus 'slim'	*takurus* 'slimmer'
habaŋ 'red'	*tahabaŋ* 'redder'

6.2.4 SWY

SWY *tə-* (with variants *təx-*, *to-*) has the same functions as SM *tər-*. It denotes either involuntariness or feasibility of a performance when prefixed to VTRs or VDIs, and it denotes a superlative degree on the basis of VSIs. No examples of *tə-* forms with a subject in the agent role are given in Helfrich.

6.2.5 IBN

IBN *tə-* (with its variant *t-*) is equivalent to SM *tər-* when prefixed to VDIs and VTRs. Depending on the context, it denotes non-intention, ability and possibility (Asmah 1977:87). It does not occur with VSIs.

---

[186]On the basis of VTRs consisting of VSIs, VDIs, and nouns + a combination of the VTR markers *-i* + *-akan*, *ta-* can also express a request, e.g.
diam 'stay', *diami* 'stay on O', *tadiamiakan* 'ask O to stay';
kuraŋ 'less', *kuraŋi* 'diminish', *takuraŋiakan* 'ask for a bit less O';
bañu 'water', *bañui* 'give water', *tabañuiakan* 'ask for O to be given water (or to be watered)';
hatap 'roof', *hatapi* 'roof O', *hatapiakan* 'make a roof on behalf of O', *tahatapiakan* 'ask to roof (s.th.) for O'.
(No intermediate forms ⁺*diamiakan*, ⁺*kuraŋiakan*, or ⁺*bañuiakan* were found in Asfandi's data.)

6.2.6 JKT

JKT *tə-* is not productive; it is in the process of being replaced by *kə-* and *kə- -an* (which are treated in 6.6.1).[187] *tə-* occurs with VDIs and VTRs, and it expresses unintentionality or feasibility, e.g.

tidur 'sleep'	*tətidur, kətidur(an)* 'fall asleep'
ŋərasè 'feel'	*tərasè, kərasè* 'felt'
-tawè '(laughing)'	*tətawè, kətawè* 'laugh'
-lambat '(delayed, slow)'	*təlambat* '(too) late'
pandaŋ 'look at'	*təpandaŋ* 'esteemed'
bòŋkar 'take apart'	*təbòŋkar, kəbòŋkar* 'turned inside out, in disorder can be disclosed'
-plèsèt '(stumbling)'	*təplèsèt (kəplèsèt)* 'stumble, trip (by accident)'
bukè 'open', *bukain* 'open s.th.'	*təbukè, kəbukè* 'opened (by accident, involuntarily) open; can be opened'

N.B. According to Muhadjir (1981:36) *tərasè, təplèsèt,* and *təbukè* are probably alternants coming from Indonesian SM.

6.2.7 DISCUSSION OF SM *tər-* AND ITS CORRESPONDENCES

All isolects have a cognate of SM *tər-*. These cognates agree in their function of denoting unintentionality or potentiality of action, process or state when they are prefixed to a VTR or a VDI (or a precategorial). In SM, MIN, BH, and IBN, the subject of a *tər-/ta-/tə-* form is (involuntarily) agent or object of the performance. In SWY, the subject of a form with *tə-* is always the object (in the examples in Helfrich and in Aliana et al.). In JKT (and in Indonesian SM) the subject of a *tə-(/ter-)* form is always object if the base is a VTR. JKT *tə-* forms are not productive.

I presume that the PM ancestral form of *tər-* etc. was a prefix denoting unintentionality and potentiality of a performance, and that the active or passive meaning of this ancestral form depended on the context.[188] That the semantics of *tər-* etc. should not be associated with passivity is even more clear from its application in Classical and Malaysian SM, where sentences like *saya tərmakan lalat* 'I swallowed a fly (by accident)' (with a subject in the agent role) are accepted. In Indonesian SM the same message would be expressed as *Lalat tərmakan oleh saya*.

The prefix *tər-* and its cognates are also prefixed to VSIs: in SM and SWY they form superlatives, in MIN excessives, and in BH comparatives. In IBN and JKT *tə-* does not occur with VSIs. There are two possible underlying causes of this situation: either (1) *tər-/ta-/tə-* extended its valency to VSIs in SM, MIN, BH, and SWY, or (2) constructions of IBN and JKT *tə-* + VSIs were replaced by periphrastic constructions. If (1) applies, then *tər-* + VSIs is innovative. In case (2) a PM form which also occurred with VSIs must be reconstructed. Such a form may have denoted a higher degree, which later on developed into clearcut meanings such as comparative, superlative, or excessive. But it could also have

---

[187] According to Ikranagara (1980:137) no cognate of SM *tər-* occurs in JKT.

[188] Wouk (1980:84) stresses the unintentional and involuntary aspect of *tər-* in Indonesian SM: "The implications of state and unexpressed outside agency cause this construction – and by extension *tər-* in general – to be associated with the concepts of stativity and passivity, leading to some confusion about its underlying significance".

had an excessive meaning, which later on weakened into superlative in SM and SWY, and into comparative in BH. Positing an original excessive meaning has the advantage of putting the origin of *tər-* etc. + VSIs more in line with *tər-* etc. + VTRs, VDIs or precategorials (i.e. with *tər-* with an unintentional meaning). This means that only one PM prefix would need to be reconstructed for both, with an unintentional and excessive meaning combined with a potential one; this prefix would have occurred with all verbs (VSIs, VDIs, VTRs, precategorials). But with the evidence at hand it is safer to limit oneself to what can actually be reconstructed with certainty. That is a PM prefix **tAr-*, which occurred at least with VTRs and VDIs (including precategorials), and which conveyed the notion of unintentionality and potentiality. The subject of a **tAr-* form was either object or (involuntary) agent of the performance.

6.2.8 RECONSTRUCTION OF **tAr-*

I reconstruct **tAr-*, which contributed the notion of unintentionality or feasibility to the VTR or VDI to which it was affixed. It is unclear whether **tAr-* also occurred on the basis of VSIs (with the meaning of a degree marker).

6.3 THE FOCUS-MARKING AFFIXES *məN(1)-*, *ø-*, AND *di-*, AND THE INTRANSITIVE VERBAL PREFIX *məN(2)-*, AND THEIR CORRESPONDENCES

6.3.1 SM

Prentice distinguishes *məN(1)-* and *məN(2)-* in SM: *məN(1)-* is an (inflectional) transitive verbal prefix indicating orientation towards the agent of an action. *məN(1)-* corresponds to *ø* and *di-*:

ø   applies when the verb is object-oriented and the agent is a first or second person (which includes verbs in the imperative mood);[189]

di-   is prefixed to object-oriented verbs, when the agent is a third person.[190]

Examples:

(ikut) məŋikut 'follow (O)': *ikut(lah)!* 'follow (O)!' (*-lah* is a clitic denoting emphasis), *(ku)ikut, (kaw)ikut* 'be followed (by me, you)', *diikut* 'be followed (by third person)';
(pərhatikan) məmpərhatikan 'observe (O)', *pərhatikan(lah)!* 'observe (O)!', *(ku)pərhatikan, (kaw)pərhatikan* 'be observed (by me, you)', *dipərhatikan* 'be observed (by third person)';
(bayar) məmbayar 'pay (O)', *bayar(lah)!* 'pay (O)!', *(ku)bayar, (kaw)bayar* 'be paid (by me, you)', *dibayar* 'be paid (by third person)';
(makan) makan 'eat (O)' (with frequently occurring verbs beginning with *m-* the prefix is usually omitted), *makan(lah)!* 'eat (O)!', *(ku)makan, (kaw)makan* 'be eaten (by me, you)', *dimakan* 'be eaten (by third person)'.

məN(2)- is a derivational prefix forming VDIs on the basis of VDIs, VSIs, nouns and precategorials. On the basis of VDIs it does not add to their meaning in a systematic way; it is sometimes omitted in colloquial speech. In some cases it is interchangeable with *bər-*, e.g.

---

[189]In Indonesian SM *ø-* occurs when the agent is not a common noun (i.e. it may also occur with a third person pronoun in the agent role).
[190] In Indonesian SM *di-* may also occur when the agent is a first or second person pronoun preceded by *oleh* 'by'.

tuŋgu 'wait'	*mənuŋgu (bərtuŋgu)* 'wait'
lompat 'jump'	*məlompat* 'jump'

With VSIs it forms inchoative verbs, e.g.

kuniŋ 'yellow'	*məɲuniŋ* 'turn yellow'
tətap 'fixed, regular'	*mənətap* 'settle, establish oneself'
jauh 'far'	*məɲjauh* 'withdraw'
gəlap 'dark'	*məŋgəlap* 'become dark'

With nouns it forms verbs denoting 'behave like, resemble (noun), move towards (if (noun) is a place or direction), collect or produce (noun), use or consume (noun)', e.g.

batu 'stone'	*məmbatu* 'turn to stone, be rock-hard'
puñcak 'peak'	*məmuñcak* 'peak, rise to a climax'
səmut 'ant'	*məɲəmut* 'swarm (like ants)'
təpi 'edge'	*mənəpi* 'go to the edge, move aside'
kiri 'left'	*məŋiri* 'move, keep to the left'
darat 'dry land'	*məndarat* 'go ashore, land'
rotan 'rattan'	*mərotan* 'collect rattan'
sajak 'poem'	*məɲajak* 'compose poetry'
rumput 'grass'	*mərumput* 'cut grass; graze (of cattle)'
tuba 'fish poison'	*mənuba* 'fish with tuba poison'
kopi 'coffee'	*məŋopi* 'drink coffee'

It occurs with precategorials, where it is also sometimes interchangeable with *bər-*, e.g.

-taŋis '(weeping)'	*mənaŋis* 'weep'
-tari '(dancing)'	*mənari* 'dance'
-ñala '(blazing, flaring)'	*məɲala, bərñala* 'blaze, flare up'
-ñañi '(singing)'	*məɲañi, bərñañi* 'sing'

6.3.2 MIN

MIN *maN(1)-, ø-, di-* have the same functions as SM *məN(1)-, ø-, di-,* but their distribution is different in one detail: a second person agent expressed by the pronouns *aŋ, awaʔ* or *kito,* may follow an object-oriented verb with *di-* prefixed (no examples are given in Van der Toorn). *maN(2)-* is similar to SM *məN(2)-,* but prefixed to a noun denoting a place, it conveys the meaning '(be) on (noun)', e.g.

pasiᵊ 'sand, beach'	*jalan mamasiᵊ* 'a journey along the beach'
tapi 'edge, rim'	*manapi* 'be on/go along an edge'
rumpuyʔ 'grass'	*marumpuyʔ* 'sit on the grass'

Prefixed to VSIs it sometimes yields causative VTRs, e.g.

sirah 'red'	*mañirah* 'make (O) red'
itam 'black'	*maitam* 'blacken (O)'

Prefixed to VDIs it denotes a state or movement which is brought about by the subject's own will, cf. the following sentences:

(21) *Tupay mañjatuᵊh.*
squirrel *maN-* + fall `
The squirrel let itself fall.

(22) *Di bukiʔ   uraŋ   lah   manajun.*
at mountain person particle *maN-* + jump
People jumped down from the mountains.

(23) *Teñju-no lapeh   den mandududᵊʔ.*
fist-his   let loose I   *maN-* + sit
When he hit me, I went and sat down.

cf. also *tiduᵊ* 'sleep, lie down' and *maniduᵊ* 'go and lie down'.

6.3.3 BH

BH *maN(1)-*, *ø-*, *di-* are equivalent to SM *məN(1)-*, *ø-*, *di-*. They can also form a VTR on the basis of a noun referring to a tool, e.g.

gargaji 'saw (n)'   *maŋgargaji* 'saw (O)'   *digargaji* 'be sawn'
pahat 'chisel (n)'   *mamahat* 'chisel (O)'   *dipahat* 'chiselled'
kuñci 'key'   *maŋuñci* 'lock (O)'   *dikuñci* 'locked'

maN(2)- also forms VDIs on the basis of VSIs and nouns, but the resulting forms sometimes differ in meaning from their SM equivalents.

Prefixed to VSIs *maN(2)-* conveys the meaning 'become as if, act as if (VSI)', e.g.

pintar 'clever'   *mamintar* 'act as if clever, try to be clever'
kuniŋ 'yellow'   *maŋuniŋ* 'turn yellow'
tuli 'deaf'   *manuli* 'play deaf, act as if deaf'

Prefixed to nouns it means 'use or work with (noun), trade in, make, look for (noun) as a means of living', or 'look like (noun), be oriented towards (noun)', e.g.

taŋguk 'landing net'   *manaŋguk* 'use a landing net'
banih 'rice'   *mambanih* 'trade in rice'
waday 'cake'   *mawaday* 'make cake'
iwak 'fish'   *maiwak* 'look for fish'
tampiray 'basket trap'   *manampiray* 'use a basket trap'
rumah 'house'   *marumah* 'like home, be a stay-at-home'
raja 'king'   *maraja* '(be) like a king'
tikus 'mouse'   *manikus* '(be) like a mouse'
hatap 'roof'   *mahatap* 'fit roofs (as a profession)'

6.3.4 SWY

SWY *(mə)N(1)-*, *ø-*, *di-* are equivalent to SM *məN(1)-*, *ø-*, *di-*. According to Helfrich, the clitic *-la(h)* as a rule[191] is added to verbs in the imperative mood, e.g.

---

[191] In SM *-lah* is also often postcliticised to verbs in the imperative mood, but it is not a marker as such of the imperative mood, as is suggested by Aliana et al. for SWY *-la(h)*.

ambiᵊʔ 'take (O)'	*ambiᵊʔla(h)!* 'take (O)!'
bərsiᵊ(h)ka 'clean (O)'	*bərsiᵊ(h)kala(h)!* 'clean (O)!'
xumputi 'weed (O)'	*xumputila(h)!* 'weed (O)!'

According to Aliana et al., the imperative form may also be preceded by *di-*, as in the following examples:

minum 'drink (O)'	*kupi itu diminumla(h)!* 'drink that coffee!'
bataʔ 'carry, take (O)'	*baraŋ itu dibataʔla(h)!* 'take those goods!'

(mə)N(2)- is equivalent to SM *məN(2)-*.

6.3.5 IBN

IBN *N-*, *ø-*, *di-* are equivalent to SM *məN(1)-*, *ø-*, *di-*. (*N-* is realised as homorganic nasal substitution, see 2.5.2). Some transitive verbs (especially those with initial *c*) can also have *əN-* (in free variation with *N-*) prefixed, e.g.

cabaw 'cut, mow'	*ñabaw, əñcabaw* 'cut, mow'
cabik 'tear to pieces'	*ñabik, əñcabik* 'tear to pieces'

əN- (nasal accretion, with an epenthetic *ə* preceding the resulting consonant cluster) is used to form VDIs (Asmah 1977:82). It forms VDIs on the basis of VDI bases, nouns and precategorials (according to the examples given by Asmah); e.g.

səput 'breath, life'	*ənsəput* 'breathe'
pəkap 'cackling (of hens)'	*əmpəkap* 'cackle'
-kalik, kalikkalik 'dangle or swing continuously'	*əŋkalik* 'dangle or swing (once)'

əN- is to some extent comparable with SM *məN(2)-* although apparently it does not occur on the basis of VSIs, and alternates with the transitive marker *N-* in a few cases (see above).

6.3.6 JKT

JKT *N(1)-*, *ø-*, *di-*, are equivalent to SM *məN(1)-*, *ø-*, *di-*. *N(1)-* has several variants: *məN(1)-*, *N(1)-*, and *ŋə(1)-*. *məN(1)-* is in most cases interchangeable with *N(1)-* and seems to be favoured in formal speech (Muhadjir 1981:46). *N(1)-* and *ŋə(1)-* are partly in complementary distribution and partly in free alternation: *ŋə(1)-* is a morphophonemic alternation of *N(1)-* before a liquid or a semivowel, or if the base is monosyllabic, and it is in free variation with it before voiced stops (cf. 2.6.2). *di-* is also favoured in imperative sentences, e.g.

(24)   *Tu   sayur   jaŋan   digadòin!*
      that   vegetable   don't   *di-* + eat.s.th.without.rice
      Don't eat those vegetables without rice!

(25)   *Diminum   tu   té-ñè!*
      di- + drink   that   tea-your (polite)
      Please drink your tea!

The distribution of *N(2)-* and its variants *məN(2)-* and *ŋə(2)-*) is equivalent to that of SM *məN(2)-* and its variants.

6.3.7 DISCUSSION OF SM *məN*(1)-, *ø-, di-,* AND *məN*(2)-, AND THEIR CORRESPONDENCES

In Malayic isolects, there occur correspondences of SM *məN*(1)-, or simply nasalisation of the beginning of a stem, or a combination of both. In all isolects but IBN, there are correspondences of SM *məN*(2)-, which are formally identical to correspondences of SM *məN*(1)-, and which all form intransitive verbs. Further, the morphophonemic alternations of SM *məN(1)-,* SM *məN*(2)-, and their correspondences, and of mere nasalisation, also differ from one isolect to the other. So there are three problems involved in the reconstruction of PM ancestors of SM *məN*(1)- etc. and SM *məN*(2)- etc.:

(1) did they have the form *mAN-* or **N-* ?
(2) what morphophonemic alternations did they exhibit?
(3) were there any formal or functional differences between them?

These three problems will be considered in turn.

(1) SM *məN*(1)- and MIN, BH *maN*(1)- reflect PM **mAN*(1)-; SWY has a correspondence *məN*(1)- as well as *N*(1)-, but the sources do not state explicitly what the distribution of these allomorphs is. Apparently, *məN*(1)- occurs before liquids and is favoured before nasals (where it alternates with *ø*, as in SM in certain cases). JKT has *N*(1)-, *ŋə*(1)-, and *məN*(1)-. *ŋə*(1)- may be due to SUN influence, where *ŋa-* as an alternant of *N-* occurs in similar circumstances (Ikranagara 1980:135). JKT *məN*(1)- either reflects an unproductive older affix now being replaced by *N*(1)-/*ŋə*(1)- or it is a loan morpheme from SM.

IBN is the only isolect that does not have a reflex of **mAN*(1)-; it is also the only isolect in which a different process for the nasalisation of VDIs and VTRs is involved (see below).

I presume that **mAN*(1)-, rather than only **N-* was the original proto-form from which emanated the contemporaneous reflexes. First, if one proceeded from PM **N*(1)-, there would be no way to account for the occurrence of *məN*(1)-/*maN*(1)- in the contemporaneous isolects. Moreover, formally an original **mAN*(1)- fits in better than **N*(1)- in a pattern **pAr-* : **(mb)Ar-*//**pAN-* : x (see 6.1.2 and 6.7). The reduction of **mAN*(1)- to **N*(1)- can be explained by the tendency to disyllabicity, and by the fact that no loss of functional load or danger of homonymy was involved in this reduction.

Correspondences of both *məN*(1)-/*maN*(1)- and *N*(1)- are found in many other Austronesian languages, although correspondences of *məN*(1)- / *maN*(1)- are restricted to languages of Indonesia, the Philippines and Madagascar. In languages from this area correspondences of *məN*(1)-/*maN*(1)- or *N*(1)- developed into a verbal morpheme. Dahl takes this as a subgrouping argument for a western Austronesian branch, against a northern (= Formosan) and an eastern (= Pacific) Austronesian branch. For the different correspondences he finds an original Proto Western-Austronesian **maN-* more likely than **N-* because it matches better with Proto Western-Austronesian **paN-*. Wolff (1973:72) claims that many PAN affixes are in their deep structure composed of combinations of derivational and inflectional affixes. In this way PAN **maR-* can be analysed as a combination of a derivational **paR-* with an (inflectional) **-um-*, thus **maR-* < **paR-* + **-um-* (with regular loss of first syllable if the initial consonant is a **b* or **p*); and **maN-* < **paN-* + **-um-* (the last example is not given explicitly by Wolff).

Wolff's assumptions are based on evidence from Philippine and ·Formosan languages, where reflexes of PAN **maR-* and **paR-* are still analysable as deep-structure combinations of (reflexes of) **paR-* and **paN-* with (reflexes of) **-um-*. The same situation is also found in other languages (cf. Tondano, Sneddon 1975:208ff.; Timugon Murut, Prentice

1971:128). If this view is sound, then an original (and complex) *maN- must be reconstructed for PAN, and hence also for Dahl's PWA.

With the present knowledge of PAN grammar it is not possible to take any view with certainty, but scholars agree that *maN- must be reconstructed for proto-languages which at least include Indonesian and Philippine languages. On the basis of the correspondences of SM məN(1)- in other Malayic isolects I reconstruct PM *mAN(1)-. For the exact function of this proto-phoneme, see (3) below.

(2) The following chart shows the various morphophonemic changes in the isolects of məN(1)-/maN(1)-/N(1)- (cf. Chapter 2 for a fuller treatment):

initial phoneme	SM	MIN	BH	SWY	IBN	JKT
p	məm	mam	mam	m	m	(mə)m
t	mən	man	man	n	n	(mə)n
c	məñc	mañc	mañc	ñc	ñ	(mə)ñc/ñ (2.6.2)
k	məŋ	maŋ	maŋ	_	_	(mə)ŋ
b	məmb	mamb	mamb	mb/m	m	(mə)mb/ŋəb
d	mənd	mand	mand	nd	n	(mə)nd/ŋəd
j	məñj	mañj	mañj	ñj	ñ	(mə)ñj/ŋəj
g	məŋg	maŋg	maŋg	ŋg	ŋ	(mə)ŋg/ŋəg
m	məm	mam	mam	(mə)m	m	(mə)m
n	mən	man	man	(mə)n	n	(mə)n
ñ	məñ	mañ	mañ	(mə)ñ	ñ	(mə)ñ
ŋ	məŋ	maŋ	maŋ	(mə)ŋ	ŋ	(mə)ŋ
l	məl	mal	mal	(mə)l	ŋəl	ŋəl
r	mər	mar	mar	(mə)r	ŋər	ŋər
s	məñ	mañ	mañ	ñ	ñ	ñ
h	məŋh	--	mah	--	--	--
V	məŋV	maV	maV	ŋV	ŋV	ŋV

N.B. (1) In Aliana et al.'s description of SWY all initial stops undergo nasal substitution;
(2) IBN ŋə- and JKT ŋə- are also prefixed to monosyllabic roots.

The morphophonemic changes of məN(2)- etc. are identical to those of məN(1)- etc. except for IBN, where əN- has the following alternations: əm- before initial p-, ən- before initial t- and s-, əñ- before initial c-, and əŋ- before initial k- (see 2.5.2 C).

From the above chart it appears that all isolects agree in showing homorganic nasal substitution for initial p, t, k, and palatal nasal substitution for s, and no change at all before initial nasal. Except for IBN, they all agree in showing homorganic nasal accretion before initial voiced stops (in SWY homorganic nasal accretion is still found in Helfrich's description, whereas voiced stops are replaced by homorganic nasals in Aliana et al.; cases with homorganic nasal replacement of voiced stops in Helfrich are incidental, 2.4.2). One could argue that nasalisation originally did not occur before liquids, and that it appeared in this position in IBN and JKT because of the need for a formal indication of agent-orientedness.

But it is not possible to reconstruct a sound set of morphophonemic alternations for *mAN(1)- apart from that of homorganic nasal accretion before voiced stops, homorganic nasal substitution for initial *p, *t, *k, palatal nasal substitution for initial *s, and loss of the *N in *mAN(1)- before an initial nasal.

As for *mAN(2)-, it apparently had the same morphophonemic alternations as *mAN(1)- : only IBN has a different set of alternations, but since the IBN distinction between nasal substitution and nasal accretion as an indication of transitivity and intransitivity respectively is not found elsewhere among Malay isolects, it will not play a part in the reconstruction of PM. The difference between PM *mAN(1)- and *mAN(2)-, then, was only a functional one.

(3) It seems that the function that must be attributed to *mAN(1)- is that of agent-oriented verb marker, since this is the function of its reflexes in the isolects. By the same token, the function of intransitive verb marker must be attributed to *mAN(2)-. As the PM ancestor of *di*- was not yet a prefix (as will be seen below), there is no reason to consider it as the object-oriented counterpart of PM *mAN(1)-. The only difference between *mAN(1)- and *mAN(2)- was that one was prefixed to VTRs and the other to VDIs. For this reason it is simpler to reconstruct only PM *mAN- as an active verb marker which was prefixed to both VDIs and VTRs, and to regard the functional differentiation as a later development.

However, it is relevant in this respect that in KD and SD, the function of N(1)- (corresponding to SM *məN*(1)- etc.) is not primarily that of marking agent-orientedness. N- also occurs in object-oriented verbs: N- conveys here the realis mood, that is, it indicates that an act is really taking place, or has really taken place. Conversely, the absence of N- in object-oriented verbs indicates that the act has not (yet) taken place, or will not take place (which includes negative sentences, imperatives and subjunctives). (N- is also absent in series of consecutive verbs which are used for describing a process.) In the examples below all sentences are in the realis mood except for (27), which expresses an intention. This irrealis-realis distinction may be an exclusive innovation in KD and SD, but it is also reminiscent of the PAN morphological distinction between independent and dependent verbal forms (cf. Wolff 1973).

SM *di*- and its correspondences is prefixed to object-oriented verbs when the agent is a third person. The correspondence of SM *di*- in the other isolects is also *di*-. There are, however, some important reasons not to reconstruct a prefix for this correspondence set:

(a) *di*- is the only prefix that did not undergo unexpected antepenultimate neutralisation in SM, SWY, IBN and JKT;

(b) in KD and SD *di*- is not only cliticised before the verb itself, but also before the agent if it precedes the verb. Compare the following examples (in KD and SD object-oriented verbs, as indicated above, nasalisation indicates realis mood):

KD (Dunselman 1949:70; I have modified Dunselman's spelling in the following way: ng : ŋ, ' : ʔ, nj : ñ; j : y, dj : j; tj : c; oe : u; é : e)

(26) *(Kamudaʔ) di-ña-nurun-an   ka-tanah.*
       child   by-him-go.down-causative to-ground
       The child was put down on the ground by him.

(27) *S-ekoʔ   jiba   di-ku-kurukŋ.*
       one-(classifier) soul by-me-cage
       One soul I hold captive.

(28) *di-ujatn najar*
       by-rain attacked
       be caught in the rain

SD (Ina Anak Kalom & Hudson 1970:287-288; I have modified Ina Anak Kalom & Hudson's spelling in the following way: ng : ŋ; ' : ʔ; ny : ñ; ch : c)

(29)   *Oraŋ Kaya Atoi di-parintah   naŋkap.*
       Oraŋ Kaya Atoi by-government catch
       Oraŋ Kaya Atoi was arrested by the government.

(30)   *Sobat   namu   saduaʔ piŋatn-ne dah   di-Selako*
       Chinese discover some   plate-their already by-Selako

       macah-atn-iʔ.
       N.broken-causative-diffuseness.marker
       The Chinese discovered that some of their plates had been broken by the Selako.

N.B. The locative preposition *di* is also used as an agent marker in MIN, e.g.

(31)   *Di kawan-ño di-ciloʔ   pitih.*
       by friend-his be-stolen money
       Money was stolen by his friend.

(32)   *Si   Amin di-paŋgiᵊ di tuan.*
       (personal article) Amin be-called by lord
       Amin is summoned by his master.

In SWY it may occur (instead of *liᵊ(h)*) as an agent marker (particularly in imperative constructions; Helfrich 1904:211).

(c) *di-* has only a limited spread outside the Malayic group: it occurs in Lampung and Batak isolects, in JV and SUN, and apparently also in languages of Sulawesi (Teeuw 1959:143). In JV *di-* is an innovation: Old Javanese does not have it (it has *-in-* instead), and if it were inherited in JV it would have undergone antepenultimate neutralisation.

(d) OM does not have a prefix *di-*: it has *ni-* instead (see N.B. below).

Summarising the above arguments, I conclude that the PM ancestor of *di-* was not yet a prefix. Two possible explanations present themselves. It could have been a reduced form of a (post-PM) pronoun *+d-ia* which was cliticised to (SM?) object-oriented verbs, then became a marker of object-orientedness, and finally was borrowed into other Malayic and non-Malayic isolects later on. This explanation is supported by the fact that *di-* was originally only compatible with a third person agent. It may also have developed from the (SM) locative preposition *di,* which then extended its function to that of a marker of object-orientedness, and gradually became cliticised to the following verb. In MIN, SWY, KD and SD, it also became an agent marker. The main support for this assumption would come from KD and SD. No reconstruction is made on the basis of *di-*.

N.B. OM has an object-oriented verb marker *ni-*. In contrast to the traditional use of *di-*, *ni-* is also compatible with agents of the first and second person. De Casparis (1956:24) believes that *ni-* is a cognate of *di-*, and that its initial consonant was denasalised in other Malayic isolects. Teeuw (1959:141-144), although he does not accept the evidence provided by De Casparis, basically agrees with him and favours the possibility of denasalisation in the Malayic contemporaneous isolects of the first consonant of the ancestral forms of both OM *ni-* and *mar-* (thus: PM **ni-* > OM *ni-* SM etc. *di-*; PM **mAr-* > OM *mar-*, SM etc. *bər-*, *ba-*, *bəx-*, *bə-*). Teeuw rejects Aichele's assumption that *di-* developed from a preposition and replaced an inherited *ni-*, and that *mar-* was borrowed from Batak isolects. He criticises

Aichele for taking SM too much as a point of reference in his study, and thereby for using Batak influence as an explanation for the occurrence of characteristics in OM which are not found in more recent forms of Malayic. I agree with Teeuw's last two criticisms. OM *mar-* seems a regular reflex of PMP *maR-*, and the fact that contemporaneous Malayic isolects have *bər-* etc. is no decisive reason to take *mar-* for a Batak loan. It is conceivable that OM had retained PMP *maR-* (and hence PM *mAr-*), while in other forms of Malayic *m* became *b* under the influence of the following *r*. As in a large majority of cases *mAr-* was prefixed to disyllabic lexemes and was never stressed, it may have been realised as a consonant cluster (as is also often the case with SM *bər-* etc.). This may have given rise to a *b,* possibly through an intermediate stage where +m was still realised but had acquired an epenthetic +b, that is, *mAr-* > +m(A)r- > +mbAr- > +bAr- > *bər-, ba-,* etc. The development of an epenthetic stop is common enough in the history of the Malayic isolects, cf. excrescent stops in IBN, and PM *tim(ə)rah* > MIN *timbarah*, AR *jumlah* > JKT *jumbəlè*. But I also agree with Aichele that *di-* could have originated from a former preposition. The KD and SD evidence and the use of *di* as an agent marker in MIN and SWY, added to the fact that *di-* is the only prefix in the contemporaneous Malayic isolects with a vowel other than *ə* (MIN, BH *a*), give some ground for this assumption. As to OM *ni-*, this may be an inherited prefix of which the cognates were lost in the contemporaneous Malayic isolects. I will reconstruct PM *(*ni-)*, an object-oriented verbal prefix which seems to reflect PMP *ni-*, which had an allomorph *-in-* and was a perfective aspect marker. As none of the other Malayic isolects have a reflex of this proto-affix, I write it between brackets.

6.3.8 RECONSTRUCTION OF PM *mAN-*

On the basis of the SM agent-oriented verb marker *məN(1)-*, the SM VDI marker *məN(2)-*, and the cognates of these prefixes in the other isolects, I reconstruct PM *mAN-*. *mAN-* was apparently an agent-oriented verb marker prefixed to VTRs and VDIs. It formed VDIs on the basis of precategorials, adjectives, and nouns. The KD and SD reflex of *mAN-* suggest that its use in some cases involved a modal (realis-irrealis) distinction as well (but this requires further investigation).

mAN- underwent morphophonemic alternations according to the initial phoneme of the verb to which it was affixed. The exact alternations are difficult to trace, but it is likely that the *-N-* of this prefix was realised in at least the following ways:

(a) homorganic nasal accretion before initial voiced stops;
(b) homorganic nasal substitution for initial *p, *t,* and *k*;
(c) palatal nasal substitution for initial *s*;
(d) ŋ before initial nasals and liquids.

6.4 EVIDENCE FROM OLD MALAY AND MALAYIC DAYAK

6.4.1 RECONSTRUCTION OF A PM SUBJUNCTIVE MARKER *-a?*

KD has a suffix *-a?* to which Dunselman (1949:61) attributed a 'voluntative' function. Compare examples (8) and (9) (Dunselman 1949:62-63):

(33) *Kade? kita? ñuman-a? jukut k-aku baik kita? suman ka-dapur lain.*
      if   you cook + -a? thing at-I   better you cook   at-kitchen other
      If you intend to cook that stuff here, you'd better go to another kitchen.

(34)   *Dah*   *habis hal-ña*   *ñian Ne⁷ Do⁷akŋ minta balajar-a⁷*
        already done event-its this Ne⁷ Do⁷akŋ ask for learn + *-a⁷*

    bajalatn-a⁷ laki ka apa⁷-ña.
    go + *-a⁷*   man to father-his
    After this had happened, Ne⁷ Do⁷akŋ asked his father for permission to learn to
    go his way as a man (= to go hunting).

This suffix also occurs in SD, compare examples (35) and (36) (Ina Anak Kalom & Hudson 1970:289, 293):

(35)   *Ia*   *mau muru-a⁷*   *Baruk ari hujatn darakŋ.*
        He (= *ŋambakŋ*) want chase-away + *-a⁷* Baruk sun rain   hot
        Ngambakŋ wanted to chase Baruk away but it was still raining (with sunshine).

(36)   *Jaji s-eko⁷ paŋalima kayo*   *aŋkoa ... mau tarajutn-a⁷ kadaapm kubu.*
    then one[192] warrior enemy that   want jump + *-a⁷* into   fort
    Then one enemy warrior...was about to jump down into the fort.

Finally, the suffix *-a* in OM expresses future or irreality (De Casparis 1956:344), as is seen in the following sentence from the Telaga Batu inscription in South Sumatra (De Casparis 1956:33 line 10):

(37)   ... *athavā kadāci kāmu māti* malūn *mamrurū-a:* *athavā kāmu*
        or   if   you   dead not.yet succeed + *-a* or   you

    larī-ya   *mamlarīya*   *lai*   *kāmu nivunuḥ kāmu sumpaḥ.*
    run.away + *-(y)a* let.flee + *-(y)a* other(s) you   be.killed you   curse
    ...or if you die before having succeeded (in destroying my palace) or flee or help
    others to flee, you will be killed by the curse.

Wolff (1973:90) reconstructs a PAN subjunctive suffix **-a* on the basis of evidence from Atayal and JV: Atayal *-a* expresses subjunctivity, and JV *-a* expresses the subjunctive and imperative mood. In these languages subjunctivity 'ranges in meaning from the optative and hortatory...to the concessive and even resultative'.[193] So there are two witnesses within the Malayic group for the reconstruction of a PM subjunctive suffix, and this reconstruction is supported by evidence from outside the Malayic group. (KD and SD are too closely related to be considered as separate witnesses.) I reconstruct PM **-a⁷*, a subjunctive-marking suffix.

---

[192]*S-eko⁷* ('s-eko' in Ina Anak Kalom & Hudson) derives from *eko⁷*, the SD word for 'tail'. A numeral classifier for counting people on the basis of a word for 'tail' seems unusual, but it is also found in KD (cf. Dunselman) and in other Bornean languages. It probably represents a generalisation of an earlier classifier for animals to all living beings.

[193]Reflexes of PAN **-a* with divergent meanings occur in Malagasy and Javanese and in languages of Formosa, the Philippines, Sulawesi, and Borneo. Wolff reconstructs its subjunctive meaning on the basis of the semantically agreeing JV and Atayal reflexes (both being subjunctive markers). He reconstructs PAN **-a* (> JV *-ɔ*, Atayal *-a*) and PAN **-ay* (> JV *-(n)é*, Atayal *-ay*) as subjunctive markers which are active and local passive respectively. Reflexes of PAN **-a* and **-ay* with divergent meaning are for instance Timugon Murut *-o⁷* (< PAN **-a*) and *-i⁷* (< PAN **-ay*) being suffixed to respectively object-oriented and referent-oriented verbs in atemporal aspect (which is used with imperatives, narrative mode, and serial verbs; the referent-focus includes beneficiary- and location-orientedness in Timugon Murut; Prentice 1971:218).

6.4.2 RECONSTRUCTION OF A TRANSITIVE MARKER *(*maka-)*

KD and SD have a prefix *maka-* which forms transitive verbs, both with a causative meaning. It is found in only two examples:

KD *maka-lalu molot* (SD *maka-lalu moot*) 'keep one's promise, act according to what one has said';
SD *maka-rehetn* 'make light (particularly a punishment)', cf. *rehetn* 'light'.

A corresponding prefix *maka-* with a usually (but not always clearly) causative meaning is found in Old Malay, cf.

maka-ləŋit 'make disappear'[194]
maka-gīla 'make crazy' (De Casparis 1956:39 fn.24)

A corresponding prefix with different meaning is found in other Austronesian languages, cf. Philippine languages *maka-*, Malagasy *maha-* 'able to do [base]', and Old Javanese *maka-* 'consider/have/use as [base]' (Zoetmulder 1983). I reconstruct PM **maka-*, a transitive marker which occurred on the basis of VSIs and VDIs. The causative notion of *maka-* + VSI derivations seems to be a secondary effect of the transitivisation of VSI bases.

6.5 SM *-an*(1), SM *-an*(2), AND THEIR CORRESPONDENCES

The functions of *-an* are rather diverse, and, on the basis of what is generally found in the isolects, I prefer to distinguish two suffixes *-an*:

-an(1): a (nominal and verbal) suffix denoting (with verbs) diffuse action, plurality of subject, reciprocity, and (with nouns) collectivity and similarity;
-an(2): a noun-forming suffix occurring with VSIs, VDIs, and VTRs.

6.5.1 SM

Verbs with *-an*(1) always have *bər-* prefixed. With VDIs, *bər- -an*(1) denotes diffuse action, or plurality of object. The base is sometimes doubled in order to put more emphasis on the notion of diffuse action or plurality, e.g.

lari 'run'	*bərlari(lari)an* 'run (of many people, or in different directions)'
hañut 'float'	*bərhañutan* 'float (of many things); float around'
hambur 'scatter'	*bərhamburan* 'scattered all around'
suka 'be cheerful'	*bərsuka(suka)an* 'be cheerful together, celebrate'
bərgantuŋ 'hang'	*bərgantuŋan* 'hang (of many things)'

Affixed to (derived or underived) VTRs, *bər- -an*(1) forms reciprocal verbs. The notion of reciprocity may already be conveyed by *bər-* only, in which case competing forms may occur (cf. *bərkirim* or *bərkiriman* 'correspond with each other', *bərtembak* or *bərtembakan* 'shoot at each other'). Doubling of the base is used to put more emphasis on the reciprocal action. *-i* and *-kan* are deleted when *bər- -an*(1) is suffixed, e.g.

---

[194]De Casparis (1956:347) translates this derivation as 'making impotent' on account of JKT *ləŋit* 'indolent, lazy (with the underlying implication of making money by sly means or without working)', but I prefer 'make disappear' on account of Old Javanese *ləŋit* 'vaguely visible in the distance, vanishing from sight' and SD *aŋit*, KD *ləŋit*, SUN *ləŋit* 'disappeared', and I reconstruct PM **ləŋit* 'disappeared, vanished' on the basis of this evidence.

bunuh 'kill (O)'	*bərbunuh(bunuh)an* 'kill each other'
sahut 'answer (O)'	*bərsahut(sahut)an* 'respond to each other'
suapi 'feed (O)'	*bərsuap(suap)an* 'feed one another'
panahi 'shoot arrows at (O)'	*bərpanah(panah)an* 'shoot arrows at each other'

On the basis of a noun (often reduplicated), *-an*(1) adds the notion of collectivity or similarity, e.g.

darat 'shore, land'	*daratan* 'mainland'
laut 'sea'	*lautan* 'ocean'
kayu 'tree; wood'	*kayukayuan* 'trees (collective)'
daun 'leaf'	*daundaunan* 'foliage'
rambut 'hair of the head'	*rambutan* 'k.o. fruit with a hairy skin'
oraŋ 'human being'	*oraŋoraŋan* 'statue; puppet; pupil (eye)'
jambaŋ 'vase'	*jambaŋan* 'flowerpot, flower stand'
anak 'child'	*anakan* 1. 'puppet '
	2. 'interest (e.g. on loan)'

-an(2) occurs with VSIs, VDIs, and VTRs. Suffixed to VSIs it forms nouns with the meaning 'something that has the quality of (VSI)', e.g.

manis 'nice, sweet'	*manisan* 'sweetmeats'
luar 'outside, out'	*luaran*, in *oraŋ luaran* 'foreigner'
kuniŋ 'yellow'	*kuniŋan* 'brass'

Suffixed to VDIs (and precategorials) it forms verbal abstracts, that is, nouns referring to the performance as such denoted by the underlying form, e.g.

-taŋis '(weep)'	*taŋisan* 'crying, weeping'
-ñañi '(sing)'	*ñañian* 'singing, song'
roboh 'fall, crash'	*robohan* 'collapse (n)'
bərtəkan 'lean, press'	*təkanan* 'pressure, suspense, stress, emphasis'

From VTRs it forms nouns referring to (1) the object of an act, (2) the place where the act is performed, (3) the instrument used to perform the act. *-i* and *-kan* are deleted when *-an*(1) is suffixed, e.g.

kirim 'send'	*kiriman* 'parcel, present'
kənal 'know (a person)'	*kənalan* 'acquaintance'
aŋkat 'lift, raise'	*aŋkatan* 's.th. raised, e.g. generation, troops of an army'
pukul 'hit, strike'	*pukulan* 'blow, strike'
kumpulkan 'collect (O)'	*kumpulan* 'collection'
timbaŋ 'weigh'	*timbaŋan* 'balance, weighing machine'
sumbaŋ 'contribute'	*sumbaŋan* 'contribution'
sindir 'mock'	*sindiran* 'mockery, satirical poem'
suruk 'hide, conceal, by crouching or drawing back'	*surukan* 'hiding place'
gantuŋ 'hang'	*gantuŋan* 'hanger'
	anak -- 'gallows bird'

6.5.2 MIN

MIN -*an*(1) on the basis of verbs co-occurs with *ba-*, and has the same functions as SM *bər- -an*(1), e.g.

(38) *Ba-a kalian   ba-caliᵊʔ-an   juo?*
       why you (plural) *ba- -an*(1) + look only
       Why are you all looking (like that)?, Why do you keep looking?, or
       Why are you looking at each other?

(39) *Uraŋ   tu   lah   lamo   baintayan.*
       person that already long (time) spy
       Those people have been spying on each other for a long time.

(40) *Baru tiŋaran musuᵊh ka dataŋ, lah   baintayan*
       just audible enemy   will come, already *ba- -an*(1) + spy

       yo   kadalam sama?   nan-tun.
       3PL into   bushes those
       As soon as they heard the enemy coming, they went spying in the bushes.

-*an*(1) + a noun refers to 'a place where (noun) is found in great store', or 'a place which is entirely occupied by (noun)', (with reduplication) to 'a diversity of (noun)', or to 'something resembling (noun)', e.g.

tupay 'squirrel'	*tupayan* 'place full of squirrels'
uraŋ 'human being'	*uraŋan* 's.o. who receives many visitors'
buŋo 'flower'	*buŋobuŋoan* 'diversity of flowers'
duri 'thorn	*durian* 'the durian (i.e. a fruit with a thorny skin)'

-*an*(2) is suffixed to VSIs, VDIs, and VTRs. The resulting forms have the following meanings:

With VSIs they refer to 'something that has the quality of (VSI)'. They have the same meaning as corresponding forms in SM, although Van der Toorn (1899:4-5) describes them as "the place where (base) is found", e.g.

manih 'sweet'	*manisan* 's.th. sweet, sweetmeat, sweetness'
dalam 'deep'	*darah dalaman* 'coagulated blood under the skin, blood blister'

With VDIs they refer to the place where the action is performed, e.g.

tiŋgiə 'perch (of birds)'	*tiŋgiᵊran* 'place where birds perch'
lumpeʔ 'jump'	*lumpeʔan* 'place where one jumps'
basanda 'lean'	*sandaran* 'place to lean, support'

With VTRs they refer to the result, goal, or instrument of the action, or to the place where the action is performed, e.g.

tambah 'add'	*tambahan* 'addition, appendix'
saso? 'drink (impolite)'	*saso?an* 'drinking place for animals'
rundiᵊŋ 'talk, discuss'	*rundiᵊŋan* 'conversation, matter discussed'
kukuᵊ 'rasp (v)'	*kukuᵊran* 'rasp (n)'

6.5.3 BH

BH -*an*(1) suffixed to verbs has the same functions as SM *bǝr-* -*an*(1): with VDIs it denotes plurality of subject or diffuse action, and with VTRs[195] (in combination with *ba-*) plurality of the subject or reciprocity. Transitive suffixes (-*i* and -*akan*) are elided, e.g.

dataŋ 'come'	*datanan* 'come (of many people)'
bukah 'run'	*bukahan* 'run (of many people); run in all directions'
badiam 'be quiet'	*badiaman* 'all be quiet'
sariki 'be angry at (O)'	*basarikan* 'get angry with each other, all get angry'
jual 'sell'	*bajualan* 'sell to each other'
iñjam 'borrow'	*baiñjaman* 'borrow from each other'
babulik 'go home'	*babulikan* 'all go home' (Asfandi); 'come back to each
buliki 'come to, visit (O)'	other (of a divorced couple)' (Abdul Jebar)
baradiu 'have a radio'	*baradiuan* '(all) having a radio'

Sometimes there is no difference in meaning between forms with -*an*(1) and those without (although they are listed separately in Asfandi, see p.52 and p.56). This is the case with some forms with a noun as their primary base, e.g.

baju 'shirt'	*babaju, babajuan* 'wear a shirt'
darah 'blood'	*badarah, badarahan* 'bleed'

There are also -*an*(1) forms that have *siŋ-* cliticised. These forms have an intensive meaning, and are usually preceded by *kada* 'not', e.g.

dataŋ 'come'	*kada siŋdatanan* 'just not coming'
guriŋ 'sleep'	*kada siŋgurinan* 'not being able to sleep'
baduit 'have money'	*kada siŋduitan* 'be without a penny'
badarah(an) 'bleed'	*kada siŋdarahan* 'not bleeding at all'
habaŋ 'red'	*siŋhabanan* 'very red'
parak 'close by'	*siŋparakan* 'very close'

Suffixed to a noun, -*an*(1) forms a noun denoting 'something resembling (noun)', 'a collectivity of (noun)' or 'an area where a collectivity of (noun) is found', e.g.

hutan 'forest'	*hutanan* 'forested area, jungle'
gunuŋ 'mountain'	*gununan* 'mountainous area; s.th. resembling a mountain'
kampuŋ 'village'	*kampunan* 'area where people live'
kayu 'wood, tree'	*kayuan* 1. 'tree'
	2. 'many kinds of wood'

-*an*(2) is suffixed to VTRs and to some VSIs denoting a colour: with a VTR,[196] -*an*(2) forms a noun referring to the place where the act is performed, or to the goal or result of the act, e.g.

ulah 'do'	*ulahan* 'product'
lipat 'fold (v)'	*lipatan* 'fold (n)'

---

[195]In Asfandi -*an*(1) is described as a suffix occurring with VDIs and denoting plurality of actor, but from the examples as presented here (some of which are taken from Abdul Jebar) it appears that -*an*(1) has a wider application.

[196]Asfandi does not specify whether -*an*(2) is suffixed only to VTRs or also to VDIs.

tabuk 'dig'	*tabukan* 1. 's.th. that is dug up'
	2. 'pit, canal'
antas 'take a short cut'	*antasan* 'short cut; canal to short cut, meanders'

Asfandi gives three examples of *-an*(2) on the basis of a VSI denoting a colour; the resulting forms are nouns referring to entities which somehow agree in colour with (VSI), e.g.

kuniŋ 'yellow'	*kuniŋan* 'brass'
habaŋ 'red'	*habaŋan* 'menstruation'
hiraŋ 'black'	*hiraŋan* 'k.o. black monkey'

6.5.4 SWY

From Aliana et al. it appears that *-an*(1) in combination with *bə-* is equivalent to SM *bər- -an*(1). But *bə- -an*(1) also occurs on the basis of VSIs, and then means 'up to, until becoming (VSI)', e.g.

rusaʔ 'broken'	*bərusaʔan* 'until broken'
xəmas 'broken'	*bəxəmasan* 'until broken'

With nouns *-an*(1) forms nouns or VSIs. The resulting nouns refer to a variety of (noun); the first syllable of these nouns is reduplicated (with antepenultimate neutralisation of the reduplicated syllable), e.g.[197]

bua(h) 'a fruit'	*bəbua(h)an* '(all kinds of) fruit'
buŋo 'flower'	*bəbuŋoan* '(all kinds of) flowers'
buni 'noise'	*bəbunian* '(all kinds) of noises'

The resulting VSIs have the meaning 'to be affected by (noun)', e.g.

dəbu 'dust'	*dəbuan* 'dusty'
duxi 'thorn'	*duxian* 'thorny'
daxa(h) 'blood'	*daxa(h)an* [daxa:n] 'bloodstained'

-an(2) is suffixed to VSIs and VTRs: suffixed to VSIs it forms nouns with the notion of 'having the quality of (VSI)', e.g.

kuniŋ 'yellow'	*kuniŋan* 'brass'
manis 'sweet'	*manisan* 'sweets, sweetness' (Aliana et al.)
cabiᵊʔ 'torn'	*cabiᵊʔan* 'second-hand textile' (Aliana et al.)

Suffixed to VTRs it forms nouns referring to the result or goal of an act, or to the place where the act is performed; when *-an*(2) is suffixed, the transitive suffixes *-i* and *-ka, -kan* are deleted, e.g.

kixim 'send (O)'	*kiximan* 'present'
bəli 'buy (O)'	*bəlian* 'purchase (n)'
puliᵊ(h) 'obtain (O)'	*puliᵊ(h)an* 'profit, revenue, acquisition'
puput 'blow away (O)'	*puputan* 'bellows; forge, smithy'
pipis 'grind (O) between two stones'	*pipisan* 'stone on which is ground'

---

[197]No examples with an initial phoneme other than *b* were found in Aliana et al.

gantuŋka 'hang (O)'   *gantuŋan* 'gibbet, gallows'
kukux 'scratch, scrape off (O)'   *kukuxan* 'grater'

6.5.5 IBN

IBN no longer has a living suffix *-an*. However, a fossilised *-/an* still occurs, which appears to have been suffixed to verbs and nouns. The examples I was able to find all correspond to SM *-an(2)* etc. except for *gaam, gəm/an*, and *saraŋ, səraŋ/an*. Compare:

ləmah 'weak, soft (used in songs)'   *ləma/an* 'mishap, trouble'
rəbah 'fall, be knocked down'   *rəba/an* 'felled timber not yet fired'
əŋkan/iʔ 'feed (animals)'   *əŋkan/an* 'rice for eating'
main 'do, act; game, sport, play'   *pin/an* 'a pet '(cf. 3.1.3.2 N.B.)
gaam, in ŋəliʔ -- 'back teeth'   *gəm/an, in ŋəliʔ* -- 'back teeth'
kuruŋ 'enclose, shut up'   *kəruŋ/an* 'cage for chickens, coop'
saraŋ 'container'   *səraŋ/an* 'sheath, nest'
tugal 'dibble, sow'   *təgal/an* 'farmland after burning and before sowing'

6.5.6 JKT

In JKT the suffix *-an* has many applications. It is possible to distinguish a verbal *-an(1)*, with a far wider application than *-an(1)* in the other isolects, and a noun-forming *-an(2)*, which is comparable with *-an(2)* in the other isolects.

-an(1) occurs on the basis of VSIs, VDIs, VTRs and nouns. On the basis of VSIs it denotes a comparative degree, e.g.

gədé 'big'   *gədéan* 'bigger'
pintər 'smart'   *pintəran* 'smarter'
pəlit 'stingy'   *pəlitan* 'stingier'

According to Ikranagara (1980:137, 141) this *-an(1)* - construction (which has no parallel in the other isolects) is probably due to influence from Balinese or SUN. It is also often used in imperative sentences, e.g. *rajinan dòŋ!* 'work harder!' (*rajin* 'industrious', *dòŋ* '(emphatic particle)').

-an(1) with VSIs also denotes a reciprocal act, e.g.

baèk 'good'   *baèkan* 'be good to each other'
marè 'angry'   *marèan* 'be angry with each other'

Suffixed to VDIs it adds an aspect of stativity and durativity, e.g.

rəbè 'lie down'   *rəbaan* 'lying down (continuously)'
tidur 'sleep'   *tiduran* 'sleeping continuously, lying around'
cəmburu 'be jealous'   *cəmburuan* 'having a jealous character'

On the basis of VDIs which can have a complement introduced by the preposition *amè,* it adds an element of reciprocity to the meaning (Muhadjir calls these VDIs semi-transitives, cf. 6), e.g.

dəmən (amè) 'like, love'	*dəmənan (amè)* 'be in love with each other'
kənal (amè) 'be acquainted with, know'	*kənalan (amè)* 'get to know each other'

Suffixed to VTRs it adds an aspect of stativity and continuity, e.g.

jual 'sell (O)'	*jualan* 'sell (O) (habitually)'
pəgaŋ 'hold (O)'	*pəgaŋan* 'hold on to (O)'

Suffixed to nouns it forms VDIs with the meaning 'produce, grow (noun)', or, if the noun denotes a physical condition, 'suffer from (noun)', e.g.

jèŋgòt 'beard'	*jèŋgòtan* 'have or grow a beard'
iŋus 'snot'	*iŋusan* 'have a runny nose'
panò[198] 'skin disease'	*panò*ʔ*an* 'suffer from a skin disease'
kòrèŋ 'sores'	*kòrèŋan* 'have sores'

It also forms adverbs with the meaning 'more to the (noun)' if the noun refers to a place or direction, e.g.

dəpan 'front'	*dəpanan* 'more to the front'
piŋgir 'edge'	*piŋgiran* 'more to the edge'

Finally, on the basis of nouns it forms nouns with the notion of collectivity, e.g.

sayur 'vegetable'	*sayuran* 'various vegetables'
pərabòt 'tool'	*pərabòtan* 'equipment'

N.B. In a few cases *-an*(1) occurs in combination with *bə(r)-*. It does so with VDIs, VSIs, and precategorials, and it yields reciprocal verbs with a stative meaning. *bə(r)- -an*(1) forms are often in free variation with *bə(r)-* forms and *-an*(1) forms, e.g.

bəjubəl 'crowd (v)'	*bəjubəlan* 'id.' (Muhadjir 1981:49-50)
bərgaul 'associate (v)'	*bərgaulan* 'id.' (Muhadjir 1981:49-50)
kənal 'know, be acquainted with'	*bəkənalan, kənalan* 'get to know each other'
dəkət 'close'	*bədəkətan* 'be near (to each other)'
laèn 'other'	*bəlaènan* 'be different (from each other)'

(One of Muhadjir's examples is actually derived from a noun: *(bə)pacaran,* 'be in love with each other' is derived from *pacar* 'girl- or boy-friend'.)

-an(2) is suffixed to VSIs, VDIs, and VTRs. Suffixed to VSIs it yields nouns referring to objects which somehow have the quality of (VSI). These nouns are few in number and have a rather lexicalised meaning, e.g.

manis 'sweet'	*manisan* 'sweetened fruit candy'
asin 'sour, salty'	*asinan* 'k.o. food made of vegetables with peanut sauce'

---

[198]Abdul Chaer gives *panòʔ* whereas Muhadjir gives *pano,* without glottal stop (Muhadjir does not distinguish *ò* and *ó* cf. 2.6.1).

Suffixed to VDIs it yields abstract nouns or nouns referring to the place where the act is performed. When the resulting noun is an abstract noun, it is always followed by a nominal or a nominal phrase in the agentive role (Muhadjir 1981:60), e.g.

tərèak 'yell'	*tərèakan (lu)* '(your) yelling'
bañòl 'make jokes'	*bañòlan(-ñè)* '(her/his) joking'
pèŋkòl 'turn'	*pèŋkòlan* 'intersection'

On the basis of VTRs it forms abstract nouns or nouns referring to the goal, result, place, or instrument of the action. Here too, if the resulting form is an abstract noun, it is followed by a nominal or a nominal phrase, e.g.

dòròŋ 'push (O)'	*dòròŋan (lu)* '(your) pushing'
jait 'sew (O)'	*jaitan(-ñè)* '(his,her) (way of) sewing'
tanəm 'plant (O)'	*tanəman* 'plant (n)'
minum 'drink (O)'	*minuman* 'drink (n)'
pəgaŋ 'hold (O)'	*pəgaŋan* 'handle (n)'
gòsòk 'rub, iron (O)'	*gòsòkan* '(clothes-)iron'

6.5.7 DISCUSSION OF SM -*an*(1), SM -*an*(2), AND THEIR CORRESPONDENCES

SM, MIN, BH, and SWY, have a suffix -*an*(1) denoting plurality of subject, or diffuse action when suffixed to a VDI, reciprocity when suffixed to a VTR, and collectivity and/or similarity when suffixed to a noun. In MIN, -*an*(1) may also denote plurality of subject or diffuse action when suffixed to a VTR; moreover, in MIN and BH it may convey the combined meaning of location and collectivity when suffixed to a noun. When affixed to verbs, -*an*(1) often (and in SM and MIN always) co-occurs with *bər-, ba-, bə-*. This SM, MIN, BH, SWY suffix -*an*(1) is only to a certain extent comparable to JKT -*an*(1): the latter can also denote reciprocity on the basis of VSIs and VDIs ('semi-transitives', see 6), and it denotes stativity and durativity with (other) VDIs and with VTRs. Finally, it may form a comparative degree when suffixed to VSIs. JKT -*an*(1) in some ways agrees more with -*an* in neighbouring non-Malayic languages, as has already been pointed out by Ikranagara for -*an*(1) forming a comparative degree on the basis of VSIs. But in other ways too JKT -*an*(1) seems to agree more with non-Malayic languages: JV and SUN have VDIs consisting of a noun + -*an* with the meaning 'wear, have (noun)' (cf. JKT *jèŋgòtan*, etc.), and JV has VDIs consisting of a noun + -*ən* meaning 'suffer from (noun)' (cf. JKT *iŋusan, panò'an, kòrèŋan*, although one would expect an ending -*ən* instead of -*an*(1)). Like JKT, JV also has VDIs denoting stativity or durativity (cf. JKT *rəbaan, tiduran, cəmburuan*, and *jualan, pəgaŋan*) or reciprocity (cf. JKT *dəmənan, kənalan*) which are formed on the basis of a VDI or VTR + -*an*. As yet it is uncertain whether JV, SUN and/or Balinese had the most influence on the JKT morphology in this respect, but it seems very likely that the differences between SM, MIN, BH, SWY -*an*(1) on the one hand, and JKT -*an*(1) on the other, are due to influence from non-Malayic languages on JKT. In IBN a living suffix -*an* no longer occurs; as far as IBN has fossilised forms with /-*an*, these are evidence for an ancestral form of SM -*an*(2) etc., but do not yield decisive evidence for an ancestral form of SM -*an*(1) (except maybe for IBN *saraŋ* 'container' vs *səraŋ/an* 'sheath, nest', and *ŋəli gaam* vs *ŋəli gəm/an* 'back teeth').

-*an*(2) occurs in all isolects; in IBN it only occurs in fossilised forms. Suffixed to VSIs it agrees in all isolects in denoting 'something that has the quality of (VSI)'. Suffixed to VTRs it agrees in all isolects in denoting goal or result of an action, or place where the action takes

place; in SM it may also denote the instrument, and in JKT it may denote the instrument, or form an abstract noun.

On the basis of what is commonly found in SM, MIN, BH, and SWY, it is possible to reconstruct for -an(1) a PM ancestor which denoted plurality of subject, or diffuse action on the basis of VDIs, reciprocity when suffixed to VTRs, and collectivity and/or similarity when suffixed to nouns. One could argue that this ancestor did not have the meaning of reciprocity, because in the above four isolects -an(1) usually co-occurs with bər- etc. when forming reciprocal forms, and reciprocity is then already indicated by the prefix. (If this line of reasoning holds, then -an(1) probably has the same function whether suffixed to VDIs or VTRs, since plurality of subject is a consequence of reciprocity.) On the basis of -an(2) in all isolects, a PM ancestral form can be reconstructed which was suffixed to VSIs and VTRs. Suffixed to VSIs it denoted something with the quality of (VSI); suffixed to VTRs it referred to the goal, result or location of the act. Such an ancestral form, however, would correspond to two PMP suffixes: *-an (referring to the place where an act is performed), and *-en (referring to the object of an act), which would have yielded JKT +-an and +-ən. But JKT only has one nominalising suffix -an. There are four possible solutions for this apparent inconsistency: (1) JKT does not distinguish PMP *a and *e in final syllables, (2) PM merged PMP *-an and *-en to *-an, (3) JKT has merged PM *-an and *-ən to -an, or (4) JKT -an(2) is innovative. As far as solution (1) is concerned, it was seen in Chapter 3 that the JKT distinction of a and ə in final syllables must be inherited. Solutions (2) and (3) are also unlikely, because if the distinction between JKT final-syllable a and ə is inherited (and therefore reflects PM *a and *ə), there is no a priori reason why PMP *-an and *-en would have merged in PM, or in JKT. Solution (4) yields more perspectives, because it fits in with the fact that JKT -an(1) also seems to be borrowed. Furthermore, JKT has only two suffixes (-in and -an(1,2)) of which one, -in, is not inherited in any case and has replaced an earlier *-i (cf. 6.1.1). It is quite possible that at an earlier stage JKT lost all PM suffixes, and that JKT -an(1,2) is also innovative. But then again, if JKT -an(2) (forming nouns denoting goal or result on the basis of VTRs) is a loan morpheme, this time its source must be a Malayic isolect on account of its vowel (and not Balinese, as with -in, or JV, SUN, or Balinese, as with other forms of -an).

As it is, it is not possible to make definite statements about the history of JKT -an(1,2) and its reflex (or reflexes) in PM. I will maintain the (artificial) distinction I made between -an(1) and -an(2) for PM, and reconstruct separate proto-forms on the basis of them.

6.5.8 RECONSTRUCTION OF PM *-an(1), *-an(2), AND *-An

On the basis of -an(1) in the isolects I reconstruct a PM ancestral form *an(1), which denoted diffuse action or plurality with VDIs, reciprocality with VTRs, and collectivity, place where a collectivity is found, and/or similarity with nouns. It is possible that the notion of reciprocity on the basis of VTRs was already included in *(mb)Ar-, with which it must have co-occurred. In that case, this suffix probably only added the notion of diffuse action or plurality of subject (as with VDIs). On the basis of -an(2) in the isolects I reconstruct two different PM ancestral forms. As all isolects agree in having nouns on the basis of a VTR + -an referring to the goal or result of a VTR, and as JKT as a rule distinguishes between PMP *a and *e in final syllables, and JKT -an(1,2) may be innovative, I reconstruct for this agreement *-An. This was a noun-forming suffix occurring on the basis of VTRs and denoting the goal or result of an act. (The vowel of this suffix is

174

*A since it is assumed to be a continuation of synonymous PMP -*en). On the basis of the other functions of -an(2) I reconstruct PM *-an(2), a suffix which formed nouns denoting the quality of (VSI) on the basis of VSIs, and denoting the place where (VTR) was performed on the basis of VTRs.

6.6 SM kə- -an(1), SM kə- -an(2), AND THEIR CORRESPONDENCES

6.6.1 SM kə- -an(1) AND ITS CORRESPONDENCES

6.6.1.1 SM

SM *kə- -an*(1) is a circumfix denoting an unintentional action or state, or a potential action. As such it is equivalent to *tər-* + *a* VDI or VTR, but it occurs less frequently, and it is also circumfixed to nouns and VSIs; in the latter case it has the meaning 'unintentionally being affected by (noun) or by the quality of (VSI)'. It occurs most frequently on the basis of *dəŋar* 'hear' and *lihat* 'see' viz. *kədəŋaran* 'be heard, audible; sound (v)' and *kəlihatan* 'be seen, visible; seem, appear'. (These forms are largely synonymous with *tərdəŋar* and *tərlihat*, although the latter do not share the notions 'sound (v)' and 'seem, appear' respectively). Other examples:

dataŋ 'come'	*kədataŋan* 1. 'accessible'
	2. 'be visited, attacked'
hilaŋ 'disappear; lost'	*kəhilaŋan* 'lose'
tahu 'know'	*kətahuan* 'be found out, discovered'
habis 'be finished'	*kəhabisan* 'run out of'
masuk 'go in'	*kəmasukan* 1. 'penetrable'
	2. 'be entered, broken into'
hujan 'rain'	*kəhujanan* 'caught by the rain'
raja 'king'	*kərajaan* 'made, proclaimed a king'
tulaŋ 'bone'	*kətulaŋan* 'swallow a bone'
sakit 'ill, sick; suffering'	*kəsakitan* 'in pain; agonising; tormented, worried'
lupa 'forget'	*kəlupaan* 'forgotten'

N.B. On the basis of (often doubled) VSI- and noun-bases *kə- -an*(1) sometimes also forms VSIs denoting 'behaving like or resembling (VSI/noun)', e.g.

perak 'silver'	*kəperak(perak)an* 'silverish'
ibu 'mother'	*kəibu(ibu)an* 'motherly'
kuniŋ 'yellow'	*kəkuniŋan* 'yellowish'

6.6.1.2 MIN

MIN *ka- -an*(1) is equivalent to SM *kə- -an*(1), but differs from it in one respect: circumfixed to VSIs it also denotes a comparative degree, e.g.

(41) *Lai sa-heto kapanjaŋan tuŋkeʔ ko pado tuŋkeʔ tu.*
more a-cubit *ka- -an* + long stick this than stick that
This stick is a cubit longer than that one.

(42)　*Kabareʔan*　　　*jawi naŋko sapuluᵊh kati.*
　　　　ka- -an + heavy　cow　this　ten　　kati
　　　　This cow is ten katis heavier (1 kati is 617 grams).

6.6.1.3 BH

BH *ka- -an*(1) occurs with VSIs and nouns. On the basis of VSIs it denotes an excessive degree, e.g.

lamak 'fat'　　　　　　　*kalamakan* 'too fat'
handap 'short'　　　　　　*kahandapan* 'too short'
habaŋ 'red'　　　　　　　*kahabaŋan* 'too red'

Circumfixed to nouns (of which the first syllable is reduplicated) it conveys the meaning 'to suffer from the effects of (noun)', e.g.

samut 'ant'　　　　　　　　*kasasamutan* 'have pins-and-needles'
sitan 'devil'　　　　　　　　*kasisitanan* 'be pestered by, suffer from, a devil'

6.6.1.4 SWY

SWY *kə- -an*(1) is equivalent to SM *kə- -an*(1).

6.6.1.5 IBN

No affix corresponding to SM *kə- -an*(1) occurs.

6.6.1.6 JKT

Apart from the prefix *tə-* (treated in 6.2.6), JKT has two other affixes expressing unintentionality or feasibility, viz. *kə-* and *kə- -an*(1). The latter is usually in free variation with the former, but is the obligatory variant if the base is a VTR consisting of a VDI + the transitivising suffix *-in*, e.g.

campur 'mix'　　　　　　　*kəcampur, kəcampuran* 'be mixed, involved'
dəŋər 'hear'　　　　　　　*kədəŋəran* 'be heard, audible'
dudukin 'sit upon'　　　　　*kədudukan* 'sat upon (unintentionally)'
jatòin 'fall on'　　　　　　*kəjatòan* 'be struck by s.th. falling'
alaŋin 'prevent'　　　　　　*kalaŋan* 'prevented'

(cf. also 6.2.6 for *kə-* and *kə- -an*(1) forms in free variation with *tə-* forms.)

6.6.2 SM *kə- -an*(2) AND ITS CORRESPONDENCES

6.6.2.1 SM

SM *kə- -an*(2) is circumfixed to VSIs, VDIs and nouns. With VSIs it forms nouns referring to the quality of (VSI) as such, e.g.

merah 'red'	*kəmerahan* 'redness'
rajin 'industrious'	*kərajinan* 'industry, application'
bəsar 'big, large'	*kəbəsaran* 'largeness'

With VDIs it usually forms abstract nouns, but it may also form nouns referring to the place where the act is performed, e.g.

dataŋ 'come'	*kədataŋan* 'coming, arrival'
pərgi 'go'	*kəpərgian* 'going (n)'
mati 'die; dead'	*kəmatian* 'dying, death'
	e.g. *rumah kəmatian* 'house of the deceased, house of mourning'
tidur 'sleep'	*kətiduran* 'sleeping place'
naik 'climb; increase'	*kənaikan* 1. 'increase (n)'
	2. 'what is mounted, e.g. a boat, cart '
	3. 'ascension (of Christ)'

On the basis of nouns it forms nouns referring to the place where (noun) is found, to a collectivity of (noun), or to the quality as such of (noun). *kə- -an*(2) derivations referring to the quality as such of (noun) are as a rule neologisms; other *kə- -an*(2) derivations on the basis of nouns must be older, e.g.

hiaŋ 'divinity'	*kəhiaŋan* 'dwelling of gods' (usually written 'kayangan')
raja 'king'	*kərajaan* 'kingdom'
pulaw 'island'	*kəpulawan* 'archipelago'
binataŋ 'animal'	*kəbinataŋan* 'bestiality'
buaya 1. 'crocodile'	
2. 'lecher, scoundrel'	*kəbuayaan* 'lechery'

N.B. In a few cases *kə- -an*(2) seems also to occur with a VTR, e.g. *putuskan* 'decide (O)' and *kəputusan* 'decision'; *simpulkan* 'conclude (O)' and *kəsimpulan* 'conclusion'. *Kəputusan*, however, is possibly derived from its primary base *putus,* which is a VSI meaning 'broken off, ended; disposed of, decided'. *Kəsimpulan* and *simpulkan* must be neologisms, as they do not occur in pre-war dictionaries.

6.6.2.2 MIN

MIN *ka- -an*(2) is equivalent to SM *kə- -an*(2) but it does not occur with nouns.[199] (In Van der Toorn's material the impression is given that MIN *ka- -an*(2) is not as frequent as SM *kə- -an*(2). Also, Van der Toorn treats *ka- -an*(1) and *ka- -an*(2) as a single category, which makes their description somewhat confusing.)

6.6.2.3 BH

BH *ka- -an*(2) with VSIs forms nouns referring to the quality of (VSI), and with VDIs forms nouns referring to the place of an action, e.g.

---

[199] In Moussay's description *ka- -an*(2) does occur with nouns, which may be due to SM influence in the MIN he described (Moussay 1981:118).

baik 'good'	*kabaikan* 'goodness'
buŋul 'stupid'	*kabuŋulan* 'stupidity'
ñaman 'tasty'	*kañamanan* 'tastiness (good taste)'
diam 'stay'	*kadiaman* 'dwelling place, residence'
titi 'cross s.th. narrow'	*katitian* 'place to cross; bridge'
piŋkut 'hold, grip'	*kapiŋkutan* 'handle, hinge'

6.6.2.4 SWY

SWY *kə- -an*(2) forms occur on the basis of VSIs, and denote the quality expressed by (VSI). Only two possible *kə- -an*(2) forms on the basis of non-VSIs are found (in Helfrich) viz. *kayaŋan* 'dwelling of gods' and *koidupan* '(means of) living, livelihood' (cf. *idup* 'live'). But *kayaŋan* does not have a corresponding base; *kayaŋan* and *koidupan* are probably loans from SM.

6.6.2.5 IBN

IBN does not have a corresponding affix.

6.6.2.6 JKT

JKT *kə- -an*(2) on the basis of VSIs[200] forms nouns referring to the quality of (VSI), e.g.

pandé 'clever'	*kəpandéan* 'cleverness'
kuat 'strong'	*kəkuatan* 'strength'
untuŋ 'lucky'	*kəuntuŋan* 'profit'

In Muhadjir there is one example of a *kə- -an*(2) form with a VTR base:

bakar 'burn (O)'	*kəbakaran* 'fire'

Not all VSIs take this circumfix: some (e.g. *gədé* 'big', *tiŋgi* 'high', and colour terms) form equivalent nouns through suffixation of *-ñè*. Most *kə- -an*(2) forms in Muhadjir (including the ones presented here) also occur in SM,[201] and JKT *kə- -an*(2) is probably not inherited.

6.6.2.7 DISCUSSION OF SM *kə- -an*(1), SM *kə- -an*(2), AND THEIR CORRESPONDENCES

In SM, MIN, and SWY, *kə- -an*(1), *ka- -an*(1) denote unintentionality or feasibility of an action when they are circumfixed to a VDI or a VTR. When circumfixed to a noun or a VSI, they mean 'unintentionally affected by (noun) or by the quality of (VSI)'. In BH, *ka- -an*(1) circumfixed to nouns means 'to suffer from the effects of (noun)'. In JKT, *kə- -an*(1) alternates with *kə-* and *tə-* except when affixed to a VTR consisting of a VDI + *-in,* i n which case only *kə- -an*(1) applies. In IBN no correspondence of *kə- -an*(1) etc. occurs. Apart from these meanings BH *ka- -an*(1) also forms excessives on the basis of VSIs, while MIN *ka- -an*(1) with VSIs (apart from denoting unintentionality) can also form a comparative

---

[200]According to Muhadjir (1981:65-66, section 6.10.2-3) this circumfix also forms abstract nouns from VDIs ('intransitive verbs') and 'semi-transitive' verbs (see 6), but the four examples which he gives (viz. *jahat, jail, pəlit,* and *ribut*) are all labelled as adjectives in his wordlist.

[201]Although there are exceptions, which are all loanwords, mostly from JV, viz. *kədəmənan* 'love (n)', *kədòyanan* 'liking', *kəmacətan* 'blockage, hold-up', (on the basis of *dəmən* 'love (v)', *dòyan* 'like (v)', *macət* 'jammed', all from JV), and *kətòma?an* 'greed' (on the basis of 'greedy', from AR *tammāʕ* 'greedy, desirous').

degree. The valency of *kǝ- -an*(1) etc. differs from one isolect to the other, and the only feature common to all isolects (except IBN) is that they have *kǝ- -an*(1) etc. affixed to nouns. But *kǝ- -an*(1) etc. also occurs with verbs in SM, MIN, SWY, and JKT, and I will reconstruct a PM ancestor which occurred on the basis of verbs and nouns, and which denoted unintentionality and feasibility. As to its function of forming comparatives with VSIs in MIN and of forming excessives with VSIs in BH, this is an even smaller base for a PM reconstruction than in the case of *tǝr-* etc. with VSIs (cf. 6.2.2). Since *kǝ- -an*(1) and *tǝr-* etc. are to a great extent similar in function and meaning, a difference in distribution would be expected between them. However, I have not been able to discover any such difference that would hold for all isolects.

All isolects except IBN have a correspondence of SM *kǝ- -an*(2) occurring with VSIs and forming nouns referring to the quality denoted by (VSI). JKT *kǝ- -an*(2), however, may not be inherited. SM *kǝ- -an*(2) and MIN *ka- -an*(2) also occur with VDIs (forming abstract nouns, and nouns referring to the place where the action is performed). In BH, *ka- -an*(2) is found with VDIs, but here the resulting noun exclusively refers to the place where the action is performed. It is therefore possible to reconstruct a PM ancestral form of SM *kǝ- -an*(2) which on the basis of VSIs formed nouns referring to the quality of (VSI) as such, and which on the basis of VDIs formed nouns at least referring to the place where (VDI) takes place.

In none of the isolects do *kǝ- -an*(1) and *kǝ- -an*(2) etc. seem to have a connection, in spite of their identical shape. They will therefore yield separate reconstructions in PM. Comparative research involving non-Malayic languages may show a historical relationship between *kǝ-_-an*(1) and *kǝ- -an*(2) etc. With the material at hand, however, it seems likely that such a possible relationship was already historical at the PM level.

6.6.2.8 RECONSTRUCTION OF PM **kA- -an*(1) AND **kA- -an*(2)

On the basis of SM, SWY, JKT *kǝ- -an*(1), MIN, BH *ka- -an*(1), I reconstruct **kA- -an*(1), which contributed the notion of unintentionality or feasibility to the VTR or VDI to which it was affixed. **kA- -an*(1) was also circumfixed to nouns and VSIs, to which it added the notion of 'unintentionally being affected by (noun) or by the quality of (VSI)'. It is not certain whether **kA- -an*(1) was also a degree marker when affixed to VSIs.

On the basis of SM, SWY, (and JKT?) *kǝ- -an*(2), MIN and BH *ka- -an*(2), I reconstruct PM **kA- -an*(2), a circumfix forming nouns with VSIs and VDIs. When circumfixed to VSIs, **kA- -an*(2) referred to the quality as such of (VSI); when circumfixed to VDIs, it formed nouns referring to the place where (VDI) is performed.

6.7 THE NOUN-FORMING AFFIXES *pǝN-, pǝr-, pǝN- -an, pǝr- -an,* AND THEIR CORRESPONDENCES

These four noun-forming affixes are defined by their meaning, and by their formal agreement and valency with certain classes of verbs. Generally speaking, *pǝN-* and *pǝr-* and their correspondences both refer to the actor of a performance, to the instrument with which the action is performed, or, with VSIs, to someone or something having the quality of (VSI) as a characteristic. It is more difficult to describe the meanings which *pǝN- -an* and *pǝr- -an* usually have in the isolects. Their meanings are broader than those of *pǝN-* and *pǝr-*

(including the action or event itself (verbal abstract), or the actor, instrument, or place, depending on the particular isolect). As far as the valency and the formal agreement of these affixes are concerned, SM *pəN-* and *pəN- -an* and their correspondences form nouns on the basis of VTRs that do not have the VTR marker *pər-* etc. prefixed, and VDIs that have *məN(1)-* etc. prefixed, whereas *pər-* and *pər- -an* form nouns with VDIs that have *bər-* etc. prefixed, and with VTRs that have the VTR marker *pər-* etc. prefixed (cf. 6.1.2).

6.7.1 SM *pəN-* AND ITS CORRESPONDENCES

6.7.1.1 SM

SM *pəN-* occurs with verbs, and forms nouns usually referring to the actor of a performance, to the instrument with which the action is performed or, with VSIs, to someone or something having the quality of (VSI) as a characteristic. Some *pəN-* forms have a complement. If the underlying verb has a derivative affix, this is elided; *pəN-* forms (especially with VSIs) are often used attributively and can even function as VSIs if the base is also a VSI, e.g.

kirim(kan) 'send (O)'	*pəŋirim* 'sender'
bantu 'help, assist (O)'	*pəmbantu* 'assistant'
layani 'serve (O)'	*pəlayan* 'waiter'
kumpulkan 'collect (O)'	*pəŋumpul* 'collector'
mənari 'dance'	*pənari* 'dancer'
tidur 'sleep'	
tidurkan 'put (O) to bed, send (O) to sleep'	*pənidur* 'sleepyhead'
obat 'medicine'	*obat pənidur* 'narcotic, sleeping-pill'
pijit 'massage (O)'	
kaki 'foot'	*pəmijit kaki* 'masseur of the feet'
tuŋgu "guard (O)'	
pintu 'door'	*pənuŋgu pintu* 'porter, doorkeeper'
ikat 'bind (O)'	
tali 'rope'	*tali pəŋikat* 'rope for binding'
lari 'run'	
sahaya 'slave'	*sahaya pəlari* 'runaway slave; slave with the tendency to run away'
takut 'afraid'	*pənakut* 'coward; timid, cowardly'
mabuk 'drunk'	*pəmabuk* 'drunkard; (someone) with a tendency to drink, addicted to alcohol'
malu 'shy'	*pəmalu* 'chaste, modest (person)'

There are, however, many *pəN-* forms that do not fit into this description. Sometimes they are nouns that do not (or, at any rate, do not clearly) denote actor, instrument or (with VSIs) a characteristic; in modern usage these nouns are exceptional, but in Classical Malay they occur more often, e.g.

(h)antar 'convey, lead (O)'	*pəŋantar* 'introduction (in books, etc.)'
dapat 1. 'be able to (VDI)'	
2. 'find, acquire (VTR)'	*pəndapat* 'opinion'
bəri 'give'	*pəmbəri* 'gift' (Classical Malay)

ajar 'teach' *pəŋajar* 'instruction' (Classical Malay and Malaysian)
sakit 'ill' *pəñakit* 'illness'

(Indonesian SM has *pəmbərian* for 'gift' and *(pəŋ)ajaran* for 'instruction'.)

In Classical Malay *pəN-* forms also occur in prepositional phrases, e.g.

(43) *air   akan pəmbasuh   kaki suami-ña*
      water for   *pəN-* + wash foot husband-her
      water for washing her husband's feet

(44) *Uaŋ   itu   saya bəlañjakan akan pəmbayar   hutaŋ saya.*
      money that I   spend   for   *pəN-* + pay debt I
      That money I will use to pay my debts.

They denote a measure of time or space, or a point in time, when *sə-* is prefixed (many of these forms belong to Classical Malay), e.g.

pəluk 'embrace' *səpəməluk* 'fathom, the diameter of an embrace'
ludah 'saliva'
məludah 'spit' *səpəludah* 'as far as one can spit (as a primitive measure of distance)'
tiŋgalkan 'leave (O *səpəniŋgal ayah* 'at (the time of) father's departure'

6.7.1.2 MIN

MIN *paN-* is prefixed to VTRs, VDIs, and VSIs. The resulting forms are nominals which may occur independently with the meaning 'the habit of doing (base)', or 'something to perform (base) with'; but they are often used attributively or predicatively with a noun as referent. They mean 'having the habit, inclination or disposition to do (base), being able to do (base) or be used for doing (base)', and are sometimes followed by a complement, e.g.

kali 'dig' *paŋali* 'spade'
toko? 'hit, knock' *panoko?* 'k.o. small hammer'
lambe? 'slow' *palambe?* 'the habit of being slow'

(45) *Panduduə? di pintu sañjo ari inda? elo?.*
      paN- + sit at door dusk day not good
      The habit of sitting in the doorway at dusk is bad.

(46) *uraŋ   pamaliəŋ*
      human.being *paN-* + steal
      s.o. with the tendency to steal

(47) *uraŋ   palari*
      human.being *paN-* + run (away)
      a runaway

(48) *kabaw palalo?*
      buffalo *paN-* + sleep
      sleepyhead (said of a buffalo) (cf. *uraŋ palalo?* 'sleepyhead (of a person)')

(49)   *padati pambao   baban*
       cart   *paN-_ +* carry   load
       cart used to carry a load (cf. *uraŋ pambao baban* 's.o. who carries a load')

(50)   *aiᵊ   pambasuᵊh   muko*
       water   *paN- +* wash   face
       water to wash one's face

(51)   *añjiᵊŋ pamburu   ruso*
       dog   *paN- +* hunt   deer
       a dog for hunting deer

(52)   *Pariuᵊʔ ko   pananaʔ   katan.*
       pot   this   *paN- +* cook glutinous.rice
       This pot is used for cooking glutinous rice.

(53)   *Caŋkiᵊ tu   paminum   ubeʔ.*
       cup   that   *paN- +* drink   medicine
       That cup is for taking medicine.

(54)   *Awaʔ-ño palarikan   anaʔ-biñi uraŋ.*
       3SG   *paN- +* run.away.with child-wife human.being
       He likes to run off with other people's wives.

(55)   *Tambaʔ-kan lah   ubeʔ   ko di kañiᵊŋ   aŋ:*
       put   particle medicine this on forehead you

       pandiŋini   kapalo aŋ saketeʔ.
       paN- + cool.off head   you little.bit
       Put this remedy on your forehead: it will cool off your head a bit.

6.7.1.3 BH

BH *paN-* forms nouns with VSIs, VDIs, and VTRs. With VSIs *paN-* forms nouns referring to 's.o. having the characteristic or inclination of being (VSI)', e.g.

kulir 'lazy'	*paŋulir* 'lazybones'
gamat 'late'	*paŋgamat* '(s.o.) always late'
mauk 'drunk'	*pamauk* 'a boozer, s.o. with a disposition to drinking'

With VDIs and VTRs they refer to the actor performing (VDI/VTR) or to the instrument used for performing (VDI/VTR), e.g.

tulis 'write (O)'	*panulis* 'writer, clerk'
sadap 'tap a tree'	*pañadap* 'tapper; tapping knife'
kayuh 'row (v)'	*paŋayuh* 'oar, paddle'
tambal 'patch (v)'	*panambal* 'glue, s.th. used for patching'

6.7.1.4 SWY

SWY *pəN-* forms nouns with VDIs, VTRs, precategorials, and VSIs. Some of these nouns refer to the actor or to the instrument used to perform the act, but due to the scarcity of examples it is not possible to make an adequate semantic description of them, e.g.

*tidu*ᵊ*ʔ* 'sleep'	*pənidu*ᵊ*ʔ* 'sleepyhead'
-gəriwit '(eat or steal sweets)'	*pəŋəriwit* 'a sweet-toothed sneakthief'
kikir 'file (v)'	*pəŋikir* 'file (n)' (Aliana et al.)
*poti*ᵊ*(h)* 'white'	*pəmoti*ᵊ*(h)* 's.th. used to whiten' (Aliana et al.)
iriŋ 'accompany, escort (O)'	*pəŋiriŋ* 'escort, company' (Aliana et al.)

But cf. also:

*əñ ju*ᵊ*ʔ* 'give (O)'	*pəŋəñ ju*ᵊ*ʔ* 'gift'
sakit 'ill, sick'	*pəñakit* 'illness'

6.7.1.5 IBN

IBN *pəŋ-* forms nouns with VSIs, VDIs, and VTRs. It is the only noun-forming affix in IBN, and the meanings of the resulting nouns are quite diverse: they can refer to actor, object, place, or instrument, and they can also be abstract nouns, e.g.

insap 'smoke'	*pəŋinsap* 'smoker'
saup 'help (v)'	*pəñaup* 'helper'
*indi*ʔ 'tread on'	*pəŋindi*ʔ 'mat used in treading sago'
balut 'wrap, bandage (O)'	*pəmalut* 'anything used for wrapping or bandaging'
makay 'eat'	*pəmakay* 'food'
təmu 'know'	*pənəmu* 'knowledge'
manah 'beautiful'	*pəmanah* 'beauty'
arap 'hope, trust, believe'	*pəŋarap* 'faith, religion'
*pandi*ʔ 'bathe'	*pəmandi*ʔ 1. 'bathing place'
	2. 'bather'
	3. '(the act of) bathing'
pəkul 'enclose, drive into a corner'	*pəməkul* 'enclosure; game net'

6.7.1.6 JKT

JKT *pəN-* is unproductive (the functions of actor of a performance or bearer of a characteristic are usually circumscribed by *tukaŋ* + verb, e.g. *còpèt* 'steal, pickpocket', and *tukaŋ còpèt* 'pickpocket (n)'; *jait* 'sew', and *tukaŋ jait* 'tailor'); it occurs on the basis of VSIs, VDIs, (precategorials) and VTRs. When occurring with VDIs and VSIs, the derivational affixes of underlying forms are elided. With VSIs *pəN-* yields nouns referring to a person having (VSI) as a characteristic, e.g.

diəm	quiet	*pəndiəm*	quiet person
jahat	bad, wicked	*pəñjahat*	crook, wicked person

With VDIs, it yields nouns referring to the actor of the performance, e.g.

məñañi	sing	*pəñañi*	singer
maèn	play	*pəmaèn*	player

With VTRs (and some precategorials), it yields nouns referring to actor or object of the act, to the instrument used to perform the act, or it yields an abstract noun; the suffix *-in* of some underlying VTRs is lost, e.g.

kasi 'give'	*pəŋasi* 'gift'
-tòntòn '(watching)'	*pənòntòn* 'spectator'
alaŋin 'obstruct, hinder'	*pəŋalaŋ* 'block, obstruction'
pèlèt 'enchant'	*pəmèlèt* 'charm, trick, trap'
laŋkain 'overstep'	*pəlaŋkè* 'undertaking, step'
layanin 'serve'	*pəlayan* 'servant'
-èjèk '(waiting)'	*pəŋèjèk* 'waiter at a party'

6.7.1.7 DISCUSSION OF THE NOUN-FORMING AFFIXES *pəN-*, *pər-*, *pəN- -an*, *pər- -an*, AND THEIR CORRESPONDENCES

At first sight there is much difference between the function and valency of *pəN-* and its correspondences in the six isolects. In SM *pəN-* occurs with VSIs, VDIs, and VTRs; forms with VSIs denote a characteristic. Forms on the basis of VDIs or VTRs usually denote an actor or instrument, but (particularly in Classical Malay) they also denote a goal or result, or they form an abstract noun. Furthermore *pəN-* forms are used attributively, and, on the basis of VSIs, they can function as VSIs. In Classical Malay they may also occur following a preposition. Finally, when they have *sə-* prefixed they denote a measure of time or space, or a point in time (these forms are also more usual in Classical Malay).

In MIN, *paN-* occurs with VSIs, VDIs, and VTRs. The resulting nouns are used attributively and predicatively with the meaning 'having the quality or habit of (base)', and they may have a complement; they are also used as a noun denoting a habit, an instrument, or an actor.

In BH *paN-* forms nouns referring to the actor or instrument, or it forms abstract nouns, when prefixed to VDIs and VTRs, and it refers to someone having a characteristic or inclination when prefixed to VSIs.

Only a few examples are available of SWY *pəN-* forms: they refer to actor or instrument, but in one case they form an abstract noun on the basis of a VSI *(pəñakit)*, and in another case a noun referring to the object of an act on the basis of a VTR *(pəŋəñjuʔ)*.

In IBN *pəN-* is the only nominaliser, and its meaning is quite general (including actor, object, instrument, place, and abstract noun).

JKT *pəN-* refers to the actor if the base is a VDI, it refers to the actor, object, instrument, or it forms an abstract noun, if the base is a VTR, and it refers to a person who has (VSI) as a characteristic if the base is a VSI.

There are a few general remarks that can be made on the above picture. Firstly, *pəN-* and its correspondences do not refer to the location where the action or event takes place. An exception to this is IBN *pəN-* which, as the only nominalising affix in IBN, has generalised its meaning in a maximal way. Secondly, all isolects agree in forming nouns or VSIs referring to an inclination or characteristic by prefixing *pəN-* etc. to a VSI. Thirdly, in all isolects *pəN-* etc. in some cases also forms nouns referring to the actor or to the instrument, and abstract nouns; MIN has abstract nouns only with the connotation of the action 'as a habit'; *pəN-* forms referring to the object are found in SM, SWY, IBN, and JKT (i.e. the word for 'gift' in SM, SWY, JKT). Fourthly, the syntactical role of SM *pəN-* and MIN *paN-* is quite different from the (more restricted) role of their correspondences in other isolects. This fact is very important for the reconstruction of a PM ancestor, because for SM

and MIN we have recurrence to grammars that reflect these isolects in a relatively archaic form (they were written at the end of the nineteenth century, and are partly based on language material found in old manuscripts). The material for them is at any rate more complete than the descriptions available for BH, SWY, and IBN. (JKT is sufficiently well described, but in its grammatical structure it is probably most affected by external influences.) In view of this it is not unwarranted to put extra weight on the SM and MIN evidence.

It is possible to reconstruct for *pəN-* a PM ancestral form that referred to an inclination or a characteristic when prefixed to a VSI, and that denoted the actor of a performance or the instrument with which an act was performed, when prefixed to a VDI or a VTR; this PM prefix did not refer to the location where the action or event takes place. But from all other regular and sporadic functions which *pəN-* etc. has in the isolects, it seems that its PM ancestor had a wider application. The explanation that I propose for the diffuse picture of *pəN-* etc. is that its PM ancestor was used attributively and predicatively with a noun as antecedent. (This situation is preserved best in MIN.) Later on the derived forms also occurred without antecedent noun, and were reinterpreted as independant deverbal nouns. Abstract nouns and deverbal nouns referring to object or place were already available in PM (viz. **pAN- -an/*pAr- -an, *-An,* and **-an*(2) respectively, cf. 6.7.3-4 and 6.5). This explains why nouns formed with the ancestor of *pəN-* would more often than not refer to actor or instrument, for which no other formant was available.

6.7.2 SM *pər-* AND ITS CORRESPONDENCES

6.7.2.1 SM

SM *pər-* is prefixed to the base of VDIs with *bər-*, and of VTRs with the VTR marker *pər-*. The resulting forms have the same range of function and meaning as *pəN-* forms, and refer to actor or instrument. They may have a complement; if the underlying VTR has a transitivising suffix, this is elided. *pər-* forms are not frequent and sometimes have a more current variant with *pəN-* or with *pə-* (see below). Moreover, in some morphophonemic environments (as before liquids), *pər-* and *pəN-* both have a regular allomorph *pə-*, so that the distinction is lost, e.g.

bərjudi 'play dice'	*pərjudi* (also *pənjudi*) 'dice-player'
bərtənuŋ 'predict'	*pərtənuŋ* (also *pənənuŋ*) 'fortune-teller'
pərtuñjukkan 'show'	*pərtuñjuk* (more often *pətuñjuk*) 'indication'
bərtanak 'cook rice'	*səpərtanak nasi* (also *səpənanak nasi*) '(the time needed to cook rice =) somewhat less than half an hour'

As with *pərtuñjuk* and *pətuñjuk,* there also occurs a prefix *pə-*, which, according to some scholars (Gerth van Wijk p.162), would correspond to object-oriented VTRs, e.g. *pətaruh* 'pawn' vs *pənaruh* 's.o. who takes or gives in custody, possessor' (*taruh* 'keep, have, harbour (O)'). There is not much evidence for a separate meaning of *pə-*, which is in fact much more common than *pər-* (and has become a productive prefix). It is difficult to determine the difference between *pər-* and *pə-*, and I consider *pə-* as a variant of *pər-*, which presumably originated through interdialectal borrowing. Many Malayic isolects (including the peninsular ones, which had most influence on SM) lost *+r* in their cognates of the

prefixes *bər-*, *pər-*, and *tər-*. Moreover, the fact that both *pəN-* and *pər-* have allomorphs *pə-* (see 2.1.2) may have added to the expansion of *pə-* at the cost of *pər-* (and even *pəN-*), e.g.

bərjuaŋ 'fight, clash (esp. of large animals)'   *pəjuaŋ* 'fighter, struggler'
bərdagaŋ 'trade (v)'   *pədagaŋ* 'trader'
bərtugas 'work, have a task'   *pətugas* 'functionary'

6.7.2.2 MIN

MIN *pa-* occurs with the bases of VDIs that have *ba-* or *ø-*, and with the bases of VTRs that have (the transitivising prefix) *pa-*. It has the same meaning as *paN-* + VDIs or VTRs. In many cases it is not possible to determine whether a particular form has *paN-* or *pa-* prefixed (e.g. when the base verb has an initial vowel, nasal, or liquid, in which case both prefixes appear as *pa-*), e.g.

bacukuᵊ 'shave'   *sabun pacukuᵊ*   shaving soap
kasiᵊh 'love, like'   *pakasiᵊh* 'love potion'
bacaruyʔ 'use foul talk'
pacaruyʔ 'call (O) names, use foul
   language to (O)'   *pacaruyʔ* 'a foul-mouthed person'
pasalaŋi 'give (O) in loan'   *pasalaŋ* 's.th. which is lent'
baburu 'go/be hunting'   *paburu*   1. 'hunter'
                                                           2. 's.th. used for, or s.o. devoted to,
                                                             hunting'

(56)   *Uraŋ*   *nan tun paburu*   *banay.*
         human.being which that *pa-* + hunt true
         That man is devoted to hunting. (or That man is a true hunter.)

According to Van der Toorn, *pa-* is not prefixed to underlying object-oriented forms with *ba-*, for example, *bajua* 'be sold' (as in *kudo bajua* 'the horse is sold') does not have a corresponding *pa-* form (Van der Toorn 1899:4). Furthermore he says that *pa-* also occurs with nouns, but these forms are probably derived from VDIs consisting of *ba-* + noun rather than directly from the noun, cf. *kayu* 'wood', *bakayu* 'look for wood', and *pakayu* 'wood for construction work'; cf. also:

(57)   *uraŋ*   *pakudo*
         human.being *pa-* + horse
         someone who is often on horseback, someone devoted to horses
         (cf. also *bakudo* 'be on horseback; have a horse')

(58)   *ula*   *patikuyh*
         snake *pa-* + mouse
         a snake looking for mice
         (cf. also *batikuyh* 'with mice, have mice')

Van der Toorn points out that there are sometimes variants with *paN-* and *pa-*, but from his examples it seems that the difference between these variants is that those with *paN-* have a complement (and hence are probably derived from VTRs, and not from VDIs), e.g.

(59)　*Pariuᵊʔ ko　panananʔ　katan.*
　　　pot　　that　*paN-* + cook　kətan (glutinous rice)
　　　This pot is for cooking glutinous rice.
　　　(cf. also *pariuᵊʔ patanak* 'cooking pot', Van der Toorn 1899:6)

(60)　*añjiᵊŋ pamburu　ruso*
　　　dog　*pa-* + hunt　deer
　　　a dog for hunting deer
　　　(cf. *anjiᵊŋ paburu* a 'hunting dog', Van der Toorn 1899:6)

6.7.2.3 OTHER ISOLECTS

In Asfandi no mention is made of a BH form corresponding to SM *pər-* and MIN *pa-*. I was able to find only one example with a prefix corresponding to SM *pər-* etc. in the material on SWY isolect versus *pərmakan* 'food supply' (cf. *makan* 'eat'); the (apical) *r* in this example points to borrowing. IBN and JKT do not have a corresponding prefix.

6.7.2.4 DISCUSSION OF SM *pər-* AND ITS CORRESPONDENCES

SM *pər-* occurs with VDIs that have *bər-* prefixed and with VTRs that have (the transitivity marker) *pər-* prefixed. It is unproductive, and apparently being replaced by *pə-* and *pəN-*. Its meaning is equivalent to that of *pəN-* with VTRs and VDIs. MIN *pa-* occurs with VDIs that have *ba-* or *ø-* prefixed and with VTRs that have (the transitivity marker) *pa-* prefixed. Its meaning is equivalent to that of *paN-* with VDIs and VTRs.

Other isolects do not have a corresponding prefix: *pər-/pa-* is in a paradigmatic relation with *pər- -an/pa- -an, bər-/ba-*, and (transitive) *pər-/pa-* on the one hand, and with *pəN-/paN-* on the other, so it is very likely that the other isolects had corresponding forms which were lost. PM must have had an ancestral form which occurred with VDIs which had **(mb)Ar-* prefixed, and with VTRs which had the transitivity marker **pAr-* prefixed; this ancestral form must have had a meaning equivalent to that of the ancestral form of *pəN-* etc. occurring with VDIs and VTRs.

6.7.3 SM *pəŋ- -an* AND ITS CORRESPONDENCES

6.7.3.1 SM

SM *pəŋ- -an* occurs on the basis of VDIs that are formed with *məN(2)-* and of VTRs that are not formed with *pər-*. The resulting forms are usually abstract nouns, but they can also refer to the place where an action is performed, and, with VTRs, to the goal or result of an action. If the underlying VTR has a transitivising suffix (*-i* or *-kan*), this is elided, e.g.

məndarah 'bleed heavily'	*pəndarahan* 'haemorrhage'
məŋəmbara 'wander, travel'	*pəŋəmbaraan* '(act or place of) wandering'
məñəbəraŋ 'go across'	*pəñəbəraŋan* '(act or place of) crossing'
dəŋar 'hear'	*pəndəŋaran* '(sense of) hearing'
kirim 'send'	*pəŋiriman* 'consignment'
kətahui 'know'	*pəŋətahuan* 'knowledge'
pulaŋkan 'give back, bring or send home'	*pəmulaŋan* 'restitution, repatriation'

pandaŋ 'see, look'　　　　　*pəmandaŋan* 'seeing, observation; view, outlook'
bəkalkan 'provide, supply (O)'　　*pəmbəkalan* 'supply, provisioning'
tawarkan 'bargain'　　　　　*pənawaran* 'bargaining; offer, bid'

6.7.3.2 MIN

As Van der Toorn does not distinguish in a systematic way between *paN -an* and *pa- -an*, it is not always clear what the difference between them is. From his examples it seems that *paN- -an* is circumfixed to VTRs that do not have the transitivising prefix *pa-*, (and that *pa- -an* is circumfixed to VTRs formed with *pa-*, to VDIs, and to nouns, cf. 6.7.4.2).

paN- -an usually forms nouns referring to the place where an action is performed, but it also forms abstract nouns and nouns referring to instrument or goal of an action. In some cases it depends on the context which of the above meanings apply to a particular *paN- -an* form, e.g.

jamuᵊ 'put to dry (in the sun)'　　*panjamuᵊran* 'drying field'
guntiᵊŋ 'cut with scissors'　　*paŋguntiᵊŋan* 'cutting, clipping'
kikiᵊ 'file (v)'　　　　*paŋikiᵊran* 'filings'
pandaŋ 'look at'　　　*pamandaŋan* 'view'
bari 'give'　　　　*pambarian* 'gift'
jamu 'receive guests'　　*panjamuan* 'festive meal (to welcome guests)'

When prefixed with *sa-*, *paN- -an* can denote a measure of distance, for example, *badiᵊ* 'shoot', and *sapambadiᵊlan* (or, by false analogy, *sapambadiᵊran*, cf. SM *bədil*) 'the distance of a gunshot'; *bae* 'throw', and *sapambaean* 'the distance of a stone's throw'.

6.7.3.3 BH

BH *paN- -an* occurs with VSIs, VDIs, and VTRs. When affixed to a VSI, *paN- -an* refers to someone with (VSI) as a characteristic (the resulting forms are similar to *paN-* forms with VSIs, but they are more intensive), e.g.

kulir 'lazy' (cf. *paŋulir* 'lazybones')　　*paŋuliran* 'a very lazy person'
sarik 'angry'　　　　　*pañarikan* 's.o. who often gets angry'

When it is affixed to a VDI or a VTR, it refers to the actor or location of a performance, e.g.

bawa 'carry'　　　*pambawaan* 'carrier, porter'
urut 'massage'　　*paŋurutan* 'masseur'
simpaŋ 'branch off'　　*pañimpaŋan* 'bifurcation, branching-off'
lipat 'fold (v)'　　*palipatan* 'fold, folded spot; bend of the knee'
　　　　　(SM *pəlipatan kaki*)

6.7.3.4 SWY

From the very few examples found in Helfrich and Aliana et al., it seems that *pəŋ- -an* is equivalent to SM *pəŋ- -an*, e.g.

ajax 'teach'　　*pəŋajaxan* 'instruction'
suŋgut 'arrive'　　*pəñuŋgutan* 'arrival, anchoring'

6.7.3.5 IBN

IBN does not have a corresponding circumfix, *pəŋ-* being the only noun-forming affix in this isolect.

6.7.3.6 JKT

JKT *pəŋ- -an* is circumfixed to VSIs, VDIs and VTRs. With VSIs it forms nouns meaning 'someone who or something which has (VSI) as a characteristic', e.g.

malu 'shy, embarrassed'	*pəmaluan* 'shy person'
dəŋki 'jealous'	*pəndəŋkian* 'jealous person'

It can also form a VSI meaning 'always be (VSI), have a (VSI) character', e.g.

(61) *Pəmabòkan   bənər diè.*
      pəN- -an + drunken true 3SG
      She's a real drunkard.

(62) *òraŋ   pəŋikutan*
      human.being *pəN- -an* + follow
      a person who always follows

With VDIs it forms nouns referring to the place where the action is performed, or to the action as a process; it also forms nouns meaning 'someone who or something which often performs (VDI)',[202] e.g.

brənti 'stop'	*pəmbrəntian (bis)* '(bus)stop'
ŋaji 'recite the Koran'	*pəŋajian* 'place to recite the Koran'
maèn 'play, act '	*pəmaènan* 'game'
diri 'stand'	*pəndirian* 'founding; building; place to build'
bəlajar 'study'	*pəlajaran* 'lesson'
bòhòŋ 'lie'	*pəmbòhòŋan* 'liar'

With VTRs *pəN- -an* yields: (1) nouns referring to place or goal of the action, or abstract nouns. If the underlying form has the transitivising suffix *-in*, this is elided. (Abstract nouns formed with *pəN- -an* forms are not common), e.g.

buaŋ 'throw away'	*pəmbuaŋan* 'garbage container'
gòrèŋ 'fry'	*pəŋgòrèŋan* 'frying pan'
gəbug [gəbuk] 'hit'	*pəŋgəbugan* 'hammer'
jait 'se	*pəñjaitan* 's.th. to be sewn; sewing'
alamin 'experience (v)'	*pəŋalaman* 'experience (n)'
bəli 'buy'	*pəmbəlian* 'buying'
liat 'see'	*pəŋliatan* 'sight'

(2) VSIs denoting a characteristic or feature; these forms are limited in number, e.g.

còlòŋ 'steal'	*pəñòlòŋan* 'thievish, prone to stealing'

---

[202]On pp. 105 and 108 in Muhadjir *aji* and *diri* are refer ed to as precategorials, but on pp.113 and 94 *ŋaji* (with a nasal) and *diri* are given as VDIs.

(63)   *Susè   anak mudè jadi   pəŋgayaan.*
        troublesome child   young become *pəN- -an* + beg
        A youth who likes to beg is troublesome.

6.7.3.7 DISCUSSION OF SM *pəŋ- -an* AND ITS CORRESPONDENCES

SM, MIN, and (as far as shown by the only two known examples) SWY, agree in having a circumfix *pəN- -an* or *paN- -an* which occurs with VTRs (i.e. primary VTR bases). The resulting form is a noun forming a verbal abstract (particularly in SM) or a noun referring to place (particularly in MIN) or, to a lesser extent, to instrument and object.

BH *paN- -an* forms nouns with VSIs, VDIs, and VTRs. These nouns refer to actor or place when their base is a VDI or VTR, and they refer to someone with (VSI) as a characteristic when their base is a VSI.

IBN does not have a corresponding affix.

JKT *pəN- -an* forms nouns or VSIs on the basis of VSIs, VDIs, and VTRs. These nouns refer to actor or place when their base is a VTR, or, sometimes, a VDI; with other VDIs, they form an abstract noun or refer to place or to someone who often performs (VDI), and with VSIs they refer to someone who has (VSI) as a habit.

From the above picture it appears that the PM ancestral form of *pəN- -an* and *paN- -an* was a nominalising affix occurring at least with VTRs. The ancestral form may have been a combination of affixes (each with its own functions) but as yet there is no evidence for such an analysis. The resulting nouns were at least abstract nouns and nouns referring to the location where the action takes place. They may also have been nouns referring to the goal or instrument of the action (as such nouns are also found in some of the isolects). Whether the PM ancestral form had other applications is difficult to tell. BH *paN-_-an* forms also refer to actor. They are close in meaning to *paN-* forms, the only difference being that a notion of intensity occurs in the former. Here *paN- -an* is possibly analysable as a combination of *paN-* (cf. 6.7.1.3) and *-an*(1) (cf. 6.5.3); this may also be assumed for BH *paN- -an* on the basis of VSIs. These BH forms are on the other hand similar to JKT *pəN- -an* forms denoting 's.o. who / s.th. which often performs (VSI)'. However, the JKT forms occur with VSIs, whereas the BH forms do not . At this stage it is best to limit ourselves to the reconstruction of a PM ancestral form for *pəN- -an* that in the first place formed abstract nouns or referred to the location of an action, and that possibly also formed nouns referring to the object of an action, and to the instrument used to perform an action.

6.7.4 SM *pər- -an* AND ITS CORRESPONDENCES

6.7.4.1 SM

SM *pər- -an* occurs with VDIs that have *bər-* prefixed, with VTRs that have *pər-* prefixed, and with nouns. The resulting forms on the basis of verbs are equivalent in meaning with *pəN- -an* forms: they are usually abstract nouns, but they can also refer to place, instrument, goal or result. *pər- -an* forms on the basis of nouns were originally in a paradigmatic relationship with VDIs consisting of *ber-* + a noun. The transitivising suffixes of underlying forms are elided. On the basis of nouns, *pər- -an* denotes a collectivity and/or place, e.g.

bərjañji 'promise, make an agreement'   *pərjañjian* 'promise, agreement, testament'

bərhias 'decorate'	*pərhiasan* '(s.th. used for) decoration'
bərtana 'ask, enquire'	*pərtañaan* 'question'
bərburu 'hunt'	*pərburuan* 1. 'hunting'
	2. 'game'
	3. *alat pərburuan* 'tool for hunting';
	padaŋ pərburuan 'hunting field',
	añjiŋ pərburuan 'hunting dog'
pərkəcil 'make smaller'	*pərkəcilan* 'minimising, reduction in size'
pərbaiki 'repair, ameliorate'	*pərbaikan* 'amelioration'
pərtuñjukkan 'show, demonstrate'	*pərtuñjukan* 'show, demonstration'
gunuŋ 'mountain'	*pəgunuŋan* (Malaysian SM *pərgunuŋan*)
	'mountains, mountain range'
dusun 'village'	*pədusunan* 'countryside, rural area'
induk 'mother, dam'	*pərindukan* 'children of one mother, family'
kubur 'tomb'	*pəkuburan* 'graveyard'
sawah 'ricefield'	*pəsawahan, pərsawahan* '(complex of) ricefields'

Just as there exist derivations with *pə-* (instead of expected *pər-* or *pəN-*, see 6.7.2.1), so there are also many derivations with *pə- -an* without nasalisation or *-r-* in the first component, as seen in some of the examples above. In the same way, some grammarians (Gerth van Wijk p. 168) consider these forms as derived from object-oriented VTRs: thus the underlying form of *pəsuruhan* 'messenger, delegate', would thus be *disuruh* 'be sent, ordered', and not *məñuruh* 'send'. I do not find much support for this view. The existence of doublets such as *pəgunuŋan, pərgunuŋan* and *pəsawahan, pərsawahan* suggests that the *pə- -an* forms were originally dialectal variants of *pər- -an* forms, and that they are now expanding at the cost of the latter (though to a smaller extent than with *pər-* and *pə-* forms, cf. 6.7.2.1). There are a very few verbs without *bər-* or *pər-* which have a corresponding deverbal noun on the basis of *pər- -an*, for example, *toloŋ* 'help, rescue O (v)', and *pərtoloŋan* 'help (n)'; *minta* 'ask for O', and *pərmintaan* 'request'. *Pərtoloŋan* is an anomaly: in Classical Malay *toloŋ* is still a noun, and does not yet occur as a VTR. But no such explanation is available for *pərmintaan* (Classical Malay has a base *pinta* 'request' which is now rare, having been largely replaced by *minta*), nor for, for example, *pərcetakan* 'printing office', *pərkawinan* 'wedding', or *pərtumbuhan* 'growth' (cf. *cetak* 'print O', *kawin* 'get married', *tumbuh* 'grow').

6.7.4.2 MIN

MIN *pa- -an* is prefixed to VDIs with *ba-*, to VTRs with *pa-*, and to nouns. The resulting forms are similar in meaning to *paN- -an* forms: they usually refer to the place where the action is performed, and to a lesser extent to the action itself (forming an abstract noun), to the instrument, or to the goal or result. *pa- -an* forms on the basis of nouns denote the place

where (noun)[203] is found, or they have a particular lexicalised meaning (as in the case of *pataunan*, see below), e.g.

batamu 'meet, visit each other'	*patamuan* 'meeting point'
bajalan 'walk, go'	*pajalanan* 'way of walking; journey'
baara? 'walk in procession'	*paara?an* (also *parara?an*) 'procession; seat or other construction used to carry s.o. in a procession'
bareh '(uncooked) rice'	*pabarehan* 'rice basket, rice box, place where rice is stored'
taun 'year'	*pataunan* 'harvest; rice cycle'

Sometimes -*r*- is found between *pa*- and a following initial vowel, as above in *parara?an:* this is due to SM influence, or it may be a sporadic retention from PM (**pAr*- > MIN *pa*-, see 6.7.5).

6.7.4.3 BH

Asfandi does not mention a circumfix *pa*- -*an*, but from the derivations in Abdul Jebar it appears that it exists, and that it agrees in meaning with SM *pər*- -*an*. It occurs on the basis of VDIs and nouns. With nouns it refers to the place where (noun) is found. With a VDI it forms nouns referring to the place where an act is performed, and, in one case, *pabukaan*, to the act itself, but the examples in Abdul Jebar are too few in number to consider these meanings as the only possible ones, e.g.

buka 'open(ed)'[204]	*pabukaan* 'meal before fasting'
batapa 'live as an ascetic'	*patapaan* 'hermitage'
kajaŋ 'bamboo plaitwork'	*pakajaŋan* 'roof (of a boat)'
kucur 'saliva of a betel chewer'	*pakucuran* 'spittoon'
galaŋ 'bracelet'	*pagalaŋan* 'wrist, ankle'

In one case -*r*- occurs between *pa*- and the following initial vowel, cf. *paraduan* 'competition', from *baadu* 'compete (in sports etc.)'. This must be a SM loan (or a retention from PM **r?*).

6.7.4.4 SWY

Only two forms corresponding to SM *pər*- -*an* etc. are found:

bay 'dam, mother animal'	*pəxbayan, pərbayan* 'brood, nest'
kayu '(piece of) wood'	*pəkayuan* '(concentration of) trees'

6.7.4.5 IBN

No corresponding affix occurs.

---

[203] According to Moussay (1981:120) *pa*-[noun]-*an* forms denote a collectivity.
[204] *Babuka* 'break the fast' would be expected along with *pabukaan*, cf. also SM *bərbuka(puasa)* 'break the fast'.

6.7.4.6 JKT

JKT has a circumfix *pər- -an* forming nouns with VDIs (including 'semi-transitives', see 6) and VTRs. According to Muhadjir, these nouns refer to place, instrument,[205] or actor, or they form an abstract noun. It does not occur frequently. Most of the *pər-_-an* forms found in Muhadjir also occur in SM, and is is likely that this class of derivations as a whole is borrowed from SM, e.g.

bəranak 'give birth'	*pəranakan, pranakan* 'womb, offspring'
bəjudi 'gamble'	*pərjudian* 'place for gambling'
bəlari 'run'	*pəlarian* 1. 'escape (n)'
	2. 'fugitive'
jalan 'go, walk'	*pərjalanan* 'journey'

6.7.4.7 DISCUSSION OF SM *pər- -an* AND ITS CORRESPONDENCES

SM *pər- -an* and MIN *pa- -an* agree in occurring on the basis of VDIs which have *bər-/ba-* prefixed, on the basis of VTRs which have (the VTR marker) *pər-/pa-* prefixed, and on the basis of nouns. They also agree in meaning, forming abstract nouns (particularly in SM) or nouns referring to place (particularly in MIN), goal, result, or instrument.

BH *pa- -an* occurs with VDIs forming nouns referring to place (and, in one case, forming an abstract noun), and it occurs with nouns forming nouns referring to the place where (noun) is found.

SWY has only two examples of *pəx- -an:* both are based on nouns, and one refers to a place where (noun) is found, and the other to a collectivity of (noun).

IBN has no corresponding affix.

JKT has a restricted number of *pər- -an* forms: these are nouns referring to actor, place or instrument, or they are abstract nouns. The JKT *pər- -an* forms may all be loanwords.

On account of the above material, and the paradigmatic relationship of *pər- -an* etc. with *bər-* etc., (nominal) *pər-* etc., and (transitive verbal) *pər-* etc. on the one hand, and with *pəN- -an* on the other, a PM ancestral form should be reconstructed for *pər- -an* etc. which occurred on the basis of VDIs that had **(mb)Ar-* prefixed, VTRs that had **pAr-* prefixed (cf. 6.1.2), and nouns. It is quite probable that **pAr-*[noun]*-an* forms were not derived from nouns properly, but rather from **(mb)Ar-* forms on the basis of nouns (thus: **pAr-* [(mb)Ar-noun]*-an* + deletion of **(mb)Ar-*). On the basis of VDIs and VTRs, this ancestor at least formed abstract nouns and nouns referring to the place where an action or event takes place.

6.7.5 RECONSTRUCTION OF PM **pAN-, *pAr-, *pAN- -an,* AND **pAr- -an*

PM **pAN-* is reconstructed on the basis of SM, SWY, IBN, JKT *pəN-,* MIN, BH *paN-.* It occurred with VSIs, with VDIs that did not have **(mb)Ar-* prefixed, and with VTRs that

---

[205] I question the validity of Muhadjir's single example of a *pər- -an* form referring to the instrument of an action: *pəriŋatan* 'remembrance'. Compare the following sentence:

Ni	*guè*	*kasi*	*tandèmatè*	*buat*	*pəri t n.*
this	I	give	souvenir	for/in.order.to	*pər--an* + remember

I give you this for remembrance.

did not have *pAr-* prefixed. PM *pAr-* is reconstructed on the basis of SM *pər-*, MIN *pa-*. It occurred with VDIs that had *(mb)Ar-* prefixed, and with VTRs that had the transitive verbal prefix *pAr-* prefixed. *pAN-* and *pAr-* formed deverbal nouns that were used attributively, predicatively, and in prepositional phrases, and that had a nominal as head or subject. They denoted a purpose or instrument when prefixed to VDIs and VTRs. Moreover, *pAN-* denoted an inclination or characteristic when prefixed to VSIs.

PM *pAN- -an* is reconstructed on the basis of SM, JKT, *pəN- -an,* (sporadically) SWY *pəN- -an,* and MIN, BH *paN- -an.* It occurred with underived VTR bases, and with VDIs that had *mAN(2)-* prefixed. PM *pAr--an* is reconstructed on the basis of SM (and JKT?) *pər- -an,* MIN, BH *pa- -an,* and, sporadically, SWY *pəx- -an.* *pAr- -an* occurred with VDIs that had *(mb)Ar-* prefixed, and with VTRs that had the VTR marker *pAr-* prefixed. *pAN- -an* and *pAr- -an* formed abstract nouns and nouns referring to the location where an action or event takes place. They may also have formed nouns referring to the goal, result, or instrument, but these roles are not well attested in the isolects.

6.8 EVIDENCE FROM FOSSILISED AFFIXES

A number of PMP affixes have become fossilised in the Malayic isolects. The evidence for these affixes is fragmentary and scattered over the individual isolects, and it is insufficient for the reconstruction of PM living affixes. If a set of cognate affixes occurs in fossilised form in each of the members of a linguistic group, their comparison would only lead up to the reconstruction of a proto-affix if their presence in a substantial list of lexemes in one isolect could be set off against their absence in a corresponding list in another isolect. This is evidently not the case in the Malayic isolects.

Possible PM reflexes of PMP *ma-* (a VSI marker) and PMP *-in-* (a marker of object-orientedness) are restricted to a very few cognate sets of which each member moreover agrees in showing the reflex in question, cf. SM *ma/buk, ma/kan* (3.6.1.1; 4.5), *m/erah* (3.1.2.5), *b/in/ataŋ* (which is unlikely to be inherited), *m/ən/antu* (3.1.3.1) and their correspondences. Under such circumstances it is quite unlikely that the affixes in question were not already fossilised in the proto-language. (A prefixed allomorph of PMP *-in-* was probably inherited as PM *(*ni-),* cf. 6.3.7 last N.B.). PM reflexes of the PMP VDI marker *-um-* / *(u)m-* are more substantial in number. Nevertheless, the evidence is still too fragmentary to give grounds for the reconstruction of an affix which was still alive at the PM level. It is possible that PM still had a living affix reflecting PMP *-um-* / *(u)m-,* but it remains uncertain, and it is equally possible that PM had a corresponding affix which was already fossilised, or at least highly unproductive.

The following cognate sets have members some or all of which reflect PMP *-um-* / *(u)m-;* their members are VDIs except for *tali-t/əm/ali* etc., *tipis* etc. and (possibly) *turun-t/əm/urun* etc. :

All isolects reflect PMP *-um-* / *(u)m-* :

m/asak 'cooked, done, ripe' (5.7 lemma 39) (< PMP *m/asak* 'done, ripe, cooked', and *tasak* 'done, ripe, accomplished', cf. Toba *tasak* 'accomplished', Old Javanese *tasak* 'ripe; accomplished');

m/ati 'die, be dead' (5.7 lemma 75) (< PMP *matey* 'id.', which derived from *-um-* + *patey,* with regular loss of initial labial);

m/andiʔ 'bathe' (3.5.2) (< PMP *anDuy* 'id.');

*m/əntah / *m/atah 'raw, unripe' (3.6.2) (< PMP *m-a(n)taq / *m-e(n)taq);
*m/u(n)tah 'vomit (v)' (< PMP *u(n)taq 'id.').

The isolects invariably reflect PMP *-um- / *(u)m-, but some also have a doublet without this affix:

*t/um/buh 'grow' (< PMP *t-um-ubuq 'id.') and *tubuh 'body' (4.6; 5.6.4) (< PMP *tubuq);

*m/impi / *impi 'dream (v)' > a.i. mimpi 'dream (v, n)' and SM, JKT impi 'dream, or hope for (v)' (4.3.1; 5.7 lemma 50)';

*turun 'go down, descend' > SWY tuxun, o.i. turun (3.4.1.2), and SM turun-t/əm/urun 'continuous descent, whether of rain or of an ancient family', MIN turun-t/um/urun, BH turun-t/am/urun (with unexplained a), SWY tuxun-t/əm/uxun '(succeed, go down) through the generations';

*tali 'rope, string' (3.4.2.6) and SM, SWY tali-t/əm/ali, MIN tali-t/um/ali 'cordage of all sorts';

*m/uda? 'young, unripe' > IBN muda? 'id.', SM muda, MIN, SWY mudo 'young, unripe; light (of colour)' and SM uda (as in ma? uda 'mother's younger sister'), MIN udo 'older brother';

*m/ampus / *ampus 'wiped out, gone' > SM, BH, JKT m/ampus, MIN m/ampuyh 'dead, wiped out (coarse)' and SM (h)ampus 'id.', MIN ampuyh 'wiped out' (4.3.1); cf. also KD and SD ampus 'go (away)';

*nipis / *mipis 'thin' > SM, SWY tipis, nipis, MIN, IBN nipih, BH nipis, JKT tipis, and MIN mipih, IBN mipis, mipaw, mipih (5.7 lemma 156);

*udi/k 'upstream, upriver area' (5.2) and *m/udi/k 'go upstream, go back against the current' (3.5.2); both etyma are derived from PM *udi '(part) behind', which is still reflected in SM k/əm/udi/an 'then, subsequently', SM kəm/udi 'rudder (of a boat); croup (of a horse)', IBN udi 'after, later, follow after'; cf. also SD d-udi-e 'afterwards'.

The isolects disagree in reflecting the affix:

*m/inum / *inum 'drink' > SM, SWY, JKT m/inum, MIN m/inun, BH k/inum (5.7 lemma 40).

Of the above sets, *m/asak, *m/ati, *m/andi?, *m/əntah (*matah) and *m/u(n)tah provide no evidence for a living PM VDI marker *um- / *(u)m- : as none of the isolects has a doublet without a reflex of the affix, they may as well already have been lexicalised forms in PM. The same applies to *t/um/buh: there are related forms without the affix, but these forms reflect PM *tubuh which occurred along with *t/um/buh. It is quite possible that *t/um/buh and *tubuh were not felt any more as belonging to the same regular morphological paradigm, a circumstance which may have been stimulated by syncope of *u in *t/um/buh (cf. 4.6). In SM m/inum, BH k/inum etc., the isolects disagree in showing m/-. Provided that BH k/inum is not due to borrowing or backformation (cf. 5.7 lemma 40), this set represents the strongest evidence for a living PM affix *um- / *(u)m-. The other sets show that there was an affix, but, as already stated above, they do not provide unambiguous evidence for a living affix at the PM stage. PM may already have had the doublets represented in these sets.

THE CHANGES FROM PMP TO PM

In this chapter the various changes are described that occurred in the evolution from PMP into PM. Austronesianists have gained a fair amount of insight into the course of sound changes between PMP and the contemporaneous AN languages. On the lexical level there is enough reconstructed material available to trace the developments from PMP to PM as regards terms for times of the day, directions, numerals, kinship terms, pronouns, and body parts, and to compile a Swadesh 200-item basic wordlist for PMP (as is done by Blust, see Blust forthcoming) which can be contrasted with the equivalent list for PM. The reconstruction of PAN/PMP morphology, however, has been attempted only recently, and has not met the same agreement among scholars as the reconstruction of PAN/PMP phonology and lexicon. Moreover, attempts to reconstruct PAN/PMP morphology have so far been mainly concerned with the verbal system and, more particularly, with the verbal focus system. This focus system, although reasonably well reflected in the languages of Formosa, the Philippines, northern Sulawesi, north-east Borneo and Madagascar, had undergone some radical changes in the western Indonesian languages. I restrict myself here to a comparison between PMP and PM phonology (including word structure) and lexicon.

7.1 PHONEMIC DEVELOPMENTS

The sound changes from PMP to PM are as follows:

	PMP	:	PM
(a)	*a*	:	*a*
	e	:	*ə, a*l __ *h*
	i	:	*i*
	u	:	*u*

Examples:

PMP *waRi* 'day; sun' > *hari* (3.1.1.1); PMP *baRah* 'live coal' > *baraʔ* (3.1.1.1); PMP *tinuʔun* 'weave' > *tənun* (3.1.1.2); PMP *seŋet* 'sting' > *səŋət* (3.1.1.2); PMP *liqeR* 'neck' > *lihər* (3.1.2.1); PMP *DaReq* 'blood' > *darah* (3.7.3); PMP *quluh* 'head' > *hulu(ʔ)* (3.4.2.5).

	PMP	:	PM
(b)	*-ay*	:	*-ay*
	-ey		
	-uy	:	*-i*
	-iw		
	-ew	:	*-u*
	-aw	:	*-aw*

Examples:

PMP *$suŋay$ 'river' > *$suŋay$ (3.2.1); PMP *$baŋkay$ 'corpse' > *$baŋkay$ (3.2.1); PMP *$qatey$ 'liver' > *$hati$ (3.2.3); PMP *$pajey$ 'rice plant' > *$padi$ > SM, MIN, SWY, IBN, JKT *padi* 'id.'; PMP *$babuy$ 'pig' > *$babi$ > a.i. *babi* (but cf. also IBN *babuy*, which is probably a loan); PMP *$hapuy$ 'fire' > *api (5.7 (143)); PMP *$laRiw$ 'run, flee' > *$lari$ > SM, BH *lari*, MIN, IBN *lari, rari*, SWY *laxi, lari*, JKT *lari?, rari* 'id.'; PMP *$baRiw$ 'spoiled' > *$bari?$ > SM, BH *bari/bari* 'fruit-fly', SWY *padi bə/bari/an* 'rice that has been stored away for more than two years' (*bə/baxi* 'owl-moth': also a cognate?), IBN *bari?* 'musty', *bə/bari* 'fruit-flies'; PMP *$buRew$ 'chase, run (away), hunt' > *$buru$ (5.7 (69)); PMP *$pulaw$ 'island' > *$pulaw$ > JKT *pulo*, o.i. *pulaw* 'id.'; PMP *$panaw$ 'white spots on the skin' > *$panaw$ (3.2.2).

N.B. In two cases PM shows an unexpected *-*? after *i reflecting a PMP final diphthong: PMP *$beRey$ 'give' > *$bəri?$, and PMP *$anDuy$ 'bathe' > *$m/andi?$ (3.5.2).

	PMP	:	PM
(c)	*y*	:	*y*
	w	:	ø-, -*w*-

Examples:

PMP *$qayam$ 'domesticated; play' > *$hayam$ 'domesticated animal, pet-animal, plaything' (3.3.1); PMP *$bayaD$ 'pay (v)' > *$bayar$ (3.3.1); PMP *$sawah$ 'python' > *$sawa?$ (3.3.2); PMP *$ñawa$ 'soul, life' > *$ñawa$ (3.3.2); PMP *$wakaD$ 'root, creeper' > *$akar$ (3.1.1.3); PMP *$waDa?$ '(not) exist, there is (not)' > *$ada(?)$ (3.4.2.5).

N.B. In one case PM reflects *h for PMP initial *w: PMP *$waRi$ 'sun; day' > *$hari$ (3.1.1.1). The correspondence PMP *$waRi$ > SM *hari* induced Dempwolff to posit a sound law PMP *w- > SM *h*, notwithstanding the fact that in other cases SM has ø for PMP *w-. Nothofer reconstructed two PMJ phonemes to account for the SM reflexes for PAN *w-: PMJ *w_1- > SM *h*, and PMJ *w_2- > SM ø. PMJ *w_1aRi (> SM *hari*) is the only proto-phoneme containing *w_1- (except for PMJ *w_1aiR 'water' > *air, which is probably a misprint for *w_2aiR, see Nothofer p.165). Rather than taking SM *h*- as a regular reflex of PMP *w- (and PMJ *$w_{(1)}$-), I assume that PMP *w- became ø in PM (and SM), and that in some isolects an epenthetic -*h*- originated between like vowels in the compound *$mata$ + *$hari$. From Blust's fieldnotes on SAR, it appears that this isolect has *ari* 'day', but *matahari* 'sun'. A comparable development is seen in Dairi, where PMP *w- was lost in *ari* 'day', but was maintained (between like vowels) in *matawari* 'sun' (Adelaar 1981:13).

	PMP	:	PM
(d)	*p*	:	*p*
	t	:	*t*
	c	:	*c*
	k	:	*k*

Examples:

PMP *$puluq$ 'ten' > *$puluh$ (3.4.1.1); PMP *$qatep$ 'roof, roofing thatch' > *$hatəp$ (3.4.1.2); PMP *$taliŋa$ 'ear' > *$t.Aliŋa(?)$ (5.7 lemma 43); PMP *$m/atey$ 'dead' > *m/ati (5.7 lemma 75); PMP *$hepat$ 'four' > *$əmpat$ (3.4.1.1); PMP *$kutu$ 'head louse' > *$kutu$ (3.4.1.4); PMP *$pucuk$ 'top, summit' > *$pucuk$ (3.4.1.3); PMP *$cukup$ 'enough' > *$cu(ŋ)kup$ (3.4.1.3); PMP *aku 'I' > *aku (3.4.2.6).

	PMP	:	PM
(e)	b	:	b, -p
	d ⎱	:	d, -t
	j ⎰		
	D	:	d, -r
	z ⎱	:	j
	Z ⎰		
	g	:	g, -k

Examples:

PMP *beRey > 'give' > *bəri? (3.5.1; *-? unexplained, see (b) above); PMP *tebuh 'sugarcane' > *təbu (3.4.2.6); PMP *kaban 'companion' > PM *kaban 'companion, follower; herd, group' (3.5.1); PMP *DuRi 'thorn' > *duri? (3.5.2); PMP *(iu)juhuŋ 'nose' > *hiduŋ (3.5.2; *h- unexplained); PMP *anDuy > *m/andi? (3.5.2, 7.1b N.B.); PMP *tazem 'sharp' > *tajəm (3.5.3); PMP ·*Zual 'sell' > *jual (3.5.3); PMP *Zaqet 'evil-hearted' > *jahət (3.1.1.5 IC); PMP *pajey 'rice plant' > *padi (see (b) above); PMP *deŋeR 'hear' > *dəŋər (5.7 lemma 44); PMP *kunij 'curcuma' > *kunit 'curcuma; yellow' (3.6.1.2IC; see 3.4.2.2 for PMP *-j > JKT r in lalər, pusər, and ulər, and JKT d/t in anud/anut); PMP *bayaD 'pay (v)' > *bayar (3.3.1); PMP *wakaD 'root, creeper' > *akar (3.1.1.3); PMP *tuhed 'knee' > *tuøət (3.10); PMP *lebleb 'immerse, inundate' > *lələp 'submerge, disappear' > SM ləlap 'disappear, be sound asleep', MIN lalo? 'be asleep', JKT lələp 'sunken, submerged'; PMP *zegzeg 'have firm ground under one's feet' > *jəjək 'step on; footprint' > SM jəjak, MIN jaja?, BH jajak, SWY jəja?, (JKT jəjak 'footprint' must be a loan; *jəjək is possibly related to *səjək, cf. 4.3.2).

	PMP	:	PM
(f)	m	:	m
	n	:	n
	ñ	:	ñ
	ŋ	:	ŋ

Examples:

PMP *(i)kami '(1st pers. pl. excl.)' > *kami (3.6.1.1); PMP *mata 'eye' > *mata (3.4.1.2); PMP *tazem 'sharp' > *tajəm (3.5.3); PMP *naneq 'pus' > *nanah (3.6.1.2); PMP *ñamuk 'mosquito' > *ñamuk (3.6.1.3); PMP *añam 'weave, plait' > *añam (3.6.1.3); PMP *ŋaŋa 'open (mouth)' > *ŋaŋa(?) 'agape' (3.6.1.4); PMP *laŋit 'sky' > *laŋit (3.6.1.4); PMP *tanem 'plant (v)' > *tanəm (3.6.3.1); PMP *bulan 'moon, month' > *bulan (3.6.3.2); PMP *DiŋDiŋ 'wall' > *dindiŋ (3.6.3.3).

	PMP	:	PM
(g)	r ⎱	:	r
	R ⎰		
	l	:	l

Examples:

PMP *limah 'five; hand' > *lima? 'five' (3.7.1); PMP *tuqela(nŋ) 'bone' > *tulaŋ (3.7.1); PMP *Rumaq 'house' > *rumah (3.1.1.3); PMP *beReqat 'important, heavy' > *bərat (3.7.3); PMP *rantaw 'inlet, bay; travel along the coast' > *rantaw 'coastland, inlet, foreign country' (3.2.2); PMP *ukur 'measure (v)' > *ukur 'measure, test, divine' > SM, BH, JKT

ukur SWY *ukur* 'measure (v)', MIN *uku^ə* 'measure (n)', IBN *ukur* 'luck, chance', and *ŋ-ukur* 'test (one's luck), read (someone's fortune)'.

	PMP	:	PM
(h)	*s*	:	*s*

Examples:

PMP **salaq* 'error' > **salah* 'at fault, amiss' (3.1.1.1); PMP **tasik* 'sea' > **tasik* 'sea' (3.4.2.3); PMP **a ta^ʔas* 'on top of, above' > **atas* (3.8.2).

	PMP	:	PM
(i)	*q*	:	*h*
	ʔ	:	*ø, -ʔ*
	h	:	*h-/ø-, -ø-, -ʔ/-ø*

Examples:

PMP **qihu* 'shark' > **hiu^ʔ* (3.9.1); PMP **taqun* 'year' > **tahun* (3.4.1.2); PMP **suqsaq* 'trouble, worries' (3.9.3) > **susah* 'difficult, troublesome' (3.9.3); PMP **daqan* 'branch' > **dahan* (5.7 lemma 113); PMP **haliq* 'go, move' > **alih* 'move, change' > SM, BH *alih*, MIN *ali^əh*, SWY *ali^ə(h)* 'id.', IBN *alih* 'turn over'; PMP **hepat* 'four' > **əmpat* (3.4.1.1); PMP **tuhed* 'knee' > **tuøət* (3.10); PMP **Duha* 'two' > **dua(ʔ)* (3.1.1.4); PMP **lahud* 'towards the sea' > **laut* 'id.' (5.7 lemma 124); PMP **tebuh* 'sugarcane' > **təbu* (3.4.2.6); PMP **qumah* 'farm(land)' > **huma(ʔ)* (3.2.3, 3.4.2.5); PMP **kitah* '(1st pers. pl. incl.)' (cf. Zorc, 3.4.2) > **kita^ʔ* (3.2.3, 3.4.2.4); PMP **paRih* 'rayfish' > **pari^ʔ* (3.4.2.4); PMP **ha(n)teD* 'deliver, escort' > **hantə(rt)* > SM *hantar*, SWY *antat, antar*, JKT *antər*; PMP **hiRup* 'sip, slurp' > **hirup* (3.9.1); PMP **ʔenem* 'six' > **ənəm* (3.1.1.5); PMP **a ta^ʔas* 'on top of, above' > **atas* (3.8.2); PMP **be(n)ti^ʔis* 'calf (leg)' > **bətis* 'part of leg between knee and ankle' (5.6.5); PMP **Datu^ʔ* 'chief' > **datu^ʔ* 'head of a clan' (3.4.2.4); (cf. also PMP **mata* 'eye' > **mata* (3.4.1.2); PMP **kutu* 'head louse' > **kutu* (3.4.1.4)).

7.2 PHONOTACTIC DEVELOPMENTS

7.2.1 DEVOICING OF PMP FINAL VOICED STOPS IN PM

In PM no final voiced stops or palatals occurred: PMP **-b*, **-d*, and **-g* became devoiced, PMP **-D* became **r*, and PMP **-j* became **t* (7.1e).

7.2.2 REDUCTION OF PMP CONSONANT CLUSTERS

(a) Consonant clusters were reduced to the second component, unless the first component was a nasal:

Examples:

PMP **zegzeg* > **jəjək* (7.1e);
PMP **bejbej* 'wind (v)' > **bəbat* 'bandage' > SM *bəbat*, BH *babat* 'bandage', MIN *babe^ʔ* 'waist belt', SWY *bəbat* 'bandage; waist belt';
PMP **buRbuR* 'gruel' > **bubur* (3.1.2.4);

PMP *suqsaq 'trouble, worries' > *susah 'difficult, troublesome' (3.9.3);
PMP *tektek 'cut off' > *tətək 'cut up' (3.4.2.3).

(b) Heterorganic nasals became homorganic to following stops:

Examples:

PMP *DemDem 'think, be quiet' > *dəndəm 'foster (secret) feelings of revenge; pine, long for' > SM dəndam, SWY dəndam, dəmdam (medial m unexplained), MIN, BH dandam, JKT dəndəm 'id.', IBN dəndam 'feud';
PMP *DiŋDiŋ 'wall' > *dindiŋ (3.6.3.3);
PMP *gemgem 'hold (in the fist)' > *gəŋgəm 'hold (in the fist); closed hand, fist' (cf. 7.3.7, (89));
PMP *paŋuDan 'pandanus' > *pandan (4.6);
PMP *n(iu)pi 'dream (n)' > *impi 'dream (v)' (3.4.1.1).

7.2.3 SYLLABLE REDUCTION

Syllable reduction occurred in lexemes of more than two syllables (this has been an ongoing process from PMP to the contemporaneous isolects). There are three ways in which syllable reduction took place:

(a) contraction of vowels in lexemes of more than two syllables: see 4.5;
(b) syncope of penultimate vowels in trisyllables: see 4.6;
(c) loss of PMP initial *(h/q/ʔ/ø)a: see 3.1.3.3.

7.2.4 MERGERS OF PMP *a AND *e IN PM *a BEFORE *-h

Examples:

PMP *salaq 'error' > *salah 'at fault, amiss' (3.1.1.1);
PMP *suqsaq > *susah (see 7.2.2a);
PMP *Dareq 'blood' > *darah (3.7.3);
PMP *naneq 'pus' > *nanah (3.6.1.2);
PMP *taneq 'land, soil, earth' > *tanah > JKT tanè, SWY tana(h), o.i. tanah.

7.2.5 CHANGES OF PMP SEQUENCES OF *(h, q, ø)ə- +STOP TO PM SEQUENCES OF *(h)əN- + STOP

PM did not have initial *(h)ə- sequences of + stop (4.2 and fn. 128). Corresponding to PMP initial sequences consisting of (laryngeal +) *e(N)- + stop, PM always has *(h)əN- + stop.

Examples:

PMP *hepat 'four' > *əmpat (3.4.1.1);
PMP *e(N)bun 'dew' > *əmbun (3.1.2.4);
PMP *e(N)taq 'raw' > *m/əntah (3.6.2UIC);
PMP *qe(N)tut 'fart (n)' > *k/əntut > SM, SWY, JKT kəntut, MIN kantuyʔ, BH kantut 'id.', IBN kəntut 'break wind'.

7.2.6 LOSS OF PMP SEMIVOWELS IN INITIAL POSITION

PMP initial semivowels were lost; intervocalically, PMP *y was maintained in the position *(a,u) _ (*a,u), and PMP *w was maintained in the position *a _ *a. Lexeme finally, PMP semivowels occurred as part of a diphthong (see 7.1b).

7.3 LEXICAL REPLACEMENT

The developments in lexical changes that took place between PMP and PM will be investigated according to the semantic fields to which the respective lexemes belong.

7.3.1 TIMES OF THE DAY

Two PMP terms referring to times of the day have been reconstructed on the basis of sets of widespread cognates: PMP *Rabii 'evening' and PMP *beRŋi 'night'. PMP *Rabii and *beRŋi were replaced by PM *kə-lə(hø)əm hari and *ma-lə(hø)əm respectively. No sound PMP reconstructions have yet been proposed for 'morning' or 'afternoon'.

7.3.2 DIRECTIONAL TERMS

Blust (1980c:220) proposes four PMP directional terms, two of which are actually names of winds:

PMP *lahud 'towards the sea' > *laut 'id.' (5.2.2; 5.7 lemma 124);
PMP *habaRat 'north-west monsoon' > *barat 'id.' (> SM barat 'west');
PMP *hatimuR 'south-east`monsoon' > *timur 'id.' (> SM timur 'east');
PMP *Daya 'towards the interior' > *daya? 'id.' (5.2.2).

PM *barat and *timur were still names of monsoon winds: SM barat and timur do not have trustworthy cognates in the other isolects, which makes it impossible to assign the meanings 'west' and 'east' to their PM proto-form (see 5.2.2).

7.3.3 NUMERALS

In the numeral system PM replaced the original PMP terms for 'seven', 'eight', and 'nine', cf.

PMP *pitu 'seven' vs *tujuh 'id.';
PMP *walu 'eight' vs *dua(?) alap-an 'id.';
PMP *siwa 'nine' vs *(ə)sa? ambil-an, *(ə)sa? alap-an 'id.'.

Otherwise, the PM numerals and numeral system derived regularly from PMP (5.3.1-2).

7.3.4 KINSHIP TERMS

The following kinship terms are characteristic for PM, and are not derived from PMP forms with widespread reflexes:

PMP --- vs *(əm)pu - *hiaŋ 'ancestor';
PMP *e(ŋ)pu 'reciprocal term for grandparent and grandchild' vs *nini? 'grandparent;

PMP *ama, *(t)-ama 'father' vs *apa(ʔ) 'id.';
PMP *ina, *(t)-ina 'mother' vs *(ə)ma(ʔ) 'id.';
PMP *mak(a,e,i,u)mpu 'grandchild' vs *cucuʔ 'id.'.

7.3.5 PRONOUNS

No lexical replacements took place in the personal pronoun system, although the PMP personal markers (*i and *si) were partly lost, cf.

PMP *(i)aku '(1st pers. sg.)' > *aku, *ku-/*-ku 'id.';
PMP *(i)kita '(1st pers. pl. incl.)' > *kitaʔ 'id.';
PMP *(i)kami '(1st pers. pl. excl.)' > *kami 'id.';
PMP *(i)kahu '(2nd pers. sg.)' > *kauʔ 'id.';
PMP *(i)kamu '(2nd pers. pl.)' > *kamu(ʔ); *-mu(ʔ) (sg. and pl.);
PMP *(si)ia '(3rd pers. sg.)' > *ia, -*ña 'id.';
PMP *(si)iDa '(3rd pers. pl.)' (Blust 1977a:11) > *sidaʔ 'id.'.

PM retained the PMP demonstratives:

PMP *qi-ini 'this' > *(i)ni(ʔ) 'id.';
PMP *qi-tu 'that (near hearer)' > *(i)tu(ʔ) 'id.';
PMP *qi-na 'that (distant)' > *(i)na(n),*(a)na(ʔ) 'id.'.

The loss of PMP *q in the PM reflexes is unexplained.

The locative pronouns are also basically reflexes of PMP locatives, although they do not agree semantically for the second person, and they do not match in the third person. In some ways the IBN locatives agree better with the PMP locatives than the ones reconstructed for PM (cf. IBN dituʔ 'here', diaʔ 'there', and diin 'yonder'). Compare:

PMP *di-ni 'here' > *(ʔ)(i)ni(ʔ) 'here';
PMP *di-tu 'here' > *(ʔ)(i)tu(ʔ) 'there';
PMP *di-a 'there' (Blust 1970) > *(ʔ)-(i)na(n)/ *(ʔ)-(a)na(ʔ) 'yonder' (5.5.2.2).

Of the five interrogative pronouns reconstructed, two are innovations, and one has no corresponding PMP reconstruction:

PMP *mana 'how?' > *mana(ʔ) 'which?';
PMP *i-nu 'where?' vs *-mana(ʔ) 'id.';
PMP *apa 'what?' > *apa 'id.';
PMP *(i)sai 'who?' vs *si-apa, *sai 'id.'.

7.3.6 PARTS OF THE BODY

In the field of body parts the following PM reconstructions are not reflexes of PMP terms with the same meaning:

(internal body parts)
PMP *tian 'belly' vs *pərut (*tian 'uterus');
PMP *t-in-aqi 'guts' vs *pərut 'id.';
PMP *pusuq 'heart' vs *jantuŋ 'id.'.

(head)
PMP *bahaq-bahaq 'mouth' vs *mulut 'lips; mouth'
PMP *(l,n,ŋ,q)ipen, *(ŋ)isi 'tooth' vs *gigi 'id.';
PMP *buh(ue)k 'head hair' vs *rambut/*buø(ue)k 'id.'.

(trunk)
PMP *likud 'back' vs *bAlakaŋ 'id.';
PMP *lambuŋ 'side, flank' vs *rusuk 'id.'.

(limbs)
PMP *qaqay/*waqay 'foot, leg' vs *kaki 'id.';
PMP *(qa-)lima/*kamay 'hand' vs *taŋan 'id.'.

(body parts of animals)
PMP *panij 'wing' vs *sayap 'id.';
PMP *suŋut 'snout' vs *juŋur 'id.'.

N.B. (1) A semantic shift occurred in PMP *tian 'belly' > *tian 'uterus' (2) Dempwolff reconstructed PMP *peliR on the basis of only two correspondences, SM pəlir and JV pəli, both of which he labelled 'penis'. In fact only pəli has this meaning: SM pəlir means 'testicle' and reflects *pəlir 'id.' (3.1.2.3). (3) For two concepts PM reconstructions have been made, but no well established PMP correspondences are available: *puø(uə)t 'vagina', and *dagu? 'chin'.

7.3.7 200-ITEM BASIC WORDLIST

In the Swadesh 200-item basic wordlist for PM (5.7) the following lexical replacements took place:

	English	PMP	PM
1.	hand	*(qa-)lima/*kamay	*taŋan
4.	leg, foot	*qaqay/*waqay	*kaki
5.	walk, go	*laka(dt)/*lakaw/	
		*panaw/*lampa(ŋ?)	*(mb)Ar-jalan
8.	turn (v)	*biliŋ/*ileŋ/*li(u)liu	*biluk
9.	swim	*laŋuy/*naŋuy	*(mb)A-rənaŋ (but cf. 5.7.1 lemma 9)
10.	dirty	*cemeD/*daki/*ma-ilaŋ	*kamah/*kumuh
13.	belly	*tian	*pərut
16.	guts	*t-in-aqi	*pərut
30.	mouth	*bahaq-bahaq	*mulut
31.	tooth	*(l,n,ŋ,q)ipen, *(ŋ)isi	*gigi
41.	bite (v)	*kaRat/*katkat/*ketket/ *kitkit/*kutkut/*kete(b,p)	*gigit
46.	see	*kita	*lihat
51.	sit	*untud/*tubaŋ	*duduk
53.	person, human being	*tau/*tau-mataq	*uraŋ (but cf. 5.7.1 lemma 53)
57.	husband	*qasawa/*bana	*laki
58.	wife	*qasawa	*bini
59.	mother	*(t-)ina	*(ə)ma(?)/*indu?(*ina)

60.	father	*(t-)ama	*apa(?)
69.	hunt	*qanup	*buru
70.	shoot (an arrow)	*panaq	*timbak/*panah
73.	steal	*takaw	*maliŋ
77.	scratch (an itch)	*kaRaw, *garut	*garu?/*garut/*garuk
82.	dull, blunt	*pu(n)dul/*dumpel	*tumpul
87.	swell (as an abscess)	*baReq	*bəŋkak
89.	hold (in the fist)	*gemgem	*pəgaŋ
94.	throw (as a stone)	*tudaq	*limpar
97.	bird	*manuk	*buruŋ
100.	wing	*panij	*sayap
101.	fly (v)	*Rebek	*tAr(ə)baŋ
102.	rat	*labaw	*tikus
104.	fat, grease	*meñak/*miñak	*ləmək
107.	(earth)worm	*kalati/*(qali-)wati	*hulət/*caciŋ
118.	grass	*baliji	*rumput
121.	sand	*qenay/*benaqi	*pasir
130.	star	*bi(n)tuqen	*bintaŋ
131.	cloud (not a rain-)	*Rabun	*a(bw)an
134.	thunder	*kuDug/*ru(ŋ)guŋ	*guntur/*guruh
140.	dry	*(ma-)Raŋaw	*kəriŋ
145.	smoke (of a fire)	*anus/*qasu/*ebel	*asəp
151.	green	*(ma-)iselem	*hijaw
152.	small	*Dikiq/*keDi/*keDik	*kəcil/*kəcik
155.	long (of objects)	*anaduq/*adaduq	*pañjaŋ
159.	wide	*(ma-)lawa	*libar
161.	shy, ashamed	*(ma-)hiaq	*malu
164.	good	*ma-)pia/*dia	*baik
167.	night	*beRŋi	*ma-lə(hø)əm
179.	near	*azani	*dəkət/*(h)ampiŋ
181.	where (interrogative)	*i-nu	*-mana(?)
190.	other	*duma	*bukən
191.	all	*amin	*habis
195.	no, not	*diaq/*qazi	*-da?

In twelve cases, two reconstructions occur in the PM list, one of which does not reflect the corresponding PMP reconstruction(s) in the Swadesh wordlist:

	English	PMP	PM
26.	hair of the head	*buh(ue)k	*rambut/*buø(uə)k
38.	chew (v)	*mamaq	*mamah/*kuñah
39.	cook (v)	*nasuk/*tanek/*Zakan	*m/asak/*tanək
71.	stab, pierce	*suksuk	*tusuk/*tikəm
72.	hit (with a stick)	*palu	*pukul/*palu?
78.	cut, hack	*tektek/*taRaq	*tətək/*putuŋ
93.	pound, beat	*tuktuk/*bayu	*tumbuk/*tutuk
95.	fall (as fruit)	*ka-nabuq/*ma-nabuq	*jatuh/*labuh
103.	meat, flesh	*hesi/*isi	*isi?/*dagiŋ
144.	burn (s.th.)	*tunu	*bakar/*tunu

| 153. | big | *(ma-)Raya | *bəsar/*raya |
| 189. | who (interrogative) | *(i)sai | *si-apa/*sai |

N.B.

(1) The *t- in PM *tusuk is unexplained (due to regressive dissimilation?)

(2) Some PMP basic proto-lexemes have PM reflexes, but these were probably not part of the PM basic vocabulary:

7. PMP *maRi 'come' > *mari(?) 'come; hither' > SM mari 'come here!', SM, JKT kə/mari, MIN, BH ka/mari 'hither' (possibly connected with SM d/ari (d- < *di-?), SWY g/axi (g- unexplained), IBN ari 'from');
10. PMP *cemeD 'dirty' > *cəmər > SM cəmar, MIN cama; PMP *daki 'body dirt' > *daki? (3.4.2.4);
11. PMP *abuk 'dust' > *abuk > SM, IBN abuk, SWY abuᵊ?;
44. PMP *diŋa 'hear' > *diŋa > IBN diŋa;
88. PMP *peReqes 'squeeze' > *pərəs > JKT pərəs;
89. PMP *gemgem 'hold (in the fist)' > *gəŋgəm 'hold (in the fist); closed hand, fist' > SM gəŋgam, MIN gaŋgam, JKT gəŋgəm '(grasp, grip, hold in) the closed hand or fist', MIN gaŋgam, SWY gəŋgam, 'the fist, closed hand', IBN gəŋgam 1. 'the width of the fist' 2. 'a handful';
112. PMP *buRuk 'rotten' > *buruk 'stinking, rotten' (5.7 lemma 112);
148. PMP *burak 'white' (Blust 1970:119) > *burak > IBN burak;
168. PMP *qalejaw 'day' > *andaw 'day, daylight' (5.1.2).

(3) Some PMP basic proto-lexemes have reflexes that underwent a semantic shift:

14. PMP *tian 'belly' > *tian 'uterus' (7.3.6);
21. PMP *DemDem 'think, be quiet' > *dəndəm 'foster secret feelings(of revenge, etc.)' (7.2.2b);
54. PMP *(ma-)Ruanay 'man, male' > *mA(r)(w)anay 'brother (woman speaking)' (5.4.5);
55. PMP *b-in-ahi 'woman' > *bini 'wife' (5.7 lemma 55);
70. PMP *panaq 'shoot' > *panah 'bow' (5.7 lemma 70); JKT panè, o.i. panah;
87. PMP *baReq 'swell (an abscess)' > *barah 'abscess' > SM, MIN barah, SWY baxa(h);
97. PMP *manuk 'bird' > *manuk 'domestic fowl' > BH, IBN manuk;
104. PMP *miñak 'fat, grease' > *miñak 1. 'oil' 2. 'fat';
123. PMP *aluR 'flow' > *alur 'hollow or current (?) in a river' > SM alur, MIN aluᵊ, SWY alur/an 'groove, channel, hollow', BH alur 'long line', IBN alur 'current in a river';
194. PMP *mana 'how' > *mana(?) 'which' (5.5.4, 5.5.6).

(4) Loss of morpheme boundary and other formal changes occurred in the following PM lexemes:

2. PMP *kiwa 'left' > *kiba? (with unexplained *-b-);
35. PMP *(m-)utaq 'vomit' > *m/u(n)tah (5.7 lemma 35);
36. PMP *luZaq 'spit' > *ludah (with *d for expected *j, 5.7 lemma 36);
37. PMP *ka?en 'eat' > *ma/kan ((3.6.1.1);
50. PMP *(mi-)hepi, *nipi, *nupi 'dream (v)' > *impi (3.4.1.1);
76. PMP *quDip 'live' > *hidup (4.4 and 5.7 lemma 76);
137. PMP *hiup 'blow' > *t/iup (5.7 lemma 137).

CONCLUSIONS

8.1 RESULTS

In the previous chapters I have attempted a reconstruction of PM. I was able to point out that some of the isolects still show retentions that were lost in SM. Among other things, JKT final-syllable schwa was shown to be a retention from PMP, and a corresponding schwa was reconstructed for PM. A further consequence of the fact that JKT schwa is a retention (rather than a loan phoneme) is that if lexemes in Malayic isolects have a corresponding form with a final syllable schwa in Javanese, this form may have been borrowed from JKT into Javanese. There are at any rate no grounds for automatically assuming that the Javanese correspondence must be inherited on account of its final-syllable schwa. It may turn out to be an original Malayic lexeme that came into Javanese through JKT, or through another Malayic isolect which retained schwa in an earlier stage. This may eventually throw new light on the mutual influence that Malayic isolects and Javanese exercised on one another. It was also demonstrated that, although there is a *r/x* distinction in SWY, this does not reflect the alleged PMP **r/*R* distinction, and that SWY *r* is an innovation.

The study of PM word structure has yielded some insights which I hope will prove to be useful as a test for inheritedness of lexical items. For instance, there is a strong tendency for articulation-type harmony of homorganic consonants at the beginning of syllables in (disyllabic) lexemes. Furthermore, it is very likely that PM lexemes with final **-im* or **-ip* sequences did not occur. This would explain the metathesis in **hidup* (< PMP **quDip*) which was suggested as a criterion for subgrouping Malayic isolects with some other languages (Blust 1981:463).

On the morphological level it appears that (besides the single transitive suffix **-i*) PM probably had a subjunctive marker **-a*. A patient-oriented prefix ancestral to *di-* in each of the isolects did not occur.

The lexical reconstructions which I have made are meant to show sound correspondences between the isolects, and to present a sufficient corpus of PM basic vocabulary. In a number of cases I have made corrections to higher order (PMP) lexical reconstructions.

8.2 SUBCLASSIFICATION

It is difficult to make a detailed subclassification of the Malayic isolects on the basis of this study. The major differences between the isolects are either retentions, or innovations that are not exclusive to a particular group of isolects. Such differences cannot provide evidence for a subgrouping argument. This is particularly clear from the phonological comparison of the isolects. Some features (e.g. JKT *ə|_* C#, IBN *b|a _ a*, MIN and BH antepenultimate

i/u) are retentions from PM and PMP. Other features are idiosyncratic innovations (e.g. SWY *r* vs *x*, IBN diphthongisation of final syllables) or seemingly shared innovations that are quite recent and, after close examination, turn out to be independent (e.g. diphthongisation of final-syllable vowels, which works in a different way for MIN and SWY, and which does not occur in all MIN subdialects, or not in the same form). Comparison on the lexical and morphological level yields a somewhat different picture: IBN is singled out as an isolect that underwent a rather different development from other isolects. Being originally an interior Bornean isolect, however, IBN was least affected of all the six isolects by the long-standing and variegated influences (Sanskrit, Javanese, Arabic, Portuguese) that had such a converging effect on the Malayic isolects in general.[206] In IBN, interference from SM has begun to play an important role only in the last century and a half. It is therefore not surprising that IBN shows some morphological differences, and that it scores low in cognate percentage with other Malayic isolects. And it is inversely not surprising that SWY and JKT seem to be much closer than BH or IBN to SM. For the speakers of SWY and JKT, SM has been a normative dialect to a much higher extent than for the speakers of IBN, MIN and BH.

It is quite possible that the Malayic-Dayak isolects form a separate branch within the Malayic linguistic group, and that SWY and JKT are more closely related than the other isolects to SM. However, in view of the distortional effects of the influences to which the other isolects have been jointly exposed, and summarising the above considerations, it seems too early to make an internal classification of Malayic isolects.

8.3 THE PM HOMELAND

Nothing definite is known about the original homeland of the Proto Malayic speakers. Kern (1917:119-120) proposed the Malay Peninsula as the most probable homeland of the Proto Malayic people (the 'Malays' in his publication), ruling out Sumatra on account of the SM word *səlat/an* (meaning both 'strait' and 'south', cf. 5.2.1). He considered an earlier migration from Borneo to the Malay Peninsula as improbable, mainly because it leaves open the question why a people living in a part of Borneo would bother to migrate overseas before it had colonised the remainder of the sparsely populated and fertile island. One would rather expect an external pressure for such a spread overseas, and Kern assumed that this pressure existed in mainland Southeast Asia.

The peoples living there must have pushed out an earlier Austronesian stock. That such a stock did live there is still witnessed by a large number of Austronesian loanwords in Kampuchean, Annamite, and Thai, "much more than can be explained from the current state of mainland Malayo-Polynesian [read: Austronesian] languages" (p.120). In other words, Kern believed that the Austronesian peoples originally lived in mainland Southeast Asia, and were pushed out from there by the peoples that are living there now. Some of them, the ancestors of the present speakers of Malayic isolects, went to the Malay Peninsula, from where some of them migrated to other areas such as Sumatra and Borneo.

There are several reasons for not accepting Kern's migration theory. As regards his assumption that the Austronesian homeland was on the Southeast Asian mainland, I refer to

---

[206]However, to some extent even the Malayic-Dayak isolects show these influences. It is very likely that most western (including SKT and AR) influence reached SM first, and affected the other isolects via the latter. This may also have happened with JV influence, although it is quite evident that BH, SWY, and JKT, were also directly influenced by JV.

more recent studies on this subject (Dyen 1965a; Dahl 1976:123-129). Suffice it to say that Austronesianists now generally believe that the early Austronesians spread from Taiwan to the rest of the Austronesian world, and not from the Southeast Asian mainland to insular Southeast Asia and Oceania. The arguments against Kern's assumption that the Malay Peninsula was the original Malayic homeland before migrations took place to Sumatra and Borneo are as follows. The term *səlat/an* replaced an older directional term **daya?*, which actually had a quite opposite meaning and, if anything, suggests a Malayic homeland which had the interior to the south, and the sea to the north (cf. 5.2.2). Furthermore, the way most Malayo-Polynesian peoples migrated or expanded their territories was by sea and not by land. Speakers of Malayic isolects were probably no exception to this, to judge from the seafaring tradition of many of them. There is no reason why a Proto Malayic people living in part of Borneo would have preferred territorial expansion in Borneo's almost inaccessible interior to maritime expansion. The Malay Peninsula shows the typical demographical pattern of a technically advanced coastal people (the Malays) in the process of pushing further into the interior a technologically less developed older stock (the Orang Asli, some of whom are still speakers of Austro-Asiatic languages).

In Borneo, a similar demographical movement can be observed, although here the Malays and Chinese have not penetrated as deeply into the interior. The important difference between the Malay Peninsula and Borneo, however, is that some of the autochthonous Dayaks of the interior are also speakers of Malayic isolects, viz. the Malayic-Dayak speakers. (The Ibans with their recent expansion to the coast are rather exceptional in this respect.) So, whereas in both cases Malays (and Chinese) have populated the coast and are slowly expanding towards the interior, some of the oppressed interior people in West Borneo are themselves speakers of Malayic isolects. Their Malayic isolects are autochthonous, and not the result of language shift, as for instance is the case with several Malayic isolects spoken in eastern Indonesia. Their authenticity is testified by their typically interior Borneo culture and by their languages. These show much variation among themselves, and have undergone hardly any Sanskrit or Arabic influence, which is in contradistinction to other Malayic isolects. The above facts show that Borneo deserves serious consideration as a possible Proto Malayic homeland.[207] Prentice (1978:19) believes that the core of the Malay language lay in the area around both sides of the Strait of Malacca, and he considers the coastal Malay isolects of Borneo a later offshoot. This does not contradict the possibility of Borneo as the Proto Malayic homeland, since Prentice's use of the term 'Malay' does not include the Malayic-Dayak isolects. It is likely that some of the coastal Malay isolects of Borneo are the result of back-migration, and that they were introduced from the Malay Peninsula.[208]

8.4 SUGGESTIONS FOR FURTHER RESEARCH

More material is needed in order to make a more accurate comparative-historical study of Malayic isolects. There is hardly any Malayic isolect that is not in need of a full grammatical and lexical description, and an immense task awaits those who are interested in the description, classification or comparison of Malayic isolects of the Malay Peninsula, the Sumatran, Borneo, or Javanese coasts, or eastern Indonesia. But even so it is clear that

---

[207]The idea that Borneo was the original PM homeland was first suggested to me by Robert Blust.
[208]cf. also Bellwood (in press). Clear evidence for a late introduction from the Malay Peninsula is the Malay spoken in and around Pontianak (West Kalimantan) (Fokker 1895). Jack Prentice informs me that Brunei-Malay speakers were originally speakers of the Bisayan language (still spoken in Brunei and neighbouring areas).

isolects like the Malayic-Dayak ones or Bacan show more retentions than others. These isolects stayed out of the cultural mainstream which determined the shape of most Malayic isolects, and therefore were able to retain PM characteristics that were lost elsewhere. The Malayic-Dayak ones are the more interesting because, apart from their independent development from other Malayic isolects, they also differ considerably from each other.

Another way to amplify the historical picture of Malayic isolects is to draw data from older texts, grammars and dictionaries into the comparison. There are many pitfalls in this sort of material, but with a critical approach to the drawbacks of text corruption or early linguistic and lexicographical deficiency one can benefit greatly from it.

An understanding of the history of Malayic isolects is also largely dependent on a comparative-historical study of Javanese. Javanese and Malayic isolects have been influencing one another for as much as a millennium, and by studying the history of Malayic isolects without simultaneously studying the history of Javanese isolects (or, for that matter, studying the history of Javanese isolects without at the same time taking the history of Malayic isolects into account), one is left with too many loose ends. Furthermore, a better understanding of this mutual influence would be of great interest to higher order reconstructions. Given the great influence which Javanese and Malayic isolects must have had on one another, and given the intensive influence from Malayic isolects on other Indonesian and Philippine languages, it is likely that the study of this Javanese and Malayic mutual influence will prove to yield some corrective viewpoints vis-à-vis PMP vocabulary (especially the part reconstructed by Dempwolff).

Finally, the affinity of PM with other languages should be tested in order to make a further classification of PM within the branch of Western-Malayo-Polynesian languages. Malayic shows a particularly great number of similarities in sound changes and in lexicon with Achehnese and Chamic.[209] Another apparently close relationship which deserves serious attention is that between Malayic and Balinese.

---

[209] cf. their reflexes of PMP consonants, their numeral system and their vocabulary for body parts. Many of these similarities are not shared by JV, SUN, and Madurese, which, together with SM, constitute the Malayo-Javanic subgroup on the basis of which PMJ was reconstructed (Nothofer 1975).

This appendix contains a list of disyllabic lexemes that do not conform to the tendency to consonant-harmony. The lexemes are grouped per isolect, and are subdivided according to their pattern (cf. 4.3.1). Where possible (and with an emphasis on the SM ones), the lexemes are provided with historical information. The following list is exhaustive, except for examples exhibiting combinations of *d*, *t* and *n*: these examples are all inherited from PMP and are designed to show that such combinations were permitted in the history of the Malayic isolects.

SM

N.B. A degree sign (°) at the upper left corner of a lexeme indicates that this lexeme is not found in Iskandar, and that it is of doubtful status in SM. The following lexemes occur in Iskandar and are not found in Wilkinson (1959): *bawel* 'talkative, quarrelsome' (< JV), *bempǝr* 'car bumper' (< DU), *kagok* I 'disturbed' (< JKT), *kagok* II 'differing from the general speech' (< JV), *mǝmpan* 'vulnerable; efficacious (medicine)', *mepet* 'squeezed' (< JKT), *mopit* 'Chinese writing brush' (< CHI).

Pattern I

*b*V*p*V(C):
bapa, bapa/ŋ, bapa/k 'father'
bepaŋ, bipaŋ, bepa, in *kueh bepaŋ*, '(k.o.) sweetmeat' (possibly < CHI, according to Wilkinson);
bopeŋ 'pockmarked' < CHI (Leo 1975:8);

*p*V*b*V(C):
°*pabu*, in *main pabu* 'somersaults and other tricks by Chinese tumblers' < CHI;
°*pǝbin* 'k.o. teetotum' < CHI;
°*pobien* (disyllabic?) 'wharf' < CHI;

*j*V*c*V(C):
jicuy 'opium dross doctored a second time for consumption' (*jiciŋ* in Iskandar) < CHI;

*c*V*j*V(C): --------

*g*V*k*V(C):
ge?kok, go?kek 'gecko, house lizard', an onomatopoeia from JV;

*k*V*g*V(C):
kaga 'no' (which is JKT and must be *kaga?*, cf. Abdul Chaer);
kaget 'startled' (a JKT lexeme borrowed from JV);
kagum 'astonished' (a JKT lexeme borrowed from JV);
kugah 'k.o. shrub'.

Pattern II

*b*V*mp*V(C):
bimpaw, bimpo 'handkerchief, towel' < CHI;

*p*V*mb*V(C):
°*pombak* 'dove, pigeon' < POR;

*j*V*ñc*V(C):
°*jiñceŋ* 'grateful' < CHI;

*c*V*ñj*V(C): ---------

*g*V*ŋk*V(C): ---------

*k*V*ŋg*V(C): ---------

Pattern III

*b*V*m*V(C):
bami (ba?mi) '(k.o.) noodle dish' < CHI;
°*beman*, in *kabur* °*beman* 'tax for the support of royal bandsmen in Old Perak', probably <
   ENG 'bandsman';
bima '(a proper name)', < SKT;
bomo, bomoh, bomor 'mage, sorcerer';
bumi 'earth' < SKT;

*p*V*m*V(C):
°*pama* 'police informer', probably < ENG 'informer';
pa?ma (with a medial cluster) 'the plant *Rafflesia hasseltii*' < JV;
pamah 'low-lying (land)';
paman 'uncle' < JV;
pamer 'talking big, boasting' < JV;
pamit 'beg leave or depart' < JV;
pamor, pamur 'alloyed iron' < JV;
°*pomaŋ* 'timber tree', probably a Jakun loan (cf. Adelaar 1983);

*j*V*ñ*V(C):
jaña 'say, think' < MIN (most likely < +*(u)jar* + *-ña*);

*c*V*ñ*V(C):
°*cuña* 'k.o. boat' (= JKT, < CHI);

*g*V*ŋ*V(C):
°*giŋin* 'k.o. herb';

*k*V*ŋ*V(C):
°*kaŋa*, in *damar kaŋa* 'k.o. tree';
kaŋar, in *laŋ kaŋar*, 'bird of prey';
kaŋən 'long or pine for' (JKT, borrowed from JV);
koŋek '(belonging to the) public' (JKT, borrowed from CHI).

Pattern IV

*m*V*b*V(C):
°*mabaw* 'evil spirit of disease';
°*mabay* 'k.o. tree';
mabir, in *tabir mabir,* 'all kinds of wall-drapery' (a derivation from *tabir*);
mabuk 'intoxicated';
mabub '(a proper name)' < AR;
mabur 'fly (v)' < JV;
mobil 'car' < DU;
mubal 'shoot up' < JV;
°*mubeŋ* 'circle, revolve' < JV;
°*mubyar* (with a consonant cluster; written 'mubiar' in Wilkinson) 'strike the eye, scream (of colour)' < JV;

*m*V*mb*V(C):
mambaŋ I 'spirit of the Indonesian pantheon'; II *ikan mambaŋ* 'a snapper', derived from *bambaŋ*; III '(title given occasionally to negrito headmen)';
mambu I *səmambu* 'Malacca cane'; II see *mabaw*; III 'smell (v)' < JV;
mambul 'k.o. climber';
mimbar 'pulpit in a mosque' < AR;
mumbaŋ 'coconut in its earliest stage of growth';
mumbul 'k.o. climber, *Milletia sericea*';

*m*V*p*V(C):
mapag, mapak 'go out and meet and then escort to one's house' (JKT, borrowed from JV or SUN);
°*mapar*, in *gaŋ mapar* 'flat-ended brazier's chisel or punch', derived from *papar*;
°*mapat* 'k.o. tree';
°*mapuk* 'k.o. herb';
°*məpas, məmpas* 'fish with the fly';
meper, mipir 'edge away under a blow', derived from *peper*;
mipis 'thin, tenuous', a variant of *nipis* and *tipis*;
°*mopeŋ* see *bopeŋ*;
°*mupar, ular təduŋ mupar* 'black cobra', derived from *upar*;
°*mupuh, mupus* 'k.o. tree';

*m*V*mp*V(C):
mampat I 'tightly packed'; II 'k.o. tree';
mampir 'stop; touch at', < JV (the usual SM term is *siŋgah*);
mampu(h) 'having the means for', a JKT lexeme (< SUN);
mampuŋ 'light and spongy in texture' (from *(h)ampuŋ* (?));
mampus 'die (vulgar); be wiped' (from *(h)ampus*);
°*məmpas* I see *məpas*; II *məmpas jantan* 'k.o. tree';
məmpat 'k.o. tree';
mempər 'bear some resemblance to' < JV;
°*məmpoŋ* 'k.o. tree';
mimpi 'dream' (derived from *impi*);
°*mompoŋ* 'used up (*nipah, pandan,* etc.) of which the fronds have been taken';
°*mumpun* 'blunt, stumpy' < MIN;

ñVjV(C): -------

ñVñjV(C): -------

ñVcV(C): -------

ñVñcV(C): -------

ŋVgV(C): -------

ŋVŋgV(C): -------

ŋVkV(C):

°*ŋakak* 'cackle (fowl)' from *kakak*, a JV onomatopoeia;
ŋəkek 'giggle', derived from *kek* (~ *kek, kekek, kekek-kekek, məŋgəlekek*);

ŋVŋkV(C): -------

Pattern V

(*bVwV*(C)):

bawa, bawaʔ, bawak 'convey' < PMP *baba;
bawab 'gatekeeper, porter' < AR;
bawah 'position under or below' < PMP *babaq (PM *bah);
bawal, in *ikan bawal* 'k.o. fish, pomfret' (<TAM? Wilkinson);
°*bawan* 'comrade, playfellow' (probably < ⁺*bau* + ⁺*an*, cf. *sə-baw-* 'alike');
bawaŋ 'bulb (more specifically onion)';
°*bawar* I 'customs barrier' (~ *gawar*); II 'sword of office' (~ *baur*);
bawat I *payuŋ bawat* 'state umbrella'; II 'drooping, inclining downwards'; III *tali bawat* 'braces';
°*bewah* 'feast for the dead (v)' (Kedah) (< ⁺*bər-arwah* according to Wilkinson);
°*bewak* (Kedah and Pattani) 'monitor lizard' (~ *biawak*);

pVwV(C):

paʔwa 'eldest uncle' (with a consonant cluster, < ⁺*bapak* ⁺*tua*);
pawah I *pawahkan* 'hire or lend on the metayer system, debtor and creditor sharing the proceeds'; II *rəmpah pawah* 'all kinds of curry stuff';
°*pawan* I 'title or appellation for Malay-speaking Indians'; II 'k.o. plant';
pawaŋ 'expert in any art believed to need the use of magic; guide; navigation officer, shipmaster; conductor' (< **(əm)pu* + **a(bw)aŋ*);
°*pawas* 'k.o. freshwater fish' < MIN;
°*pawat*, in *payuŋ* °*pawat* 'k.o. umbrella' (~ *bawat*);
paway 'insignia borne after a prince; insignia-bearers';
°*pawon* 'kitchen' < JV;

mVwV(C):

°*mawa* 'gibbon' (~ *waʔwaʔ*);
mawar I 'rosewater' < AR; II *tawar mawar* 'harmless, nullified' (derived from *tawar*);
mawas 'orang hutan' (~ *mayas* (in Kalimantan));
maway 'k.o. shrub';
mawin, in *kawin mawin* 'marriage festivities of all sorts', (derived from *kawin*);
mewah 'plenteous'

mewek 'pursing up the mouth, pouting'; a JKT lexeme according to Iskandar, and ultimately from JV according to Kähler;
°*mewer* 'sob' < JV;

*j*VyV(C):
jaya 'triumphant' < SKT;
jayeŋ 'victorious over' < JV (< *jaya* + *iŋ*);

*c*VyV(C):
caya 'lustre, glow, brightness' < SKT;
°*cayah* 'careless, neglectful' (~ *cuay*) < MIN;
cayar 'diluted, watery (of viscous things)' (~ *cair*, < PMJ *caiR*$_{13}$, Nothofer 1975:165);
°*cayu* 'sitting-mat' (~ *siu*);

*ñ*VyV(C):
°*ñaya* 'injustice' (~ *aniaya*, < SKT);
°*ñeyag* 'snap (as a dog)', (JKT, ultimately < SUN).

Inherited combinations of *d, t* and *n*:

Pattern I

dataŋ 'come' < *datəŋ*;
datar 'level, flat' < PMP *DataR*;
datu, datuʔ, datuk, datuŋ 'chief, head of the family; grandfather, ancestor' < *datu*;
dətik 'ticking sound' < PMP *detik* (Blust 1970);
təduh 'abated, calm (wind)' < PMP *te(n)duq*;
tidur 'sleep' < PMP *tiDuR*;
tudiŋ 'aslant, at an angle' < PMP *tudiŋ* 'indicate';
tuduh 'accuse' < PMP *tuduh* 'indicate'.

Pattern II

tanda 'sign' < PMP *ta(n)da*;
tandas 'state explicitly' < PMP *(CtT)a(n)(dDj)es* (Blust 1970);
tandiŋ 'division into equal parts, compare' < PMP *(Ct)anDiŋ* 'equal, compare' (Blust 1970);
tandu 'hammock-litter' < PMP *tandu*;
tindas 'crush' < PMP *tiDes* (Blust 1970);
tindih 'lie in heaps' < PMP *ti(n)diq*;
tunda 'drag (v)' < PMP *tunDa*;
tunduk 'bend down, bow' < PMP *tu(n)Duk*;
tondoŋ 'chase away, oust' < PMJ *tunḍuŋ*.

Pattern III

danaw 'lake, pool' < PMP *Danaw*;
tanah 'land, earth' < PMP *taneq*;
tanak 'cook rice' < PMP *tanek*;

tanam 'plant (v)' < PMP **tanem*;

tənar 'be publicly known' < PMP **teneR* 'voice' (Blust 1970);

tənun 'weave' < PMP **tinuʔun*;

tənuŋ 'gaze fixedly, diagnose illness, divine' < PMP **(Ct)e(nN)uŋ* 'find by divination' (Blust 1970);

tuna 'eel' < PMP **tuna*;

tunas 'shoot, bud' < PMP **tunas*;

tunay 'cash, ready money' < PMP **tu(nN)ay*;

tunu(n) 'burn up' < PMP **tunu* 'burn, fry'.

Pattern IV ---------

MIN

Pattern I

*b*V*p*V(C):
bapo, bapaʔ 'father'.

Pattern II ------

Pattern III

*b*V*m*V(C):
bumi 'earth, land' < SKT;

*j*V*ñ*V(C):
ja/ño (~ *ja/no*) 'he/she says, said' (= *ja* + -*ño*).

Pattern IV

*m*V*b*V(C):
mabaw, si- 'k.o. evil spirit';
mabuʔʔ 'drunk, intoxicated';

*m*V*mb*V(C):
mambaŋ 'ghost, spirit';
mambu, in simambu 'k.o. rattan';
mimba 'pulpit in a mosque' < AR;
mumbaŋ 'young coconut';

*m*V*p*V(V): -------

*m*V*mp*V(C):
mampeʔ 'in order, fine, excellent';
mampuyh 'dead';
mimpi 'dream';
mumpun 'blunt'.

Pattern V

*b*V*w*V(C):

bawa I 'k.o. fish (*Stromateus*)'; II 'wrapping around the hilt of a chisel, a hoe or a stick to
   protect it from cracking or breaking';

bawah 'under, below';

bawaŋ 'bulb';

*p*V*w*V(C):

pawa 'ruminate';

pawaŋ 'expert in magic; guide, shipmaster, shaman, trapper';

paweh 'k.o. sweetwater fish';

*m*V*w*V(C):

maweh 'orangutan';

*c*V*y*V(C):

cayah 'forgetful, indifferent';

cayo 'shine, lustre, glitter' < SKT.

Combinations of *d, t* and *n*:

Pattern I

data 'flat, level';

data? 'a cracking sound';

dataŋ 'come';

dateh 'become weak and slow (e.g. of breathing)';

datiᵊ? 'ticking sound of watch or clock';

dato? 'clear';

datu I 'knowledgeable about medicinal herbs'; II 'palm-fibre thread';

datuᵊ? I 'traditional chief'; II 'grandfather (in Simabur)'; 'Sir (in Payakumbuh)';

datuy?, in *datuy?kan jari* 'flick or stretch one's fingers to make a sound';

datuyh I 'popping or snapping sound'; II *ati badatuyh* 'have doubts';

deta 'head-cloth' < Persian;

doto 'doctor, physician' < DU;

tadah I 'saucer'; II 'be visible, clear';

tadi 'recently, just a while ago';

tadiᵊ 'wall of plaited lath';

tadin (in Bonjol) see *tadi*;

tado 'thrash, beat';

taduᵊh 'calm, quiet';

taduᵊŋ 'k.o. snake';

tida? 'not';

tido '(there is) no, not';

tiduᵊ? 'sleep (v)';

todaᵊ? 'a swordfish';

todaŋ 'disappeared, gone, lost';

tudi, in *basitudi* 'behave in a disturbing way, touch everything';

tuduᵊh 'accusation';

tuduŋ 'anything used to cover or protect'.

Pattern II

dantam 'roar, boom (like the sound of a cannon)';
dantaŋ 'the sound of a gun';
danteŋ 'a sound (lighter than *dantaŋ*)';
dantiᵊŋ 'a sound (lighter than *danteŋ*)';
danto 'ivory';
dantuᵊŋ 'sound of a heartbeat';
tandah 'deck (of a ship, etc.)';
tanday I 'lavatory, latrine'; II 'sign, token, indication';
tandan 'cluster of fruit';
tandaŋ 'visit (v)';
tandeh 'all gone, clean, all spent';
tandiᵊ, in *tandiᵊ majaiᵊ* 'courtiers carrying the king's mirror';
tandiᵊŋ 'equal, match; counterpart, partner, opponent';
tando 'sign, token, indication';
tandu 'palanquin';
tanduᵊʔ 'horn';
tenda 'tent' < POR;
tendaʔ, in *batendaʔ* 'hold a competition';
tendeʔ 'addition, supplement';
tendoh see *tundiʔ*;
tindaŋ, in *batindang* 'disappear';
tindeh 'pinch, squash (with the nails), suppress';
tindiᵊh 'press heavily on, oppress';
tindiᵊʔ 'pierce (an earlobe)';
tindin I *manindin jo taŋan* 'use one's fists often'; II 'low tones heard between high ones (e.g. as in the gamelan)';
tindo (Koto Tuo) 'last fruits (often small); youngest children';
tondiʔ, *tundiʔ* I '(of a married couple) going somewhere together, go out together';
tundiʔ I see *tondiʔ*; *tundiʔ* II 'bore, pierce in upward direction with s.th. pointed';
tundo 'push forth';
tunduᵊʔ 'bow, stoop; submit to';
tunduŋ 'neck'.

Pattern III

danaʔ I, *denaʔ* 'dwarfish, stunted';
danaʔ II 'nearest point to the target in certain throwing games';
danaw 'lake, pool';
dano 'result';
danuʔ, in *badanuʔan darah iduᵊŋ-ño* 'have nosebleed';
denay (den, deyen) 'I';
denaʔ see *danaʔ* I;
dunia 'world, earth' < AR;
tanah 'land, ground';
tanay 'carry on the palm of the hand';
tanaʔ 'cook rice';

tanam 'plant';
tanaŋ 'calm, quiet, silent';
tanaw 'k.o. parakeet';
tane⁷ I, *tani⁷* 'spin around like a top';
tane⁷ II, *taŋe⁷* 'impede, obstruct, stop s.o. from doing s.th.';
tanuᵊ⁷ 'tapir';
tanuᵊŋ 'divine (v)'; .
tanun 'weave';
tano⁷ 'deadly quiet, desolate';
teno⁷ 'aim, point at';
tenoŋ 'k.o. basket';
tino (an abbreviation of *ati-no*);
tone⁷ 'try to please, avoid vexing or annoying people';
tunuᵊŋ 'desperate, flurried'.

Pattern IV

nata 'point, dot' < AR;
nanta⁷ (in Suliki) 'visible';
nanti, nanti⁷ 'soon, later';
nantun 'that' (from *naN + tu(n)*).

BH

Pattern I

bVpV(C):
bapa 'father';

pVbV(C): -------
pabrik 'factory' < DU;

(jVcV(C)): -------

(cVjV(C)): -------

(gVkV(C)): -------

(kVgV(C)):
kagum 'astonished' < JV.

Pattern II ------

Pattern III

bVmV(C):
bima '(name of a character from the *wayang*)' < SKT;
bumi 'world, earth' < SKT;

*p*V*m*V(C):
paman 'uncle' < JV; (BK *pamor* 'luck, fortune; magical power');
pamuŋ 'caretaker, supporter, mentor, teacher' < JV;

*j*V*ñ*V(C):
jaña '(s)he says, said' (from +*ja* + +*ña*);

*c*V*ñ*V(C): -------

*g*V*ŋ*V(C):
gaŋan 'k.o. curry';

(*k*V*ŋ*V(C)):
kuŋris 'congress' < DU.

Pattern IV

*m*V*b*V(C):
mabuk 'crazy about s.th.';

*m*V*mb*V(C):
mambu 'I 'put on airs'; II 'properly speaking, after all';

*m*V*p*V(C):
mapa 'how';

*m*V*mp*V(C):
mampan 'have effect, hit, strike';
mampu 'rich, well-to-do';
mampus 'dead';
mimpi 'dream (n)';

*ñ*V*j*V(C): -------

*ñ*V*ñj*V(C): -------

*ñ*V*c*V(C): -------

*ñ*V*ñc*V(C): -------

*ŋ*V*g*V(C): -------

*ŋ*V*ŋg*V(C): -------

*ŋ*V*k*V(C): -------

*ŋ*V*ŋk*V(C): -------

Pattern V

*b*V*w*V(C):
bawa 'carry';
bawah 'under, below';
bawaŋ 'bulb';

*p*V*w*V(C):
pawa 'room, gap, open space';

*m*V*w*V(C):
mawah 'worried, afraid';
mawar 'rosewater'.

Combinations of *d, t* and *n*:

Pattern I

dataŋ 'come';
datar 1. 'level, flat' 2. 'same, equal';
datik 'a ticking sound';
datu 'head, chief, elder; grandfather, ancestor';
tada 'poisonous, strong, vigorous';
tadah 'cistern, tank, reservoir';
tadas 'have hold on, have effect on, hit, strike';
tadi 'recently, a while ago';
tadih see *tadi*;
taduh 'calm down; stop (rain)';
taduŋ 'k.o. snake';
tuduh, in *manuduhi* 'advise, lead, give indications';
tuduŋ 'cover, lid'.

Pattern II

dintu 'in such a way';
duntu 'tea';
tanda 'sign, token, indication';
tandak 'deposit, sediment';
tandar 'move, shift, rub, grate';
tandik, in *batandik* 'jump up and down';
tandiŋ 'equal, peer, match; opponent, counterpart';
tandu 'palanquin';
tinda 'tent' < POR;
tindas 'suppress, oppress';
tindih 'lie on a heap, on top of each other';
tindik 'pierce earlobes';
tunda 'drag, postpone';
tunduk 'bend, bow, submit to';
tundun 'a bunch of bananas'.

Pattern III

dana 'gift, alms' < SKT;
danak 'short and fat';
danaw 'lake, pool';
dinar 'dinar' < AR;
dinas 'service' < DU;
dini 'this';

dini/hari 'dawn';
tanah 'ground, land';
tanay 'catch, intercept, receive';
tanak 'cook rice';
tanam 'plant (v)';
tanaŋ 'calm, quiet, silent';
-tani '(farming)' < SKT;
tanis, in *batanis,* 'humid, wet, watery';
tanun 'weave'.

Pattern IV

nadar 'vow to Allah' < AR;
natal 'Christmas' < POR;
natu (~ *na/itu* 'that one, the one over there'));
nitral (with a consonant cluster) 'neutral' < DU;
nitu (~ *na/itu*) 'that one, the one over there'.

SWY

Pattern I

*b*V*p*V(C):
bapa?, bapaŋ, bapo 'father';

*p*V*b*V(C): ------

*c*V*j*V(C): ------

*j*V*c*V(C): ------

*g*V*k*V(C): ------

*k*V*g*V(C):
(BSM *kagul* 'confused').

Pattern II

*b*V*mp*V(C): ------

*p*V*mb*V(C): ------

*j*V*ñc*V(C): ------

*c*V*ñj*V(C): ------

*g*V*ŋk*V(C): ------

Pattern III

*b*V*m*V(C):
bumi 'earth' < SKT;

*p*V*m*V(C):
pama(h) 'swamp';
pamur 'alloyed iron' < *j*V;

*j*V*ñ*V(C):   ------

*c*V*ñ*V(C):   ------

*g*V*ŋk*V(C): ------

*k*V*ŋg*V(C): ------

Pattern IV

*m*V*b*V(C):
mabuᵊʔ 'drunk, intoxicated';
mibar 'whirl, flutter about (e.g. bats)';

*m*V*mb*V(C):
mumbaŋ 'very young, of fruits'; the coconut in its first stage of development';
mumbo 'k.o. edible plant';

*m*V*p*V(C):
mupur 'ruffle the feathers in defensive position (of hens etc.)';

*m*V*mp*V(C):
mampus 'wasteful, dissipating';
mimpi 'dream (n)';

*ñ*V*j*V(C):   ------

*ñ*V*ñj*V(C): ------

*ñ*V*c*V(C):   ------

*ñ*V*ñc*V(C): ------

*ŋ*V*g*V(C):
ŋiguᵊʔ 'walk like a duck or a goose'.

Pattern V

*b*V*w*V(C):
bawa(h) 'under, below';
bawaŋ 'onion';
bawaʔ 'skin';
bawo 'carry, bring (along)'.

Combinations of *d, t* and *n*:

Pattern I

dataŋ 'come';
datax 'level, flat';
dəta?, bədəta? 'knock with a hammer';
dətas 'the sound of paper or leaves being cut';
tada(h) I 'catch from below, intercept'; II 'pervade with force (e.g. the wind)';
tadi 'recently, now';
tədo 'ask for (literary style)';
təduᵊ(h) 'calm, quiet (wind, waves)';
tiduᵊ? 'sleep (v); set (of oil)';
tuduᵊ(h), in *bətuduᵊ(h)*, 'be on first names';
tuduŋ 's.th. used to cover or protect'.

Pattern II

dəntam 'heavy crashing sound as of a tree falling';
dəntaŋ 'onomatopoeia for the sound of the town crier's gong';
dəntum 'a thud, as of a heavy body falling';
tandan I 'a cluster of fruit'; II 'rope to tie up horned stock';
tandaŋ 'visit, meet without particular purpose or aim';
tandi 's.th. put next to s.th. else in order to compare';
tando 'sign, token, indication';
tandu 'palanquin';
tanduᵊ? 'horn';
tindiᵊ(h) 'lie in heaps';
tinda [sic] 'k.o. cotton curtain' (< POR);
tindan/an 'parrot's perch';
tindi? 'pierce through';
tunday 'follow';
tundo 'show, demonstrate s.th.';
tunduŋ 'cede, withdraw, go away';
tunduᵊ? 'bent down; bow (head, body)'.

Pattern III

danaw 'lake, pool';
dənaŋ 'swim' < *(mb)A-rənaŋ*;
tana(h) 'ground, land';
tana? 'cook rice';
tənaŋ 'calm, quiet (water)';
tənun 'weave';
tənuŋ 'divine, prophesy';
tuna? 'be/stay together';
tunaŋ, in *nunang gadis* '(of a boy to the girl he wants to be engaged to) give five dollars as a guarantee and a token';
tunas 'bud of a plant';

tunay 'cash';
tunu, in *nunu* 'be afire; set fire to';
tunun see *tunu*.

Pattern IV

nədas 'make a crackling sound';
nantu 'child-in-law' < PMP **b/in/antu*.

IBN

Pattern I

(*bVpV*(C)):
bapaʔ 'term of address for father-in-law';

pVbV(C): -------

jVcV(C): -------

cVjV(C): -------

gVkV(C): -------

kVgV(C):
kigal 'bounce (v)'.

Pattern II

bVmpV(C): ------

pVmbV(C):
pambam 'the game of hide-and-seek';
pambar I 'shattered'; II 'dispersed, shattered';
pambuʔ 'fruit bad inside; (in songs) dead';
pambur 'burst';
pambus 'punctured';

jVñcV(C): ------

cVñjV(C): ------

gVŋkV(C): ------

kVŋgV(C):
kaŋgan 'black cotton cloth';
(*kəŋgay* 'complain, be unwilling to' < *k-əŋgay*, cf. 3.1.1).

Pattern III

b-umay 'farm (v)', derived from *umay*;
bumi 'earth' < SKT;
pama 'good, fine';

pama⁷ 'k.o. frog';
pamur 'cloudy (of water)';
pəmai⁷ I 'inheritance'; II 'disability'; III 'choice, decision';
p-umay, in *di-p-umay* (an object-oriented form of *b-umay*);
kaŋaw, ŋaŋaw 'call, shout'.

Pattern IV

mVbV(C):
mabuk 'drunk';
mabuŋ, in *təkuyuŋ mabuŋ* 'the snail has left its shell (died)';
mabu⁷ 'shallow';

mVmbV(C):
mambam 'dull (weather)';
mambuŋ 'k.o. plant, burnt to drive off insects';
mumban 'k.o. water plant';

mVpV(C):
mapap 'foolish, inconsequent';

mVmpV(C):
mimpi 'dream (v,n)';

ñVɟV(C): ------

ñVñɟV(C): ------

ñVcV(C): ------

ñVñcV(C): ------

ŋVgV(C):
ŋagay 'to(wards)', derived from *gagay*;
ŋigal 'bounce', derived from *kigal*;
ŋigaw 'walk, grope about (in sleep)';

ŋVŋgV(C): ------

ŋVkV(C):
ŋəkal 'gasp in paroxysm of crying';
ŋəkuk 'make the noise of a *kok-lir*';
ŋikil 'giggle, titter (v)';

ŋVŋkV(C):
ŋuŋkat 'have a relapse'.

Pattern V

bVwV(C):
bawa⁷ 'k.o. ant';
baway I *igi baway* 'swollen glands' II 'tired, stiff';
bawaŋ 'onion';

*p*V*w*V(C):
pawik, in *mawik* 'put or hold hands behind back' (also *pəlawik, məlawik*);

*m*V*w*V(C):
mawa 'unsettled, wandering';
mawah 'blow away';

*j*V*y*V(C): -------

*c*V*y*V(C):
caya 'brilliance, brightness' < SKT;
cuyukcuyuk 'cower, flinch';

*ñ*V*y*V(C):
ñaya 'spill'.

Combinations of *d, t* and *n*:

Pattern I

datay I 'come, arrive'; II *natay* 'report';
datas 'above', from *di atas*;
datu? '(in songs) nobleman, chief';
d/itu? 'here';
tada 'a cock's spur';
tadi? 'indicating recent occurrence or previous mention';
təda 'remains, s.th. left over';
təduh 'calm (of the sea), ceased (of rain)';
təduŋ 'a cobra';
tudah '(preceding a name) poor', e.g. *tudah Iñjat* 'poor Iñjat';
tuduh I 'instructions (for doing s.th.)'; II 'leaky, leaking (of houses)';
tuduŋ 'a cover, lid'.

Pattern II

tanda I 'a sign, mark'; II *tanda sirat* 'the embroidered end of a loin-cloth';
tanda? 'dances (of various kinds)';
tandan 'a bunch, cluster, the whole bunch of bananas';
tandaŋ, in *nandaŋ* 'take on a visit, tour (cock fighting etc.)';
tandas I 'chopsticks'; II 'close ([of] cut[ing])';
tandin, in *nandin* 'hold the foot and kick with the knee (game)';
tanduh 1. 'spout' 2. 'penis' 3. 'term of endearment for boys';
tanduk I 'horn'; II *nanduk* 'cup (v)';
tandu? 'do s.th., use s.th. again';
tənday 'the warp beam (weaving)';
təndaŋ, in *nəndaŋ* 'kick';
təndu?, təndur 'slack (rope, etc.)';
tinda '(in songs) *tinda nuan* 'you (respectful)';
tinduh 'edible maggots';
tinduk 'sleep';

tunda? 'imitate, mimic, follow (advice)';
tundan, in *taŋga tundan* 'a platform erected for felling a tree';
tundi? I 'play, joking, fooling'; II *nundi?* 'try to wheedle, cajole';
tunduk I 'yield, submit; bend (the neck)'; II 'wheedle'.

Pattern III

dana '(male proper name)';
danan 'k.o. rattan';
danaw 'pool, puddle';
dani 'awake';
dini 'where?';
dini/hari 'the time before dawn, 3-5 a.m.';
tanah 'earth, land';
tanak 'fry in oil';
tanam 'plant (v)';
tənun 'weave';
tənuŋ, in *nənuŋ* 'divine';
tunay 'ready money, cash down';
tunaŋ 'a betrothal gift';
tunu 'burn, roast'.

Pattern IV

naday '(have, there is) not, no';
nandan '(in songs, of time) exactly'.

Pattern V ---------

N.B. The following entries are from Bruggeman (see appendix in Scott 1956):

təday see *təda;*
ŋagu 'tremble, shiver with cold';
mawaŋ 'a species of mango'.

JKT

Pattern I

bVpV(C):
bapa? 'father';
ba?pao (with consonant cluster) 'k.o. food preparation';
bapèt 'without money';
baplaŋ 'thick and large (of a moustache)'
bipak 'barracks', < DU;
bòpèŋ 'pockmarked' < CHI (Leo 1975);

bòpòŋ I 'back' (~ *bòkòŋ*); II 'hold (a baby etc.) in one's arms';
bupèt 'k.o. cupboard', < DU;

*p*V*b*V(C): --------

*j*V*c*V(C):
jicap 'twenty (rupiah)' < CHI;
jicapgò 'twenty-five (rupiah)' < CHI;

*c*V*j*V(C): --------

*g*V*k*V(C): --------

*k*V*g*V(C):
kaga?, kagè 'no, not';
kagèt 'startled' < JV (< +*ka-giat*);
kagòk 'difficult, impeded'.

Pattern II

*b*V*mp*V(C):
bèmpər 'car bumper' < DU < ENG;

*p*V*mb*V(C): --------

*j*V*ñc*V(C): --------

*c*V*ñj*V(C): --------

*g*V*ŋk*V(C): --------

(*k*V*ŋg*V(C)): --------

Pattern III

*b*V*m*V(C):
bémò 'motorised pedicab', from *bècak mòtòr*;

*p*V*m*V(C):
pamèr, in *pamèrin* 'exhibit', *pamèran* 'exhibition';
pamòr 'lustre, splendour, shine';

*j*V*ñ*V(C): --------

*c*V*ñ*V(C): --------

*g*V*ŋ*V(C): --------

*k*V*ŋ*V(C):
kaŋən 'long, pine for' < JV.

Pattern IV

*m*V*b*V(C):
mabòk 'drunk, intoxicated';
mabruk (with a consonant cluster) 'ugly, of inferior quality';

mèbəl 'furniture' < DU;

mòblòŋ (with a consonant cluster) 'open (clothes)';

*m*V*mb*V(C):

mambu 'stink';

məmbal 'springy, elastic';

məmbèl 'bent down, weighed down, lowered because pressed down';

mimbar I 'pulpit in a mosque'; II 'podium, platform, forum';

mumbul 'rise, get up, boom';

*m*V*p*V(C):

mèpèl 'knead dough for cake, make patterns in dough';

mèpèr 'rub, clean one's dirty hands with a cloth, towel etc.';

mòpit 'k.o. Chinese writing brush';

*m*V*mp*V(C):

mampət 'stopped up, stagnated', < *pampət*;

mampu 1. 'able, capable' 2. 'rich, well-to-do';

məmpan I *ŋgè?məmpan* 'invulnerable'; II 'efficacious (medicine)';

məmpət 'feel annoyed' (from *əmpət*);

məmplək (with a consonant cluster) 'keep or heap up in a random way';

mimpi I 'dream (n)'; II 'dream, thought';

mumpuŋ 'as long as';

*ñ*V*j*V(C): -------

*ñ*V*ñj*V(C): -------

*ñ*V*c*V(C):

ñicip 'taste, try a bit (food)' (from *cicip*);

*ñ*V*ñc*V(C): ------

*ŋ*V*g*V(C):

ŋègaŋ 'walk in a staggering way and with legs apart';

ŋègòt 'walk like a duck' < *ègòt*.

Pattern V

*b*V*w*V(C):

bawè 'carry';

bawa?an '1. s.th. usual 2. usually';

bawè 'below, under';

bawèl 'quarrelsome, talkative';

('bowès' I and II have a non-phonemic glide and should phonemically be analysed as *boès*);

*p*V*w*V(C): -------

*m*V*w*V(C):

mawar, aèr mawar 'rosewater';

mèwa [sic] 'wealthy, luxurious';

mèwèk 'cry, weep';

*j*V*y*V(C): -------

*c*V*y*CV(C):
caya⁷ 'believe' < SKT;

*ñ*V*y*V(C): -------

Combinations of *d, t* and *n*:

Pattern I

datar 'level, flat';
datəŋ 'arrive; from';
datu⁷ 'grandfather';
dətik I 'ticking sound of a watch'; II 'second (n)';
dustè (with a consonant cluster) 'lie, cheat (v)';
tada, tadè I 'cistern, reservoir, tank'; *tukaŋ tadè* 'receiver of stolen goods'; II *tadè ujan* 'topmost hand of bananas in bunch';
tadé, tadi 'lately, just now';
tèdèŋ 'cover-up, shield (for a secret or s.th. bad)';
tədu 'become quiet, calm';
tida⁷ 'no, not';
tidur 'sleep (v)';
tòdòŋ 'threaten with a weapon';
tudahan, tudəhan [sic] 'used goods';
tudiè 'that's her/him, there he/she is' (< *itu diè*);
tudiŋ 'point one's finger at s.th.';
tudu 'accuse';
tuduŋ 1. 'hat' 2. 's.th. used to wear, protect or cover'.

Pattern II

dəntum 'sound of cannon fire';
tandak, in *nandak* 'dance (v)';
tandan I *tandanan* 'long central stalk in bunch of fruit (e.g. bananas, coconuts, areca nuts)'; II 'family, relatives';
tandè 'sign, indication, symbol';
tandəs I 'all gone, clean'; II *tandəsin* 'say s.th. explicitly'; III *kətandəsan* 'exaggerated, overdone, beyond the limit';
tandiŋ I 'arrange according to size, make s.th. match in size, place in lots (one's merchandise)'; II *bətandiŋ* 'measure one's strength in games';
tandòn 'guarantee, surety';
tanduk 'horn';
tandur, in *nandur* 'grow rice (move it from a nursery bed to a ricefield)';
təndaŋ 'kick (v)';
tèndè 'tent' < POR;
tindak I 'step, pace';II *tindaktanduk* 'behaviour, acting';
tindəs I 1. 'lying one on the other, superincumbent'; 2. 'crush, oppress'; II *indəsan* 'copy (of s.th. written)';
tindi 'lying on one another';
tindik, in *nndik* 'bore an earlobe';

tundanan 'railway track'? (cf. Abdul Chaer: *'jalan kereta api rem* [sic] *dsb.'*);
tundaŋan 'fiancé(e)';
tundè 'postpone'.

Pattern III

dinəs 'service, duty' < DU;
dènòk 'short and fat, corpulent';
tanè 'earth';
tanəm 'plant (v)';
tani, in *p aʔtani* 'farmer';
tənaŋ 'be quiet';
tènòŋ 'k.o. basket' (of JV make, see Wilkinson);
tənun 'weave';
tənuŋ 'divine (v)'.

Pattern IV

nanti 'soon, later';
natar I *natar(an)* 'front-yard of a house'; II 'ground colour (of a design)';
nòŋtòt 'emerge, protrude (from a pocket or hole)';
nòtòk 'knock hard', from *tòtòk*;
nòtès 'notebook' < ENG (?);
nòtun 'two rupiah' < CHI;
nutug, nutuk I 'full'; II 'true, perfect'.

APPENDIX II

The following list contains all disyllabic lexemes with a combination of *s* and a palatal in initial and intervocalic position occurring in the isolects.

SM

N.B. a degree sign (°) at the upper left corner of a lexeme indicates that this lexeme is not found in Iskandar, and that it is of doubtful status in SM. The following lexemes were taken from Iskandar, and are not found in Wilkinson (1959): *sicə* 'set of table and chairs', *siñal* 'signal' (both < DU), *secaŋ* (~ *sacaŋ*, see below), and *señjoŋ* (~ *səñjiŋ*, see below).

(*jVsV(C)*):
jasa I 'loyal service, doing duty zealously' < SKT; II (~ *jaksa*) 'prosecutor' < SKT;
jasad, jasat 'body' < AR;
jisim 'physical body, body (in the mathematical sense)' < AR;
jose, in *kain jose* 'Chinese silk crape' < CHI;
jusuh '(colloquial form of the proper name Yusuf)' < AR;

(*sVjV(V)*):
saja (~ *sahaja*) 'intentionally; only' (< SKT);
sajak I (also spelled *saja*) 'assonance, cadence, melodious harmony, rhythm' < AR; II see *səjak*;
sajaŋ 'spirituous liquor' < JV;
saji I 'served up, dished up, dressed (of food courses arranged on dishes)'; II *main saji* 'a form of entertainment';
°*sajuk* see *səjuk*;
səjak 'since';
səjam see *suñjam*;
səjat 'be squeezed or otherwise disposed of (liquid)';
səjuk 'coolness, pleasant lowering of the temperature';
səjud see *sujud*, < AR;
sijil 'scroll, certificate, written record < AR;
soja, sojah 'bowing the body in salutation (as is done by the Chinese)' < CHI;
sujen 'small pike, spit, splinter' <JV;
suji I 'k.o. granular meal' < Urdu; II 'embroidery' < Old Javanese;
sujud 'kneeling and bowing the head to the ground' < AR;
sujut, see *sujud*;

(*sVcV(C)*):
°*sacaŋ*, in *kayu* °*sacaŋ* 'sappanwood' < JV;
suci 'pure, holy' < SKT;

231

232

(sVñjV(C)):
°sañja I 'look in on a person' < JV; II see sənja;
sañjay 'well-made, well-proportioned (men)' < MIN;
sañjuŋ 'flattery, praise, making much of' < MIN;
sənja 'evenfall, about 5.30 p.m.' < SKT;
°sənjah 'grab, snatch angrily';
sənjak (səjak, səmənjak) 'since';
sənjaŋ 'differing, unlike' < MIN;
sənjiŋ 'small metal saucer';
sənjuŋ I 'bar of a balance'; II see sənjiŋ;
°siñjaŋ 'a Javanese sarong';
°siñjuh 'nudge (v)';
suñjam, in tərsuñjam 'fallen or held head downwards';

sVñcV(C)):
°sañcoŋ 'bury, inter' (JKT, and borrowed from CHI (Wilkinson));

(sVñV(C)):
saña, bahwasaña 'verily, of a truth' (short for bahwa səsuŋguhña);
°sañok 'still, lonely' < MIN;
sənak see sənap I;
sənap I suñisənap 'quite deserted, lonely in the extreme'; II sənaptidur 'sound asleep' III
    'shut up! be silent!' (to a child);
sənar 'tingling; the sensation when the funny bone is knocked';
sənuh 'snatch hastily';
señor, señur 'Mr, Sir' < POR;
sənum 'smile, smiling';
siño see señor;
siñuh I see señor; II see siñjuh;
suñi 'lonely, desolate';

yVsV(C):
yasin 'one of the chapters of the Koran' < AR;

sVyV(C):
saya I (~ sahaya) 'I, me; servant' < SKT II 'sarong waxed to a billowy form < JKT (< POR
    saia 'skirt');
sayak I 'a hemispherical bowl of coconut shell' < MIN; II 'stiff pleated sarong' (~ saya II);
sayaŋ I 'pining, longing, pitying; longing, affection'; II tiaŋ sayaŋ 'derrick'; III 'k.o. tree'
sayap 'wing';
sayat 'slicing off';
sayip (in Pahang) 'of royal blood on both sides < AR;
sayu I 'melancholy, plaintive'; II '(title for Vaisya ladies in Bali, short for gusti ayu)';
sayuŋ I 'doing a thing unevenly or crookedly'; II 'k.o. oar'; III sayuŋ tikus 'k.o. grass'; IV
    '(originally: name of a place in Acheh)';
sayup 'barely perceptible; just fading away in the distance';
sayur 'green food, edible vegetables';
soyak, soyat 'rending from end to end, tearing in two';
soyok (written 'soyo') see suyuk;
suyak see soyak;

suyar (~ *laŋsuyar*) 'vampire';
suyuk 'projecting part of a roof at the narrow ends of a Malay house' < MIN.

MIN

jVsV(C):
jasa 'prosecutor' < SKT;
jaso 'service' < SKT;

(sVjV(C)):
sajak I 'since, from'; II 'rhyme, verse';
saji (in speeches) 'course, dish, food';
sajo 'only, but';
sajuᵊʔ 'cold, chilly';
suji 'embroidery' < Old Javanese;
sujuyʔ 'prostration';

sVcV(C):
suci 'pure, clean, sacred';

sVñjV(C):
sañjay 'tall but not slim, well-built';
sañjo 'dusk';
sañjuᵊŋ 'praise, flatter';
señjaŋ 'aslant, not parallel';

sVñV(C):
sañun 'smile';
siño 'Indo-European boy' < POR;
suñi, suni 'silent, lonely, deserted';

sVyV(C):
sayaʔ '(piece of a) coconut shell';
sayaŋ 1. '(it is a) pity'; 2. 'love';
sayo (in Lintau and Buo) 'I, me' < SKT;
sayuyʔ 'indistinct, vague; scarcely, hardly';
sayoʔ 'wing';
suyaʔ (~ *soeʔ*, ~ *kuyaʔ*) 'torn';
suyuᵊʔ, soyoʔ 'projecting gable'.

BH

jVsV(C):
jasa 'service' < SKT;
jasat 'body' < AR;

sVjV(C):
saji I (~ *sadi*) 'ready, willing'; II *sasaji* 'sacrifice' < SKT;
sujut 'prostration' < AR;

*s*V*c*V(C):
suci 'pure, clean, sacred' < SKT;
suciŋ 'honest, fair (in fighting, games etc.)';

*s*VñV(C)
(BK *señap* 'lonely, desolate, quiet');
suñi 'silent, lonely, deserted';

*s*V*y*V(C):
sayaŋ '1. (it is a) pity; 2. love';
sayat 'cut fine, notch';
sayuŋ 'quiet, listless';
sayup '(too) late';
sayur 'vegetables';
suyuk 'be submissive, humble oneself, lack inspiration or vigour'.

SWY

*s*V*j*V(C):
sajo I 'only, merely' < SKT; II 'purpose (intention)'; III *sajoka(n)* 'meet a request';
səjax I 'do in parts, do little by little, work on s.th. slowly'; II 'seedlings transplanted from nursery bed';

*s*VñjV(C):
siñjo 'dusk, sunset glow' < SKT;
(BSM *sañjə* 'lie in an ambush at a village of the enemy');

*s*VñV(C):
sənap 'dizzy, stupefied, numb';

*s*V*y*V(C):
sayaŋ 'feel sorry or pity for, regret';
sayap 'wing';
sayo I 'servant; I, me'; II *sayo*... *sayo*... 'the more..., the more...';
sayiʔ(h) 'look after, take care of';
sayaʔ (written 'saj'jaq' in Helfrich) 'half a coconut, shell of a coconut';
sayup 'not reaching one's aim; fail (e.g. a harvest)';
suyun 'group, cattle, shoal'.

IBN

*s*V*j*V(C):
sajaʔ 'only, simply' (see *ajaʔ*) < SKT;

*s*V*c*V(C):
tuci (< ⁺*suci*) 'clean, pure' < SKT;

*s*VñV(C):
suñi 'lonely, deserted' < SKT;

*s*V*y*V(C):
sayap 'wing';

sayat 'cut, saw';
sayaw 'having a strong desire that...not';
sayuk 'flowerbuds of palms, maize';
sayur 'vegetables';
suyam 'brushy, thick (of beard)';
suyuŋsuyuŋ 'swollen (of the stomach)'.

JKT

jVsV(C):
jasad 'body' < AR;
jasa 'service, duty' < SKT;

sVjV(C):
səjəg, səjək 'since, from ... on';
sòja 'bow in salutation' < CHI;
sujèn I 'hollow in cheek when laughing or smiling'; II 'bamboo skewer for meat';
suji I (~ *sugi*) 'k.o. tree'; II 'make embroidery';
sujud 'prostrate' < AR;

sVcV(C):
sacap 'thirty (rupiah)';
səcaŋ 'k.o. tree, *Caesalpina Sappan*';
sècəŋ 'syphilis';
səcèŋ 'a thousand (rupiah)';
sicə 'set of comfortable table and chairs' < DU;
suci 'pure, clean, sacred' < SKT;

sVñjV(C):
sañjuŋ 'praise';

sVñcV(C):
sañcaʔ, sañcè 'python' < SUN;

sVñV(C):
səñap 'quiet, desolate';
səñum 'smile (v)';
siñal, siñar 'signal (for trains)' < DU;
siño 'Dutch boy' < POR.

REFERENCES

LIST OF ABBREVIATIONS

BKI	*Bijdragen tot de Taal-, Land- en Volkenkunde,* The Hague
BSOAS	*Bulletin of the School of Oriental and African Studies,* London
DBP	Dewan Bahasa dan Pustaka, Kementerian Pendidikan Malaysia
JMBRAS	*Journal of the Malayan Branch of the Royal Asiatic Society,* Singapore
OL	*Oceanic Linguistics,* University of Hawaii, Honolulu
PL	*Pacific Linguistics,* Canberra: Australian National University
Nusa	*Nusa linguistic studies of Indonesian and other languages in Indonesia,* Jakarta: Badan Penyelenggara Seri Nusa
PPPB	Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan
SMJ	*Sarawak Museum Journal,* Kuching
VBG	*Verhandelingen van het Koninklijk Bataviaasch Genootschap van Kunsten en Wetenschappen,* Bandung
VKI	*Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde,* The Hague
WPLUH	*Working Papers in Linguistics of the University of Hawaii,* Honolulu
ZES	*Zeitschrift für Eingeborenen-Sprachen,* Berlin

N.B. In view of the various onomastic devices used in Indonesia, in this list of references I order most Indonesian names according to the first full name given in the original source. Thus, if the original sources give Abdul Jebar Hapip and Hoesein Djajadiningrat, the bibliographical reference to these authors will be found under A and H respectively. If the original sources give M. Asfandi Adul, the bibliographical reference will be found under A (from Asfandi). I deviate from this principle when original sources include titles (e.g. Teuku Iskandar is referred to under I), or when the author has a marga [clan] name (e.g., Anggur P. Tambunan is referred to under T).

Abdul Chaer Mad'ie, 1976, *Kamus dialek Melayu Jakarta-Bahasa Indonesia.* Jakarta: Nusa Indah.

Abdul Jebar Hapip, 1977, *Kamus Banjar-Indonesia.* Jakarta: PPPB.

Abdurachman Ismail et al., 1979, Bahasa Bukit. Jakarta: PPPB.

Adelaar, K.A., 1981, Reconstruction of Proto-Batak phonology. *Nusa* 10:1-20.

    1983, Malay consonant-harmony: an internal reconstruction. *Nusa* 16:57-67.

    1984a, Some Proto-Malayic affixes. *BKI* 140/4:402-421.

    1984b, Review of Muhadjir (1981) in *BKI* 140/4:522-525.

    1985, Proto-Malayic: the reconstruction of its phonology and parts of its lexicon and morphology. (PhD dissertation, University of Leiden.) Alblasserdam (Netherlands): Kanters.

    1988, More on Proto-Malayic. In Mohd. Thani Ahmad and Zaini Mohamed Zain, eds, 59-77.

    1989, Malay influence on Malagasy: linguistic and culture-historical inferences, *Oceanic Linguistics* 28/1:1-46.

    in press a, The classification of the Tamanic languages, to appear in *Language contact and change in the Austronesian world,* 1-40. (T.Dutton and D. Tryon eds). Berlin: Mouton de Gruyter.

in press b, The history of thing, animal, person and related words in Malay, to appear in the proceedings of the conference on Austronesian terminologies (Canberra, 18-22 October 1990).

in press c, The importance of Selako for the reconstruction of Proto-Malayic and for the interpretation of the Old Malay inscriptions, to appear in the proceedings of the First Extraordinary Session of the Borneo Research Council (Kuching, Sarawak, 4-9 August 1990).

forthcoming, The internal classification of the Malayic isolects, paper presented at the 6th International Conference on Austronesian Linguistics (Honolulu, May 1991).

Aichele, W., 1942-43, Die altmalaiische Literatursprache und ihr Einfluss auf das Altjavanische. *ZES* 33:37-66.

Aliana, Z.N., et al., 1979, *Bahasa Serawai*. Jakarta: PPPB.

Anceaux, J.C., 1965, Austronesian linguistics and intra-group comparison. *Lingua* 14:306-314.

1982, Towards a typological reconstruction of the verbal system in Proto-Austronesian. In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics,* vol.2: *Tracking the travellers,* 101-110. *PL,* C-75.

Anttila, Raimo, 1972, *An introduction to historical and comparative linguistics.* New York: Macmillan.

Antonissen, A., 1956, *Kadazan-English and English-Kadazan dictionary.* Canberra: Government Printing Office.

Asfandi Adul, M., 1976, Morfologi Bahasa Banjar. Unpublished MS.

Asmah Haji Omar, 1964, Bahasa Iban: laporan penyelidekan ilmu bahasa. Kuala Lumpur: Universiti Malaya, Jabatan Pengajian Melayu. Xeroxed MS.

1975, The nature of Tamil loanwords in Malay. In *Essays on Malaysian linguistics,* 303-335. Kuala Lumpur: DBP.

1977, The Iban language. *SMJ* 25/46:81-100.

1981, *The Iban language of Sarawak: a grammatical description.* Kuala Lumpur: DBP.

1985, *Susur Galur Melayu.* Kuala Lumpur: DBP.

Aymonier, E., and **A. Cabaton**, 1906, *Dictionnaire cam-français.* Paris, Publications de l'École Française d'Extrême Orient, Volume 7.

Bellwood, Peter, in press, Cultural and biological differentiation in peninsular Malaysia: the last 10,000 years (to appear in the Proceedings of the 2nd Conference on Malay Civilisation held in Kuala Lumpur in August 1989).

Blust, Robert A., 1970, Proto-Austronesian addenda. *OL* 9/2:104-162.

1972, Additions to 'Proto-Austronesian addenda' and 'Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages'. *WPLUH* 4/8:1-17.

1973, Additions to 'Proto-Austronesian addenda' and 'Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages' -II. *WPLUH* 5/3:33-61.

1974a, The Proto-Austronesian word for 'two': a second look. *OL* 13, 123-161.

1974b, The Proto-North Sarawak vowel deletion hypothesis. Unpublished PhD dissertation, Honolulu: Department of Linguistics, University of Hawaii.

1976, Dempwolff's reduplicated monosyllables, *Essays in honor of Samuel H. Elbert.* *OL* 15/1-2:107-130.

1977a, The PAN pronouns and Austronesian subgrouping: a preliminary report. *WPLUH* 9/2:1-15.

1977b, Sketches of the morphology and phonology of Bornean languages, I: Uma Juman (Kayan). *Papers in Borneo and Western Austronesian Linguistics* No. 2, 9-122. *PL,* A-33.

1978a, *The Proto-Oceanic palatals.* Memoir 13. The Polynesian Society, Wellington.

1978b, Eastern Malayo-Polynesian: a subgrouping argument. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings,* Fascicle 1, *Western Austronesian,* 181-234. *PL,* C-61.

1979, Proto-Western Malayo-Polynesian vocatives. *BKI* 135/2-3:205-251.

1980a, *Austronesian etymologies. OL* 19/1-2.

1980b, Iban antonymy: a case from diachrony? In D.J. van Alkemade et al., eds *Linguistic studies offered to Berthe Siertsema.* Amsterdam: Rodopi.

1980c, Early Austronesian social organization: the evidence of language. *Current Anthropology* 21/2, 205-226 (Comments 227-247, 415-419).

1981, The reconstruction of Proto-Malayo-Javanic: an appreciation. *BKI* 134/4:456-469.

1982a, An overlooked feature of Malay historical phonology. *BSOAS* 45/2:284-299.

1982b, The Proto-Austronesian word for 'female'. In Carle et al., eds 1982:17-30.

1984a, Austronesian etymologies - II. *OL* 22/23:29-149.

1984b, On the history of Rejang vowels and diphthongs. *BKI* 140/4:422-450.

1986, Austronesian etymologies - III. *OL* 25/1-2:1-123.

1988, Malay historical linguistics: a progress report. In Mohd. Thani Ahmad and Zaini Mohamed Zain, eds, 1-33.

forthcoming, Variation in retention rate among Austronesian languages.

Borneo Literature Bureau, 1967, *English-Iban phrase book (Bup jako Inglis-Iban).*

Brandes, Jan Lourens Andries, 1884, *Bijdrage tot de vergelijkende klankleer der westersche afdeeling van de Maleisch-Polynesische taalfamilie.* (PhD dissertation of Leiden.) Utrecht: Stoomdrukkerij P.W. van de Weijer.

Brown, C.H., 1982, [An article on the semantics of directional terms in various languages] Unpublished typescript.

Burkill, I.H., 1966, *A dictionary of the economic products of the Malay peninsula.* 2 vols. Kuala Lumpur: the Ministry of Agriculture and Co-operatives.

Burrow, T., and M. B. Emeneau , 1961, *A Dravidian etymological dictionary.* Oxford: Clarendon Press.

Carle, Rainer, et al., eds., 1982, *Gava': studies in Austronesian languages and cultures, dedicated to Hans Kähler.* Berlin: Dietrich Reimer.

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

Cense, A.A., and E.M. Uhlenbeck, 1958, *Critical survey of studies on the languages of Borneo.* The Hague: Nijhoff.

Chrétien, C. Douglas, 1965, The statistical structure of the Proto-Austronesian morph. *Lingua* 14:243-270.

Çoedès, G., 1930, Les inscriptions Malaises de Çrīvijaya. *Bulletin de l'École Française de l'Extrême Orient* (Paris) 30:29-80.

Collins, James T., 1983, *The historical relationships of the languages of Central Maluku, Indonesia. PL,* D-47.

1985, The phonology of Tioman Malay and the reconstruction of Proto-Malay. In Suriya Ratanakul et al., eds *Southeast Asian linguistic studies presented to Andre G. Haudricourt,* 541-566. Bangkok: Mahidol University.

1986a, Penggolongan Bahasa Bacan, in *Antologi kajian dialek Melayu,* 133-167. Kuala Lumpur: DBP. (Also published in1983 in *Nusantara* 10:86-125).

1986b, Kajian dialek daerah dan rekonstruksi Bahasa Purba, in *Antologi kajian dialek Melayu,* 171-194. Kuala Lumpur: DBP. (Also published in 1986 in *Dewan Bahasa* 30:344-365).

1987, *Dialek Melayu Sarawak.* Kuala Lumpur: DBP.

Conant, Carlos Everett, 1911, The RGH law in Philippine languages. *Journal of the American Oriental Society* 31:70-85.

Coolsma, S., 1913, *Soendaneesch-Hollandsch woordenboek* (2nd edn). Leiden: Sijthoff.

Dahl, Otto Christian, 1938, Le système phonologique du proto-malgache. *Norsk Tidsskrift for Sprogvidenskap* 10:189-235.

1951, *Malgache et Maanjan: une comparaison linguistique.* Avhandlinger utgitt av Instituttet 3. Oslo: Egede Instituttet.

1976, *Proto-Austronesian* (2nd rev. edn; 1st edn, 1973. Lund: Studentlitteratur), Scandinavian Institute of Asian Studies Monograph Series No. 15. London: Curzon Press.

1977, La subdivision de la famille Barito et la place du malgache, Acta Orientalia (Copenhagen) 38:77-134.

De Casparis: see Casparis

De Josselin de Jong: see Josselin de Jong

Dempwolff, Otto, 1934-38, *Vergleichende Lautlehre des Austronesischen Wortschatzes,* I (1934): *Induktiver Aufbau einer Indonesischen Ursprache;* II (1937): *Deduktive Anwendung des Urindonesischen auf austronesische Einzelsprachen;* III (1938): *Austronesisches Wörterverzeichnis. ZES* Beihefte 15, 17, 19. Berlin: Dietrich Reimer.

Duden, 1976-81, *Duden: das Grosse Wörterbuch der deutschen Sprache* (re-edited and revised by Günther Drosdowski et al.). Mannheim/Wien/Zürich: Bibliographisches Institut (Dudenverlag).

Dunselman, P. Donatus, 1949, Bijdragen tot de kennis van de taal en adat der Kendajan-Dajaks van West-Borneo. *BKI* 105/1:59-105, *BKI* 105/2-3:147-218.

1950, Bijdragen tot de kennis van de taal en adat der Kendajan-Dajaks van West-Borneo II. *BKI* 106/3-4:321-73.

1955, *Kana Sera: Zang der Zwangerschap. VBG* 17. The Hague: Nijhoff.

Durdje Durasid and Djantera Kawi, 1978, *Bahasa Banjar Hulu.* Jakarta: PPPB.

Dyen, I., 1949, On the history of Trukese vowels. *Language* 25:420-436.

1951, Proto-Malayo-Polynesian *Z. *Language* 27:534-540.

1953, *The Proto-Malayo-Polynesian laryngeals.* Baltimore: Linguistic Society of America.

1956, The Ngaju-Dayak 'old speech stratum'. *Language* 32, 83-87.

1965a, *A lexicostatistical classification of the Austronesian languages. International Journal of American Linguistics,* Memoir 19. Baltimore: Waverly Press.

1965b, Formosan evidence for some new Proto-Austronesian phonemes. *Lingua* 14:285-305.

Eggink, H.J., 1936, *Angkola- en Mandailing-Bataksch - Nederlandsch woordenboek. VBG* 72/5.

Elkins, Richard E., 1968, *Manobo-English dictionary. OL* Special Publication No.3. Honolulu: University of Hawaii Press.

Emeis, M.G., 1945, Vorm en functie in klassiek en modern Maleisch: de verbale constructies. PhD dissertation, Utrecht.

1955, Vocaal-harmonie in de Maleise kern van de Indonesische vocabulaire. *BKI* 111:191-201.

Ferrell, Raleigh, 1969, *Taiwan aboriginal groups: problems in cultural and linguistic classification.* Taipei: Institute of Ethnology, Academia Sinica.

1982, *Paiwan dictionary. PL,* C-73.

Fokker, A.A., 1895, *Malay phonetics.* Leiden: Brill.

Fudiat Suryadikara et al., 1979, *Bahasa Kutai.* Jakarta: PPPB.

1981, *Geografi dialek Bahasa Banjar Hulu.* Jakarta: PPPB.

Gericke, J.F.C., and T. Roorda, 1901, *Javaansch-Nederlandsch handwoordenboek* (vermeerderd en verbeterd door A.C. Vreede). Amsterdam: Johannes Muller.

Gerth van Wijk, D., 1889, *Spraakleer der Maleische taal.* Batavia: G. Kolff.

Gonda, J., 1973, *Sanskrit in Indonesia.* New Delhi: International Academy of Indian Culture.

Goris, R., 1938, *Beknopt Sasaksch-Nederlandsch woordenboek.* Singaradja (Bali): Publicatie Kirtya Liefrinck-van der Tuuk.

Grace, George W., 1959, *The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) family.* Baltimore.

1969, A Proto-Oceanic finder list. *WPLUH* 1/2:39-84.

Grijns, C.D., 1980, Some notes on Jakarta Malay kinship terms: the predictability of complexity. *Archipel* 20, 187-212.

1981, Jakartan speech and Takdir Alisjahbana's plea for the simple Indonesian word-form. In Nigel Phillips and Khaidir Anwar, eds *Papers on Indonesian languages and literatures*, 1-34. London: Indonesian Etymological Project, SOAS; Paris: Cahier d'Archipel 13, Association Archipel.

Haaksma, R., 1933, *Inleiding tot de studie der vervoegde vormen in de Indonesische talen*. PhD dissertation. Leiden: Brill.

Hakim Usman, A., 1985, *Kamus umum Kerinci-Indonesia*, Jakarta: PPPB.

Hardeland, August , 1859, *Dayacksch-Deutsches Wörterbuch*. Amsterdam: Frederik Muller.

Helfrich, O.L., 1904, *Bijdragen tot de kennis van het Midden Maleisch (Besemahsch en Serawajsch dialect)*. VBG 53.

1915, Nadere aanvullingen en verbeteringen op de Bijdragen tot de kennis van het Midden Maleis. *VBG* 61/3-4:1-28.

1921, Supplement op de Nadere aanvullingen.... *VBG* 63/1.

1927, *Nadere bijdragen tot de kennis van het Midden Maleisch*. VBG 68/3.

1933, *Bijdragen tot de kennis van het Midden Maleisch* (supplement to Helfrich 1927). The Hague: Nijhoff.

Henderson, Eugénie J.A., 1965, The topography of certain phonetic and morphological characteristics of South East Asian languages. *Lingua* 15:400-438.

Hoesein Djajadiningrat, 1934, *Atjèhsch-Nederlandsch woordenboek*. 2 vols. Batavia: Landsdrukkerij.

Hogan, David W., 1972, Men of the sea: Coastal tribes of South Thailand's West Coast. *Journal of the Siam Society* (Bangkok) 60/1:205-235.

Hogan, David W., in collaboration with Stephen W. Pattemore, 1988, *Urak Lawoi': basic structures and a dictionary*. PL, C-109.

Horne, Elinor Clark, 1974, *Javanese-English dictionary*. New Haven and London: Yale University Press.

Howison, James, 1801, *A dictionary of the Malay tongue*. London: the Arabic and Persian Press.

Hudson, Alfred. B., 1967, *The Barito isolects of Borneo*. Southeast Asia Program (Dept. of Asian Studies), Data Paper no. 68. Ithaca (NY): Cornell University Press.

1970, A note on Selako: Malayic Dayak and Land Dayak languages in Western Borneo. *SMJ* 18/36-37 (New Series):301-318.

1978, Linguistic relations among Bornean peoples with special reference to Sarawak: an interim report. In Sutlive et al., eds *Sarawak: linguistics and development problems* . Studies in Third World Societies (Williamsburg, VA) 3:1-45.

Ikranagara, Kae, 1980, *Melayu Betawi grammar*. Nusa 9. Jakarta: Badan Penyelenggara Seri Nusa.

Ina Anak Kalom and Alfred B. Hudson, 1970, Selako traditional history: a story on the origins of Kampong Pueh. *SMJ* 18/36-37 (New Series):281-300.

Iskandar, Teuku ed., 1970, *Kamus Dewan*. Kuala Lumpur: DBP.

Jansz, P., 1913, *Practisch Javaansch-Nederlandsch woordenboek* (2nd rev. edn). Semarang, Surabaya, The Hague: Van Dorp.

Josselin de Jong, P.E. de, 1951, *Minangkabau and Negeri Sembilan: socio-political structure in Indonesia*. Leiden: Eduard Nijdo; 2nd edn (1980), The Hague: Nijhoff.

Kähler, Hans, 1960, *Ethnografische 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 Reimer.

1966, *Wörterverzeichnis des Omong Djakarta*. Berlin: Dietrich Reimer.

Kern, H., 1917, Taalkundige gegevens ter bepaling van het stamland der Maleisch-Polynesische volkeren. In H. Kern *Verspreide Geschriften VI*, 105-120. The Hague: Nijhoff. (Originally published in 1889 in the *Verslagen en Mededeelingen der Koninklijke Akademie van Wetenschappen*, Afdeeling letterkunde (3rd Series) 6:270-287.

Kersten, J., 1978, Kamus kecil Bahasa Bali. Unpublished MS.

Klinkert, H.C., 1916, *Nieuw Maleisch-Nederlandsch woordenboek* (3rd revised and enlarged edn). Leiden: Brill.

Lafont, Pierre, Bernard, 1968, *Lexique français-jarai-vietnamien*. Paris: Publications de l'École Française d'Extrême Orient, Volume 63.

LeBar, Frank M., 1972, *Ethnic groups of insular Southeast Asia*. New Haven: Human Relations Area Files Press.

Lee, Ernest Wilson, 1966, Proto-Chamic phonologic word and vocabulary. PhD dissertation, Indiana University (Universal Microfilms International, Inc., Ann Arbor, MI).

Leo, Philip, 1975, *Chinese loanwords spoken by the inhabitants of the city of Jakarta*. Jakarta: Lembaga Ilmu Pengetahuan Indonesia, Lembaga Research Kebudayaan Nasional Seri data dasar no.7.

Li, Paul Jen-Kuei, 1982, Kavalan phonology: synchronic and diachronic. In Carle et al., eds 1982:479-495.

Marsden, William, 1812, *A dictionary of the Malay language* (reprinted in 1973 by the Books for Libraries Press, NY).

Maxwell, C.N., 1936, Light in the Malay language (with a foreword by R.J. Wilkinson). *JMBRAS* 14/3:89-154.

Milke, Wilhelm, 1968, Proto-Oceanic addenda. *OL* 7/2:147-171.

Mills, Roger J., 1975, Proto-South-Sulawesi and Proto-Austronesian phonology. 2 vols. PhD dissertation, University of Michigan, Ann Arbor.

Mohd. Thani Ahmad and Zaini Mohamed Zain, eds, 1988, *Rekonstruksi dan cabang-cabang Bahasa Melayu Induk*. Kuala Lumpur: DBP.

Moussay, Gérard, 1972, *Dictionnaire cam-vietnamien-français*. Phanrang: Centre Culturel cam.

1981, *La langue Minangkabau*. Cahier d'Archipel 14. Paris: Association Archipel.

Muhadjir

1981, *Morphology of Jakarta dialect, affixation and reduplication*. Nusa 11. Jakarta: Badan Penyelenggara Seri Nusa.

Neumann, J.H., 1951, *Karo-Batak-Nederlands woordenboek*. Jakarta: Lembaga Kebudayaan Indonésia.

Niemann, G.K., 1891, Bijdrage tot de kennis der verhouding van het Tjam tot de talen van Indonesië. *BKI* 40:27-44.

Nothofer, Bernd, 1975, *The reconstruction of Proto-Malayo-Javanic*. The Hague: Nijhoff.

1984, Further evidence for the reconstruction of *-əy and *-əw. *BKI* 140/4:451-458.

1988, A discussion of two Austronesian subgroups: Proto-Malay and Proto-Malayic. In Mohd. Thani Ahmad and Zaini Mohamed Zain, eds 1988:34-58.

Ophuysen, Ch. A. van, 1907, *Kitab logat Melajoe, handleiding voor de spelling van de Maleische taal*. Batavia.

Oxford Dictionary, 1944, *The shorter Oxford English dictionary* (3rd edn, edited and revised by C.T. Onions). Oxford: Clarendon Press.

Panganiban, Jose Villa, 1966, *Talahuluganang Pilipino-Ingles*. Manila: Kawanihan ng Palimbagan.

Pawley, Andrew, and Lawrence A. Reid, 1976, The evolution of transitive constructions in Austronesian. *WPLUH* 8/2:51-74.

Penninga, P., and H. Hendriks, 1937, *Practisch Madurees-Nederlands woordenboek* (1st edn 1913). Semarang, Surabaya, The Hague: Van Dorp.

Pigafetta, Antonio, 1972, *L'Indonesia nella relazione di viaggio di Antonio Pigafetta*. (Relazioni di viaggiatori italiani in Indonesia 2; introduzione e note di A. Bausani). Rome: Istituto italiano per il Medio ed Estremo Oriente; Jakarta: Centro italiano di cultura.

Pigeaud, Th., 1938, *Javaans-Nederlands handwoordenboek* (2nd rev. edn). Groningen, Wolters-Noordhoff.

Poerwadarminta, W.J.S., 1951, *Logat ketjil Bahasa Indonésia* (1st edn 1949). Groningen-Batavia.

Prentice, D.J., 1971, *The Murut languages of Sabah*. *PL*, C-18.

1974, Yet another PAN phoneme? *OL* 13:33-75.

1978, The best chosen language. *Hemisphere* (Melbourne) 23/3:18-23 and 23/4:28-33.

1987, Malay (Indonesian and Malaysian), a chapter in Bernard Comrie ed. *The world's major languages*. London: Croom Helm.

n.d., Lecture-notes on Indonesian and Malaysian grammar. MS.

forthcoming, Timugon-Murut-English dictionary.

Prentice, D.J., and A. Hakim Usman, 1978, Kerinci sound-changes and phonotactics. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, Fascicle I: *Western Austronesian*, 121-163. *PL*, C-61.

Ras, J.J., 1970, Lange consonanten in enige Indonesische talen (II). *BKI* 126:429-447.

Ray, Sidney H., 1913, The languages of Borneo. *SMJ* 1/4:1-196.

Richards, A., 1981, *An Iban-English dictionary*. Oxford: Clarendon Press.

Ronkel, Ph.S. van, 1902, Het Tamil-element in het Maleisch. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* Koninklijk Bataviaasch Genootschap voor Kunsten en Wetenschappen 45:97-119. Batavia.

Roolvink, R., 1953, *Dialek Melayu di Deli*. Bahasa dan Budaya, 3-40. (Jakarta: Lembaga Bahasa dan Budaya Fakultet Sastra dan Filsafat Universitet Indonesia).

1965, The passive-active per-/ber-//per-/memper-correspondence in Malay. *Lingua* 15:310-337.

Saragih, St. J.E., 1989, *Kamus Simalungun - Indonesia*. Pematang Siantar: GKPS.

Scott, N.C., 1956, *A dictionary of Sea Dayak*. London: School of Oriental and African Studies.

1957, Notes on the pronunciation of Sea Dayak. *BSOAS* 20:509-512.

Sneddon, James N., 1978, *Proto-Minahasan: phonology, morphology and wordlist*. *PL*, B-54.

Steingass, F., 1930, *Persian-English dictionary*. London: Kegan Paul, Trench, Trübner.

Steinhauer, H., 1980, *On the history of Indonesian*. Studies in Slavic and General Linguistics 1. Amsterdam: Rodopi.

Steinhauer, H., and A. Hakim Usman, 1978, Notes on the morphemics of Kerinci (Sumatra). In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, Fascicle 1: *Western Austronesian*, 483-502. *PL*, C-61.

Steuerwald, Karl, 1972, *Türkisch - Deutsches Wörterbuch*. Wiesbaden: Harrassowitz.

Stokhof, W.A.L., with Lia Saleh-Bronckhorst, eds, 1980, *Holle lists: vocabularies in languages of Indonesia 2: Sula and Bacan Islands, North Halmahera, South and East Halmahera*. *PL*, D-28.

Stresemann, Erwin, 1927, *Die Lauterscheinungen in den ambonischen Sprachen*. *ZES* Beiheft 10. Berlin: Dietrich Reimer.

Svelmoe, Gordon, and Norman Abrams, 1955, Mansaka vocabulary. Summer Institute of Linguistics, University of North Dakota (in cooperation with the Bureau of Public Schools and the Institute of Natural Language of the Department of Education, Manila).

Tambunan, Anggur P., 1977, *Kamus Bahasa Batak Toba-Indonesia*. Jakarta: PPPB.

Tamsin Medan, 1980, Dialek-dialek Minangkabau di daerah Minangkabau/ Sumatera Barat (Suatu pemerian dialektologis). Jakarta: PPPB. MS.

Teeuw, A., 1959, The history of the Malay language. *BKI* 115, 138-156.

Teeuw, A., with the assistance of H.W. Emanuels, 1961, *A critical survey of studies on Malay and Bahasa Indonesia*. The Hague: Nijhoff.

Thaib gl. St. Pamoentjak, M., 1935, *Kamoes Bahasa Minangkabau-Bahasa Melajoe-Riau*. Batavia: Balai Poestaka.

Tindi Radja Manik, 1977, *Kamus Bahasa Dairi Pakpak-Indonesia*. Jakarta: PPPB.

Toorn, van der J.L., 1891, *Minangkabausch-Maleisch - Nederlandsch woordenboek*. The Hague: Nijhoff.

1899, *Minangkabausche spraakkunst*. The Hague: Nijhoff.

Tsuchida, Shigeru, 1976, *Reconstruction of Proto-Tsouic phonology*. Study of Languages and Cultures of Asia and Africa Monograph Series No.5. Tokyo.

Tuuk, H.N. van der, 1861, *Bataksch - Nederduitsch woordenboek*. Amsterdam: Frederik Muller.

Van der Toorn: see Toorn

Van der Tuuk: see Tuuk

Van Ophuysen: see Ophuysen

Vikør, Lars S., 1988, *Perfecting spelling. Spelling discussions and reforms in Indonesia and Malaysia 1900-1972*. Verhandelingen van het KITLV 133. Dordrecht (Netherlands)/Providence (USA): Foris.

Vlis, C.J. van der, 1842, Formosaansche woorden-lijst, volgens een Utrechtsch Handschrift. *VBG* 45:437-488.

Von de Wall: see Wall

Voorhoeve, Petrus, 1955, *Critical survey of studies on the languages of Sumatra*. The Hague: Nijhoff.

Walker, Dale, 1976, *A grammar of the Lampung language: the Pesisir dialect of Way Lima*. Jakarta: Badan Penyelenggara Seri Nusa.

Wall, H. von de, 1877-97, *Maleisch-Nederlandsch woordenboek* (uitgegeven met weglating van al het overtollige door H.N. van der Tuuk) (3 vols). Batavia: Landsdrukkerij.

Warren, Charles P.,1959, A vocabulary of the Batak of Palawan, Philippine Studies Program, transcript no.7, University of Chicago.

Warneck, Joh., 1977, *Toba-Batak-Deutsches Wörterbuch* (neu herausgegeben von R. Roolvink). The Hague: Nijhoff.

Wehr, Hans, 1976, *A dictionary of modern written Arabic* (edited by J. Milton Cowan). Ithaca (NY): Spoken Language Services.

Wilkinson, R.J., 1936, Onomatopoeia in Malay. *JMBRAS* 14/3:72-88.

1959, *A Malay-English dictionary* (2 vols). London: Macmillan.

Winstedt, R.O., 1927, *Malay grammar* (2nd rev. edn). Oxford: Clarendon Press.

Wolff, J.U., 1973, Verbal inflection in Proto-Austronesian. In Andrew B. Gonzalez, ed. *Parangal Kay Cecilio Lopez*. Quezon City (Philippines): Philippine Journal of Linguistics (Special Monograph 4), 71-91.

1974, Proto-Austronesian *r and *d. *OL* 13 (Papers of the First International Conference on Comparative Austronesian Linguistics):77-121.

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

Wouk, Fay, 1980, The *ter*-prefix in Indonesian, a semantic analysis. In Buenaventura Paz Naylor, ed. *Michigan Papers on South and Southeast Asia*, 81-87. Ann Arbor: University of Michigan, Center for South and Southeast Asian Studies.

Wurm, S.A., and B. Wilson, 1975, *English finderlist of reconstructions in Austronesian languages (post-Brandstetter)*. PL, C-33.

Yuslizal Saleh et al., 1979, *Bahasa Semende*. Jakarta: PPPB.

Zoetmulder, P.J., with the collaboration of S.O. Robson, 1982, *Old Javanese-English dictionary* (2 vols). The Hague: Nijhoff.

1983, *De taal van het Adiparwa, een grammaticale studie van het Oudjavaans*. KITLV Indonesische Herdrukken. Dordrecht (Netherlands)/Cinnaminson (NJ): Foris Publications.

Zorc, R. David, 1971, Proto-Philippine finderlist, Ithaca (NY): Cornell. (English glossary added in 1973.) Unpublished MS.

1982, Where, O where, have the laryngeals gone? Austronesian laryngeals re-examined. In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, Vol.2: *Tracking the travellers*, 111-144. PL, C-75.

1983, Proto Austronesian accent revisited. *Philippine Journal of Linguistics* 14/1:1-24.

INDEX

N.B. The etyma are arranged (1) according to the proto-language for which they were reconstructed; the proto-languages are ordered from their highest to their lowest order, beginning with PAN and (2) according to the following order of symbols: *a, A, b, c, C, d, D, e, ə, g, h, ʔ, i, j, k, l, L, M, n, ŋ, ñ, N, p, q, r, R, s, S, t, u, v, V, w, x, y, z, Z.* Where brackets, slashes, question marks, zero signs, commas, hyphens or spaces (in compound etyma) occur, they are ignored for ordering purposes, and ordering is done according to the following symbol.