A grammar of Jahai

Burenhult, N. *A grammar of Jahai.*
PL-566, xiv + 245 pages. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-566.cover*

Pacific Linguistics 566

Pacific Linguistics is a publisher specialising in grammars and linguistic descriptions, dictionaries and other materials on languages of the Pacific, Taiwan, the Philippines, Indonesia, East Timor, southeast and south Asia, and Australia.

Pacific Linguistics, established in 1963 through an initial grant from the Hunter Douglas Fund, is associated with the Research School of Pacific and Asian Studies at The Australian National University. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world. Publications are refereed by scholars with relevant expertise, who are usually not members of the editorial board.

A grammar of Jahai

Niclas Burenhult

Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
Australia

First published 2005

National Library of Australia Cataloguing-in-Publication entry:

Burenhult, Niclas
A grammar of Jahai.

Bibliography
ISBN 0 85883 554 1

1. Semang languages – Grammar. 2. Semang (Malaysian people) – Languages. I. The Australian National University. Research School of Pacific and Asian Studies. Pacific Linguistics. II. Title.

495.93

The cover illustration shows a Jahai headman and spirit-medium Cheneleg Piloi beside the sacred Batu Rem, a limestone outcrop on the Pergau river, in Kelantan, Peninsular Malaysia. (Photo: Niclas Burenhult.)

Copyedited by Basil Wilson
Typeset by Jeanette Coombes
Cover design by Cirils Printers
Printed and bound by Cirils Printers, Fyshwick, Canberra

Table of contents

Maps and figures

Tables

Abbreviations

AGR	agreement
CAUS	causative
CLF	classifier
COLL	collective
CONT	continuative
CONTR	contrastive
D	dual
DEIC	deictic
DES	desiderative
DET	determiner
DIS	distant
DISTR	distributive
DP	diverse plural
EMP	emphatic
EQU	equative
EXCL	exclusive
FAM	familiar
GOAL	goal
HORT	hortative
ID	identification
IMP	imperative
IMPF	imperfective
INCL	inclusive
INT	intimate
IRR	irrealis
ITER	iterative
LOC	locative
M	affix /m/
N	noun
NEG	negative
NM	nominaliser
NP	noun phrase
P	plural
POSS	possessor
PP	prepositional phrase
PREP	preposition
PRO	pronoun
PROG	progressive
PROH	prohibitive
PROP	property
Q	interrogative
QNT	quantifier
REC	reciprocal
REL	relative
RP	root possibility
RT	relational tense
S	singular
SOURCE	source
UNIT	unitiser
V	verb
1	first person
2	second person
3	third person

Preface and acknowledgements

My first encounter with the Jahai was over in a moment. Blowpipes poised, a small hunting party emerged from the towering green wall of rainforest and stepped out onto the East-West Highway just as the Land Rover I was riding in roared by. I managed to catch only a brief glimpse of a people once made legendary by Father Schebesta's early 20th century writings. Today, several field trips later, I am fortunate enough to have had the opportunity to get to know the Jahai and to learn a little about their language. The present grammar — a revision of my 2002 doctoral dissertation — presents the findings of this linguistic inquiry. For the Jahai, language is intimately associated with the ancestral past and the 'old ways', which are still very much part of everyday life. Yet the very nature of my first encounter somehow symbolises the changes that inevitably lie ahead. The recent opening up of previously inaccessible areas of northern Peninsular Malaysia, coupled with the country's ambitious policy of national development which aims at turning Malaysia into a fully industrialised nation by the year 2020, is bound to have far-reaching consequences for the Jahai community. I hope the present work can be of some benefit to the Jahai in the process.

I would like to take this opportunity to direct my heartfelt thanks to the 'people from the eye of the forest', who have shared with me not only their language, but also their everyday life, their rainforest, their good sense of humour, and their moments of happiness and grief. For their patience, I thank Cheneleg bin Piloi, Ating bin Piloi, Cemerbak s/o Rantau, Alang bin Jelatang, Salleh bin Busu and their families. Without their enthusiastic co-operation, this study would have been an impossibility.

I also wish to thank several non-Jahai who have been important sources of inspiration to my work, including my supervisor Jan-Olof Svantesson, whose down-to-earth commenting and profound knowledge of Mon-Khmer linguistics have been of immeasurable help; as well as Geoffrey Benjamin, Gérard Diffloth and Nicole Kruspe, who have all generously shared with me their experience in the field of Aslian linguistics. My thoughts also go to the late Thore Pettersson, without whose encouragement I probably would not have pursued the field of linguistics, and to the late Cornelia van der Sluys, who initiated me into the ways of the Jahai.

My work has also benefited greatly from the kind help and input from the following people: Christian Bauer, Gösta Bruce, Norah Carp, Johan Dahl, Kirk Endicott, Anna Flyman-Mattsson, David Gil, Marianne Gullberg, John Hajek, Arthur Holmer, Hood Salleh, Merle Horne, Gisela Håkansson, Birgitta Lastow, Kristina Lindell, Ann Lindvall, Lye Tuck-Po, Jim Matisoff, Ingrid Mellqvist, Britt Nordbeck, Damrong Tayanin, Mechtild Tronnier, Åke Viberg, Wazir Jahan Karim, Caroline Willners, Zaharani Ahmad, Elisabeth Zetterholm and several others. Special thanks are directed to Kartini Ahmad and

Rogayah Bt. A. Razak at the Department of Audiology and Speech Sciences, Universiti Kebangsaan Malaysia, my academic contact in Malaysia. For their careful examination of the original manuscript for this book, I am grateful to Jørgen Rischel and an anonymous reviewer of Pacific Linguistics.

Furthermore, I would like to express my acknowledgements to the Economic Planning Unit, Putrajaya, for granting me permission to conduct research in Malaysia. I also wish to thank the helpful staff of the Jabatan Hal Ehwal Orang Asli, including Wan Said bin Wan Chek, Mohd. Sayed Abdullah, Mohd. Zamri bin Mustajap, Anuwar bin Yahaya, Mat Nizam bin Lal Mohd., Rashdi bin Mahmud and Sam bin Yusof, and I am particularly indebted to Abu Othman bin Abdul Karim and his family for their kind help and great hospitality.

The fieldwork on which this study is based was carried out with the aid of grants from Crafoordska stiftelsen, Kungl. Humanistiska Vetenskapssamfundet i Lund, Fil dr Uno Otterstedts fond, and Birgit och Gad Rausings Stiftelse för Humanistisk Forskning.

Last, I am grateful to my family for providing a childhood environment where curiosity was paramount, where borders were meant to be traversed, and where the exotic turned into the commonplace.

Niclas Burenhult

Nijmegen, January 2004

Burenhult, N. *A grammar of Jahai.*
PL-566, xiv + 245 pages. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-566.cover*

1 *Introducing Jahai*

This chapter gives an introduction to Jahai, the Mon-Khmer language which is the subject of the present grammatical description. The first section (§1.1) provides a brief ethnographic account of the Jahai-speaking community. The following sections outline the linguistic relationships of Jahai (§1.2), its sociolinguistic situation (§1.3) and previous linguistic research related to Jahai and its close relatives (§1.4). The final section (§1.5) presents the background, aims, form and outline of the present study.

1.1 The Jahai

The Jahai are a group of hunter-gatherers, traders and occasional swidden cultivators inhabiting the montane rainforests of the Malay Peninsula. Their territory covers an area on both sides of the main Titiwangsa watershed around the upper reaches of the rivers Perak, in Perak state, and Pergau, in Kelantan state, Peninsular Malaysia, as well as the adjoining parts of Yala and Narathiwat provinces of southernmost Thailand. Being mobile foragers until recently, many Jahai today lead a settled or semi-settled life in regroupment programs established by the Malaysian government following the construction of the Temenggor hydroelectric dam in the late 1970s, notably that of Air Banun, in Hulu Perak district, Perak state. A regroupment village has also been established at Sungai Rual, near Jeli, in Kelantan. Some Jahai still pursue a mobile existence; in 1993 their number was estimated at 150 (van der Sluys 1996:2, 1999:310). The total number of Jahai is usually estimated at around 1000, the latest official figure from the website of the Jabatan Hal Ehwal Orang Asli (Department of Aboriginal Affairs) being 1049.

The Jahai form the largest ethnic group of a cluster of hunter-gatherer populations in the Malay Peninsula referred to generically as 'Semang'. Because of their physical characteristics, the Semang are often also referred to as 'Negritos', a racial label given to a number of small groups of Southeast Asian hunter-gatherers which are found in the Malay Peninsula, the Philippines and the Andaman Islands, and which are sometimes believed to be descendants of the original population of the region. However, as pointed out by van der Sluys (1999:307), the present-day Jahai do not make up a physically homogeneous population. The indigenous populations of Peninsular Malaysia, of which the Semang only form a small portion, are referred to generically in Malaysia as *Orang Asli* 'aboriginal people'.

Burenhult, N. *A grammar of Jahai*.
PL-566, xiv + 245 pages. Pacific Linguistics, The Australian National University, 2005. DOI:*10.15144/PL-566.1*

The traditional subsistence system of the Jahai is flexible and opportunistic. It is based on hunting, fishing and the collecting of wild tubers and vegetables. Occasionally they make small swiddens where they grow mainly cassava and dry rice. They are also engaged in the collecting of commercial forest products, mainly rattans, gaharu (*Aquillaria* spp.) and honey, and recently some groups have taken up commercial fishing in Lake Temenggor. Traditionally, the Jahai live in mobile groups of 15–50 people, sheltering in windbreak huts and moving camp every one to two weeks. When engaged in swidden cultivation and commercial collecting of forest products, a group may settle down for a month or so in more permanent houses (van der Sluys 1999:308–309).

Jahai society is egalitarian, with a strong emphasis on individualism, sharing and non-aggression (van der Sluys 1996, 1999:310, 2000). Authority, often associated with spirit-mediums, is based on charisma and the art of persuasion, but society is otherwise non-stratified. Nowadays the Malaysian authorities appoint headmen (*penghulu*) among the Jahai for dealings with officials.

The exact etymology of the ethnonym *Jahai* is unknown. It is suggested by van der Sluys (1999:307) that it is a compound of *ja* 'time before' and *hai* 'to walk in single file along the forest trails' and roughly means 'we who walk the trail of our ancestors'.[1] This is also the explanation given by some Jahai. Alternative spellings are *Jahay* and *Jehai*. The Jahai frequently also refer to themselves as *mendraq* 'people', as opposed to *gop* 'strangers', 'outsiders'.

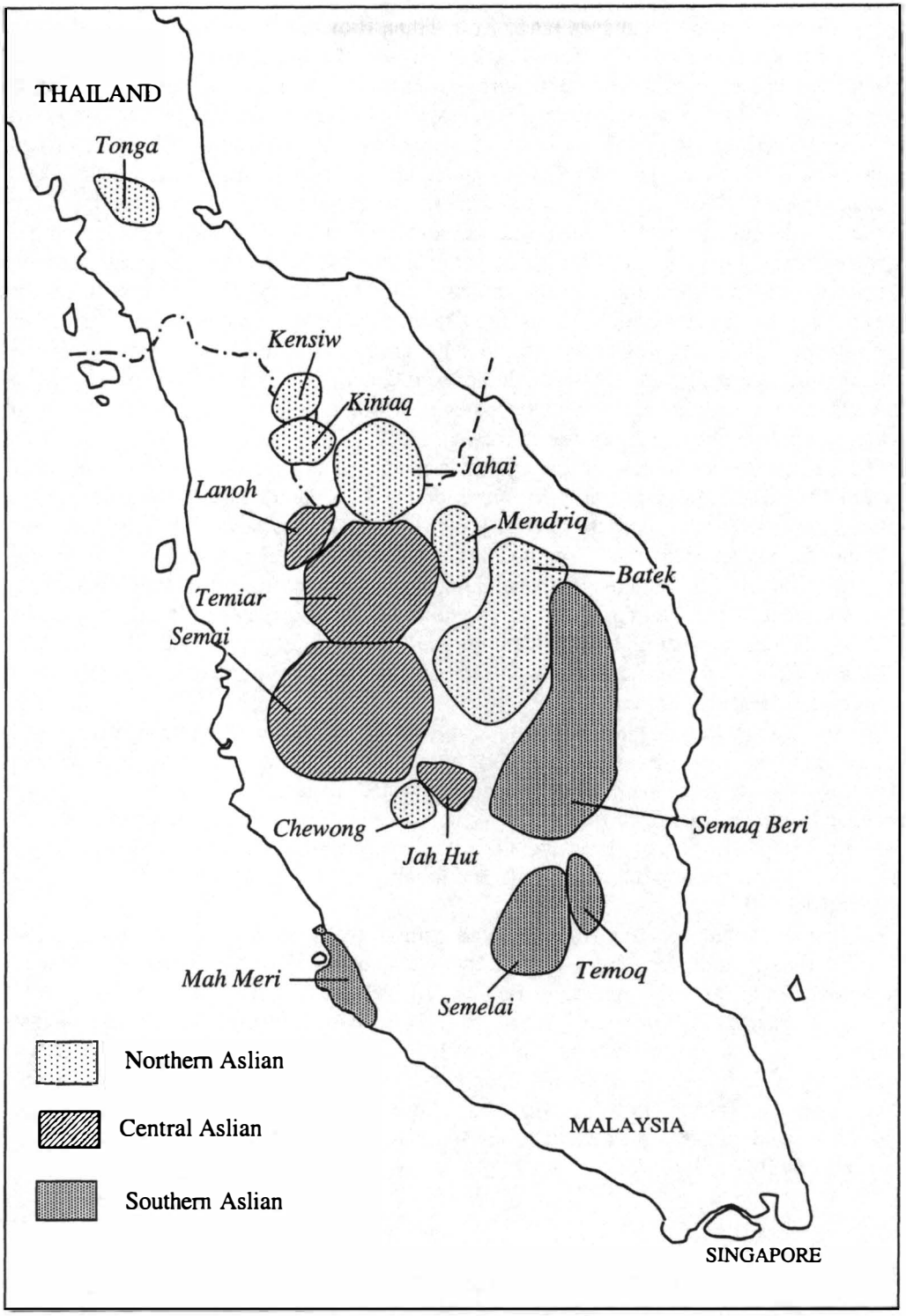
Early ethnographic accounts of the Jahai are found e.g. in Skeat and Blagden (1906), and more systematic anthropological work was later carried out by Schebesta (1928b, 1952, 1954, 1957). In the early 1990s, van der Sluys (1996, 1999, 2000) conducted anthropological research among the still mobile groups of the Temenggor area in Perak.

1.2 Linguistic classification and history

The language of the Jahai, referred to by the same name, is a member of the Northern Aslian subgroup of the Aslian languages, a geographically and genetically distinct branch of the Mon-Khmer language family, and, ultimately, the Austroasiatic stock. The Aslian languages (from Malay *Orang Asli* 'aboriginal people'), spoken by some 60,000 tribal people in the interior parts of the Malay Peninsula, are conventionally divided into three separate subgroups, Northern Aslian being one, and Central and Southern Aslian representing the other two. The characterisation of Aslian linguistic classification presented here is based on Diffloth (1975) and Benjamin (1976a, 2001, in press). The spelling conventions of language names are drawn from the latest writings of Benjamin (2001).

---

1 In the present work, this would correspond to the relational tense proclitic /ɟa=/ (see §4.10.1.1) and the verb /haj/ 'to follow'. The relational tense proclitic has not been found to attach to dynamic verbs in the present study of contemporary Jahai, so this interpretation of the ethnonym finds no synchronic linguistic support.



Map 1.1: Approximate distribution of Aslian languages (adapted from Benjamin 1976a:46, in press).

The Northern Aslian languages range from Trang Province in southern Thailand down into the Malaysian states of Kedah, Perak, Kelantan, Trengganu and Pahang. Northern Aslian may be further subdivided into three groups: a western one containing the closely related Kensiw and Kintaq languages, which are spoken on the border between Thailand and the Malaysian states of Kedah and Perak, and possibly also varieties spoken exclusively in Thailand, such as Tonga; an eastern one, including Jahai, Mendriq and several varieties of Batek, which together form a continuum of dialects ranging from southern Thailand down to Pahang; and the southerly and more distantly related outlier Chewong, spoken in western Pahang. To some extent, the Northern Aslian languages are associated with the Semang cultural sphere, introduced in §1.1. However, there are exceptions to this generalisation. Thus, the Chewong language, unquestionably a member of Northern Aslian, is spoken by people who are not Semang. And, inversely, some Semang groups speak languages which do not belong to the Northern Aslian subgroup. The total number of Northern Aslian speakers is estimated at a little more than 3000, of which speakers of Jahai represent about one-third.

The Central Aslian languages are spoken in a continuous area covering the interior of western Peninsular Malaysia, including parts of Perak, Kelantan and Pahang. The group includes at least four languages: Lanoh (which itself contains several distinct varieties, including Semnam and Sabüm), Temiar, Semai and Jah Hut. However, the status of the latter as a Central Aslian language is uncertain, and it has been suggested that it may form an independent, fourth branch of Aslian. Central Aslian languages are largely associated with populations subsisting on swidden horticulture; speakers of Lanoh, however, are usually considered part of the Semang cultural sphere. The total number of Central Aslian speakers is estimated at around 45,000.

The southern Aslian languages are found in two separate areas of Peninsular Malaysia. Three of the four languages — Semaq Beri, Semelai and Temoq — are spoken in an area stretching through much of central Pahang and adjacent parts of Trengganu and Negri Sembilan, whereas Mah Meri (or Besisi) is spoken in the coastal parts of southern Selangor. Southern Aslian is associated with populations primarily engaged in a subsistence system based on collecting and trading. The total number of speakers is approximately 9000.

The Aslian languages are firmly placed within the Mon-Khmer language family. Relatives thus include languages belonging to a number of branches scattered widely throughout Southeast Asia, including Bahnaric, Kammuic, Katuic, Khasi, Khmer, Mon, Nicobarese, Palaungic, Pearic and Viet-Muong. In the wider Austroasiatic context, distant relatives also include the Munda languages of India. However, the exact relation between Aslian and other branches of Mon-Khmer remains unclear. Some evidence appears to suggest a rather close relationship with Mon (Diffloth 1984), and that Mon, Aslian and Nicobarese may make up a Southern Mon-Khmer constellation. However, this is questioned by Bauer (1992:537).

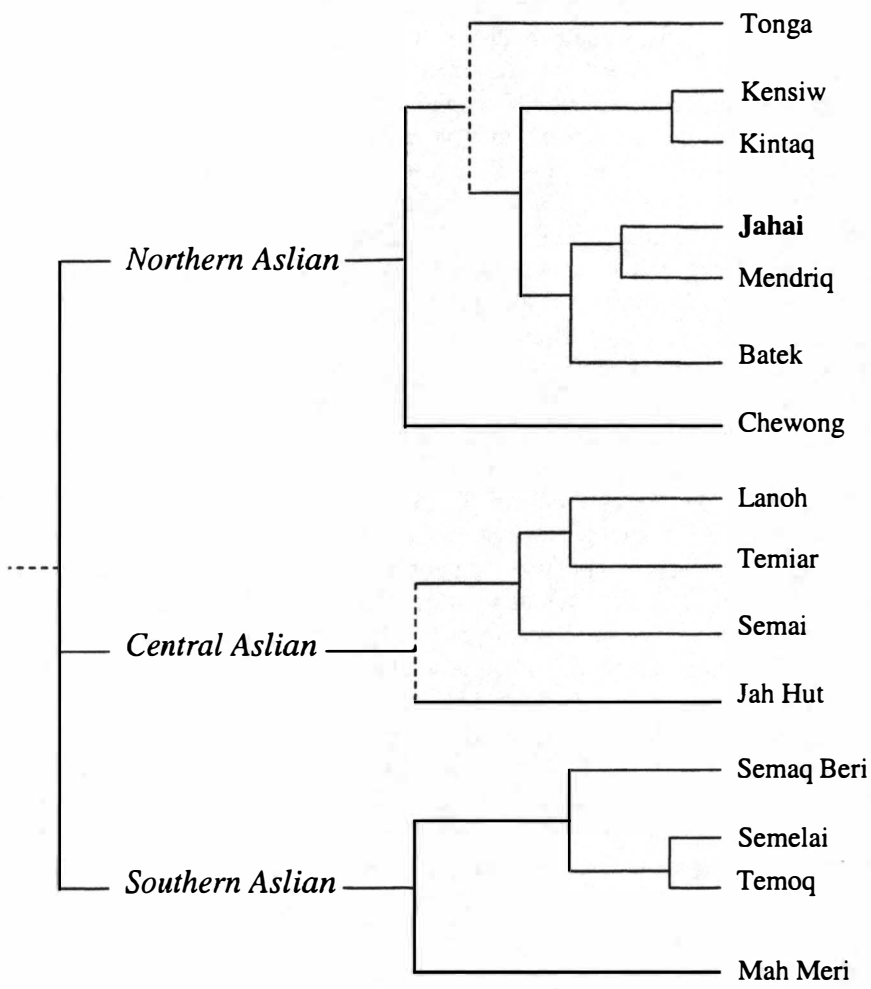


Figure 1.1: Aslian genetic relationships (simplified from Benjamin, in press)

Going beyond genetic relationships, there is evidence to suggest substantial secondary borrowing of vocabulary between the Aslian languages, and such intra-Aslian loans appear to have been particularly common among the Northern Aslian languages (Benjamin 1976a:74). For instance, Mendriq has exchanged vocabulary with both Batek and Kensiw, whereas the latter has a high loan rate with Lanoh, which, in turn, has exchanged loans with Jahai and Central Aslian Temiar. Moreover, Mendriq and some Batek varieties display a fairly significant loan rate with Southern Aslian languages (notably Semaq Beri) and, more surprisingly, Kensiw shows traces of vocabulary exchange with geographically distant Chewong (Benjamin 1976a:76–81; Bauer 1991:313).

In addition to such intra-Aslian loans, Aslian speakers have also borrowed extensively from non-Aslian languages. For instance, Malay, the unrelated Austronesian majority language in the peninsula, has had a considerable influence on the Aslian vocabulary. This was noted already by Blagden (1902, 1906a:435), and Benjamin (1976a:72–73) provides

figures of the loan rates from Malay in several Aslian languages. These show that, among the Northern Aslian languages, some varieties of Batek display the largest number of Malay loans, whereas Kensiw and Kintaq have the lowest rates.

Furthermore, some scholars have noticed that certain loan words in Aslian look conspicuously Austronesian but are clearly not of Malay origin, giving support to the idea of a pre-Malay Austronesian stratum in the peninsula (Blagden 1902, 1906a:435–438; Benjamin 1987:130–131; Bauer 1991:313). It has also been proposed that such loans provide evidence of some historical connection between Aslian and the Austronesian languages of Borneo (Adelaar 1995:87–91). Yet another source of loans is Thai, especially among some Northern Aslian languages, although Bauer (1991:313) comments on the scarcity of such loans in Kensiw. Bauer (1992:536–537) also points to the possibility of loans due to historic contact with Mon.

It should also be mentioned that some early writers pointed out that Northern Aslian languages contain quite a few lexical elements of unknown origin, that is words that appear to be neither Mon-Khmer cognates nor secondary loans of any sort. These were interpreted as remains of a non-Mon-Khmer substratum language spoken by the Negritos before the arrival of Aslian. On account of the physical and cultural similarities between the peninsular Negritos and the indigenous population of the Andaman Islands, in the Bay of Bengal, attempts were made to link these supposed substratal remains to the distinctive and genetically isolated Andamanese languages according to what Zide and Pandya (1989:648–650) and others have labelled the 'Proto-Negrito Hypothesis' (see e.g. Trombetti 1923:64). The ideas of a Negrito substratum and an Andamanese 'link' have remained popular and are cautiously maintained by e.g. Benjamin (1976a:83) and Matisoff (2003:8–9). However, the number of corresponding lexical items has not been significant enough to warrant a clear genetic connection (Bloch 1952:512), and some researchers, notably Gérard Diffloth, have argued that much of the supposed substratum vocabulary is indeed of Mon-Khmer origin (Diffloth, pers. comm.; Zide and Pandya 1989:649). The 'Proto-Negrito Hypothesis' must therefore be regarded as highly speculative.

1.3 The sociolinguistic situation

1.3.1 Idiolects, dialects and multilingualism

Anthropologists have frequently commented on the high degree of idiolectal variation and change in Northern Aslian languages (Endicott 1990; Benjamin 1976a:76, 1985a:234–235; 1987:114, in press a). The mobile lifestyle of Northern Aslian speakers, manifested in their system of intermarriage between individuals of widely dispersed groups, as well as in their marked pattern of group disintegration and regrouping into new constellations as an adaptive response to ever-changing subsistence conditions, has implications on the language of individual speakers. A speaker may move through several linguistic environments throughout his or her lifetime, leading to an overt and unusually high rate of idiolectal change. At the same time, the manifold linguistic origins of the members of a group also lead to marked variations in the language use of different individuals. Benjamin (in press) speaks of 'a mesh-like relation between different varieties of Northern Aslian, which is as much idiolectal as dialectal'. As a result of the mix of language varieties, it is difficult to identify clear-cut language boundaries within the Northern Aslian group. A hint of the intricacy of the problem is provided by Bishop and Peterson (1993:1), who report

that in one Semang settlement of southern Thailand they came upon six languages and/or dialects among the 13 adults present.

A related characteristic of Northern Aslian speech communities is the multilingualism exhibited by their speakers. It is not unusual for them to speak two or more languages fluently (Benjamin, in press). This is due to their frequent contact with speakers of neighbouring Northern Aslian languages, as well as languages belonging to other branches of Aslian, and majority languages like Malay and Thai.

These typically Northern Aslian patterns of idiolectal variation and multilingualism largely apply to the Jahai speech community. Although speakers claim that different Jahai groups speak differently, it is very difficult to find systematic dialectal variation between groups, idiolectal differences within one and the same group being more prominent. Manifold linguistic origins of members of the Jahai speech community are evident, and it is not uncommon to find individuals whose native language is not Jahai but Mendriq, Batek, Lanoh or Temiar. It is consequently also common for speakers to have parents from two different speech communities. Temiar holds a special position in this respect, since intermarriage between Jahai and Temiar is common.

As to multilingualism, many Jahai speak both Temiar and the local dialect of Malay fluently as second languages. Temiar is the *lingua franca* used when talking to speakers of other Aslian languages in the area; Malay is used in dealings with the Malay-speaking majority. Gender differences may exist, women generally being less proficient in these second languages. This is most probably due to the fact that it is the men who are engaged in wage-labour and trade and for whom a *lingua franca* is necessary. Native Jahai speakers who have remained within the Jahai speech community generally do not speak additional Aslian languages other than Temiar. Some individuals who trace their origin to northerly groups can speak Southern Thai. Reportedly, northerly groups of Jahai (notably in the remote Belum River area) are less proficient in Temiar and Malay.

1.3.2 Schooling and literacy

Adult Jahai lack formal education and are generally non-literate. However, an increasing number of children living in the regroupment areas receive primary education locally in Malay. Jahai is not used in education, and it is not a written language.

1.3.3 Endangerment

Northern Aslian languages have very few speakers, figures varying from about 150 (Mendriq) to a little more than 1000 (Jahai). Some Northern Aslian languages are known to have died out during the past two centuries, including the 'Bila', 'Wila' or 'Lowland Semang' language spoken on the coast opposite Penang in the early 1800s (Blagden 1906a:390–391; Benjamin 1976a:50), as well as varieties spoken in Kedah and lowland Perak until the 1920s (Benjamin, in press). This is in part connected to a general long-term trend of cultural and linguistic assimilation of Aslian speakers, either to the Malay community or to a larger Aslian-speaking group. Using population figures gathered during the past century, Wazir (1996:9–11, 2001) argues that the Semang show a discouraging pattern of population growth and expresses concern about the future of their languages. According to the population figures given, the number of speakers of some languages (notably Mendriq) has dropped by almost two-thirds between 1924 and 1994, whereas the

number of speakers of Batek has increased by over 700 per cent during the same period. The number of speakers of Jahai, for comparison, has increased by approximately 26 per cent. The total population growth for all Semang groups in Malaysia was 41 per cent between 1969 and 1994.

Benjamin (in press), on the other hand, suggests that the Northern Aslian speech communities are not necessarily in immediate danger of becoming extinct and points out that their small numbers of speakers have been maintained for a very long period of time as a consequence of their small-scale and mobile lifestyle. He also argues that Northern Aslian speakers have long been used to linguistic contact and linguistic non-uniformity and are therefore likely to be well-prepared to maintain their own linguistic identity.

Jahai, being the largest Northern Aslian language, does not appear to be in immediate danger of extinction. For children of Jahai parents, Jahai always represents the first language, and, in the absence of people who do not speak Jahai fluently, it is the language used in everyday situations by most Jahai. It is possible, however, that increased permanent settlement and contact with the outside world, facilitated by the recently constructed East-West Highway, which runs straight through Jahai territory, may pose a long-term threat to the language. Perhaps a more immediate linguistic threat to some southerly groups of Jahai, who are increasingly settling down in villages also populated by speakers of Central Aslian Temiar, is the frequent intermarriage with the Temiar. Children of such mixed descent almost invariably grow up with Temiar as their dominant language. It is possible that an expanding Temiar language represents a greater long-term threat to the existence of Jahai than does Malay.

1.4 Previous research

Section 1.4.1 describes previous linguistic sources pertaining specifically to Jahai. Section 1.4.2 summarises research and references related to other Aslian languages or to the Aslian sub-branch as a whole.

1.4.1 Research on Jahai

The earliest available sources containing linguistic data which can be possibly identified as Jahai are word lists collected during the 1800s and early 1900s, e.g. Miklucho-Maclay (1878) and Savage (1926).[2] However, the first account explicitly concerned with Jahai language is a brief grammatical sketch by the ethnographer Father Paul Schebesta (1928a). This work is described in detail in §1.4.1.1. More recently, Geoffrey Benjamin and Gérard Diffloth have collected Jahai lexical items for lexicostatistical and comparative studies of Aslian languages (see e.g. Benjamin 1976a; Diffloth 1975). Benjamin has also used such lexical material to study intra-Aslian and Malay loan rates. He notes for Jahai a conspicuous absence of vocabulary exchange between Jahai and its Northern Aslian neighbours Kensiw and Mendriq but a high loan rate between Jahai and Central Aslian Lanoh. This leads him to believe that Jahai has expanded to its present area of distribution from the south in relatively recent times (Benjamin 1976a:77). Furthermore, brief reference

2 Geoffrey Benjamin (pers. comm.) suggests that Semang songs recorded early on by W.W. Skeat in Kedah and Pattani and transcribed by R.J. Lloyd in Skeat and Blagden (1906:627–629) are in Jahai. Although much of this transcribed material does bear some similarity to the variety of Jahai described in the present work, it has not been possible to confirm this suggestion.

to the Jahai vowel system, based on unpublished field-notes, is made in Benjamin (1986:6). Previous papers on Jahai by the present author include Burenhult (2000, 2001a, 2001b, 2003, 2004a and 2004b).

1.4.1.1 Schebesta's grammatical sketch

Father Schebesta's (1928a) brief description of Jahai, translated and reworked by Charles Otto Blagden, provides one of the first grammatical accounts of any Aslian language. It introduces some aspects of the sound system, the word classes and word formation in Jahai, and it also includes three short texts. It is richly illustrated with what appears to be examples of spontaneous Jahai language use, but the various grammatical phenomena are discussed only summarily. It is not stated whether the description covers a particular variety of Jahai, nor is it said where Schebesta's informants lived. However, according to Geoffrey Benjamin (pers. comm.), Schebesta did most of his ethnographic work on the Jahai in Bersiak (present-day Kampung Bersia), a village on the Perak river about 10 kilometres upstream from Gerik, in Hulu Perak district, Perak state. It is therefore likely that the account is based on information from that area. Reportedly, the area is nowadays not inhabited by Jahai speakers, the nearest present-day Jahai settlements being located on Lake Temenggor, a further 30 kilometres or so upstream.

The description begins with an introductory note by C.O. Blagden on the orthography employed; apparently the Jahai examples in Schebesta's original version were written in the Anthropos alphabet but transcribed into the International Phonetic Alphabet by Blagden. This is followed by a short account of the sounds of the language, including phonetic exemplification of the various vowels and consonants and some discussion on the syllable- and/or word-final segments referred to in the present work as unreleased stops and prestopped nasals (cf. §2.3.1.1 and §2.3.1.2). The latter are interpreted as word-final stops followed by 'an obscure nasal release'. A number of diphthongs are listed, exhibiting either an [i] or [u] offglide; these correspond in the present work to vowel + approximant /j/ or /w/ (cf. §2.2 and §2.3.1.5). Two of the vowels listed are described as having nasal counterparts: [i] and [ɔ]. As mentioned, the representation is phonetic and no attempt is made to systematise the sounds phonemically. Syllable structure is only briefly touched upon, and it is stated that initials and finals may be either a vowel or a consonant. Stress is said to fall on the last syllable of a word. It is further suggested that there are tonal differences in a limited set of lexical items, and a short list of minimal or near-minimal pairs is given (see §2.5.2).

The description of word classes treats numerals, pronouns, nouns, adjectives and adverbs. Numerals are only briefly introduced, and it is stated that Jahai has only one true numeral, *nai* 'one', and that other numerals are borrowed from Malay (cf. §4.5.1). The system of pronouns involves singular, dual and plural number, with first, second and third person distinction for singular pronouns, but only first and non-first distinction for dual and plural pronouns. Inclusive and exclusive forms are given for first person plural. Two forms of first person dual are given, both labelled 'inclusive' and apparently in free variation. Two forms of second person dual are labelled 'exclusive'. It is not entirely clear what this terminology signifies, but it is stated that one of the forms is used to refer to two persons who are not present (cf. §4.3). Pronominal use is amply illustrated for subject and object position as well as in possessive constructions. The same pronominal forms are used for these three functions. As for demonstratives, three basic distinctions are given, corresponding to 'this', 'that' and 'that yonder' (cf. §4.4). These are said to be linked to

the pronoun or noun that they follow by means of *-t-*. Additional locatives include two forms signifying location upstream and downstream, and four forms correspond to the four cardinal points. Relative pronouns are said to be usually absent, 'the relative clause being merely co-ordinated with the principal clause' (Schebesta 1928a:810). However, a particle *na* is considered to be used occasionally to mark relativisation (cf. §4.12). Interrogatives are introduced briefly and exemplified by forms corresponding to 'who/whose?', 'what/why?', 'how many?' and 'where?' (cf. §4.6). An interrogative particle *ha* is also exemplified (cf. §4.10.3).

The section on nouns begins with a description of morphology related to number. Several plural-marking affixes are introduced, most of which are clearly associated with human nouns. Plural forms are said to be usually identical to the singular ones, and plural marking gives the impression of being optional (cf. §4.1.1). However, one of the affixes listed as a plural marker, usually involving an infixed *-n-* after the first consonant, is said to occur when the noun follows a numeral or the word for 'many'. This affix is not restricted to human nouns but may occur with most nouns (cf. §4.1.3). In a subsection on case, the syntactic behaviour of nouns is richly exemplified and a set of case-marking prepositions is introduced (cf. §5.2). The *nominative* is said to come first in a sentence, usually repeated by a pronoun. If occurring after the predicate, the nominative is expressed by putting the preposition *ka*, or occasionally *de*, before the noun. The *genitive* is expressed by placing the possessor after the possessed noun, sometimes possibly with the preposition *de* before the possessor. *Dative* is marked by the preposition *kɛ* or *de*, or without any preposition at all. *Locative* and *instrumental* are marked with the preposition *kɛ*. The *accusative*, finally, is described as usually occurring after the verb without any preposition, but occasionally it is introduced by *kɛ*.

Adjectives are said to follow the noun. With a few exceptions, adjectives used as attributives receive the prefix *t-* (cf. §4.12.2); adjectives used as predicates are unmarked (cf. §4.7). In the section on verbs, usage is exemplified with paradigms involving different subject pronouns. Several preverbal particles are introduced, including a past tense marker (cf. §4.10.1.1) and a particle *ťa* (*uťa* in third person singular) which expresses desire, will or future (cf. §4.7.4.1). A prefix *pi* is considered to form a sort of causative (cf. §4.7.2). Finally, examples of different moods are given, including conditionals, interrogatives, imperatives, negatives and prohibitives. A short section on adverbs suggests that these may be formed by means of a prefix *le*.

A brief section on word formation exemplifies reduplication, which is said to be particularly common with verbs but not leading to any significant change in meaning. The examples given are not systematically analysed but include full as well as partial reduplication, sometimes with vowel alternation (cf. §3.2).

Although short, laconic and tentative, Schebesta's description provides an interesting and enlightening early glimpse of Jahai language. It forms an excellent piece of comparative material for the present-day Aslianist, its main virtue perhaps being the great number of authentic examples given.

1.4.2 Other research related to Aslian

This section outlines research related to Aslian languages other than Jahai. An overview of research pertaining to the Aslian branch as a whole is given in §1.4.2.1, whereas sections §1.4.2.2, §1.4.2.3 and §1.4.2.4 describe work carried out on the respective subbranches of Aslian.

1.4.2.1 General

The term 'Aslian' (from Malay *Orang Asli* 'aboriginal people') as a generic designation of the group of Mon-Khmer languages spoken in the Malay Peninsula was first coined by Gérard Diffloth and later introduced in print by Diffloth (1974) and Benjamin (1976a). The three subbranches of Aslian were labelled Jahaic, Senoic and Semelaic by Diffloth, though the alternative, geographical terms proposed by Benjamin — Northern, Central and Southern respectively — have gained wider acceptance.

One of the first to recognise that several of the minority languages in the Malay Peninsula were related to Mon-Khmer was Schmidt (1901, 1903), who made a detailed classification of Aslian based on vocabulary. This was followed by a refined classification by Blagden (1906b), based not only on vocabulary but also on phonological features. Other classifications include that of Pinnow (1959:4–5) and rely heavily on earlier work.

In recent decades, two scholars have been particularly active within the field of Aslian linguistics: Geoffrey Benjamin and Gérard Diffloth. Benjamin, a linguistically trained anthropologist, made an extensive collection of lexical samples based on the Swadesh list from a large number of Aslian languages. Using lexicostatistical methods, he put forward a detailed genetic classification of Aslian and its sub-branches and advanced an historical interpretation of the linguistic data (Benjamin 1976a:37–94). Apart from his work on Central Aslian Temiar (see §1.4.2.3), Benjamin has also produced several recent papers on general aspects of Aslian (Benjamin 2001, 2004). Furthermore, he has been very active in accentuating the need for urgent scientific attention to several topics of Aslian linguistics (see e.g. Benjamin 1989:20–23). He is also the author of a guide to the pronunciation and transcription of Aslian languages for anthropologists and other non-linguists working among speakers of Aslian languages (Benjamin 1985b, 1986).

Diffloth, who conducted extensive linguistic fieldwork among speakers of Central Aslian Semai and Jah Hut, Northern Aslian Chewong and Southern Aslian Semelai in the 1960s and 1970s (see §1.4.2.3), made a detailed genetic classification of the Aslian languages on the basis of comparative phonology (Diffloth 1968, 1975, 1977, 1979), the results of which are in accordance with those produced in Benjamin's lexicostatistical study. Diffloth's general work on Aslian also includes a study of numerals (Diffloth 1976c).

A summary of research carried out on the Aslian languages up until the early 1980s is given in a recently published paper by Matisoff (2003). Drawing on the works of Asmah, Benjamin, Diffloth and others, Matisoff provides an account of the phonological, morphosyntactical, semantic and lexical characteristics of Aslian, paying special attention to those features that are regarded as peculiar to this branch of Mon-Khmer.

Other general work incorporating data from Aslian languages include e.g. Adams's (1989) study of numeral classifiers in the Austroasiatic language family, in which some older Aslian material plays an important role, as well as a short paper by Bauer (1992) on the relationship between Aslian and Mon.

As far as bibliographical work is concerned, a rather aged list of works on Peninsular Mon-Khmer can be found in Shorto, Jacob and Simmonds (1963:30–32). This was compiled before Benjamin and Diffloth took to the field and therefore obviously lacks references to the more modern developments in Aslian linguistics. More recent bibliographies include those of Parkin (1991:152–159), Bishop and Peterson (1995) and Burenhult (1999), and quite a few references to work on Aslian appear in Huffman (1986). Also, an extensive list of references relating to Orang Asli research in general, including linguistic works, has recently been compiled by Lye (2001).

1.4.2.2 Northern Aslian

Northern Aslian languages attracted some attention in the 1800s and early 1900s, and several vocabularies were collected. For example, see Crawfurd (1820:125–191), Hewett (1880), Swettenham (1880), Evans (1915), Evans (1927:8–12), Miklucho-Maclay (1878), Savage (1926) and Ogilvie (1949) for various Northern Aslian languages. Later collections include Carey (1970) for Mendriq. The first grammatical description was Schebesta's (1928a) account of Jahai, summarised in §1.4.1.1.

Comprehensive work was carried out in the 1960s on Kintaq by Asmah (1964, unseen), although only a short essay based on parts of this work has been published (Asmah 1976). This deals exclusively with the Kintaq verb, notably a set of affixes signalling aspect (desiderative, perfective and imperfective) and causative.

More recently, several scholars working on the Thai side of the border have produced works on various aspects of Kensiw, closely related to Kintaq. For example, Bishop (1992, 1996a) provides accounts of the phonology of the Kensiw dialect spoken at Bansakai, in Yala Province, southern Thailand, including extensive descriptions of phonemes, prosodic features and word/syllable structure. Of particular interest is the complex vowel system, displaying five tongue heights (atypical for Aslian) and distinctive nasality. Pitch differences are said to be contrastive in a small set of lexical items, a feature similar to the one noted by Schebesta (1928a:805) in Jahai (see §1.4.1.1, §2.5.2). Previous accounts of Yala Kensiw phonology include that of Phaiboon (1984, unseen). Also, a phonological description of the northerly Kensiw dialect spoken in Trang Province, Thailand, has been produced by Bauer (1991). Kensiw vocabulary is treated in e.g. Bishop (1996b) and Bishop and Peterson (1994), and Peterson (1993, unseen) describes the use of spatial locatives in Kensiw. Bishop (2001) describes sociolinguistic aspects of Kensiw.

On the basis of Schebesta (1928a), Bishop (1996a) and Phaiboon (1984), Hajek (2003) claims there is enough evidence to conclude that Northern Aslian languages exhibit some kind of tonal activity.

1.4.2.3 Central Aslian

Like Northern Aslian, languages of the Central subbranch of Aslian received some attention in the late 1800s and early 1900s. Early descriptions include that of Clifford (1891), and for vocabularies, see e.g. Daly (1880) and Wilkinson (1915). Schebesta (1931) provided a short description of the grammar of the so-called 'Ple-Temer' language. On the basis of this description, Benjamin (in press) states that this language is notably different from Temiar proper. Failing to find evidence for it in the field, he suggests that it is an extinct intermediate dialect between Temiar and Lanoh. Schebesta's account, which is preceded by a note by the translator C.O. Blagden on transcription, treats word classes and also includes two texts.

Three Central Aslian languages have been subjects of scholarly attention in recent years. Temiar has been studied by Carey (1961, unseen) and Benjamin (1976b, 1996), whereas Diffloth has produced several articles on Semai (see e.g. Diffloth 1968, 1972, 1974, 1976a, 1976d, 1977). Benjamin's (1976b) study involves a concise account of Temiar grammar, including phonetics, phonology and the morphosyntax of nominal and verbal elements. His unpublished essay *The anthropology of grammar: self and other in Temiar* (Benjamin 1996) is an extension and partly a revision of this study, involving an attempt at explaining significant features of Temiar grammar in light of a cultural notion of a Self/Other

distinction present in the Temiar-speaking society. Benjamin claims that this distinction is manifested through phonetic iconicity in a large number of linguistic forms, including e.g. deictic categories like pronouns and demonstratives, verbal affixes signalling voice, modal particles, role-marking particles as well as number affixes on human nouns. Self-associated forms are shown to display sounds that can be characterised as 'front' and 'closed', whereas Other-associated forms are phonetically open. Thus, Benjamin argues that phonetic iconicity based on a cultural notion imbues the whole Temiar grammatical system and links seemingly disparate areas of grammar.

Benjamin's writings on Temiar have opened the eyes of the linguistic community to the morphological complexity exhibited by Aslian languages. Matisoff (2003:22) has since suggested that the morphological resources of Aslian are 'among the richest in all of Southeast Asia, unrivalled even in most of the rest of Austroasiatic'. The Temiar verbal paradigm — with its many distinctions of voice, aspect and nominalisation — is sometimes described as the most regular and productive within Mon-Khmer. Furthermore, processes of reduplication and infixation described in Benjamin 1976b have played an important role in the development of various prosodic approaches to non-concatenative morphology (McCarthy 1982:208–221; Broselow and McCarthy 1983:38–43; Ter Mors 1984:279–295). They have also been treated within the framework of Optimality Theory (Gafos 1998).

Diffloth's work on Semai covers various linguistic domains, although much of the literature deals with historical phonology and the reconstruction of Proto Semai (Diffloth 1968, 1977). Diffloth (1972) discusses the ambiguity of certain Semai morphemes, and his 1974 paper describes how transitive sentences involving body movement behave differently from other transitive sentences with regard to word order shift. Diffloth (1976d) describes the peculiar word class known as expressives, and the Semai data presented has shaped much of the subsequent discussion of this category of words. The reduplicative features of Semai expressives illustrated in Diffloth's article have recently been analysed within the framework of the Compression Model of Optimality Theory (Hendricks 2001).

Diffloth (1976b) provides a concise and comprehensive description of Jah Hut that includes discussion on its genetic affiliation, phonology, morphology and syntax. The syntactic section contains among other things an account of expressives and also a description of the distribution of an agentive particle which suggests that Jah Hut displays an unusual type of ergativity.

1.4.2.4 Southern Aslian

For a long time, scholarly work specifically devoted to Southern Aslian was markedly scanty compared to that of Northern and Central Aslian. Early descriptions are rare and include notes on Temoq and Semelai (Collings 1949). A short description of Semaq Beri is to be found in Nik Safiah and Ton (1979). Diffloth's fieldwork on Semelai remains unpublished. However, in recent years extensive work has been carried out on Semelai by Kruspe (1999, 2004). Her (2004) grammar forms the most comprehensive description of any Aslian language, containing exhaustive treatment of phonology, morphology and syntax. The work examines in detail several aspects of Semelai which are relevant to the Aslian branch as a whole but which have been touched on only briefly in earlier descriptive work, e.g. its intricate morphological processes of reduplication and infixation, as well as the distribution of case-marking prepositions.

1.5 The present study

1.5.1 Background and aims

The present work is a descriptive study of contemporary Jahai. It was conceived in 1997 and initiated in 1998 with the primary aim of charting the morphology of the language. During the course of the data collection, the need for a more general linguistic description became increasingly evident, and a wider approach incorporating phonology, morphology and syntax was assumed. The purpose of this work, however, is not to provide a complete and exhaustive linguistic description; in that sense it is not to be viewed as a reference grammar. Rather, it should be regarded as a general introduction to the subject and a basis for further research.

Although the study largely aims at providing an unbiased description of Jahai on the language's own merits, certain general problems have inevitably served as starting-points and sources of inspiration for the line of work. One such problem is the question of morphological complexity. As we have seen, Aslian languages, especially those of the Central subbranch, are known to exhibit some of the most regular and productive systems of inflectional and derivational morphology in Mainland Southeast Asia. Northern Aslian languages, on the other hand, are occasionally described as exceptions to this pattern. For example, Benjamin (pers. comm.) suggests that Northern Aslian morphology is largely fossilised and unproductive, and Bauer (1991:313) describes 'progressive loss of morphology without extensive reanalysis' as a typical Northern Aslian feature. The writings of Schebesta (1928a) and Asmah (1976) provide indications to the contrary, and an important initial issue for the present study was to elucidate the status of Northern Aslian morphology on the basis of data from Jahai. Another question at issue was previous claims that some Northern Aslian languages exhibited contrastive tone, at least in small sets of lexical items. Tones are otherwise not a feature of Aslian languages, and one could speculate that the proximity of Northern Aslian to tonal Thai would have led to the adoption of this feature. Thus, Schebesta (1928a:805) believed he had identified tonal differences in some words in Jahai, as does Bishop (1996a:238–239) for Yala Kensiw. Phaiboon (1984, unseen) treats Yala Kensiw as fully tonal, which is rejected by Benjamin (in press). Bauer (1991) does not identify tones in Trang Kensiw. New research on the sound system of Jahai was likely to shed further light on this problem.

The present work is the first grammar of a Northern Aslian language to appear since Asmah's unpublished 1964 description of Kintaq. As such, it is intended to expand our knowledge of Aslian as a whole, grammatical descriptions of which have so far been largely concerned with Central Aslian languages, notably Temiar (Benjamin 1976b) and Jah Hut (Diffloth 1976b). In this respect, its appearance is timely because it coincides with the publication of the first comprehensive description of a Southern Aslian language, Kruspe's (2004) grammar of Semelai. It is also intended as a contribution to Mon-Khmer and Southeast Asian language studies in general, and, hopefully, also to a wider linguistic context. Furthermore, it is hoped that it may serve as a source of linguistic information for anthropologists and others working among the Jahai and other Northern Aslian speakers.

Finally, a central aspect of any study of a language with few speakers is the sheer documentation. With their limited number of speakers, Jahai and the other Northern Aslian languages represent some of the many endangered languages of the world. Although Jahai does not appear to be in immediate danger of extinction, it is clear that the gradually changing lifestyle of the Jahai presents a long-term threat to the existence of their language. Every linguistic extinction represents a serious loss of information about

linguistic variation, and the documentation of small and endangered languages is therefore of great scientific interest.

1.5.2 Fieldwork, methodology and data

The data on which the present description of Jahai is based was collected intermittently in the field among Jahai speakers during the period 1998–2000. The total amount of time spent in direct interaction with Jahai speakers was about six months, distributed over four field trips of one to two months' duration each. In the intervening periods, the data was systematised and continuously analysed. This periodic approach furthered the data collection in that the systematisation away from the field served as a useful foundation for each new round of fieldwork and facilitated the author's acquisition of Jahai.

The fieldwork was carried out in a regroupment program area, or Rancangan Pengumpulan Semula (RPS), called Air Banun, located on the eastern shores of Lake Temenggor, south of the East-West Highway, and administered by the Jabatan Hal Ehwal Orang Asli (JHEOA, Department of Aboriginal Affairs). The RPS is situated in an area of mostly primary Dipterocarp rainforest and contains a few permanent villages located along the lake or the major rivers, including Sungai Banun, Sungai Raba, Permai and Damai, all of which are predominantly inhabited by Jahai speakers. It also has an administrative centre with a school, a JHEOA office and basic medical facilities. An additional eight permanent or semi-permanent villages are scattered throughout a large area outside the RPS, from which they are administered. Most of these villages are also inhabited by Jahai speakers, although three of them are inhabited predominantly by Temiar speakers.

Fieldwork was mainly restricted to the village of Sungai Banun, located within the RPS at the confluence of the Banun and Raba rivers, not far from where the Banun empties into Lake Temenggor. The village, inhabited by some 80–100 Jahai speakers, is permanent and constantly inhabited, but its population can be characterised as semi-settled as parts of it may spend longer or shorter periods on the move or in other settlements. Although the inhabitants of the village are likely to originate from various scattered groups, many of them (including most of the informants of this study) claim to be members of a band which originally roamed an area centred on the Mangga and Kelap rivers, two northern tributaries of the Singor located some 15–25 kilometres east of RPS Air Banun. If questioned about the ethnonym of this particular group, some of its members will refer to it as Jahai To', 'ancestral Jahai'. However, there does not appear to be a generally agreed-upon system of ethnonyms for the different subgroups of Jahai. The present linguistic description rests almost entirely on data collected from members of this group; however, comparative data has occasionally been obtained from inhabitants of other villages, including Sungai Raba, Permai, Damai, Pulau Tujuh, Sungai Tekam and Ulu Tiang, as well as visitors from RPS Sungai Rual, in Kelantan, and members of still mobile groups east of RPS Air Banun. However, no systematic comparison of different speech varieties has been made for the present study.

Apart from what could be gathered from Schebesta's (1928a) sketch of Jahai, the author had no prior knowledge about the language. At first, questioning was done in basic Malay but after a few weeks increasingly in the author's own hesitant Jahai. During the last period of fieldwork, in October–November 2000, there were also opportunities to make interviews with a Jahai man who knows basic English. For the most part, however, data collection has been conducted in Jahai.

At the beginning, work focussed on the collecting of an extensive Jahai word list. The starting-point for this was the list of items used in the vocabulary compiled for Kammu by Lindell (1974), which in turn is based on Egerod (1965), a list specifically developed for Southeast Asian languages. The Jahai list has been continuously expanded and currently contains 1730 items (see Appendix I). Collection of faunal terminology was made with the help of the photographically illustrated works of Whitten (1998), Strange (1998) and Cox et al. (1998). Also, translation of numerous lexical items was made via Malay with the help of Coope (1993). The lexical material has since served as a basis for the phonological analysis presented in Chapter 2. Much of the subsequent fieldwork was concerned with the identification of bound morphemes, and extensive lists of derived word forms were elicited in order to chart the allomorphy of specific morphemes. Initially, Schebesta's 1928a description provided helpful hints for the discovery of some such morphemes, and several more were identified and investigated as the material grew larger. The morpheme inventory was also occasionally cross-checked with that of other Aslian languages, and, in at least one case, descriptive material from another Aslian language (Asmah's 1976 account of the Kintaq verb) has helped to reveal a morpheme in Jahai which would otherwise possibly have gone unnoticed. Elicitation of syntactic and morphological phenomena was also largely inspired by the Lingua Descriptive Studies questionnaire in Reichling et al. (1977:11–57). Such elicitation typically involved the presentation by the author of potential, artificial Jahai forms to the informant for judgement of acceptability and possible correction.

Thus, elicitation has been an important tool in the field for the detection and identification of various linguistic phenomena. However, it became clear early on that elicited material was not entirely reliable, partly because informants tended to equate acceptability of linguistic forms with comprehensibility rather than grammaticality. Also, elicitation sometimes proved to result in misleading over-generalisations on the informants' part. So whereas elicitation has been invaluable as a primary means of detecting patterns and tendencies, it was decided that the final analysis would rather rest mainly on recordings of authentic language use. These recordings, made continuously during the fieldwork periods, involve eleven hours of video material and fifteen hours of audio material and include long sequences of spontaneous conversation as well as narratives[3]. The latter includes e.g. hunting-stories and myths, particularly from genres referred to by the Jahai as /cnɛl/, which explain the genesis of animal species (van der Sluys 1996:13–15). Unless stated otherwise, the Jahai examples given throughout this work are drawn from such authentic text materials.

Much of the elicitation was done with a single informant, the middle-aged headman of Sungai Banun. Thus, the word list and the phonological analysis rests almost exclusively on his idiolect of Jahai, although pronunciation has been frequently cross-checked with other speakers. The recordings of spontaneous language use, however, involve an additional number of mostly male speakers of different ages. Analyses based on such recordings thus draw on the language use of several speakers. The idiolectal variation described in §1.3.1 has generally not presented any significant difficulties in the analysis of linguistic forms. An important exception is the class of expressives (see §4.8 and Chapter 6), some manifestations of which are intimately associated with specific individuals who are not originally members of the group studied. Thus, only two persons made use of what is referred to here as 'expressive elaboration', both of whom originated

---

3 I am grateful to Dr Marianne Gullberg, Max Planck Institute for Psycholinguistics, Nijmegen, for equipping me with a video camera.

from distant groups and both of whom were unusually knowledgeable about other Aslian languages. Such expressive elaboration was rejected by other speakers. Furthermore, some discrepancies encountered in the vowel system (see §2.2) may be attributed to similar individual variation. The present study is concerned with the language use of most of the Jahai speakers of Sungai Banun and does not claim to describe other varieties of Jahai. For the sake of completeness, however, deviant linguistic behaviour such as that encountered in expressive elaboration is also described but treated separately (Chapter 6) and not considered a general feature of Jahai.

The overriding aim of the present work is linguistic description, and it is therefore to be regarded as essentially non-theoretical. This means that the purpose of the study is not primarily to create, develop, modify, support, refute or in other ways evaluate particular theoretical frameworks. However, where considered particularly suitable and helpful, specific theoretical models are adopted as descriptive tools in order to provide as explicit, economic and systematic an account as possible of the material available. For example, this applies to the models of Prosodic and Template Morphology used to describe processes of affixation (§3.2). Similarly, the monostratal approach to syntactic formalisation proposed by Van Valin and LaPolla (1997) has inspired some of the terminology employed here to illustrate syntactic structures (Chapter 5).

1.5.3 Outline

This chapter has provided a brief introduction to the Jahai, their language and related research. The following chapters are concerned with the results of the present study. Chapter 2 analyses the phonological system of the language and provides the basis for the orthography used. Chapter 3 examines the units and intricate processes of Jahai morphology. Chapter 4 identifies and describes the word classes as well as the morphological categories associated with them. Chapter 5 outlines tentatively the main features of Jahai syntax. Chapter 6 provides a short account of the enigmatic phenomenon referred to here as 'expressive elaboration'. Some concluding remarks are given in Chapter 7. The appendices, finally, contain the full list of Jahai words collected for the present study.

1.5.4 Transcription, glossing and translation

Jahai is transcribed here in full accordance with the standards of the International Phonetic Alphabet (IPA). Thus, the orthographic system employed departs slightly from that of some other authors on Aslian and Mon-Khmer linguistics, including Benjamin, Diffloth and Kruspe. They write the voiced palatal stop as /j/ and the palatal approximant as /y/, whereas these are represented here as /ɟ/ and /j/ respectively. As pointed out by Benjamin (2001:120, note 10), this inconsistency of Aslian orthography may lead to some confusion. For the sake of international consistency, however, complete adherence to the IPA is preferred here.

Phonemic transcription is indicated by solidi / /, whereas phonetic transcription is indicated by square brackets [ ]. English translations of forms are indicated by apostrophes ' ', and wherever direct translation is not possible an explanatory description is given in brackets ( ). Examples of Jahai text are always given in their phonemic form, and words are consistently segmented and glossed morphemically. Following Kruspe (1999,

2004:xx), morpheme boundaries are represented by a hyphen (-) in the case of prefixes, arrows (< >) in the case of infixes and an equals sign (=) in the case of clitics. From considerations of space, glossed translations of individual words are sometimes simplified; for complete translations, see the glossary (Appendix I). Grammatical morphemes are glossed with the items given in the list of abbreviations. Translations of text examples are in colloquial English. As far as possible, the translation aims to give the reading intended in the original utterance. Parts of translated texts which do not have Jahai equivalents but which are necessary for comprehension in English are represented in square brackets [ ]. Unless stated otherwise, examples are taken from spontaneous text materials. Elicited examples are marked '(elic.)'.

2 *Phonology*

This chapter describes the phonological system of Jahai, including its phonemic inventory of vowels and consonants and the phonetic realisation of phonemes (§2.2 and §2.3), phonotactic properties (§2.4) as well as prosodic features (§2.5). In addition, a specific section is devoted to the phonological characteristics of Malay loanwords (§2.6). Some preliminary remarks are given in §2.1.

2.1 Preliminaries

This account of Jahai phonology is based on the forms listed in the glossary (Appendix I). By and large, these are citation forms of words, which generally represent synchronically minimal free forms, so-called *lexemes* (see §3.1). One significant consequence of this is that lexemes set the standard for phonotactic well-formedness. The many forms borrowed from Malay, which are included in the glossary, usually conform to the same pattern as indigenous forms and are therefore not treated separately. They occasionally differ in terms of phonotactic structure and have then not been incorporated into the general analysis.

The phonetic analysis is based mainly on auditory impression, although selective instrumental analyses have been carried out on some segments, notably final nasal consonants.[4]

The marginal phenomenon referred to here as *expressive elaboration* exhibits features which present particular difficulties with regard to phonology. These are introduced separately in §6.2 and are not incorporated into the general analysis made here.

Results are frequently compared with those of other phonological accounts of Jahai and Aslian as a whole, notably Schebesta (1928a:803–805); Bishop (1996a); Bauer (1991); Benjamin (1976b:130–153, 1985b, 1986); Diffloth (1976b:102–111); and Kruspe (2004: 32–60).

2.2 Vowels

Judging from the data available at present, the Jahai vowel system is distinguished by three degrees of vowel height for the front, central and back positions (see Table 2.1). This

[4] I am grateful to Dr Mechtild Tronnier, Department of Linguistics and Phonetics, Lund University, for conducting these preliminary acoustic analyses.

is in line with the system attributed to Jahai by Benjamin (1986:6). It is also in accordance with the systems described for some other Aslian languages, including Jah Hut (Diffloth 1976b:103), Temiar (Benjamin 1976b:131) and Trang Kensiw (Bauer 1991:316). However, it contrasts sharply with the complex pattern claimed by Bishop (1996a:228–232) for Yala Kensiw.[5]

The oral vowels contrast with a slightly smaller set of phonemically nasal counterparts, but these are not very frequent (present in about ten per cent of the lexical items collected) and do not carry a heavy functional load (see §2.2.1 for further discussion on contrastive vowel nasality). As in other Northern Aslian languages as well as Southern Aslian languages, vowel length has no phonemic significance. The back vowels display considerable rounding, especially /u/ and /o/. Front and central vowels are unrounded. As to the high central vowel, which is usually described as rounded and symbolised by /ʉ/ in other Aslian languages, minimal rounding is limited to certain environments. Thus, to mark the contrast between the rounded back and the non-rounded non-back vowels, the symbol /ɨ/ is preferred to /ʉ/.

Segments which were interpreted by Schebesta (1928a:803, 804) as diphthongs are always word-final and are more appropriately described as vowel + approximant (/w/ or /j/) in accordance with the requirement that word-final syllables are always closed (see §2.4.1). No phonemically significant diphthongs have been identified in the present material. Non-significant diphthongisation has been observed in final-syllable vowels before some palatal consonants: [sɛ^{i}c˺] /sɛc/ 'meat', [lɔ^{i}c˺] /lɔc/ 'bow', [raŋgo^{ij}ɲ] /raŋgoɲ/ 'mouth-harp'.

Table 2.1: Vowel phonemes in Jahai

ORAL			NASAL		
Front	Central	Back	Front	Central	Back
i	ɨ	u	ĩ	ɨ̃	ũ
e	ə	o		ə̃	
ɛ	a	ɔ	ɛ̃	ã	ɔ̃

5 Although the 3 x 3 vowel system may well be applied to the present material, it is not altogether unproblematic. For example, it is sometimes difficult to distinguish auditorily between high and mid vowels, particularly the front vowels /i/ and /e/, which could perhaps be interpreted as an indication of an intermediate, fourth vowel height. Similar hesitation is expressed by Bauer (1991:316) for Trang Kensiw. Furthermore, at least one speaker, originally from a different group of Jahai, made use of an additional unrounded close-mid central vowel /ə/, between /i/ and /ə/, which has so far only been attested in one word but which can be easily contrasted in the following minimal set for central vowels (cf. the full minimal set presented in §2.2):

/gis/	'to climb down'
/gəs/	'to apply make-up'
/gəs/	'to carve'
/gas/	'(a type of skin disease)'

In the idiolects of other speakers, however, this vowel appears indistinguishable from the high central /i/. It cannot be determined at this point whether /ə/ represents a marginal phoneme or if it reflects idiolectal or dialectal variation in the Jahai vowel system. Bauer (1991:316) and Matisoff (2003:16) suggest that any additions to the basic 3 x 3 vowel system in Northern Aslian languages could be explained as compensation for their lack of contrastive vowel length. For the time being, however, a 3 x 3 system will be posited for Jahai.

The following near-minimal set illustrates the full system of oral vowels:

/kis/	'to dig'	/gis/	'to climb down'	/gus/	'to be together'
/ges/	'smell'	/gəs/	'to carve'	/gos/	'belch'
/gɛs/	'to descend'	/gas/	'(skin disease)'	/gɔs/	'to live'

2.2.1 Phonemic vowel nasality

Although functionally marginal, lexically contrastive vowel nasality is very apparent in Jahai. Nasal counterparts of the oral vowels /i, ɛ, ɨ, ə, a, u, ɔ/ have been identified. These occur only in final syllables. Nasal counterparts of the mid vowels /e/ and /o/ have been found only in marginal instances of expressive elaboration and are therefore not considered part of the general phonemic inventory (see §6.2). The following lexical pairs illustrate the oral/nasal contrast:

/kis/	'to dig'	/kĩs/	'ghost'
/pɛk/	'to split'	/pɛ̃k/	'to sting'
/sit/	'honeycomb'	/sɨ̃t/	'to rub oneself'
/ttlət/	'to stare'	/klə̃t/	'to swallow'
/kawaw/	'bird'	/wãw/	'(a type of civet)'
/siruc/	'to slurp'	/grũc/	'slender-toed gecko'
/hakɔk/	'to throw'	/hokɔ̃k/	'to burn off fur'

2.2.2 Phonetic description of vowel phonemes

/i/ is a close front unrounded vowel [i]: [kitˀ] /kit/ 'buttocks', [sil] /sil/ '(a type of tortoise)'.

/ĩ/ is the rather common nasal counterpart of /i/: [hĩcˀ] /hĩc/ 'to rain', [kərĩl] /krĩl/ 'wrist'.

/e/ is a close-mid front unrounded vowel [e̞], slightly more raised than [e]: [he̞ʔ] /heʔ/ '1P.INCL', [te̞kˀ] /tek/ 'to sleep'. It has no nasal counterpart.

/ɛ/ is an open-mid front unrounded vowel [ɛ]: [ʔɛkˀ] /ʔɛk/ 'to give', [gɛs] /gɛs/ 'to descend'.

/ɛ̃/ is the rather common nasal counterpart of /ɛ/: [lapɛ̃kˀ] /lapɛ̃k/ 'mud', [pəcᶜɛ̃ʔ] /pcɛ̃ʔ/ 'to be wet'.

/ɨ/ is a close central unrounded vowel, which is slightly more backed than [ɨ]: [hɨpˀ] /hɨp/ 'forest', [sɨtˀ] /sɨt/ 'honeycomb'. Before glottal consonants it becomes a slightly lowered [ə]: [bəʔ] /bɨʔ/ 'mother'. Occasionally it is in free variation with a slightly rounded [ʉ]. Conditioned rounding is also evident following the bilabial approximant /w/: [wʉn] /wɨn/ 'to crawl'.

/ɨ̃/, the nasal counterpart of /ɨ/, is infrequent: [sɨ̃tˀ] /sɨ̃t/ 'to rub oneself', [tãʔɨ̃ʲcˀ] /taʔɨ̃c/ '(a type of large bird)'.

/ə/ is a mid central unrounded vowel [ə]: [gəs] /gəs/ 'to carve', [ʔəhəj] /ʔhəj/ 'to be small'.

/ə̃/ is the nasal counterpart of /ə/: [cᶜɔnhə̃tˀ] /cnhə̃t/ 'to be short', [səh] /səh/ 'to meet'.

/a/ is an open central unrounded vowel [a]: [ʔapˀ] /ʔap/ 'tiger', [cᶜaᵈn] /can/ 'foot'.

/ã/, the nasal counterpart of /a/, has been identified in only a few lexical items: [japãh] /japãh/ '(name of Jahai group)', [w̃ãw̃] /wãw/ '(a type of civet)'.

/u/ is a close back rounded vowel [u], frequently in free variation with a significantly advanced [ụ]: [duk˺] /duk/ 'to pounce upon', [tu^{g}ŋ] /tuŋ/ 'to fell'.

/ũ/ is the nasal counterpart of /u/: [c^{c}ũʔ] /cũʔ/ 'blind', [tũs] /tũs/ 'to grub'.

/o/ is a close-mid back rounded vowel [o]: [sop˺] /sop/ 'lung', [to^{b}m] /tom/ 'tree-trunk'. It has no nasal counterpart.

/ɔ/ is an open-mid back half-rounded vowel [ɔ]: [gɔs] /gɔs/ 'to live', [tɔ^{b}m] /tɔm/ 'water'. It is in free variation with an infrequent unrounded [ʌ].

/ɔ̃/ is the rather common nasal counterpart of /ɔ/: [ʔɔ̃t˺] /ʔɔ̃t/ 'dog', [lɔ̃r:] /lɔ̃r/ 'to hiss'.

2.2.3 Environmentally conditioned nasalisation of vowels

All nine phonemically oral vowels may be phonetically nasalised by surrounding nasal segments, mainly nasal consonants. Nasalisation is bidirectional (see §2.2.3.1) and often unpredictable (see §2.2.3.2). Under certain conditions, oral vowels may also be nasalised by phonemically nasal vowels in the same word. This phenomenon appears to be restricted to words where the final syllable onset is a glottal or an approximant, which suggests that only such consonants are transparent to nasalisation: [cũhɛ̃ʔ] /cuhɛ̃ʔ/ 'to flow', [hãʔɛ̃t˺] /haʔɛ̃t/ 'stench', [kãw̃ɔ̃t˺] /kawɔ̃t/ 'bird', [hãjɛ̃ʔ] /hajɛ̃ʔ/ 'house'.

2.2.3.1 Directionality of nasalisation

It can be suggested on the basis of the present data that nasalisation is bidirectional, i.e. both progressive and anticipatory, although it is not possible at this stage to determine whether one is stronger than the other and therefore 'primary' (see Blust 1997:150–151 for a discussion on 'primary' versus 'contragrade' nasalisation). What is clear, however, is that anticipatory nasalisation under certain circumstances is avoided in final syllables by means of prestopping of nasal consonant codas, which effectively seals off the preceding vowel nucleus from nasalising influences (see §2.3.1.2). Progressive nasalisation from the onset, on the other hand, is allowed — [mĩt˺] /mit/ 'eye', [ʔəŋũt˺] /ʔŋut/ 'throat', [mɔ̃h] /mɔh/ 'nose' — and may even be desirable, as suggested by Malay loanwords in which voiced stop onsets which follow homorganic nasals are assimilated and lost: [ləmũʔ] /lmuʔ/ from Malay *lembu* 'cattle', [rənãh] /rnah/ from Malay *rendah* 'low', [ɟ^{z}aŋũt˺] /ɟaŋut/ from Malay *janggut* 'beard' (see also §2.6.3.3).

In pre-final syllables, however, nasal codas are not prestopped and anticipatory nasalisation is allowed to operate freely, affecting not only the preceding vowel but sometimes also the consonant onset: [mũŋkɛr] /buŋkɛr/ 'to wake up', [mĩntɛ^{g}ŋ] /bintɛŋ/ 'star' (from Malay *bintang*). Anticipatory nasalisation of word-initial consonants becomes particularly apparent during infixation of /n/ (see §4.1.3 and §4.1.4.1). Word-initial /b/, /w/ and /l/ are then consistently realised as [m] and [n] respectively as the infixed /n/ forms the onset of the following new syllable: [mənawa^{i}c˺] /b<n>awac/ from /bawac/ 'pig-tailed macaque', [mənawət˺] /w<n>awət/ from /wawət/ 'rat', [nənataʔ] /l<n>ataʔ/ from /lataʔ/ 'waterfall'.

2.2.3.2 Variability in degree of nasalisation

There appear to be some differences in degree of nasalisation in the production of nasal vowels. In some lexical items nasal vowels are consistently distinctly pronounced and easily recognised, whereas those in other items are more subtle and characterised by much greater variation and unpredictability. In cases where nasal vowels occur in non-conditioned environments, and where phonemic vowel nasality thus can be posited, the former pattern is always the case. However, in conditioned environments the nasal element in vowels is commonly less salient and less predictable. Still, it cannot be made clear at this point whether this difference in degree of nasalisation really reflects a distinction between conditioned nasalisation and phonemic nasality, although potentially contrastive pairs like [ŋɔkˀ ~ ŋɔ̃kˀ] /ŋɔk/ 'to sit', in which the nasalisation of the vowel is not always apparent, and [bərŋɔ̃kˀ] /brŋɔ̃k/ '(a type of frog)', in which the vowel appears to be consistently and distinctly nasal, may point in that direction. In the phonetic examples given here, only the distinctly nasal vowels are consistently transcribed as such. Non-distinct examples are transcribed without the tilde, e.g. [mɔh] /mɔh/ 'nose', unless some point is made about their nasalisation, in which case the tilde is added for clarity: [mɔ̃h] /mɔh/ 'nose'.

2.3 Consonants

The consonant system of Jahai includes 20 phonemes (see Table 2.2).

Table 2.2: Consonant phonemes in Jahai

	Bilabial	Alveolar	Palatal	Velar	Glottal
Stop	p b	t d	c ɟ	k g	ʔ
Nasal	m	n	ɲ	ŋ	
Fricative	ɸ		s		h
Lateral		l			
Rhotic		r			
Approximant	w		j		

Features of particular interest include the common presence of 'prestopped' allophones of nasals in word-final position, the neutralisation of syllable-final stops, as well as the presence of the unusual voiceless bilabial fricative phoneme /ɸ/.

The two near-minimal sets shown in Table 2.3 contrast the consonant phonemes in initial and final position respectively.

Table 2.3: Contrastive sets for consonant phonemes in Jahai

		INITIAL CONTRAST		FINAL CONTRAST	
Stop	/p/	/pɔh/	'to fan'	/kɔp/	'to move hut'
	/b/	/bɔh/	'fruit'	—	
	/t/	/toʔ/	'ancestor'	/kɔt/	'to take'
	/d/	/dɔk/	'ipoh poison'	—	
	/c/	/cɔh/	'to bite'	/kɔc/	'to gnaw'
	/ɟ/	/ɟɔh/	'to ascend'	—	
	/k/	/kɔh/	'(a type of tortoise)'	/kɔk/	'to call'
	/g/	/goh/	'manchild'	—	
	/ʔ/	/ʔoh/	'to cough'	/makɔʔ/	'egg'
Fricative	/ɸ/	—		/kɔɸ/	'(sound of falling)'
	/s/	/soh/	'game animal'	/gɔs/	'to live'
	/h/	/hɔʔ/	'jar'	/kɔh/	'(a type of tortoise)'
Nasal	/m/	/mɔh/	'nose'	/kɔm/	'many'
	/n/	/nɔj/	'to copulate'	/ʔon/	'there'
	/ɲ/	/ɲɔp/	'to chew'	/goɲ/	'outgrowth on tree'
	/ŋ/	/ŋɔk/	'to sit'	/kɔŋ/	'to plait'
Lateral	/l/	/lɔʔ/	'(toponym)'	/gɔl/	'to carry'
Rhotic	/r/	/rɔh/	'to cut'	/tgɔr/	'to allow'
Approx.	/w/	/wɔh/	'to call'	/bakɔw/	'(a type of tree)'
	/j/	/jɔh/	'to drop'	/takɔj/	'sail-fin lizard'

2.3.1 Description of consonant phonemes and allophonic variation

2.3.1.1 Stops

Syllable-initially, voiceless /p, t, c, k/ and voiced /b, d, ɟ, g/ are usually realised as unaspirated bilabial, alveolar, palatal and velar stops. The palatals /c/ and /ɟ/ are affricated [c^{c}] and [ɟ^{ʑ}], although affrication of /c/ does not usually occur before epenthetic [i] in pre-final syllables: [cijas] /cjas/ 'hand'. Non-significant aspiration of word-initial /p/ has been noted occasionally. Word-initial /b/ is usually nasalised and realised as [m] if followed by a nasal segment: [mũŋkɛr] /buŋkɛr/ 'to wake up'. The voiced/voiceless contrast is illustrated by the following minimal pairs:

[pɔh]	/pɔh/	'to fan'	[bɔh]	/bɔh/	'fruit'
[tɛ̯ʔ]	/teʔ/	'soil'	[dɛ̯ʔ]	/deʔ/	'to make'
[c^{c}ɔh]	/cɔh/	'to bite'	[ɟ^{ʑ}ɔh]	/ɟɔh/	'to ascend'
[kɛ^{d}n]	/kɛn/	'child'	[gɛ^{d}n]	/gɛn/	'(a type of tree)'

In syllable-final position, the bilabial, alveolar, palatal and velar stops are unreleased and the voiced/voiceless distinction is lost. These unreleased stops are here transcribed phonetically as [p̚, t̚, c̚, k̚] and symbolised phonemically by the voiceless /p, t, c, k/:

[c^{c}ɨp̚]	/cɨp/	'to go'
[dut̚]	/dut/	'navel'
[sɛ^{i}c̚]	/sɛc/	'meat'
[sɔk̚]	/sɔk/	'hair'

However, speakers sometimes produce a delayed and heavily aspirated voiceless oral release of final stops after words uttered in isolation, particularly following /c/ and /k/ but occasionally also /t/: [sɛ^{i}c^{cʰ}] /sɛc/ 'meat', [wẹ̱k^{ʰ}] /wẹ̱k/ 'to return'. There is no regular pattern to suggest that this voiceless release has any contrastive significance.

The glottal stop /ʔ/ is realised as [ʔ]: [ʔap̚] /ʔap/ 'tiger', [ʄ^{ᶻ}əʔẹ̱^{ᵍ}ŋ] /ʄʔeŋ/ 'bone', [haliʔ] /haliʔ/ 'leaf', [laʔpaʔ] /lʔpaʔ/ '(eastern Negritos)'. In connected speech, the glottal stop is frequently lost syllable-finally.

2.3.1.2 Nasals

The nasal phonemes /m, n, ɲ, ŋ/ display the same places of articulation as the voiced stops and occur in both syllable-initial and syllable-final position. In all positions except some word-final environments (see below) they are realised as [m, n, ɲ, ŋ]:

[mɔh]	/mɔh/	'nose'
[nɔ̃j]	/nɔj/	'to copulate'
[ɲɔp̚]	/ɲɔp/	'to chew'
[ŋɔk̚]	/ŋɔk/	'to sit'

In word-final position, nasals are commonly initiated by a short and very characteristic voiced stop–like onset:

[ləkɛ^{ᵇ}m]	/lkɛm/	'brain'
[pətɔ^{ᵇ}m]	/ptom/	'day before yesterday'
[wɯ^{ᵈ}n]	/wɨn/	'to crawl'
[c^{ᶜ}a^{ᵈ}n]	/can/	'fʊot'
[tẹ̱^{ʲʲ}ɲ]	/teɲ/	'to plait'
[piŋlɔ^{ʲʲ}ɲ]	/pŋlɔɲ/	'to sing'
[tabə^{ᵍ}ŋ]	/tabəŋ/	'leaf monkey'
[gahu^{ᵍ}ŋ]	/gahuŋ/	'cave'

Such nasal segments have been recorded in several Aslian languages, as well as in many Austronesian languages, notably in Borneo (Blust 1997:154–169). They are varyingly described as predenasalised (Diffloth 1975:7, 10–12; Benjamin 1985b:14–16), preocclusivised (Matisoff 2003:19, fn 89), preploded (Blust 1997:154–155) and prestopped (Ladefoged & Maddieson 1996:128–129; Bishop 1996a:228, 235-236; Kruspe 2004:34–35). The term 'prestopped' will be employed here. The segments were noted in Jahai by Schebesta (1928a:805) but interpreted by him as word-final stops followed by 'an obscure nasal release', which rather makes one think of the postnasalised stops described phrase-finally by Benjamin (1976b:134) for Temiar. Also, Bishop (1996a:235) analyses similar nasals in Kensiw as the syllable-final allophones of voiced stops.

In the present Jahai material, however, there are several reasons for analysing the prestopped nasals as allophones of the simple nasals. First, the nasal portion of the segments in question is perceptually clearly more salient than the stop-like portion. The latter, which is conventionally transcribed as a homorganic stop, [^{ᵇ}m, ^{ᵈ}n, ^{ʄ}ɲ, ^{ᵍ}ŋ] or [bm, dn, ʄɲ, gŋ], is caused by a delayed and abrupt lowering of the velum simultaneously with, or following, the oral closure. Preliminary spectrogram studies of Jahai samples do not give a very uniform picture of the stop-like portion but they do indicate that there is not always an

apparent occlusive or plosive element involved in its production, homorganic or otherwise. However, the conventional way of transcribing prestopped nasals, i.e. a nasal preceded by its homorganic stop [^{b}m, ^{d}n, ^{ɟ}ɲ, ^{g}ŋ], has been retained in the present work.

Second, as also acknowledged by Bishop (1996a:235) for Kensiw, the prestopped nasals are historically developments from simple nasals and have simple nasal reflexes in other Mon-Khmer languages; compare e.g. [c^{c}a^{d}n] /can/ 'foot' with Vietnamese *chan* 'foot' and [te̯^{ij}ɲ] /teɲ/ 'to plait' with Kammu /táaɲ/ 'to weave'.

Third, there is evidence to suggest that prestopped nasals are cognitively variants of simple nasals. Notably, reduplications of prestopped nasals are always realised as the simple nasal counterpart:

[səmsɔ^{b}m]	/smsɔm/	'to buzz around a nest'
[hənhə^{d}n]	/hnhən/	'to devour'
[jiɲjɛ^{ij}ɲ]	/jɲjɛɲ/	'to dream'
[ɟ^{ʒ}əŋɟ^{ʒ}ɛ^{g}ŋ]	/ɟŋɟɛŋ/	'wide'

Furthermore, Malay loanwords which originally have final nasals are usually realised with the prestopped counterpart (cf. §2.6.1.1):

[haja^{b}m]	from *ayam*	'poultry'
[bula^{d}n]	from *bulan*	'moon'
[kuc^{c}e̯^{g}ŋ]	from *kucing*	'cat'

Fourth, there is a clear tendency for the prestopped nasals and the word-final simple nasals (the latter making up about 22 per cent of the word-final nasals) to turn up in different phonetic environments, and the nature of the environmentally conditioned nasalisation of vowels discussed in §2.2.3 may be helpful in stating these conditions. Drawing on data from several Austronesian languages in Borneo, Blust (1997:161–163) suggests that the stop portion of prestopped nasals is intended to prevent so-called contragrade, anticipatory (or coda-driven) nasalisation from spreading from word-final nasals to the preceding vowel in languages whose primary nasalisation is progressive (or onset-driven). It is further predicted that syllable-final nasals are not prestopped in final syllables whose vowel nucleus is preceded by a nasal segment, since the syllable nucleus has already been nasalised in the 'right' direction and there is therefore no need to seal off the vowel from contragrade nasalisation by means of prestopping.

Although it has not been possible to determine which direction of nasalisation is primary in Jahai (see §2.2.3.1), the model proposed by Blust appears to be largely applicable to the present material. First, spectrogram studies of prestopped nasals in Jahai confirm Blust's suggestion that prestopping prevents anticipatory nasalisation. Second, the nuclei of those final syllables that are closed by a simple nasal are indeed commonly preceded and nasalised by a nasal segment:

[kənɔ̃m]	/knɔm/	'to urinate'
[gənũn]	/gnun/	'bamboo'
[səmɛ̃ɲ]	/smɛɲ/	'to request'
[ʔamɛ̃ŋ]	/ʔamɛŋ/	'siamang'

Accordingly, the nucleus of final syllables ending with a prestopped nasal is almost always preceded by a non-nasal segment. There is a number of exceptions to this latter pattern, some of which can be explained as Malay loans in which clusters of homorganic nasal and voiced stop are reduced to the nasal: [kami^{g}ŋ] /kamiŋ/ from Malay *kambing*

'goat', [kəma^{g}ŋ] /kmaŋ/ from Malay *kembang* 'to swell', [həŋa^{g}ŋ] /hŋaŋ/ from Malay *enggang* 'hornbill' and [paŋa^{g}ŋ] /paŋaŋ/ from Malay *panggang* 'to roast' (see also §2.6.3.3). Exceptions not yet accounted for include [tama^{b}m] /tamam/ '(a type of small animal)', [luŋa^{d}n] /luŋan/ 'binturong', [həmɨ^{ɨʲ}ɲ] /hmiɲ/ '(magic word)', [bərama^{g}ŋ] /braman/ '(a type of tree)', [ʔəna^{g}ŋ] /ʔnaŋ/ 'side', [gəno^{g}ŋ] /gnoŋ/ 'wooden material' and [lamɔ^{g}ŋ] /lamɔŋ/ '(a type of fruit)'.

This explanation of the distribution of prestopped nasals should also predict that the prestopping of a syllable-final nasal is suspended if the syllable nucleus consists of a vowel with phonemic nasality (Blust 1997:172). Since the syllable nucleus in that case is already nasal, there is no need to block coda-driven nasalisation from the final nasal segment. The problem here is that there is at present no way of positively determining whether the vowel of the final syllable in words like [c^{c}əkĩm] 'peacock-pheasant', [hãw̃ẽn] 'wild boar', [hẽɲ] 'mouth' and [katɔ̃ŋ] 'claw' is phonemically nasal or if it is nasalised by the following nasal coda which for some unknown reason is not prestopped. It will be assumed here, however, that most of the words displaying such a pattern do involve a phonemically nasal vowel. More revealing is the fact that there is no example of phonemic nasality in vowel nuclei of syllables ending in a prestopped nasal. According to this analysis, the following phonetic pairs then reflect the contrast between oral and nasal vowels in words with final nasal consonants:

[c^{c}əkɨ^{b}m]	'lump'	[c^{c}əkĩm]	'peacock-pheasant'
[jɨ^{d}n]	'to hop on something'	[jĭn]	'to stop'
[go^{ɨʲ}ɲ]	'outgrowth on tree'	[gɔ̃ɲ]	'to refuse to give'
[bəli^{g}ŋ]	'upper arm'	[bərlĩŋ]	'long-tailed parakeet'

Thus, although the distribution of prestopped nasals is not entirely clear-cut, several factors indicate that an analysis of them as allophones of simple nasals is preferable to one in which they are treated as separate phonemes or as word-final allophones of voiced or voiceless stops. According to the present analysis, voiced and voiceless stops merge to become unreleased stops in syllable-final position (see §2.3.1.1) and therefore contrast with the prestopped nasals, as illustrated by the following minimal pairs:

[kɔp˺]	/kɔp/	'to move hut'	[kɔ^{b}m]	/kɔm/	'many'
[dut˺]	/dut/	'navel'	[du^{d}n]	/dun/	'to cover'
[sɛ^{i}c˺]	/sɛc/	'meat'	[sɛ^{iʲ}ɲ]	/sɛɲ/	'front'
[pɔk˺]	/pɔk/	'round object'	[pɔ^{g}ŋ]	/pɔŋ/	'to tap poison'

2.3.1.3 Fricatives

The voiceless palatal fricative /s/ is commonly an alveolo-palatal [ɕ], midway between [s] and [ç], in both syllable-initial and syllable-final position. It is in free variation with a less frequent alveolar [s] but, as described by Bishop (1996a:234) for Kensiw, individual speakers tend to use one variant consistently. For easier transcription and legibility, [s] ~ /s/ will be used here, although the characteristics of this phoneme are usually palatal. (See Kruspe (2004:35) for a discussion on the similarly problematic /s/ in Semelai). Examples of /s/ include e.g. [sɔ^{b}m] /sɔm/ 'bird's nest', [kəɲsi^{g}ŋ] /kɲsiŋ/ 'banded palm civet' and [ʔɔs] /ʔɔs/ 'fire'.

The voiceless glottal fricative /h/ occurs frequently in syllable-initial and, especially, word-final position: [hɔk˺] /hɔk/ 'to throw', [sih] /sih/ 'to pound', [paʔah] /pʔah/ 'to

kneel'. It is not always strictly glottal, as the point of friction is dependent on the surrounding vowels. However, it is consistently transcribed here as [h].

The infrequent voiceless bilabial fricative /ɸ/ has been identified in some 30 lexical items and its distribution is restricted to syllable-final position. Word-finally it is realised as a compressed and barely audible [ɸ]. In pre-final syllables it is usually realised as an unreleased bilabial stop [pˀ]. It is likely to have iconic properties, as its expiratory and fricative character appears to be largely associated with words denoting phenomena that set air in motion. Examples include [cᶜɛ̃ɸ] /cɛ̃ɸ/ 'to fan fire', [cᶜənɛɸ] /cnɛɸ/ 'tail feathers', [ɲɨɸ] /ɲɨɸ/ 'to blow smoke', [kətɔɸ] /ktɔɸ/ 'to splutter', [bəsɔɸ] /bsɔɸ/ 'to be quick', [lɔɸ] /lɔɸ/ 'to leak', and [pəlɔ̃ɸ] /plɔ̃ɸ/ '(sound of blowpipe)'. Syllable-final /ɸ/ has been identified in other Northern Aslian languages as well, including Kensiw (Bishop 1996a:23) and Batek (Benjamin 1985b:10).

2.3.1.4 Liquids

The voiced apico-alveolar lateral liquid /l/ is found in syllable-initial and syllable-final position and has approximately the same phonetic shape in all positions: [lɔ̰j] /lɔj/ 'to run', [halɐ̣h] /haleh/ 'hungry', [ʔɛl] /ʔɛl/ 'to see'.

The voiced alveolar rhotic /r/ is subject to considerable individual and free variation. Syllable-initially, it is realised phonetically as an alveolar trill [r], flap [ɾ] or approximant [ɹ]: [rɛs ~ ɾɛs ~ ɹɛs] /rɛs/ 'to fall'. After [n], it is usually realised as an approximant [ɹ] and is preceded by a short epenthetic stop transition: [sɜnᵈɹɛkˀ] /snrɛk/ 'to go out', [pənᵈɹaʔ] /pnraʔ/ '(a type of flower)', [mənᵈɹaʔ] /mnraʔ/ 'human being', [cᶜɜnᵈɹɔs] /cnrɔs/ 'finger', [manᵈɹɔᵍŋ] /manrɔŋ/ 'skink' (see Benjamin (1985b:12) for some discussion on this feature in Aslian and Malay). Syllable-finally, it is realised either as an approximant [ɹ] or as a very distinct trill [r:]: [dɔr:] /dɔr/ 'to spread something', [sajər:] /sajər/ 'herd'.

2.3.1.5 Approximants

The voiced bilabial approximant /w/ is found in both initial and final position: [wawətˀ] /wawət/ 'rat', [bəgiw] /bgiw/ 'wind', [litɔw] /litɔw/ 'boy'.

Similarly, the voiced palatal approximant /j/ is found in initial and final position. Word-finally (at least in isolation), it is subject to partial, barely audible nasality: [j̃]. This 'semi-nasal' phrase-final allophone does not appear to be a source of nasalisation as it has no effect on the preceding vowel. Examples: [jɔkˀ] /jɔk/ 'to undress', [jɨ̰j] /jɨj/ 'to carry in one's hand', [gɐ̰j] /gej/ 'to eat', [kuj̃] /kuj/ 'head'.

2.4 Phonotactics

2.4.1 Syllable structure and types

2.4.1.1 Syllable structure

The maximal syllable in Jahai consists of a simple onset, nucleus and coda: $[CVC]_{\sigma}$. Onsets are obligatory but codas are not, and the minimal syllable is therefore $[CV]_{\sigma}$. Syllable structure is represented here as a flat consonant-vowel tier (see e.g. Clements and Keyser 1983):

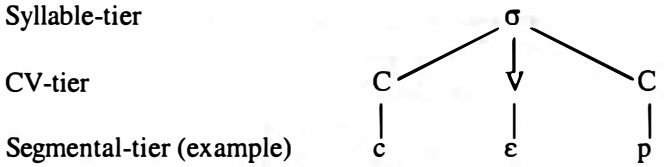


Figure 2.1: Maximal syllable structure in Jahai

This type of model is preferred to a hierarchical branching structure involving onset and rhyme, since constraints on the combination of nucleus and coda are marginal and not more significant than constraints on the combination of onset and nucleus (see §2.4.3.1). Furthermore, it can be motivated on morphological grounds, since reduplicative processes regularly involve the coda but never the nucleus and coda as a unit. For a different approach, see Kruspe's (2004:39–40) treatment of syllable structure in Semelai.

2.4.1.2 Syllable types

Syllables may be broadly divided into light syllables and heavy syllables, in consequence of the absence or presence of a coda. In their surface forms, light syllables are open, $[CV]_{\sigma}$, and heavy syllables are closed $[CVC]_{\sigma}$. In the underlying form of both light and heavy syllables, however, the vowel nucleus may be either non-predictable and prespecified (/CV/, /CVC/), or predictable and underspecified and thus omitted in phonemic transcriptions (/C/, /CC/). The following table separates the four underlying syllable types:

Table 2.4: Underlying syllable types in Jahai

	/C/	/CV/	/CC/	/CVC/
Heavy	–	–	+	+
Prespecified nucleus	–	+	–	+

A further distinction needs to be made between full syllables and half syllables. Half syllables are those which are neither heavy nor have a prespecified vowel nucleus: /C/. The other three syllable types are regarded here as full syllables: /CV/, /CC/, /CVC/. This distinction between half and full syllables is necessary in order to separate disyllabic words from sesquisyllabic words (see §2.4.2.2, §2.4.2.3). Syllable types may then be summarised as in Table 2.5.

Table 2.5: Syllable types in Jahai

	LIGHT		HEAVY	
	Half	Full		
Underlying form	C	CV	CC	CVC
Surface form	CV	CV	CVC	CVC

2.4.2 Word structure

Jahai lexemes may be monosyllabic, sesquisyllabic, disyllabic or trisyllabic. Tetrasyllabic lexemes appear to be consistently avoided (although the marginal instances of expressive elaboration deviate in this respect as well; see §6.2). Word-final syllables (including monosyllabic words) are invariably maximal, and, since all syllables require an onset, words therefore always begin and end with a consonant.

The final syllable, which is always a heavy syllable with a prespecified nucleus, /CVC/, may be regarded as the main (or *major*) syllable of a word. This always bears stress and displays the greatest phonemic variation. It thus forms a natural starting point for a leftward word analysis.

2.4.2.1 Monosyllabic words

Monosyllabic words always consist of a heavy syllable with the canonic structure /CVC/, e.g. [cᶜɛpˀ] /cɛp/ 'to catch'.

2.4.2.2 Sesquisyllabic words

Sesquisyllabic words consist of a final /CVC/ syllable preceded by a penultimate half syllable: /C.CVC/, e.g. [kənɛʲcˀ] /knɛc/ 'comb'. They are phonetically disyllabic, but the vowel of the penultimate syllable is epenthetic and predictable and therefore omitted in phonemic transcriptions.[6] However, note that a syllable boundary still separates the two consonants in the phonemic transcription. Phonetic variation in the epenthetic vowels is discussed in §2.4.3.2.1.

2.4.2.3 Disyllabic words

Disyllabic words consist of a final /CVC/ syllable preceded by a penultimate full syllable. This may be either open and then has a non-predictable vowel nucleus, usually /a/: /CV.CVC/, e.g. [kawipˀ] /kawip/ 'sun bear'; or it may be closed and then have either a predictable or a non-predictable vowel nucleus: /CC.CVC/, e.g. [təmkal] /tmkal/ 'male'; or /CVC.CVC/, e.g. [kaltoᵍŋ] /kaltoŋ/ 'knee'.

A distinction between sesquisyllabic and disyllabic words is motivated here on morphological grounds, as the two types of word structure are sometimes subject to different morphological processes. This is illustrated in §3.2. It is such morphological behaviour that motivates the treatment of words of the form /CC.CVC/ (conventionally treated as sesquisyllabic) as disyllabic in the present study. The following lexical pairs illustrate the contrast between sesquisyllabic words and disyllabic words with an open penultimate syllable:

---

6 The term 'sesquisyllable' was coined by Matisoff (1973:86) and literally means 'one-and-a-half syllable'. Most writers on Aslian (including Diffloth 1976a; Matisoff 2003; and Kruspe 1999, 2004) prefer to exclude the epenthetic vowel from phonemic transcriptions, and this is also the orthography employed here. For reasons of clarity and legibility, Benjamin (1976b:152–153) argues in favour of including it.

SESQUISYLLABIC			DISYLLABIC		
[ɟ^{ʔ}əlɔʔ]	/ɟlɔʔ/	'hole'	[ɟ^{ʔ}alɔʔ]	/ɟalɔʔ/	'casting-net'
[təwi^{j}ɲ]	/twiɲ/	'headache'	[tawi^{j}ɲ]	/tawiɲ/	'(a type of spider)'
[təgɨ^{ij}ɲ]	/tgɨɲ/	'to tear apart'	[tagɨ^{ij}ɲ]	/tagɨɲ/	'firewood'
[bəli^{g}ŋ]	/bliŋ/	'upper arm'	[bali^{g}ŋ]	/baliŋ/	'tiger'
[sijər:]	/sjər/	'to swim'	[sajər:]	/sajər/	'herd'
[ləwɛ̃j]	/lwɛj/	'honey'	[lãw̃ɛ̃j]	/lawɛ̃j/	'leech'

2.4.2.4 Trisyllabic words

Indigenous trisyllabic words consist of a final /CVC/ syllable preceded by a full penultimate syllable, which is either open or closed and has a non-predictable vowel nucleus. This, in turn, is preceded by an antepenultimate half syllable. These trisyllables are then in effect two-and-a-half-syllabic: /C.CV.CVC/, e.g. [kəlaɲis] /klaɲis/ 'heart'; /C.CVC.CVC/, e.g. [c^{c}əmalpɔk$^{˥}$] /cmalpɔk/ '(a type of millipede)'. Full penultimate syllables with a predictable vowel nucleus also exist but only in derivatives (fossilised or productive) involving coda copy (see §2.4.4, §3.2), e.g. [pərəŋgə^{g}ŋ] /prŋgəŋ/ 'pharynx'. This means that indigenous trisyllabic forms never have a half penultimate syllable. One item with a seemingly non-predictable antepenultimate nucleus has been found, [kuruhũj] /kuruhuj/ '(a type of owl)', although this may be onomatopoeic and therefore uncertain.

As described for other Northern Aslian languages (see Asmah (1976:952) for Kintaq and Bishop (1996a:240) for Kensiw), trisyllabic roots are commonly Malay borrowings.

Table 2.6: Word structure in Jahai

Word type	Canonic structure	Example	Translation
Monosyllabic	/CVC/	/cɛp/	'to catch'
Sesquisyllabic	/C.CVC/	/knɛc/	'comb'
Disyllabic	/CV.CVC/	/kawip/	'sun bear'
	/CC.CVC/	/tmkal/	'male'
	/CVC.CVC/	/kaltoŋ/	'knee'
Trisyllabic	/C.CV.CVC/	/klaɲis/	'heart'
	/C.CC.CVC/	/prŋgəŋ/	'pharynx'
	/C.CVC.CVC/	/cmalpɔk/	'(a type of millipede)'

2.4.3 Distribution of phonemes

There are clear differences in terms of phonemic variation between final syllables and pre-final syllables, the former displaying a greater variety than the latter. This pattern applies both to vowels and (to a lesser extent) to consonants. For example, the nucleus of final syllables may consist of any of the 16 vowel phonemes, whereas pre-final nuclei with occasional exceptions are restricted to oral /i/, /a/ and /u/. There is also a difference between onset and coda consonants in that onsets can generally be drawn from a greater variety of phonemes than codas. A summary of phoneme distribution is given in Tables 2.7 and 2.8, and each syllable type is described in the following sections.

Table 2.7: Vowel distribution in Jahai

		Oral									Nasal						
	Ø	i	e	ɛ	ɨ	ə	a	u	o	ɔ	ĩ	ɛ̃	ɨ̃	ə̃	ã	ũ	ɔ̃
Final	–	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Penultimate	+	+	–	–	–	–	+	+	+	+	–	–	–	–	–	–	–
Antepenultimate	+	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Table 2.8: Consonant distribution in Jahai

		p	t	c	k	ʔ	b	d	ɟ	g	ɸ	s	h	m	n	ɲ	ŋ	l	r	w	j
Coda	#CV<u>C</u># .CV<u>C</u>#	+	+	+	+	+	–	–	–	–	+	+	+	+	+	+	+	+	+	+	+
	.CV<u>C</u>. #C(V)<u>C</u>.	–	–	–	–	–	–	–	–	–	–	–	–	+	+	+	+	+	+	–	–
Onset	#<u>C</u>VC# .<u>C</u>VC# #<u>C</u>V.	+	+	+	+	+	+	+	+	+	–	+	+	+	+	+	+	+	+	+	+
	#<u>C</u>. #<u>C</u>(V)C.	+	+	+	+	+	+	+	+	+	–	+	+	?	–	–	–	+	+	–	–
	.<u>C</u>V(C).	–	–	–	–	–	–	–	–	–	–	–	–	+	+	–	–	+	+	+	–

2.4.3.1 Final syllables

Final syllables (including monosyllabic words) invariably display the canonic structure /CVC/. All consonant phonemes in Jahai can occur as the coda of such syllables except the voiced stops (see §2.3.1.1). Under certain conditions, nasals are prestopped in word-final position (see §2.3.1.2). Every consonant phoneme may occur as syllable onset, with the exception of the voiceless bilabial fricative /ɸ/. However, onsets consisting of a voiced stop seldom occur after a nasal penultimate coda (see also §2.2.3.1 and §2.3.1.2). There are also some restrictions as to the possibility of having a phonemically identical onset and coda. Thus, the palatal stops /c, ɟ/, fricative /s/ and nasal /ɲ/ almost never occupy both onset and coda position, the one exception being /cɔc/ 'scar'. A similar restriction applies to the alveolar rhotic /r/.

All 16 vowel phonemes may occur as syllable nucleus in final syllables, the oral vowels /e, ɛ, ɨ, ə/ and all nasal vowels occurring *only* in final syllables (again, the marginal instances of expressive elaboration display some exceptions; see §6.2). There are no clear constraints on the oral vowels as to their co-occurrence with initial and final consonants, although the close central and back vowels /ɨ/ and /u/ have not been found in combination with a bilabial approximant coda /-w/. Nasal vowels, on the other hand, do not usually occur with initial voiced stops /b-, d-, ɟ-, g-/. There are occasional exceptions to this pattern, e.g. /bə̃k/ 'to untie', /dadə̃l/ '(a type of reed snake)', /piɟə̃ʔ/ 'to fly up', /gɛ̃c/ 'to scratch'.[7]

---

7 Similar restrictions have been noted in other Aslian languages, e.g. Jah Hut (Diffloth 1976b:103) and Semelai (Kruspe 2004:53).

2.4.3.2 Pre-final syllables

2.4.3.2.1 The penultimate syllable in sesquisyllabic words

In sesquisyllabic words, the penultimate syllable is a half syllable /C/, consisting phonetically of a consonant onset and a predictable vowel nucleus, usually [ə]. A limited set of consonant phonemes can occur as onset: the voiceless bilabial fricative /ɸ/, the approximants /w/ and /j/, and all nasals are absent.[8] There also appear to be other restrictions in sesquisyllabic words as to which penultimate onset may be combined with the onset of the final syllable. For example, the penultimate onset is seldom identical to the onset of the final syllable, and, moreover, homorganic voiced and voiceless stops are almost never found in combination. Similarly, stops are not usually combined with their homorganic nasals. Similar restrictions have been noted in Jah Hut by Diffloth (1976b: 104–105).

As noted, the predictable, epenthetic vowel nucleus is usually realised as [ə], but it shows clear tendencies to change in response to certain phonetic environments. Thus, it is usually realised as [i] if followed by an onset consisting of the palatal approximant /j/:

[tijɔ̃k˺]	/tjɔ̃k/	'to point'
[cijas]	/cjas/	'hand'
[sijɨ^{g}ŋ]	/sjɨŋ/	'to burn'
[sijul]	/sjul/	'(a type of cobra)'

If the following onset is a glottal stop /ʔ/ or fricative /h/, the epenthetic vowel sometimes takes on the same phonetic quality as the vowel nucleus of the final syllable: [kẹʔẹp˺] /kʔep/ 'centipede', [taʔaʔ] /tʔaʔ/ 'vegetable'. This pattern is much less apparent and [ə] is equally common in this environment: [cᵉəʔəʔ] /cʔiʔ/ 'to pour'. Furthermore, if a penultimate stop onset is combined with a liquid onset in the final syllable, especially /r/, [ə] is commonly dropped altogether in fast speech: [bəraʔ ~ braʔ] /braʔ/ 'NEG', [gərɛ^{g}ŋ ~ grɛ^{g}ŋ] /grɛŋ/ '(a type of monitor lizard), [pəlɔ̃ɸ ~ plɔ̃ɸ] /plɔ̃ɸ/ '(sound of blowpipe)'. Realisation is variable, even in the same word, and cannot be stated as rules but should rather be described as tendencies.

2.4.3.2.2 The penultimate syllable in disyllabic words

In disyllabic roots with an open penultimate syllable, onset position can probably be filled by any consonant phoneme except the voiceless bilabial fricative /ɸ/, which is never an onset. However, no unequivocal examples of the syllable-initially rare nasals /ɲ/ and /ŋ/ have been found. There appear to be no clear restrictions as to which penultimate and final onset phonemes may be combined. There are several examples of words in which the two syllables have identical onsets, although it cannot be determined at this point whether these are instances of frozen morphology involving copy of the final onset:

/wawət/	'rat'
/ɲaɲuʔ/	'to die'
/mamuɲ/	'to beg'
/dadɔ̃l/	'(a type of reed snake)'
/cicar/	'(a type of tree)'

---

8 Onsets in the form of [m] occur only when the final syllable onset is a nasal, and they are interpreted here as a nasalised allophone of /b/, e.g. [mə̃nɛ̃ʔ] /bnɛʔ/ 'size' (cf. §2.3.1.1).

The non-predictable vowel nucleus of these open penultimate syllables usually consists of /a/, but there are also several examples of /i/, and /u/, whereas /o/ and /ɔ/ occur only sporadically. The remaining oral vowels, as well as the phonemically nasal vowels, are not found in this position.

/ʔalic/	'to pass'
/timɔʔ/	'hard surface'
/cuhɛ̃ʔ/	'to flow'
/pokəh/	'(a type of gecko)'
/hɔkɔ̃k/	'(sound of a leaf-monkey)'

In disyllabic roots with a closed penultimate syllable there appear to be some restrictions as to which consonant phonemes may fill the penultimate onset position: voiced stops are notably rare, and approximants and nasals are usually absent (with the possible exception of /m/, although this almost invariably turns up before a nasal segment and may then be interpreted as a nasalised allophone of /b/; cf. §2.3.1.1). The coda position is always occupied by a nasal or a liquid. Nasal codas are homorganic with the onset of the final syllable, unless the final onset is a glottal, in which case they are represented by /n/:

/ʔmpɔŋ/	'hole'
/ʔnteŋ/	'ear'
/kaɲcɔʔ/	'grandchild'
/saŋkoh/	'wreathed hornbill'
/cnhɔ̃t/	'to be short'
/knʔac/	'father-in-law'

There are a few exceptions to this pattern, e.g. /tmkɛt/ 'to be cold' and /tmkal/ 'male'. However, in derivatives involving coda copy, any consonant phoneme allowed as a final syllable coda may be reduplicated and fill the coda position of the penultimate syllable (see §2.4.4).

The vowel nucleus of closed penultimate syllables is usually a predictable epenthetic [ə]: [hərkiṭ] /hrkiṭ/ 'night'. If the surrounding onset and coda are both alveolar or palatal, the epenthetic vowel frequently becomes a slightly lowered and fronted [ɜ]:

[tɜrhiʲcˇ]	/trhic/	'(a type of small bird)'
[sɜnlɔʲcˇ]	/snlɔc/	'blowpipe dart'
[cᶜɜnᵈɹɔs]	/cnrɔs/	'nail'
[sɜltuh]	/sltuh/	'to attack'
[ʄᶻɜlmɔl]	/ʄlmɔl/	'mountain'
[ʄᶻɜrwɛj̆]	/ʄrwɛj/	'(a type of tree)'

There are also numerous examples of non-predictable /a/ in this position, e.g. /karwɔʔ/ '(a type of terrapin)' and /rampɔw/ 'long-tailed macaque'. The close vowels /i/ and /u/ are much less frequent and mainly confined to Malay loans: /lintɛs/ 'to go across', /cunʈiŋ/ 'temple'. One example of /ɔ/ has been identified, /hɔrjĩŋ/ '(a type of small animal)', but it is possible that this is in fact a compound of two free forms: /hɔr jĩŋ/. The remaining oral vowels, as well as the phonemically nasal vowels, are not found in this position.

2.4.3.2.3 Pre-final syllables in trisyllabic words

Typical examples of indigenous trisyllabic forms include the following:

/hlaŋkɛt/	'(a type of ant)'
/cmalpɔk/	'(a type of millipede)'
/crikɔ̃k/	'to jabber'
/klaŋis/	'heart'
/brubɔh/	'yellow-vented bulbul'
/hmiɾiɾ̃n/	'to extinguish by itself'
/ɟritɛw/	'to squat'

It is possible that many trisyllabic forms contain archaic morphemes which cannot be analysed synchronically. Thus, such forms are regarded here as monomorphemic.

The penultimate syllable of trisyllabic words is always full. The onset position is always filled by a sonorant: /m, n, l, r, w/; the coda position of closed penultimate syllables is always filled by a liquid or, usually, a nasal, in which case it is homorganic with the onset consonant of the final syllable. In derivatives involving coda copy, however, any consonant phoneme allowed as a final syllable coda may be reduplicated and fill the coda position of the penultimate syllable (see §2.4.4). The vowel nucleus of penultimate syllables in monomorphemic trisyllabic words is always a non-predictable /i/, /a/ or /u/; in derivatives involving coda copy it may be epenthetic and predictable.

Antepenultimate syllables are open, and their onset is typically a stop (voiced or voiceless) or a fricative (/s/ or /h/), although sporadic examples of the alveolar rhotic /r/ have been found. The vowel nucleus is always predictable, with the possible exception of /kuruhuj/ 'a type of owl'.

Occasional Malay loans fail to conform to the patterns outlined above. Thus, /dahaga?/ 'to be thirsty' (from Malay *dahaga*) and /puŋhuluh/ 'headman' (from Malay *penghulu*) have a non-sonorant penultimate onset, and the nucleus of the antepenultimate syllable may be non-predictable.

2.4.4 Reduplicative processes

Morphological processes commonly involve two types of partial reduplication which will be referred to here as *coda copy* and *onset copy*.[9] Coda copy consists of the copying of a word-final syllable coda and its infixation before the final syllable onset, thereby creating a heavy penultimate syllable. It entails that, in derivatives involving this process, any phoneme allowed as a coda of final syllables may fill the coda position of the penultimate syllable. Onset copy is restricted to monosyllabic bases and involves the copying of the onset and its prefixation to the base, commonly in combination with coda copy (see §3.2). Although several such morphological processes are productive, there are numerous examples in the present word list which give the impression of being frozen forms, the roots of which are no longer independent morphemes. However, the realisation rules outlined below apply to both productive and non-productive cases of reduplication. In the examples given, synchronically non-existent roots are marked with (*).

2.4.4.1 Phonetic realisation of coda copies

In most cases, the phonetic realisation of the copy is identical to that of the copied final syllable coda. This applies to stops /-p, -t, -c, -k, -ʔ/, the fricatives /-s, -h/, liquids /-l, -r/ and the bilabial approximant /-w/:

---

9 This is the terminology introduced by Kruspe (1999, 2004), see §3.2.

[cᶜəp˺rɛp˺]	/cprɛp/	'babbler'	*crɛp	
[kiswɛ̃s]	/k<s>wɛ̃s/	'to be sweeping'	/kwɛ̃s/	'to sweep'
[pəlpɛl]	/pl-pɛl/	'to be dripping'	/pɛl/	'to drip'
[ʔəwŋiw]	/ʔ<w>ŋiw/	'to be gazing'	/ʔŋiw/	'to gaze'

Nasals, however, which are commonly prestopped in word-final position (see §2.3.1.2), are always realised with the simple nasal counterpart; the voiceless bilabial fricative /ɸ/ is usually realised as an unreleased bilabial stop [p˺]; and the palatal approximant /j/, which is partly nasalised word-finally, is realised as a fully oral [j]:

[ləp˺lĩɸ]	/lɸ-lĩɸ/	'fontanelle'	/lĩɸ/	'fontanelle'
[həriɲpe̞ʲʲɲ]	/hrɲpeɲ/	'goose-bumps'	*hpeɲ/hrpeɲ	
[pəŋsɛᵍŋ]	/pŋsɛŋ/	'to say'	*psɛŋ	
[həlijdɨ̃j]	/hljdɨj/	'flat ground'	*hdɨj	

2.4.4.2 Phonetic realisation of epenthetic vowels

Epenthetic vowel nuclei of penultimate syllables closed by a coda copy are subject to consistent phonetic variation conditioned by the copy. Thus, if the copy is a palatal /c, s, ɲ, j/, the preceding vowel is realised as [i]:

[kəlic˺baⁱc˺]	/klcbac/	'(a type of millipede)'	*kbac	
[kiswɛ̃s]	/k<s>wɛ̃s/	'to be sweeping'	/kwɛ̃s/	'to sweep'
[piɲlɔʲʲɲ]	/pɲlɔɲ/	'to sing'	*plɔɲ	
[hijhə̃j]	/hjhəj/	'to be yawning'	/həj/	'to yawn'

In cases where the copy is a glottal /ʔ, h/, the preceding vowel is realised as [a]:

[saʔsoʔ]	/sʔsoʔ/	'blood vessel'	/soʔ/	'cubit'[10]
[baʔbɔʔ]	/bʔbɔʔ/	'to carry on one's back'	*bɔʔ	
[tahtẹh]	/thteh/	'oriental pied hornbill'	*teh	
[nahʄᶻɔh]	/nh-ʄɔh/	'height'	/ʄɔh/	'to ascend'

With other copies, the preceding vowel is realised as [ə]:

[bət˺ʔɛt˺]	/btʔɛt/	'good'	*bʔɛt	
[dək˺duk˺]	/dkduk/	'chest'	/duk/	'to pounce'
[səmsɔᵇm]	/smsɔm/	'to buzz around a nest'	/sɔm/	'bird's nest'
[səŋpɔᵍŋ]	/sŋpɔŋ/	'leafbird'	*spɔŋ	
[dəldɨ̃l]	/dldɨl/	'heel'	*dɨl	
[pərbɛr]	/prbɛr/	'lower arm'	*pbɛr	
[həwhĩw]	/hwhĩw/	'crested wood-partridge'	*hĩw	

Thus, the following rules may be set up for the realisation of epenthetic vowels preceding coda copies:

(1) $\emptyset > [i] / \_\ C_{[palatal]} +$

(2) $\emptyset > [a] / \_\ C_{[glottal]} +$

(3) $\emptyset > [ə] / \_\ C_{[-palatal, -glottal]} +$

---

10 The semantic connection between this root and the derivation is not clear.

Similar realisation rules have been described for Semelai (Kruspe 2004:46–47) and most dialects of Semai (Diffloth 1976a:237).

2.4.5 Syllabification

In surface representations, the syllabification procedure is straightforward and governed by the syllable structure as outlined in §2.4.1. Moving from right to left, since the end of the word forms the most stable and significant part, nuclei are located and linked to a syllable node to which surrounding consonants are then assigned according to the maximal syllable $[CVC]_{\sigma}$, beginning with the obligatory onset and, if possible, concluding with a coda. This is in accordance with the Onset First Principle and its associated algorithm of syllabification proposed by Clements and Keyser (1983:37–38). Such a procedure produces correctly syllabified forms like the ones shown in Figure 2.2.

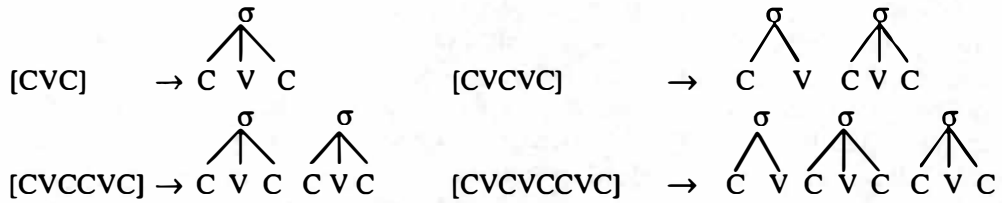


Figure 2.2: Syllabification of surface forms in Jahai

Underlying representations frequently contain sequences of unsyllabified consonants which have the potential to be syllabified as either onsets or codas. In such cases, rules of vowel epenthesis are necessary to meet the requirements of syllable structure, and the task is to predict where such epenthesis will take place and to identify syllable boundaries. As with surface representations, predictions are uncomplicated as syllabification proceeds from right to left according to the general principle of maximality. Thus, in strings of unsyllabified consonants, the syllabification procedure strives to create maximal $[CVC]_{\sigma}$ syllables, which have precedence over minimal $[CV]_{\sigma}$ syllables. Two adjacent unsyllabified consonants will therefore always be syllabified as onset and coda of a maximal syllable, and a single unsyllabified consonant will always be syllabified as onset of a minimal syllable. Epenthetic vowels can then be inserted as nuclei. In the underlying forms /CCCVC/ and /CCCCVC/, where two and three consonants respectively remain unsyllabified following syllabification of the syllable with an underlying vowel nucleus, this principle of right-to-left maximality produces the correctly syllabified forms /CC.CVC/ > [CVC.CVC] and /C.CC.CVC/ > [CV.CVC.CVC]. It excludes the unattested forms */C.C.CVC/, */CC.C.CVC/ and */C.C.C.CVC/. Recall the requirement that penultimate syllables of trisyllabic words are always full (see §2.4.3.2.3). Recall also that tetrasyllabic words are not allowed in Jahai (see §2.4.2).

In the underlying forms /CCVC/, /CCVCVC/ and /CCVCCVC/, the only consonant left unsyllabified following syllabification of the syllables with underlying vowel nuclei is the leftmost one. This is consequently syllabified as onset of a minimal syllable, producing the correctly syllabified forms /C.CVC/ > [CV.CVC], /C.CV.CVC/ > [CV.CV.CVC] and /C.CVC.CVC/ > [CV.CVC.CVC].

An analysis based on the distributional characteristics of segments, phonotactically as well morphologically determined (see §2.4.3 and §2.4.4), would yield identical results but is superfluous given the sufficient tools provided by word structure constraints.

2.5 Prosodic features

2.5.1 Stress

As also noted by Schebesta (1928a:805), Jahai has non-contrastive stress that falls invariably on the last syllable of a word, and there is no secondary stress. This applies as much to recent loans from Malay as to indigenous words.

2.5.2 Tone

Schebesta (1928a:805) believed he had identified tonal differences in a limited set of lexical items in Jahai and provided a short list of minimal pairs. Similarly, Bishop (1996a:238–239) suggests that pitch difference (mid versus high level) is lexically contrastive in a small number of words in the closely related Kensiw language. These descriptions lead Hajek (2003) to conclude that Northern Aslian languages exhibit tonal activity. However, no such distinctions can be posited for the present Jahai material. Significant pitch contrasts have not been noted, and the contrastive pairs listed by Schebesta have in general been found to contain either contrasting vowel quality or contrasting final consonants rather than tone in the variety of Jahai studied here. The following tables compare the contrastive pairs given by Schebesta (in his original orthography) with the contemporary Jahai equivalents:

Schebesta (1928a:805):

HIGH LEVEL		LOW LEVEL	
¯ĭo?	'part of a blowpipe'	_ĭo?	'bough'
¯nus	'sleeping mat'	_nus	'lip'
¯ĭɛ?	'1S'	_ĭɛ?	'1D INCL'
¯hɛ?	'1P INCL'	_he:i	'1D INCL'
¯ɛg?	'to give'	_ɛg?	'belly'
¯gus	'to rub fat into the face'	_gus	'to come down'

Contemporary Jahai:

/jo?/	'outer shaft of blowpipe'	/joh/	'crown of tree'
/nɨs/	'sleeping mat'	/nus/	'upper lip'
/jɛ?/	'1S'	/jɛh/	'1D EXCL'
/hɛ?/	'1P INCL'	/hɛj/	'1D INCL'
/?ɛk/	'to give'	/?ɛc/	'belly', 'excrement'
/gɨs~gəs/	'to apply make-up'	/gɨs/	'to climb down'

2.6 The phonology of loanwords

It has been customary in the field of Aslian linguistic research to comment on the phonological nature of Malay loanwords. For instance, see Benjamin (1976b: 147–152) for

Temiar; Diffloth (1976b:112) for Jah Hut; Bauer (1991:313) for Trang Kensiw; Bishop (1996a:234–235) for Yala Kensiw; and Kruspe (2004:55–57) for Semelai. A short summary is also presented by Matisoff (2003:52–53). Focal points of interest include the presence of intervocalic /h/, the closure of open syllables by means of glottal /h/ or /ʔ/, the realisation of the rhotic /r/, and the deletion of voiced stops in intervocalic clusters where such stops are preceded by their homorganic nasal.

Some features are considered to be retentions of historic or dialectal/colloquial Malay forms not found in present-day Standard Malay; others are viewed as Aslian innovations. An interesting example of the latter is given by Benjamin (1976b:147–152), who describes a pattern of consonantal substitution in Malay words borrowed into Temiar, which he suggests is the result of a deliberate special phonological treatment of loans. For Malay loan rates in all Aslian languages, see Benjamin (1976a:73; in press).

About one-fifth of the collected Jahai lexical items can be shown to be Malay loanwords, although such words are consistently integrated into the indigenous phonological system. The Malay loans probably represent several different strata and have been borrowed at different times in history. Many of them denote foreign phenomena which have been introduced into the Jahai community from outside, such as /kritəh/ from *kereta* 'car', /kamiŋ/ from *kambing* 'goat', and /ɟuwal/ from *jual* 'to sell'. Others represent more basic vocabulary, e.g. /batuʔ/ from *batu* 'stone', /gaɟah/ from *gajah* 'elephant', and /kuceŋ/ from *kucing* 'cat'.

A number of words can be tentatively identified as loans from Temiar, a Central Aslian neighbour, e.g. /baboʔ/ 'woman' and /ɟlmɔl/ 'mountain'. There is also a category of words which may have been borrowed historically from a non-Malay Austronesian source. Examples of such words include /kbis/ 'to die' and /bgiw/ 'wind'.[11]

Furthermore, a rather large number of English words has entered the Jahai vocabulary. Some of these have been borrowed via Malay, but many may have been borrowed directly from British troops stationed in the area during the communist insurrection 1948–60. Such words typically include military terminology such as /grinɛt/ 'grenade', /bubiʔ trɛp/ 'booby-trap', /pos/ 'post', and /trinɛŋ/ 'training'.

As indicated earlier, words borrowed from Malay and other sources appear to become fully adapted to the indigenous phonological system. There is no sign, for instance, of the type of special treatment of loans described by Benjamin (1976b:147–152) for Temiar, or of any retention of Malay features foreign to Jahai. Indeed, Jahai phonological features even pervade the Malay spoken by many Jahai as a second language. However, as pointed out by Kruspe (2004:55) for Semelai, it is sometimes difficult to establish whether Malay loanwords that do not conform to the phonology of Standard Malay involve indigenous adjustment or colloquial/archaic features already present in the Malay dialect from which they were borrowed. No systematic comparison has been made with the present-day Malay dialect spoken in Hulu Perak (predominantly Patani, see e.g. Zaharani 1991:5–7) and therefore no conclusions will be drawn about the origin of these unclear deviating patterns. Such features will be dealt with here on the basis of their divergence from Standard Malay. It is clear, however, that several patterns are indeed the result of indigenous innovation.

The phonological changes discussed have been organised below into four categories: *phonetic adaptation*, *phonemic replacement*, *reorganisation of syllabic structure* and *relocation of stress*.

---

11 I am grateful to Geoffrey Benjamin for bringing these examples to my attention.

2.6.1 Phonetic adaptation

The phonetic realisation of segments is usually very similar to that of Malay, but there are some clear examples of phonetic adaptation to indigenous realisation rules.

2.6.1.1 Prestopping of word-final nasals

Recall that word-final nasal consonants preceded by an oral vowel are prestopped, whereas those preceded by a vowel nasalised by a nasal syllable onset are not (see §2.3.1.2). This pattern is consistently applied to Malay loanwords, as illustrated by the following examples:

[tanɛ̃m]	/tanɛm/	from *tanam*	'to plant'
[tomɛ̃n]	/tomɛn/	from *toman*	'snakehead'
[kuñiŋ]	/kuniŋ/	from *kuning*	'yellow'
[ɟᶻaᵇm]	/ɟam/	from *jam*	'hour'
[bulaᵈn]	/bulan/	from *bulan*	'moon'
[gadiᵍŋ]	/gadiŋ/	from *gading*	'elephant's tusk'

2.6.1.2 Nasalisation of word-initial /b/

According to a rather regular Jahai pattern, the voiced bilabial stop /b/ becomes nasalised [m] word-initially if the following consonant is nasal: [mĩntɛᵍŋ] /bintɛŋ/ from Malay *bintang* 'star', [mĩnataᵍŋ] /binataŋ/ from Malay *binatang* 'animal'. This process occasionally results in ambiguous forms, as in the case of [mɜ̃nɛ̃ŋ], which is phonemically either /mnɛŋ/ 'to win' (from Malay *menang*) or /bnɛŋ/ 'thread' (from Malay *benang*).

2.6.1.3 Palatalisation of /s/

The voiceless alveolar fricative of Malay is usually realised as the alveolo-palatal counterpart typical of Jahai.

2.6.2 Phonemic replacement

Many loanwords contain examples of often systematic replacement of certain phonemes. Only the most conspicuous processes will be described here.

2.6.2.1 Substitution of final syllable /a/

Final syllable /a/ in Standard Malay is in Jahai often replaced by other vowel phonemes. These varied substitutions possibly reflect different stages and sources of borrowing. The most common substitute is /ɛ/:

/ʔiŋɛt/	from *ingat*	'to remember'
/buŋɛʔ/	from *bunga*	'flower'
/tanɛm/	from *tanam*	'to plant'
/pgɛŋ/	from *pegang*	'to hold'
/bilɛŋ/	from *bilang*	'to count'
/taŋɛl/	from *tanggal*	'to fall'
/lantɛj/	from *lantai*	'floor'

Sometimes /a/ is replaced by /ə/:

/tmpət/	from *tempat*	'place'
/guləh/	from *gula*	'sugar'
/katəm/	from *katam*	'to cut'
/hantər/	from *hantar*	'to send'

In other cases it is replaced by /ɔ/. This substitution is consistent before word-final /w/ as part of a reinterpretation of the word-final vowel sequence *au* in Malay. It is not phonotactically determined, however, as the word-final sequence [-aw] is allowed in Jahai, e.g. in /bukaw/ 'flower' and /blaw/ 'blowpipe'.

/kapɔʔ/	from *kapak*	'axe'
/ɟalɔʔ/	from *jala*	'casting-net'
/krbɔw/	from *kerbau*	'buffalo'
/hiɟɔw/	from *hijau*	'to be green'
/pulɔw/	from *pulau*	'island'
/gurɔw/	from *gurau*	'to jest'

However, there are also numerous examples of Malay loanwords in which final syllable /a/ has been retained:

/tiŋkap/	from *tingkap*	'window'
/kilat/	from *kilat*	'lightning'
/badaʔ/	from *badak*	'Sumatran rhinoceros'
/bras/	from *beras*	'husked rice'
/gaɟah/	from *gajah*	'elephant'
/papan/	from *papan*	'plank'

2.6.2.2 Word-final glottal stop

In Malay, the glottal stop [ʔ] is an allophone of the voiceless velar stop /k/ in word-final position. In Jahai, however, the glottal stop has full phonemic status and contrasts with /k/ word-finally: /tek/ 'to sleep', /teʔ/ 'soil'. The final /k/ in Malay is therefore consistently represented as phonemic /ʔ/ in loanwords:

/ʔiteʔ/	from *itik*	'duck'
/taseʔ/	from *tasik*	'lake'
/badaʔ/	from *badak*	'Sumatran rhinoceros'
/sɟuʔ/	from *sejuk*	'to be cold'
/lobɔʔ/	from *lobok*	'pool'
/kapɔʔ/	from *kapak*	'axe'

2.6.2.3 Voiceless labiodental fricative

Standard Malay has a rather rare voiceless labiodental fricative phoneme /f/ of Arabic origin. In Jahai, this is consistently represented by the voiceless bilabial stop /p/ in all positions. Most probably, this realisation has been taken over from the local Malay dialect.

/maʔap/	from *maaf*	'pardon'
/paham/	from *faham*	'to understand'
/pikir/	from *fikir*	'to think'

2.6.3 Reorganisation of syllabic structure

Malay forms that do not conform to the syllabic structure of Jahai are modified in different ways so that they fit the indigenous system. Such modifications include the closure of certain syllables, the reinterpretation of vowel sequences, the deletion of voiced stops in some positions, and the reinterpretation of some disyllabic words as sesquisyllables.

2.6.3.1 Closure of syllables

Recall that syllables in Jahai require a consonant onset, and word-final syllables are always closed by a consonant coda, implying that words always begin and end with a consonant. Malay words beginning with a vowel receive an initial consonant onset in the form of glottal /h/ or /ʔ/:

/hobiʔ/	from *ubi*	'root-crop'
/hajam/	from *ayam*	'poultry'
/hŋaŋ/	from *enggang*	'rhinoceros hornbill'
/ʔaŋket/	from *angkit*	'to take'
/ʔasiŋ/	from *asing*	'to be different'

Similarly, words ending with a vowel receive a glottal final coda in Jahai, usually /ʔ/ and occasionally /h/:

/nasiʔ/	from *nasi*	'cooked rice'
/limɛʔ/	from *lima*	'five'
/lataʔ/	from *lata*	'waterfall'
/kritəh/	from *kereta*	'car'
/puŋhuluh/	from *penghulu*	'headman'

Both of these processes of syllable closure may have been present in the variety of Malay from which the words were borrowed.

The final consonant requirement is also enforced on Malay words displaying the word-final vowel sequences *-au* and *-ai*, in which the offglides are reinterpreted as approximant consonants /w/ and /j/ respectively:

/halɔw/	from *halau*	'to scare'
/pulɔw/	from *pulau*	'island'
/lantɛj/	from *lantai*	'floor'

2.6.3.2 Reinterpretation of word-medial vowel sequences

Intervocalic /h/ in Malay loanwords has been described for other Aslian languages (see e.g. Kruspe 2004:56; Bauer 1991:313; Diffloth 1976b:112; Matisoff 2003:52–53) and is considered to reflect borrowing and retention of dialectal/archaic Malay forms not found in present-day Standard Malay (compare dialectal/archaic *tihang* and standard *tiang* 'house pole'). This retention is in keeping with the syllable structure of Jahai in that sequences of

vowels are not allowed and that the retained /h/ forms the obligatory consonantal onset of the final syllable. It is surely also motivated by stress, which in Jahai is always on the final syllable. However, Jahai has only sporadic examples of this phenomenon, e.g. /tuha?/ 'to ripen' (cf. *tua*), /prahu?/ 'boat' (cf. *prau*). Instead, vowel sequences of Malay words are much more likely to be broken up by the approximant consonants /w/ and /j/, as in the following examples:

/lawot/	from *laut*	'sea'
/duwa?/	from *dua*	'two'
/bawuŋ/	from *baung*	'a type of catfish'
/lajin/	from *lain*	'to be different'
/cjuŋ/	from *tiung*	'mynah'
/kajil/	from *kail*	'to fish'

Again, the inserted consonant occupies the final syllable onset position and thereby splits up the two vowels of the original vowel sequence between different syllables, in order to bring the word into conformity with Jahai word structure and stress patterns. This process appears to be productive and is likely to be a case of indigenous adaptation rather than borrowing from local Malay dialects. There is also one isolated example in which Malay intervocalic /h/ is replaced by /w/: /tawon/ from *tahun* 'year'.

2.6.3.3 Deletion of voiced stops

In Jahai, clusters of nasal consonant and its homorganic voiced stop are very rare (see §2.2.3.1, §2.3.1.2). When Malay words containing such clusters are borrowed into Jahai, the voiced stop is invariably assimilated and lost. The nasal, originally in coda position, is thereby transformed into onset of the following syllable. The same pattern of assimilation occurs in Semelai (Kruspe 2004:56).

/lmu?/	from *lembu*	'cattle'
/tmakɔw/	from *tembakau*	'tobacco'
/mah/	from *rendah*	'to be low'
/baniŋ/	from *Banding*	'(toponym)'
/ɟaŋut/	from *janggut*	'beard'
/hŋaŋ/	from *enggang*	'rhinoceros hombill'

2.6.3.4 Reinterpretation of disyllabic words as sesquisyllables

Pre-final *e* in Malay words (phonetically [ə]) is typically interpreted as a strictly epenthetic schwa and the word thus assumes a sesquisyllabic structure:

/pti?/	from *peti*	'box'
/bli?/	from *beli*	'to buy'
/pniŋ/	from *pening*	'headache'

2.6.4 Relocation of stress

Malay loanwords receive the final syllable stress characteristic of Jahai, rather than the penultimate syllable stress typical of Standard Malay: /gaˈɟah/ from *'gajah* 'elephant', /biˈlɛŋ/ from *'bilang* 'to count'.

2.7 Summary

This chapter has provided an introduction to the contemporary Jahai sound system. One obvious conclusion to be drawn is that Jahai conforms to the general phonological patterns described for most other languages of the Aslian branch of Mon-Khmer. Thus, Jahai appears to display a typical 3 x 3 vowel system, phonemically significant vowel nasality, peculiar realisations of word-final nasal consonants, closed final syllables, as well as polysyllabic lexemes. It is also characterised by features which are usually associated with Northern Aslian, such as the lack of contrastive vowel length and the presence of the unusual syllable-final voiceless bilabial fricative /ɸ/. Claims as to the presence of contrastive tone in Northern Aslian languages are not supported by the present work.

As to the phonological treatment of Malay loanwords, Jahai behaves in many ways like other Aslian languages. Patterns of phonetic adaptation, phonemic replacement, reorganisation of syllabic structure and relocation of stress bear witness to extensive adaptation of foreign elements to the indigenous phonological system.

3 *Word formation*

This chapter describes the system of word formation in Jahai, defining its morphological units (§3.1) and examining the morphological processes of affixation (§3.2) and cliticisation (§3.3). It is systematised strictly according to process rather than morphemic function and is mostly organised along a spectrum of what will be referred to below as 'base-dependence'. For descriptions of individual morphemes, their allomorphs and function, see Chapter 4. Much of the analysis rests on knowledge of the phonotactic characteristics of the language, described in §2.4. Also, several aspects of the analysis are significantly inspired by Kruspe's (1999:92–159, 2004:61–93) treatment of similar morphological processes in Semelai.

3.1 Morphological units

The following sections define the structural units of Jahai word formation that are relevant to the subsequent description, including roots (§3.1.1), lexemes (§3.1.2) and bases (§3.1.3), which are typically free forms, as well as affixes (§3.1.4) and clitics (§3.1.5), which are always bound morphemes. Terminology, definitions and notation have been adapted from Kruspe (1999:92–95, 2004:61–63).

3.1.1 Roots

Roots may be defined for Jahai as morphologically simplex words — that is, words that are both synchronically and diachronically monomorphemic and therefore unanalysable. Most roots are free morphemes and therefore also represent lexemes (see §3.1.2). However, it is also very common for roots to occur only in units which bear the hallmarks of morphological complexity but which are synchronically unanalysable (cf. §2.4.4). Such postulated, obsolete Jahai roots are marked here with an asterisk (*). Some examples follow (see also §4.1 and §4.7).

*teh	/thteh/	'oriental pied hornbill'
*twaŋ	/tŋwaŋ/	'blue coral snake'
*bɔʔ	/bʔbɔʔ/	'to carry on one's back'
*plɔɲ	/pɲlɔɲ/	'to sing'

3.1.2 Lexemes

Lexemes are synchronically minimal free forms and may occur independently. Thus, they differ from roots in that they also include morphologically complex forms that are synchronically unanalysable. Lexemes are usually represented in the citation forms of Jahai words and are therefore those forms which feature in the glossary (Appendix I). Lexemes also represent those forms on which the analysis of phonotactic restrictions is based (§2.4); lexemes thus set the standard for phonotactic well-formedness.

3.1.3 Bases

Bases are units in the form of roots, lexemes or any morphologically complex form to which a bound morpheme (i.e. affix or clitic) can be added. In the subsequent analysis of processes of affixation and cliticisation, this umbrella term is the preferred one for such units.

3.1.4 Affixes

Affixes are phonologically bound morphemes whose domain of attachment is words. They are either prefixes or infixes, never suffixes. Morpheme boundaries are marked here with a hyphen (-) in the case of prefixes and with arrows (<>) in the case of infixes. Affixes represent an inner layer of derivational morphology and are always attached to their bases prior to clitics.

3.1.5 Clitics

Clitics are phonologically bound morphemes whose domain of attachment is a phrase, clause or some other unit of words. The base to which they attach functions as a phonological host within that domain. They are invariably proclitics. Morpheme boundaries are marked here with an equals sign (=). Clitics represent an outer layer of syntactically determined post-derivational morphology and are always attached to the periphery of the base.

3.2 Affixation

When analysing processes of affixation in Jahai, three determining factors need to be addressed, including (1) the domain of attachment, (2) the structure of the base to which an affix is attached, and (3) the degree of phonologically prespecified material in the affix. These factors interplay to form a spectrum of what may be called 'base-dependence', along which the present analysis of bound morphemes is organised. Allomorphic variation is largely insensitive to base-dependence; that is, different allomorphs of the same morpheme may differ in terms of their dependence. Thus, the description is organised strictly according to process rather than morpheme. For characterisations of morphemes and their allomorphy, see Chapter 4.

Affixation involves the creation or reorganisation of a syllable. Usually a new syllable is created, which means that monosyllabic forms become disyllabic (or, unusually, sesquisyllabic; see §3.2.2.2), whereas sesqui- and disyllabic forms become trisyllabic. In

some cases, however, affixation instead involves a change in syllable type, whereby, for example, half syllables become full, and light syllables become heavy (cf. §2.4.1.2).

As to domain of attachment, affixes may be broadly divided into two types. One type, which is predominant, takes the penultimate syllable as its domain and thus occurs as prefixes or infixes which in whole or in part make up the syllable to the immediate left of the stressed final syllable of the base. The form of such affixes is typically determined by the structure of the base to which the affix is attached, and the resulting words meet the phonotactic constraints outlined in §2.4. Such processes will be referred to here as *inner affixation* (§3.2.1). The other type of affixes attaches concatenatively to the left edge of a base, thus occurring only as prefixes, and is not sensitive to the structure of the base. The resulting forms frequently violate word-structure constraints. Such processes are referred to here as *outer affixation* (§3.2.2).[12] The origins of inner and outer affixation are discussed in §3.2.3. A third type of affix, involving total reduplication of the base and displaying features of both inner and outer processes, is treated separately (§3.2.4).

Inner affixation may be further subcategorised according to the degree of phonologically prespecified material present in the affix. Many affixes are fully prespecified with regard to their segmental content, whereas others are only partly prespecified or wholly underspecified and borrow their segmental material from the base to which they are attached. Hence, in the following description, inner affixation is subdivided into three types: underspecified, partly prespecified, and prespecified affixation.

Returning to the issue of base-dependence, it is evident that the distinction between inner and outer processes of affixation reflects a fundamental difference in the degree of attachment to the base, where outer affixation represents a more independent process which shares some features with cliticisation (§3.3). In a similar way, the differences with regard to phonological prespecification in inner affixation reflect varying degrees of phonological dependence on the base, where some fully prespecified forms share features with the more independent outer affixes. The following sections are organised according to this cline of base-dependence, beginning with highly dependent processes of inner and underspecified affixation, and ending with the least dependent process of outer affixation. This, in turn, provides a natural transition to the description of cliticisation in §3.3.

3.2.1 Inner affixation

As noted, inner affixation takes the penultimate syllable as its domain. Depending on the structure of the base, it therefore involves either prefixation or infixation of segmental material which in whole or in part makes up the syllable to the immediate left of the stressed final CVC syllable of the base. This means that affixation sometimes also entails syllabic reorganisation of existing base segments.

Inner affixation produces morphologically complex forms that meet the phonotactic constraints outlined in §2.4. Thus, a fundamental feature of inner affixation is the creation of well-formed penultimate syllables, the result being that the penultimate syllable of an 'innerly' affixed form is full. In the case of sesqui- and disyllabic bases this feature is

---

12 A similar distinction of affixational processes is made for Semelai by Kruspe (2004:64–66), who employs the more conventional terms *non-concatenative* and *concatenative* to refer to inner and outer affixation respectively. The terms *inner/outer affixation* are preferred here since they are devoid of the bias which might be associated with the term *concatenative* and its negative derivation *non-concatenative*. This is in accordance with the primary, predominant and presumably autochthonous nature of inner affixation in Jahai.

perhaps only to be expected, since word structure constraints do not permit half penultimate syllables in trisyllabic forms (cf. §2.4.2.4, §2.4.3.2.3). In the case of monosyllabic bases, however, the motivation behind the phonotactic well-formedness of affixed forms must be a different one, since a half syllable added to a monosyllabic base produces a sesquisyllabic form, which is not in violation of word-structure constraints. Instead, the reason is possibly to be found in the restrictions on the distribution of phonemes that affect the penultimate onset of sesquisyllabic forms. Recall that several phonemes, notably nasals, are not allowed in this position (see §2.4.3.2.1 and Table 2.8). Given the fact that numerous affixes consist of sonorant, notably nasal, phonemes (cf. §3.2.1.2, §3.2.1.3), this restriction would seriously hamper the ability of affixes to attach to monosyllabic bases as half syllables. Furthermore, the restriction against combinations of identical final and penultimate onsets in sesquisyllabic forms (see §2.4.3.2.1) would, for example, prevent processes whereby the onset of the monosyllabic base was copied and attached to the base as a half syllable (cf. §3.2.1.2). In both cases, then, constraints on the distribution of phonemes make half syllable affixes inexpedient. Thus, the fact that penultimate syllables of affixed forms must be full may be fully attributed to phonotactic constraints.

Given the significance of syllabic structure in inner affixation, as well as its frequent use of processes of reduplication and infixation, it is convenient to describe such affixation in the light of non-linear theories of morphology, such as Prosodic Morphology and Template Morphology. The descriptive methods used here have been worked out on the basis of works like McCarthy (1982), Marantz (1982), Broselow and McCarthy (1983), and Ter Mors (1983), as well as the summary given by Spencer (1991:133–172). This approach is also significantly inspired by Kruspe's (1999:97–137, 2004:69–82) utilisation of such models in her exhaustive account of non-concatenative morphological processes in Semelai.

The fundamental idea of Prosodic and Template Morphology is that morphemes are represented on different levels or tiers. One tier, called the CV tier, consists of a prosodic template representing the canonical pattern of the morpheme without any indication of the precise identity of phonemic segments. These segments are instead represented on a second tier, the so-called phonemic melody. Melody elements are linked to the CV tier by means of association lines, and every slot in the CV tier must be associated with a melody element, and vice versa. Affixation involves the addition of material to both CV tier and phonemic melody, and every morpheme of a word is assigned to a separate tier according to the so-called Morphemic Tier Hypothesis. Such an approach to affixation has the advantage of being able to account for non-linear types of affixation, such as affixation in discontinuous morphemes or processes of reduplication. It has been of particular interest in discussions of combinations of such processes, so-called internal reduplication (Broselow & McCarthy 1983; Ter Mors 1983).

Crucial to the argument are the notions of prespecification and underspecification. In the case of 'ordinary' affixation, the added material is fully prespecified in terms of both canonical form and segmental content. In the case of reduplication, however, the reduplicated affix is specified only for its canonical form whereas its segmental slots are empty and have to be filled with segmental material copied from the base.

This section has indicated that there is a fundamental association in Jahai between inner affixation and phonotactic well-formedness. It could be argued that the location and form of inner affixes are wholly determined by syllable structure and phonotactic constraints, and, in effect, that inner affixation is *designed* to create phonotactically acceptable forms.

Its processes of copying could then be largely explained as a means of satisfying phonotactic constraints in cases where the prespecified phonemic content of an affix is insufficient for the creation of well-formed words. An analysis could thus set out from the assumption that affixes are assigned according to fixed, syllabified CV templates in which fullness of the penultimate syllable is the all-important feature. However, such an approach would result in unnecessarily circumstantial procedural accounts. For the sake of descriptive economy and elegance, therefore, the following analysis of inner affixation involves the assignment of affixes to CV templates representing unsyllabified underlying forms. Rules of syllabification and vowel epenthesis subsequently produce correct surface forms. This approach appears to presuppose that the phonotactic well-formedness of affixed forms is simply a consequence of syllabification following affixation and not that it determines the shape of affixes already from the beginning. It should be pointed out that no such standpoint is taken and that it will not be concluded here which order is the most plausible one from a psychological point of view.

As will be illustrated, the affixed melody elements associated with the CV slots of the template may be underspecified (§3.2.1.1), partly prespecified (§3.2.1.2) or fully prespecified (§3.2.1.3). Underspecified slots are filled with segmental material copied from the base, namely the consonants of the CVC string that corresponds to the stressed final syllable. Affixes involving such copied segments are always attached to the left edge of this final CVC string. Following the generalisations of Ter Mors (1983:284–288), the final CVC string, not the whole base, is regarded in the following analysis as the unit which is copied and from which consonant segments are associated with the template of the affix. This descriptive approach is adopted in order to emphasise the significance of the final CVC string as the only unit from which material may be copied, and as the only unit to which affixes involving copied material may be attached. However, an approach involving the copying of the whole base would produce identical results.

In the examples given, prespecified segmental information is represented in a melody tier *below* the CV tier, whereas reduplicated material from which segments are associated to fill underspecified slots is represented in a melody tier *above* the CV tier. Reduplicated material which is left unassociated after the template has been satisfied is subject to stray erasure.

3.2.1.1 Underspecified affixation

Jahai exhibits several affixes which are not fully prespecified phonologically. In such cases, phonemic content of the affix is copied from the base, or, more specifically, the final CVC string of the base (corresponding to the stressed final syllable), in order to make the affix phonologically complete. Phonological prespecification is commonly partial, in which case the affix is only partly made up of copied material. Such processes are described in §3.2.1.2. Some other affixes, however, are completely underspecified with regard to their phonemic content and are thus made up entirely of segments copied from the final CVC string of the base. Such affixation is characteristic of the morpheme signalling imperfective aspect in verbs (see §4.7.1.1) and the nominalising collective morpheme added to numerals (see §4.1.5.2).

Copying — which is always a form of partial reduplication that involves only the consonants of the final CVC syllable, never the vowel nucleus — is of two types. One type involves the first consonant of the CVC string, the other type involves the second consonant. Following the terminology introduced by Kruspe (1999:112–121, 2004:72–76)

for Semelai, these two processes will be referred to here as *onset copy* and *coda copy* respectively. As described by Kruspe, it is not only the segmental features which are copied, but also the 'prosodic identity' of the segment. In other words, onsets are always copied as onsets; codas are always copied as codas. As will be illustrated below, this fact can be fully accounted for by means of principles of directionality, where right-to-left association of segments forms the unmarked pattern in Jahai.

As always in Jahai inner affixation, processes of copying affect only the penultimate syllable and produce forms that satisfy phonotactic constraints, notably the requirement that the penultimate syllable of an affixed form is full. The form of the affix is determined by the structure of the base, as the degree of prespecification of segments in the penultimate syllable of the resulting form varies according to the status of the penultimate syllable of the base. The following sections describe how underspecified affixes are realised in bases of different structure through the processes of onset and coda copy.

3.2.1.1.1 Coda copy

With sesqui- and disyllabic bases, underspecified affixation invariably involves the addition of an underspecified consonant segment at the left edge of the final CVC string of the base. This underspecified slot is filled by a copy of the final consonant of the final CVC string, which corresponds to the coda of the final syllable; hence the term *coda copy* to refer to this process.[13] In the present description, coda copy is explained as the result of reduplication of the final CVC string and the right-to-left association of its consonant segments with the underspecified slot of the template of the affix.

In the following procedural description, the copying process is illustrated by the sesquisyllabic Jahai verb root /kɟeŋ/ 'to listen', the underlying structure of which is CCVC, and its disyllabic imperfective form /k<ŋ>ɟeŋ/ 'to be listening'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation:

(1)

C	C	V	C	→	C	<C>	C	V	C
|	|	|	|		|		|	|	|
k	ɟ	e	ŋ		k		ɟ	e	ŋ

ii. Copy the phonemic content of the final CVC string and associate it from right to left with the underspecified consonant slot in the template:

			ɟ	e	ŋ			
				|				
C	<C>	C	V	C				
|		|	|	|	→	k<ŋ>ɟeŋ	
k		ɟ	e	ŋ				

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /k<ŋ>ɟeŋ/ → /k<ŋ>.ɟeŋ/ → [kəŋɟˀɛ̣ᵍŋ].

With disyllabic bases with an open penultimate syllable, the process is identical. However, the vowel nucleus of the penultimate syllable is prespecified and therefore no

13 Matisoff (2003:28) refers to this process of reduplicative infixation as *incopyfixation*.

realisation rule for that vowel is necessary. This is illustrated here by the disyllabic Jahai verb root /gulɛm/ 'to carry', the underlying structure of which is CVCVC, and its imperfective form /gu<m>lɛm/ 'to be carrying':

i. Add affix to the left edge of the final CVC string of the template of the underlying representation:

(2)
```
C V C V C   →   C V <C> C V C
| | | | |       | |     | | |
g u l ɛ m       g u     l ɛ m
```

ii. Copy the phonemic content of the final CVC string and associate it from right to left with the underspecified consonant slot in the template:

```
   l ɛ m
     |
C V <C> C V C
| |     | | |   →   gu<m>lɛm
g u     l ɛ m
```

iii. Apply rules of syllabification in order to produce the correct surface form: /gu<m>lɛm/ → /gu<m>.lɛm/ → [gumlɛ^{b}m].

In the case of penultimate syllables with a vowel nucleus /a/, there is an unusual pattern of vowel suppression and replacement, whereby the prespecified penultimate nucleus is not associated with the template of the affixed form. Instead, the realisation rules for underspecified vowels are applied, as in the case of sesquisyllabic bases. The reason for this epenthetic-like behaviour and suppression of penultimate /a/ is unclear.[14] Note that it is fully non-predictable and prespecified. In the following description the affix simply replaces the /a/. The pattern is exemplified with the verb root /bakɛs/ 'to grow up' and its disyllabic imperfective form /b<s>kɛs/ 'to be growing up':

i. Add affix to the left edge of the final CVC string of the template of the underlying representation, replacing the first vowel /a/ of the base:

(3)
```
C V C V C   →   C <C> C V C
| | | | |       |     | | |
b a k ɛ s       b     k ɛ s
```

ii. Copy the phonemic content of the final CVC string and associate it from right to left with the underspecified consonant slot in the template:

```
  k ɛ s
    |
C <C> C V C
|     | | |   →   b<s>kɛs
b     k ɛ s
```

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /b<s>kɛs/ → /b<s>.kɛs/ → [biskɛs].

---

[14] One possible explanation is that speakers may interpret a penultimate /a/ as a separate affix and not part of the root (cf. the reciprocal morpheme described in §4.7.1.6). For some reason it would therefore not be licensed to form part of the base and fill the vowel slot. Affixation would thus be applied to an imagined, non-existent root reminiscent of a back-formation.

Affixation in disyllabic bases with a heavy penultimate syllable involves the same strategy, although coda copy now entails a reorganisation of segments with regard to their syllabic position. As a consequence, the resulting form is trisyllabic, and the coda of the penultimate syllable of the base instead turns up in onset position. The process is illustrated here by the disyllabic verb /krlep/ 'to forget', the underlying structure of which is CCCVC, and its trisyllabic imperfective form /kr<p>lep/ 'to be forgetting':

i. Add affix to the left edge of the final CVC string of the template of the underlying representation:

(4)
```
C C C V C     →     C C <C> C V C
| | | | |           | |     | | |
k r l e p           k r     l e p
```

ii. Copy the phonemic content of the final CVC string and associate it from right to left with the underspecified consonant slot in the template:

```
  l e p
      |
C C <C> C V C
| |     | | |      →    kr<p>lep
k r     l e p
```

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /kr<p>lep/ → /k.r<p>.lep/ → [kərəp˺lɐp˺].

3.2.1.1.2 Onset copy + coda copy

In the case of monosyllabic bases, that is bases that are tantamount to a CVC string, copying for underspecified affixation involves both consonants of the CVC string, corresponding to the onset and coda. These form a CC affix which is prefixed to the left edge of the base and creates a heavy penultimate syllable. As was the case with coda copy, onset + coda copy is explained here as the result of reduplication of the final CVC string and the right-to-left association of its consonant segments with the underspecified consonant slots of the affix template. The procedure is exemplified here with the monosyllabic Jahai verb root /pɨs/ 'to sweep' and its disyllabic imperfective form /ps-pɨs/ 'to be sweeping'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation:

(5)
```
C V C     →     C C - C V C
| | |                 | | |
p ɨ s               p ɨ s
```

ii. Copy the phonemic content of the final CVC string and associate its consonants from right to left with the underspecified consonant slots in the template:

```
p ɨ s
| |
C C - C V C
      | | |      →    ps-pɨs
      p ɨ s
```

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /ps-pis/ → /ps-.pis/ → [pispis].

Note that direction of association is irrelevant to the outcome in the case of underspecified affixation to monosyllabic bases. Right-to-left association is employed here by analogy with the process of coda copy described in §3.2.1.1.1 and processes of partly prespecified affixation described in §3.2.1.2.1.

3.2.1.2 Partly prespecified affixation

Morphological processes frequently involve the affixation of forms in which phonological prespecification is partial. In other words, the phonemic content of the affix is partly external and prespecified, and partly underspecified and then copied from the final CVC string of the base. Partly prespecified affixation is mostly applied to mono- and sesquisyllabic bases and is only marginally associated with disyllabic bases. It is the only type of morphological process that necessitates a distinction between sesqui- and disyllabic forms. The prespecified content of the affix may consist of a consonant (§3.2.1.2.1) or a vowel (§3.2.1.2.2). The underspecified content is copied according to the same principles as those described in §3.2.1.1 (although onset copy in the formation of the marginal reciprocal affix with monosyllabic bases requires a reversed direction of association; see §3.2.1.2.2). As in other types of inner affixation, processes of partly prespecified affixation affect only the penultimate syllable and produce forms that satisfy phonotactic constraints.

Partly prespecified affixation is found in such divergent morphological categories as unit (§4.1.3, §4.2.1), nominalisation (§4.1.4, §4.1.4.3, §4.1.5.1), progressive aspect (§4.7.1.2), iterative aspect (§4.7.1.3), distributive Aktionsart (§4.7.1.5), reciprocal Aktionsart (§4.7.1.6), affixation of /m/ (§4.7.1.7), and causative (§4.7.2.1).

3.2.1.2.1 Prespecified consonants

The most common process of partly prespecified affixation involves the attachment of a CC affix to the left edge of the final CVC string of the base. This affix thus forms a heavy penultimate syllable. The first consonant of this affix is prespecified and external, whereas the second consonant is underspecified and copied from the final CVC string of the base. The process is characteristic of both monosyllabic and sesquisyllabic bases. In the following procedural description, the process is illustrated by the monosyllabic Jahai noun root /te?/ 'soil' and its disyllabic unitised form /n?-te?/ '[unit of] soil', involving an affix with a prespecified /n/:

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate prespecified phonemic information:

(6)
```
C V C      →    C C - C V C
| | |            |     | | |
t e ?            n     t e ?
```

ii. Copy the phonemic content of the final CVC string and associate its consonants from right to left with the underspecified consonant slot in the template:

```
   t e ʔ
     |
C  C  -  C  V  C
|        |  |  |      →    nʔ-teʔ
n        t  e  ʔ
```

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /nʔ-teʔ/ → /nʔ-.teʔ/ → [naʔtę̱ʔ].

The same procedure may be applied to the monosyllabic verb root /pɛl/ 'to drip' and its disyllabic iterative form /ll-pɛl/ 'to drip repeatedly', which involves an affix with a prespecified /l/:

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate prespecified phonemic information:

(7)
```
C  V  C      →    C  C  -  C  V  C
|  |  |           |        |  |  |
p  ɛ  l           l        p  ɛ  l
```

ii. Copy the phonemic content of the final CVC string and associate its consonants from right to left with the underspecified consonant slot in the template:

```
   p ɛ l
     |
C  C  -  C  V  C
|        |  |  |      →    ll-pɛl
l        p  ɛ  l
```

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /ll-pɛl/ → /ll-.pɛl/ → [ləlpɛl].

The process has been documented for affixation to monosyllabic bases involving the following prespecified consonants:

/n/	unit and nominalisation (§4.1, §4.2)
/b/	progressive aspect (§4.7.1.2)
/l/	iterative aspect (§4.7.1.3)
/m/	[semantic function unclear] (§4.7.1.7)
/p/	causative (§4.7.2.1)

In the case of sesquisyllabic bases, a CC affix, identical to that added to monosyllabic bases and consisting of one prespecified and one underspecified consonant, is affixed at the left edge of the final CVC string of the base. This process is restricted to affixes involving /n/, that is those that express the categories of unit and nominalisation (§4.1, §4.2). It is exemplified here by the sesquisyllabic verb root /tbɔh/ 'to beat' and its trisyllabic nominalised form /t<nh>bɔh/ 'act of beating':

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate prespecified phonemic information:

(8)
```
C  C  V  C      →    C  <C  C>  C  V  C
|  |  |  |           |  |       |  |  |
t  b  ɔ  h           t  n       b  ɔ  h
```

ii. Copy the phonemic content of the final CVC string and associate it from right to left with the underspecified consonant slot in the template:

```
   b ɔ h
     |
C <C  C> C V C
|  |     | | |      →      t<nh>bɔh
t  n     b ɔ h
```

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /t<nh>bɔh/ → /t.<nh>.bɔh/ → [tənahbɔh].

A deviant but marginal type of process affects sesqui- and disyllabic bases in the case of the semantically unexplained affixation of /m/ (see §4.7.1.7). Here, the penultimate syllable of the base is disposed of completely in order to make way for a CC affix, resulting in forms that look like affixed monosyllabic bases of the type described in examples (6) and (7). The process is exemplified here with the sesquisyllabic base /bdɛl/ 'to shoot' and its disyllabic affixed form /ml-dɛl/:

i. Add affix to the left edge of the final CVC string of the template of the underlying representation, replacing the initial consonant of the base, and associate prespecified phonemic information:

(9)
```
C C V C      →      C C - C V C
| | | |             |     | | |
b d ɛ l             m     d ɛ l
```

ii. Copy the phonemic content of the final CVC string and associate it from right to left with the underspecified consonant slot in the CV template:

```
  d ɛ l
    |
C   C - C  V  C
|       |  |  |      →      ml-dɛl
m       d  ɛ  l
```

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /ml-dɛl/ → /ml-.dɛl/ → [məldɛl].

3.2.1.2.2 Prespecified vowels

Another type of partly prespecified affixation involves the addition of an affix in which it is the vowel that is phonologically prespecified whereas its consonants are underspecified and copied from the final CVC string of the base. This type of affixation applies to the morpheme signalling distributive Aktionsart (§4.7.1.5), involving the vowel /i/, as well as one allomorph of the marginal morpheme signalling reciprocal Aktionsart (§4.7.1.6), involving the vowel /a/.

In the following procedural description, such a process is illustrated by the sesquisyllabic verb root /sjər/ 'to swim' and its distributive form /s<ir>jər/ 'to swim [here and there]'. Here, a VC affix consisting of a prespecified vowel /i/ and an underspecified consonant is attached to the left edge of the final CVC string of the base. Segmental

material for the underspecified consonant is copied from the final CVC string according to the now familiar pattern.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate prespecified phonemic information:

(10)
```
C  C  V  C      →      C  <V  C>  C  V  C
|  |  |  |             |   |      |  |  |
s  j  ə  r             s   i      j  ə  r
```

ii. Copy the phonemic content of the final CVC string and associate it from right to left with the underspecified consonant slot in the template:

```
      j ə r
        |
C  <V  C>  C  V  C
|   |      |  |  |        →      s<ir>jər
s   i      j  ə  r
```

iii. Apply rules of syllabification in order to produce the correct surface form: /s<ir>jər/ → /s<ir>.jər/ → [sirjər].

The process is also applied to disyllabic bases with a light penultimate syllable, but, as was the case with some disyllabic bases subjected to coda copy signalling imperfective (cf. example (3)), the prespecified nucleus of the penultimate syllable of the base is not associated with the template of the affixed form. This vowel slot is instead filled with the prespecified /i/ of the affix. The process is exemplified here with the verb root /tanɛm/ 'to plant', a Malay loanword, and its distributive form /t<im>nɛm/ 'to plant [here and there]':

i. Add affix to the left edge of the final CVC string of the template of the underlying representation, replacing the first vowel of the base, and associate prespecified phonemic information:

(11)
```
C  V  C  V  C      →      C  <V  C>  C  V  C
|  |  |  |  |             |   |      |  |  |
t  a  n  ɛ  m             t   i      n  ɛ  m
```

ii. Copy the phonemic content of the final CVC string and associate it from right to left with the underspecified consonant slot in the template:

```
      n ɛ m
        |
C  <V  C>  C  V  C
|   |      |  |  |        →      t<im>nɛm
t   i      n  ɛ  m
```

iii. Apply rules of syllabification in order to produce the correct surface form: /t<im>nɛm/ → /t<im>.nɛm/ → [timnɛ̃m].

However, unlike that associated with disyllabic bases subjected to coda copy signalling imperfective, the pattern of vowel replacement associated with distributives does not appear to be restricted to a prespecified nucleus /a/ but also affects other nuclei, as suggested by occasional examples like /g<im>lɛm/, from /gulɛm/ 'to carry', and /t<iŋ>duŋ/, from /tuduŋ/ 'to cover one's eyes'.

In the case of monosyllabic bases, a CVC affix is attached to the left edge of the CVC base. The vowel slot is filled with the prespecified /i/. There is no prespecified consonantal material available for the affix, which means that both its consonant slots are filled with material from the final CVC string through the processes of onset and coda copy. The procedure is similar to that described in §3.2.1.1.2. It is illustrated here with the monosyllabic verb root /ʔɛl/ 'to look' and its disyllabic distributive form /ʔil-ʔɛl/ 'to look [here and there]':

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate prespecified phonemic information:

(12)
```
C V C      →      C V C - C V C
| | |               |     | | |
ʔ ɛ l              i     ʔ ɛ l
```

ii. Copy the phonemic content of the final CVC string and associate its consonants from right to left with the underspecified consonant slots in the template:

```
ʔ ɛ l
|   |
C V C  -  C V C
  |       | | |      →    ʔil-ʔɛl
  i       ʔ ɛ l
```

iii. Apply rules of syllabification in order to produce the correct surface form: /ʔil-ʔɛl/ → /ʔil-.ʔɛl/ → [ʔilʔɛl].

It could be argued that processes of copying constitute the first operation in the creation of these forms, after which the prespecified vowel nucleus is added to signal distributive. This would entail that distributive forms are created from imperfectives (cf. §3.2.1.1, §4.7.1.1), the latter thus serving as a base for the secondary affixation of distributive /i/. However, such a connection between imperfectives and distributives cannot be motivated semantically on the basis of present knowledge. For the sake of consistency, therefore, an analysis is preferred whereby prespecified material is associated first.

With monosyllabic bases, affixation of the marginal morpheme signalling reciprocal Aktionsart (§4.7.1.5) involves the attachment of a CV affix at the left edge of the CVC base. This affix is made up of a prespecified vowel /a/ and a preceding underspecified consonant, the segmental material of which is copied from the base. It is the first consonant of the CVC base which is copied (corresponding to the onset), and this is the only morphological process in which onset copy is not associated with coda copy. It poses particular problems to the present account of copying in terms of descriptive consistency. This is because the right-to-left association of consonants of the copied final CVC string, which is employed to account for other forms of copying, would produce incorrect forms of the reciprocal CV affix. More specifically, right-to-left association would result in coda copy rather than onset copy. Therefore, for this particular affix, an analysis involving left-to-right association is proposed. Like the morpheme itself, this procedure is to be considered a very marginal feature of Jahai morphology. It is possible that the reciprocal morpheme is in fact a recent loan from a Central Aslian language like Temiar or Lanoh, which both display productive use of a morphophonemically similar affix expressing middle voice (Benjamin 1996).

The process is exemplified here for Jahai by the monosyllabic verb root /gej/ 'to eat' and its disyllabic reciprocal form /ga-gej/ 'to eat together'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate prespecified phonemic information:

(13)
```
C V C      →      C V - C V C
| | |              |   | | |
g e j              a   g e j
```

ii. Copy the phonemic content of the final CVC string and associate its consonants from left to right with the underspecified consonant slot in the template:

```
g e j
|
C V - C V C
  |   | | |      →      ga-gej
  a   g e j
```

iii. Apply rules of syllabification in order to produce the correct surface form: /ga-gej/ → /ga-.gej/ → [gag̩j].

3.2.1.3 Prespecified affixation

Several processes of inner affixation involve the addition of affixes which consist of segments which are phonologically wholly prespecified. This entails that their attachment to the base is not associated with any form of copying. The following description of such phonologically fully prespecified affixes distinguishes between those which attach to sesqui- and disyllabic bases and which are therefore infixes, and those which attach to monosyllabic bases and are thus prefixes. Note that all such affixes take the penultimate syllable as their domain and always result in morphologically complex forms which adhere to the phonotactic constraints of the language, which is the reason for their inclusion under inner affixation.

3.2.1.3.1 Prespecified infixation

Phonologically fully prespecified affixes added to sesqui- and disyllabic bases may be either syllabic, in which case the affix makes up the whole penultimate syllable of the resulting form, or non-syllabic, in which case the affix makes up part of the penultimate syllable of the resulting form in the form of either the vowel nucleus or the onset consonant, whereas the remainder of the syllable consists of segmental material already available in the base. Non-syllabic affixes, which are described first, include collective plural /<a>/ (§4.1.1, §4.1.4.2), reciprocal /<a>/ (§4.7.1.6), and the unitising and nominalising /<n>/ (§4.1.3, §4.1.4.1, §4.1.5.1, §4.2.1).

The addition of the collective plural morpheme to bases with a heavy penultimate syllable involves the insertion of prespecified /<a>/ at the left edge of the final CVC string of the base. This results in a reorganisation of the penultimate syllable, as the original coda of the penultimate syllable of the base is syllabified as an onset in the resulting form. Also, the form becomes trisyllabic, since the onset of the penultimate syllable of the base is syllabified as onset of the antepenultimate syllable in the resulting form. The process shows similarities to that described for coda copy in disyllabic bases with a heavy

penultimate syllable (cf. example (4) in §3.2.1.1.1). It is illustrated here by the noun root /tmkal/ 'man' and its collective plural form /tm<a>kal/ 'men'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate phonemic information:

(14)
C	C	C	V	C	→	C	C	<C>	C	V	C	→	tm<a>kal	
|	|	|	|	|		|	|	|	|	|	|			
t	m	k	a	l		t	m	a	k	a	l			

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /tm<a>kal/ → /t.m<a>.kal/ → [təmakal].

The affix <a> is also an allomorph of the reciprocal morpheme and as such is added to sesquisyllabic verb roots. In such cases, the penultimate half syllable of the base is turned into a light full syllable through the insertion of the prespecified <a> at the left edge of the final CVC string of the base. This is exemplified here by the verb root /smɛɲ/ 'to ask', the underlying structure of which is CCVC, and its reciprocal form /s<a>mɛɲ/ 'to ask each other'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate phonemic information:

(15)
C	C	V	C	→	C	<V>	C	V	C	→	s<a>mɛɲ			
|	|	|	|		|	|	|	|	|					
s	m	ɛ	ɲ		s	a	m	ɛ	ɲ					

ii. Apply rules of syllabification in order to produce the correct surface form: /s<a>mɛɲ/ → /s<a>.mɛɲ/ → [samɛ̃ɲ].

Non-syllabic prespecified affixes in the form of consonants only include /<n>/, which has a variety of semantic functions and which is added to disyllabic bases (cf. the affixation of /n/ in mono- and sesquisyllabic bases, described in §3.2.1.2.1). Productively, this affix appears to be found only in onset position of the penultimate syllable. This means that the affix is not attached at the left edge of the final CVC string of the base, as has hitherto been the case. Instead it occupies a consonant position to the left of the C or V that precedes the final CVC string. This entails that the original onset of the penultimate syllable of the base is syllabified in the resulting form as onset of a new antepenultimate half syllable. The process is illustrated in the following procedural description by the disyllabic noun /babo?/ 'woman', displaying a light penultimate syllable, and its unitised form /b<n>abo?/ '[unit of] woman'.

i. Add affix to the left of the V that precedes the final CVC string of the template of the underlying representation and associate phonemic information:

(16)
C	V	C	V	C	→	C	<C>	V	C	V	C	→	b<n>abo?	
|	|	|	|	|		|	|	|	|	|	|			
b	a	b	o	?		b	n	a	b	o	?			

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /b<n>abo?/ → /b.<n>a.bo?/ → [mə̃nabo?].

Affixation of bases with a heavy penultimate syllable is illustrated here with the morphologically complex but synchronically unanalysable verb /ckwɨk/ 'to talk', where the

coda of the penultimate syllable consists of a fossilised instance of coda copy, and its nominalised form /c<n>kwɨk/ 'act of talking', 'speech'.

i. Add affix to the left of the C that precedes the final CVC string of the template of the underlying representation and associate phonemic information:

(17)

C	C	C	V	C	→	C	<C>	C	C	V	C	→	c◄n>kwɨk				
|	|	|	|	|		|	|	|	|	|	|					
c	k	w	ɨ	k		c	n	k	w	ɨ	k						

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /c<n>kwɨk/ → /c.<n>k.wɨk/ → [cᵉənəkˈwɨkˈ].

As noted, the non-syllabic prespecified affix /<n>/ typically occupies onset position of the penultimate syllable and the process affects disyllabic bases. Recall that the affixation process of /n/ into sesquisyllabic bases involves the additional strategy of coda copy (see §3.2.1.2.1). In the case of a few sesquisyllabic bases, however, a fully prespecified affix /<n>/ is attached in the usual position of affixation at the left edge of the final CVC string. This entails that it is syllabified as a coda. Consequently, the penultimate syllable becomes heavy and the resulting form is disyllabic, not trisyllabic. The formula for this affixation process is reminiscent of that describing coda copy (see example (1)), the difference being that the affixed material is phonologically prespecified. It is illustrated here by the sesquisyllabic verb root /ksɛp/ 'to adorn oneself with leaves for good luck' and its nominalised form /k<n>sɛp/ 'leaves that bring good luck'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate phonemic information:

(18)

C	C	V	C	→	C	<C>	C	V	C	→	k<n>sɛp		
|	|	|	|		|	|	|	|	|			
k	s	ɛ	p		k	n	s	ɛ	p				

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /k<n>sɛp/ → /k<n>.sɛp/ → [kənsɛpˈ].

Affixation of a prespecified coda is described as productive in some other Aslian languages, e.g. the causative <r> of Semelai (Kruspe 2004:76). Although this type of affixation cannot be regarded as productive in the present Jahai material, it is possible that codas (notably nasals and liquids, cf. §2.4.3.2.2) of some heavy penultimate syllables of what are presumed to be unanalysable disyllabic roots are in fact traces of morphemes originally assigned according to such a process.

The last type of process of prespecified infixation discussed in this section involves the addition to sesquisyllabic bases and disyllabic bases with an open penultimate syllable of a CV affix, attached at the left edge of the final CVC string of the base. The affixed material makes up the whole penultimate syllable of the resulting form. This process includes two affixes: /<ra>/, which is an allomorph of the collective plural morpheme added to human nouns and several verbs (cf. §4.1.1, §4.1.4.2), and /◄ri>/, an allomorph of the causative morpheme added to verbs (cf. §4.7.2.5). The affixation entails that the penultimate onset of the base is syllabified as onset of an antepenultimate syllable in the resulting form. Possible prespecified vowels of the penultimate syllable of the base are not associated with the template of the affixed form. The following procedural description illustrates the sesquisyllabic verb root /tbɔh/ 'to hit' and its nominalised collective plural form /t<ra>bɔh/ 'fighters'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate phonemic information:

(19)

C	C	V	C	→	C	<C	V>	C	V	C	→	t<ra>bɔh
|	|	|	|		|	|	|	|	|	|		
t	b	ɔ	h		t	r	a	b	ɔ	h		

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /t◀ra>bɔh/ → /t.<ra>.bɔh/ → [tərabɔh].

The affixation of causative /<ri>/ follows the same pattern and is exemplified here with the sesquisyllabic verb root /hgik/ 'to be afraid' and its causativised form /h<ri>gik/ 'to frighten'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate phonemic information:

(20)

C	C	V	C	→	C	<C	V>	C	V	C	→	h<ri>gik
|	|	|	|		|	|	|	|	|	|		
h	g	i	k		h	r	i	g	i	k		

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /h<ri>gik/ → /h.<ri>.gik/ → [hərigikˇ].

3.2.1.3.2 Prespecified prefixation

Prespecified prefixation involves the addition of a phonologically prespecified syllabic affix at the left edge of a CVC string corresponding to a monosyllabic base. The affix may be of the form CV or CC. Such affixes include the three allomorphs of the causative morpheme that are associated with monosyllabic roots — /pi-/, /pr-/ and /tr-/ (see §4.7.2) — as well as the rare allomorph /br-/ of noun-to-verb derivation (see §4.7.3).

This type of affixation is difficult to distinguish in a straightforward manner from outer affixation (discussed in §3.2.2), the affixes being fully prespecified and, at least superficially, prefixed to the left edge of the base. It is described here as a form of inner affixation rather than outer affixation on the basis that (1) it takes the penultimate syllable as its domain, and (2) it produces morphologically complex forms that always meet phonotactic constraints (see §3.2.1).

The process is first exemplified by the monosyllabic verb root /hit/ 'to tremble' and its causative form /pr-hit/ 'to cause someone to tremble'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate phonemic information:

(21)

C	V	C	→	C	C	-	C	V	C	→	pr-hit
|	|	|		|	|		|	|	|		
h	i	t		p	r		h	i	t		

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /pr-hit/ → /pr-.hit/ → [pərhitˇ].

The following example shows the monosyllabic verb root /ʔɛ̃m/ 'to drink' and its causative form /pi-ʔɛ̃m/ 'to suckle'.

i. Add affix to the left edge of the final CVC string of the template of the underlying representation and associate phonemic information:

(22)
```
C V C      →   C V - C V C          →   pi-ʔɛ̃m
| | |          | |   | | |
ʔ ɛ̃ m          p i   ʔ ɛ̃ m
```

ii. Apply rules of syllabification in order to produce the correct surface form: /pi-ʔɛ̃m/ → /pi-.ʔɛ̃m/ → [piʔɛ̃m].

3.2.2 Outer affixation

Outer processes of affixation are those involving affixes that attach concatenatively to the left edge of a base. Thus, unlike inner affixation, outer affixation does not specifically take the penultimate syllable as its domain and does not necessarily result in morphologically complex forms that conform to word-structure constraints (cf. §2.4). Therefore, resulting forms frequently violate such constraints. This is because outer affixation, unlike inner affixation, is applied subsequently to the syllabification of the base. This non-conformity to word structure constraints is a feature that outer affixes share with clitics (see §3.3). However, like inner affixes, outer affixes attach at word level, as opposed to clitics, which attach at phrase level. In a way, then, outer affixes may be thought of as representing an intermediate category of bound morphemes between inner affixes and clitics.

Outer affixes consist invariably of a consonant attached to the left edge of a base. Following syllabification, it is realised as the onset of a half syllable. Outer affixes include the following:

/t-/	relative (§4.12.2)
/b-/	progressive aspect (§4.7.1.2); verbalisation (§4.7.3.1, §4.7.3.2)
/l-/	iterative aspect (§4.7.1.3)
/p-/	causative (§4.7.2.1)

Note that some of these outer affixes are allomorphs of morphemes which also display allomorphs which are assigned according to inner affixation. This applies in particular to iterative /l-/, causative /p-/ and, to some extent, progressive and verbalising /b-/. Thus, allomorphs of one and the same morpheme may be attached to a base according to what is considered here to be fundamentally different processes of affixation, highlighting the necessity of keeping the notions of morpheme and process firmly apart.

Outer affixes may be subcategorised according to their flexibility in terms of domain of attachment. Members of one category — including progressive /b-/, iterative /l-/ and causative /p-/ — can occur only in antepenultimate position. They cannot be attached to monosyllabic bases, and, as noted, with such bases they display allomorphs assigned according to inner affixation, notably in connection with coda copy. Correspondingly, these affixes cannot attach to bases that are trisyllabic, which means that they adhere to the restriction that Jahai words are maximally trisyllabic. This type of outer affixes is exemplified below as *antepenultimate outer affixes* (§3.2.2.1).

Members of the second category — including relative-marking /t-/ and some instances of verbalising /b-/ — are much more flexible with regard to their domain of attachment. They are not restricted to antepenultimate position but may also be attached to

monosyllabic bases and thus turn up in the same form in penultimate position. Hence they do not display any allomorphic relationship to inner affixation. Also, there are occasional examples of these affixes being attached to trisyllabic bases, thus turning up in pre-antepenultimate position and creating tetrasyllabic forms that are in violation of the restriction that Jahai words are maximally trisyllabic. This category of affixes is therefore notably clitic-like in its behaviour (cf. §3.3.1). Such affixes are exemplified below as *flexible outer affixes* (§3.2.2.2).

Forms derived through antepenultimate outer affixation may feed further derivation in the form flexible outer affixation, but the reverse is never the case. Antepenultimate outer affixes are syllabified prior to flexible outer affixation.

3.2.2.1 Antepenultimate outer affixes

As noted, antepenultimate outer affixes include progressive /b-/, iterative /l-/ and causative /p-/ and may be attached to the left edge of sesqui- and disyllabic bases. The following examples illustrate the pattern. Note that affixation to a sesquisyllabic base produces a morphologically complex form that violates the restriction that the penultimate syllable of a trisyllabic form must be full.

(a) Sesquisyllabic bases:

/kdih/	'to say'	/b-kdih/	'to be saying'
/bdɛl/	'to shoot'	/l-bdɛl/	'to shoot repeatedly'
/hgik/	'to be afraid'	/p-hgik/	'to frighten'

(b) Disyllabic bases:

/kanər/	'to carry on shoulder'	/b-kanər/	'to be carrying on shoulder'
/bj-baj/	'to be digging'	/b-bj-baj/	'to be digging on'
/saʔot/	'to call someone'	/l-saʔot/	'to call someone repeatedly'

3.2.2.2 Flexible outer affixes

Flexible outer affixes include relative-marking /t-/ and some instances of verbalising /b-/ and may be attached to the left edge of monosyllabic, sesquisyllabic, disyllabic and trisyllabic bases. The following examples illustrate affixation to monosyllabic and trisyllabic bases. (Affixation to sesqui- and disyllabic bases would produce forms similar in syllabic structure to those exemplified in §3.2.2.1). Note that affixation to a monosyllabic base results in a sesquisyllabic form. This is the only situation in which sesquisyllabic forms may be morphologically complex. Also note that affixation to a trisyllabic base produces a tetrasyllabic form, which is in violation of the restriction that words are maximally trisyllabic.

(a) Monosyllabic bases:

/kul/	'to call'	/t-kul/	'calling'
/bəw/	'to be big'	/t-bəw/	'big'
/tɔm/	'water'	/b-tɔm/	'to have water'

(b) Trisyllabic base:

/b-tadɔʔ/	'to be waiting'	/t-b-tadɔʔ/	'waiting'

3.2.3 The origins of inner and outer affixation

Kruspe (2004:64) cautiously proposes for Semelai that non-concatenative affixation (which corresponds to inner affixation in the present account of Jahai) is particular to indigenous morphology, whereas concatenative (outer) affixation is characteristic of affixes borrowed from Malay. No such clear correlation between the origin of affixes and type of affixation can be posited for Jahai. This is because there are patterns, evident albeit not very widespread, in which affixed material of apparent Malay origin under certain circumstances is added to bases according to the principles of inner affixation. The most significant example of such a pattern is that allomorph of the progressive morpheme which is affixed to monosyllabic verb roots. Here, the borrowed Malay affix segment /b/ is added to the base in conjunction with coda copy according to a partly prespecified process of inner affixation (cf. §3.2.1.2.1, §4.7.1.2).

Other affixes may bear witness to similar use of borrowed affix segments in processes of inner affixation, one example being /l/ in combination with coda copy, which forms that allomorph of the iterative morpheme which is added to monosyllabic verbs (§3.2.1.2.1, §4.7.1.3). However, the origin of these segments, including /l/, is not known in detail. It is important in this context to reiterate that two different allomorphs of the same morpheme, although making use of the same affix segment (borrowed or not), may be assigned according to different processes of affixation. This rule applies to both the progressive and the iterative morphemes, where affixation to monosyllabic bases is inner, whereas affixation to sesqui- and disyllabic bases is outer. This allomorphic inconsistency with regard to process of affixation alone goes to prove that there is no absolute correlation between the origin of the affix and type of affixation.

This is not to say that autochthony versus foreignness necessarily is of no relevance to the distinction between processes of inner and outer affixation in Jahai. The predominant and primary nature of inner affixation may well indicate that it represents the indigenous type of affixation, and the secondary nature of outer affixation (recall its application subsequently to inner affixation and the syllabification of the base) may point to a borrowed process. If this should prove to be the case, indigenous processes in Jahai, unlike those of Semelai, are able to some extent to integrate foreign morphological elements, pointing to a high degree of adaptability and productivity of such processes.

3.2.4 Total reduplication

The last type of affixation process described here is that of total reduplication of a base. It is applied to nouns to form diverse plurals (see §4.1.2), to verbs to signal continuative aspect (see §4.7.1.4) and to some interrogatives to create indefinite forms (see §4.6). The process involves the total copying of the base and its concatenation to the left edge of the base. The syllabic structure and phonemic content of the base is thus fully retained in the reduplicated portion, the only difference being that the latter does not receive stress. In terms of segmental realisation, the reduplicated portion retains its word-like characteristics. In this respect, totally reduplicated forms bear some structural similarity to compounds (cf. §4.1).

In the present notation, the reduplicated portion is joined to its base with a hyphen and is glossed as a grammatical morpheme according to the pattern of affixes. The process is illustrated here for (a) monosyllabic bases, (b) sesqui- and disyllabic bases, and (c) a morphologically complex disyllabic base.

(a) CVC → CVC - CVC
/mɛj/ 'what' → /mɛj-mɛj/ 'whatever'
/tek/ 'to sleep' → /tek-tek/ 'to keep on sleeping'

(b) C(V)CVC → C(V)CVC - C(V)CVC
/kɟeŋ/ 'to listen' → /kɟeŋ-kɟeŋ/ 'to keep on listening'
/pagiʔ/ 'morning' → /pagiʔ-pagiʔ/ 'various mornings'

(c) CC-CVC → CC-CVC - CC-CVC
/ɟʔ-ɟiʔ/ 'to be burning' → /ɟʔ-ɟiʔ-ɟʔ-ɟiʔ/ 'to be keeping on burning'

Although this process, like many of those of inner affixation (cf. §3.2.1), utilises the strategy of copying material from the base, it is evident that it is fundamentally different from the copying processes of inner affixation, the latter targeting only the consonants of the final CVC string of the base and resulting in well-formed penultimate syllables. Structural as well as semantic features (cf. Zaharani 1991:113–119, for example) suggest that the process of total reduplication has been borrowed from Malay. A similar stance is taken by Kruspe (2004:81) on 'light syllable reduplication' in Semelai.

Totally reduplicated forms feed further derivation only in the form of outer affixation. In spite of their otherwise restricted domain of attachment, antepenultimate outer affixes (cf. §3.2.2.1) may attach to the reduplicated portion of totally reduplicated forms, e.g. /b-kdih-kdih/ 'to be keeping on saying'.

3.3 Cliticisation

The second morphological strategy involving bound morphemes is that of *cliticisation*, and the morphological units used in this strategy are *clitics*. Clitics are characterised in the present description straightforwardly as bound morphemes which attach syntactically to phrases, clauses or some other unit of words. This definition distinguishes clitics from affixes, which must attach at word-level, and from words in that they must attach phonologically to a host. The characterisation conforms to that proposed for clitics by Klavans (1982, 1985) and is explained further in §3.3.1 and §3.3.2.

However, it should be clear from the previous discussion of affixation that bound morphemes in Jahai do not admit of any uncomplicated division into affixes and clitics, some outer affixes being notably clitic-like in their phonotactic behaviour (cf. §3.2.2). The proposed continuum of base-dependence of bound morphemes appears to display a number of gradations in terms of such dependence, only one of which represents the difference between affix and clitic. This suggests that a sharp differentiation between affixes and clitics may turn out to be artificial, and perhaps a use of Klavans's term *phrasal affix* to refer to clitics would better express the undramatic transition between these two types of bound morphemes. The use of the terms *clitic* and *cliticisation* in the present account is motivated only by way of descriptive convention.

3.3.1 Phonological and phonotactic characteristics

As noted, the one factor that distinguishes clitics from words is that they must attach phonologically to a host. More specifically, clitics need a host and do not qualify for wordhood because (1) they cannot receive stress, and (2) they do not usually fulfil the

requirement that the minimal canonical structure for a word is a heavy syllable CVC with a phonemic vowel nucleus. Instead, most clitics take the form of a minimal syllable C(V). The one clitic that does represent a heavy syllable with a phonemic nucleus and which might therefore be considered a possible word on canonical grounds — the prepositional proclitic allomorph /can=/ 'SOURCE' (§4.9.3) — does not qualify as a word as it cannot receive stress and does not behave phonetically in a word-like manner, as its final nasal segment is realised phonetically as a simple nasal [-n] and not as the prestopped allophone [-^{d}n] typical of word-final position (cf. §2.3.1.2).

A problematic category of words in this context are personal pronouns (cf. §4.3), the pre-verbal subject-marking allomorphs of which (the so-called *subject particles*, see §5.1.1.1) cannot receive stress. Also, pronouns ending in a glottal stop /ʔ/ are frequently reduced phonologically in this position in the sense that the glottal stop is elided and the pronoun then takes the form of a minimal syllable CV. It might therefore be argued that pre-verbal allomorphs of personal pronouns exhibit the characteristics of dependence associated with clitics. Indeed, in most accounts of other Aslian languages, pre-verbal allomorphs of personal pronouns are described as bound, e.g. for Temiar (Benjamin 1976b:158–159, 1996), Jah Hut (Diffloth 1976b:86–87) and Semelai (Kruspe 2004:88–89). However, it is important to point out that phonological reduction of pre-verbal pronouns appears to be much greater in these languages and that Jahai does not conform to the same pattern. Thus, in Jahai, the elision of syllable-final glottal stops is a common phonetic reduction in connected speech (§2.3.1.1) and should perhaps not be considered a significant argument for the 'clitic-hood' of pre-verbal allomorphs of pronouns, especially since pronouns ending in other consonants retain their word-like characteristics, notably the form /gin/ '2/3P', the final nasal of which is realised phonetically in the typically word-final prestopped manner: [gi^{d}n]. Against this background, the only feature which distinguishes these forms from other allomorphs of pronouns (as well as words in general) is their inability to receive stress. This would suggest that they display features of both words and clitics and that they therefore hold an intermediate position between these two categories. Refutably, pre-verbal allomorphs of personal pronouns will be described here as free forms rather than clitics on the basis of the meagreness of evidence of phonological attachment to a host.

Phonotactically, clitics are by and large insensitive to the structure of the base to which they are attached. Thus, they generally do not display allomorphs determined by base structure, and they are not characterised by any distributional restrictions determined by the number of syllables of the base. For example, forms consisting of a base + a clitic frequently violate the restriction that well-formed words are maximally trisyllabic. These are features which clitics share with the flexible outer affixes discussed in §3.2.2. The following examples of irrealis clitics attached to trisyllabic bases illustrate this insensitivity to the number of syllables of the base.

(a) ja=b-ŋk-ŋɔk
IRR=PROG-IMPF-to.sit
'[I] will be sitting.'

(b) wa=b-nasiʔ
IRR.3S=PROP-rice
'He will have rice.'

However, in the case of some clitics, the vowel nucleus may be either underspecified or prespecified, and this allomorphic variation is to some extent determined by base structure

(see also §4.9). This is particularly apparent in the case of the prepositional proclitic /can=/ 'SOURCE' (§4.9.3), the phonologically reduced form of which, /cn=/, is the preferred form with disyllabic words with a penultimate nucleus /a/. The following examples illustrate this:

(a) can=hɨp
SOURCE=forest
'from the forest'

(b) can=ɟhũʔ
SOURCE=tree
'from the tree'

(c) cn=hawɛ̃n
SOURCE=pig
'from the pig'

3.3.2 Domain and location of attachment

As noted, the domain of attachment of clitics is a phrase, clause or some other unit of words. The location of attachment refers to that word within the phrase or clause which acts as the base, or host, of the clitic — that is, the word to which the clitic is phonologically attached. Jahai clitics invariably attach to the left edge of their base; hence they are always proclitics. The following description, influenced by Klavans's (1985) typology of clitics, discusses clitics from the point of view of domain and location and, in keeping with the present account of bound morphemes as a reflection of a spectrum of base-dependence, it is organised according to the degree of flexibility of clitics in relation to the categorial identity of the host, beginning with those types of clitics that are least flexible in this respect.

3.3.2.1 Clitics hosted by verbs

The modal proclitics (see also §4.7.4) include the categories of *irrealis* (/ja=/ 'IRR', /wa=/ 'IRR.3S'), *desiderative* (/ma=/ 'DES') and *hortative* (/ha=/, /ca=/, /ka=/ 'HORT'). The host of these clitics is always a verb, a feature which makes them difficult to distinguish from affixes. The irrealis proclitics are coreferential with an optional phrase denoting the subject and mutually exclusive with the subject-marking pre-verbal pronouns (cf. §5.1.1). Although their host is always a verb, they will be interpreted here as attaching at clause level. This may be formalised as follows:

[(...) **clitic**=V (...)]CLAUSE

The following clauses exemplify the pattern:

(a) **wa**=pi-ʔɛ̃m wɔŋ ʔoʔ
IRR.3S=CAUS-to.drink child 3S
'She will suckle her baby.'

(b) **ja**=cɛ̃ɸ ka=jɛʔ
IRR=to.fan.fire SUBJ=1S
'I will fan the fire.'

The syntactically determined nature of the irrealis proclitics, as well as their post-derivational attachment and insensitivity to the phonotactic structure of the base, confirm their clitic-like status. The desiderative marker /ma=/ displays similar characteristics and appears to be mutually exclusive with the irrealis markers. However, due to its marginal status, it will not be considered further here.

The hortative proclitics take the imperative clause as their domain. An example is given below.

(a) **ha=**cɨp ba=ʔə̃h
HORT=to.go GOAL=here
'Please, come here!'

Modal proclitics, always being phonologically attached to a verb, constitute the least flexible type of clitic in terms of the categorial identity of its host.

3.3.2.2 Clitics hosted by the first constituent of the NP

The prepositional proclitics (see §4.9), including the categories of *location/ instrument/subject* (/ka= ~ k=/ 'LOC'/'INSTR'/'SUBJ'), *goal* (/ba=/ 'GOAL'), *source* (/can=/ 'SOURCE'), *contrast* (/d=/ 'CONTR') and *equation* (/pn=/ 'EQU'), as well as the identification-marking proclitic /l= ~ la=/ (see §4.12.3), take the noun phrase as their domain of attachment. Their host is always the first constituent of the noun phrase and may be represented by a noun, pronoun, demonstrative, interrogative, numeral or quantifier. The following formula summarises the pattern:

clitic=[NP]

The following phrases serve as examples:

(a) **ba=**wɔŋ kɟɨh jɛʔ tə̃h
GOAL=child boy 1S this
'to this son of mine'

(b) **l=**nɛj ʔə̃h
ID=one here
'the one here'

3.3.2.3 Clitics hosted by the first constituent of the core

The proclitic signalling relational tense (/ɟa=/ 'RT', see also §4.10.1.1) takes the obligatory part of the clause, referred to here as the *core*, as its domain of attachment (cf. §5.1.1). Its host is always the first constituent of the core and may thus be represented by the preverbal subject-marking pronoun, an irrealis-marked verb or, in the case of some stative verbs, the verb itself.

clitic=[CORE]

This is illustrated in the following examples:

(a) **ɟa=**jɛʔ ʔtʔet
RT=1S to.know
'I already know.'

(b) **ɟa=**wa=cɨp
RT=IRR.3S=to.go
'Then he would leave.'

(c) **ɟa=**sɔc
RT=to.be.gone
'It's already finished.'

3.3.2.4 Clitics hosted by the first constituent of the clause

The proclitic signalling *interrogative* (/ha=/ 'Q', see §4.10.3) takes the clause as its domain of attachment (cf. §5.1.1). Its host is the first constituent of the clause and may be represented by the pre-verbal subject-marking pronoun, an irrealis-marked verb, the verb itself (all of which may also be relational tense-marked, cf. §3.3.2.3), or a clause-initial subject NP in the form of a noun or a pronoun.

clitic=[CLAUSE]

The following examples illustrate the pattern:

(a) **ha=**mɔh ʔtʔet
Q=2S.FAM to.know
'Do you know?'

(b) **ha=**wa=cɨp
Q=IRR.3S=to.go
'Is he leaving?'

(c) **ha=**btʔɛt
Q=to.be.good
'Is it good?'

(d) **ha=**tmkal wa=wek
Q=man IRR.3S=to.return
'Will the man return?'

3.3.2.5 Clitics hosted by any constituent

The relative-marking proclitic (/k=/ 'REL', see also §4.12.1) may take any relative clause or phrase as its domain. Its host is the first constituent of that domain and may be represented by a noun, a pronoun, a demonstrative, a verb or a prepositional proclitic. The following examples illustrate the pattern:

(a) **k=**babo?
REL=woman
'who [is a] woman'

(b) **k=**hapaʔ ktɔʔ wɛj
REL=to.die day past
'who died yesterday'

(c) **k=**pn=ʔəh
REL=EQU=here
'that [is] like this'

(d) **k=**jɛʔ tbɔh
REL=1S to.hit
'that I hit'

The disjunctive co-ordinating proclitic /ha=/ 'or' may be similarly attached to the first constituent of any unit (see §4.11).

3.4 Summary

This chapter has defined the units and outlined the processes of word formation in Jahai. It has been primarily concerned with the characteristics of bound morphology, the processes of which were organised according to their degree of 'dependence' on the base to which they are applied. Factors determining such base dependence include e.g. the domain of attachment, the structure of the base and the use of reduplication. Thus, apart from a conventional distinction between affixation and cliticisation, a fundamental distinction was also made between *inner* affixation and *outer* affixation. Inner affixation takes the penultimate syllable as its domain; it frequently copies phonemic information from the base; and it results in forms that adhere to the phonotactic constraints of the language. Outer affixation, on the other hand, affects the left edge of a base; it involves no copying; and it frequently produces forms that violate phonotactic constraints.

The idea that inner and outer affixation represent an indigenous and a borrowed system respectively was discussed tentatively. At any rate, it is clear that Jahai makes productive use of affix elements borrowed from Malay in both types of process. However, an additional type of morphological process, total reduplication, was suggested to have been borrowed from Malay.

By and large, the processes discussed here appear to correspond to those described for some other Aslian languages. For example, very similar processes have been described for Semelai by Kruspe (2004:61–93), and the intricate patterns of copying find equivalents in accounts of Temiar (Benjamin 1976b:168–169), Semai (Diffloth 1976a:234–237) and Jah Hut (Diffloth 1976b:105–109).

4 *Word classes*

This chapter describes word classes in Jahai and the morphological categories associated with them. The description begins with the nominal classes of nouns (§4.1), classifiers (§4.2), personal pronouns (§4.3), demonstratives (§4.4), numerals/quantifiers (§4.5), and interrogatives (§4.6). These are followed by the verbal classes of verbs (§4.7) and expressives (§4.8). These, in turn, are followed by the closed classes of prepositions (§4.9), auxiliaries and adverbs (§4.10), conjunctions (§4.11) and co-ordinating morphemes within the NP (§4.12). A summary is given in §4.13. The description focuses on the identification and semantic characterisation of word classes and their categories. Where relevant, brief reference is made to morphological processes and syntactic characteristics. However, for more extensive treatment of these areas, see Chapter 3 and 5 respectively.

4.1 Nouns

Nouns form a semantically well-defined word class in Jahai. For example, as in other Aslian languages (cf. Diffloth 1976d:249–250), very few roots have both nominal and verbal meaning. Only the following equivocal indigenous roots have been recorded so far:

Root	**Nominal meaning**	**Verbal meaning**
/ʔap/	'large feline'	'to encounter a large feline'
/ʔɛc/	'belly; excrement'	'to defecate'
/ptis/	'pain; sickness'	'to be in pain; to be sick'
/was/	'fork; confluence'	'to split'
/ʔɛ̄m/	'breast'	'to drink'
/knɔm/	'urine'	'to urinate'
/lɛn/	'loincloth'	'to wear a loincloth'
/sŋɔr/	'silence'	'to be silent'

The lexeme forms of nouns are generally morphologically simplex; that is, they are represented by monomorphemic roots and do not contain traces of morphological processes that are synchronically non-productive (see §3.1). A conspicuous exception to this pattern is the class of animal names, the members of which frequently display fossilised morphology, notably in the form of copying. Some such instances of copying may have an onomatopoeic function. Some examples are given below:

/cprɛp/	'babbler'
/wtwɨt/	'(a type of bird)'
/ktlit/	'glow-worm'
/skŋuk/	'(a type of frog)'
/ckcok/	'Diard's trogon'
/ckcɔk/	'banded palm civet'
/rksɔk/	'keelback'
/cʔcaʔ/	'common kingfisher'
/thteh/	'oriental pied hornbill'
/khkuh/	'(a type of hornbill)'
/kuhʔɔ̄h/	'(a type of turtle)'
/smsɨ̄m/	'silver-eared mesia'
/pŋpɨ̄ŋ/	'Philippine glossy starling'
/kŋkaŋ/	'Asian horned toad'
/tŋwaŋ/	'blue coral snake'
/tŋtūŋ/	'(a type of spider)'
/gŋguŋ/	'(a type of civet)'
/klɗil/	'(a type of snake)'
/brhur/	'(a type of snake)'
/hwhɨ̄w/	'crested wood-partridge'
/bjbɔj/	'(a type of insect)'

Furthermore, as in other Aslian languages, some animal names appear to contain remains of an affix involving the phoneme /l/, which is suggested by Diffloth (1976b:100–101) to represent a non-productive morpheme originally signalling erratic, 'step by step' movement (cf. §4.7.1.3 and Kruspe 2004:86). Jahai examples involve invertebrates and include the following:

/hlaŋkɛt/	'(a type of ant)'
/klutbot/	'(a type of larva)'
/klcbac/	'(a type of millipede)'
/kluktɔk/	'(a type of terrestrial gastropod)'

Morphological complexity in the form of copying may also be observed in several nouns denoting body parts, including the following:

/krtlɔ̄t/	'kidney'
/smutlɔ̄t/	'brain'
/dkduk/	'chest'
/sʔsoʔ/	'blood vessel'
/kmkə̄m/	'ankle'
/cŋcɨ̄ŋ/	'(area around the eyes)'
/dlɗil/	'heel'
/prbɛr/	'lower arm'
/ɟwɟəw/	'Achilles tendon'

A connection between fossilised reduplicative morphology and names of animals and body parts has been noted in some other Mon-Khmer languages as well, e.g. Semelai

(Kruspe 2004:85–86) and Minor Mlabri, a Kammuic language spoken in northeastern Thailand (Rischel 1995:94–95).

Nominal compounds are fairly common. These consist of a left-headed construction of two free nominal morphemes, where the modifying noun always bears stress. Names of body parts frequently occur as metaphorical heads of such compounds. Examples of compounds include the following:

(1) ʔap ʔawɛj
large.feline vine
'leopard'

(2) kit tɔm
buttocks river
'river mouth'

(3) mit ktɔʔ
eye sky
'sun'

(4) kdek ʔabuʔ
squirrel dust
'Provost's squirrel'

(5) ʔɛ̃m kajiʔ
breast bat
'(a type of thorn)'

Names of birds, fish, snakes, trees and vines frequently consist of a compound with the generic names of these classes as head. Examples include the following:

(6) ʔikəʔ bawuŋ
fish type.of.catfish
'(a type of catfish)'

(7) taɟuʔ sjul
snake type.of.cobra
'(a type of cobra)'

(8) tom taŋuj
tree rambutan
'rambutan tree'

Furthermore, a handful of locative nouns may combine with other nouns in compound-like constructions where the locative noun forms the head. Such locative nouns signal a particular location in relation to the referent of the modifying noun. These constructions are frequently best translated into English as prepositional phrases. Locative nouns include /tkɨh/ 'backside', /kjɔm/ 'underside', /sɛɲ/ 'front', /krpiŋ/ 'upper side', /klɛŋ/ 'inside', /ʔnaŋ/ 'outside' and /sir/ 'side'. Examples are given in (9) and (10).

(9) krpiŋ hajɛ̃ʔ
upper.side house
'upper side of house'/'above the house'

(10) klɛŋ bulo?
inside bamboo.tube
'inside of bamboo-tube'/'inside the bamboo-tube'

Syntactically, nouns form part of NPs, where they may function either as heads or as modifiers. As modifiers they generally represent a modifying possessor of a possessed head noun in constructions that are reminiscent of compounds, as in the following examples:

(11) hajɛ̃? gop
house stranger
'the stranger's house'

(12) haʈi? ʔap
tail tiger
'the tiger's tail'

The following sections describe the derivative categories that are marked morphologically on the noun. These include collective plural (§4.1.1), diverse plural (§4.1.2) and unit (§4.1.3), which all pertain to quantification. Nominalisations in the form of verb-to-noun derivation and numeral-to-noun derivation are described in §4.1.4 and §4.1.5 respectively.

4.1.1 Collective plural

A small set of human nouns and ethnonyms may be collectivised by means of a collective plural morpheme (COLL) which has two fully prespecified allomorphs: /<ra>/ and /<a>/. These are determined by the structure of the base. The allomorph /<ra>/ is associated with bases with a light penultimate syllable, whereas the allomorph /<a>/ is associated with bases with a heavy penultimate syllable. The following forms make up the full set of collectivised human nouns and ethnonyms recorded:

/babo?/	'woman'	/b<ra>bo?/	'women'
/tmkal/	'man'	/tm<a>kal/	'men'
/ʔanɛk/	'girl'	/ʔ<ra>nɛk/	'girls'
/kɟih/	'boy'	/k<ra>ɟih/	'boys'
/bakɛs/	'adult'	/b<ra>kɛs/	'adults'
/mɲsaw/	'daughter/son-in-law'	/mɲ<a>saw/	'daughters/sons-in-law'
/lamij/	'sister/brother-in-law'	/l<ra>mij/	'sisters/brothers-in-law'
/mnraʔ/	'person'	/mn<a>raʔ/	'persons'
/ɟahaj/	'Jahai'	/ɟ<ra>haj/	'Jahais'
/plɛh/	'Temiar'	/p<ra>lɛh/	'Temiars'
/knsiw/	'Kensiw'	/kn<a>siw/	'Kensiws'
/knta?/	'Kintaq'	/kn<a>taʔ/	'Kintaqs'
/batɛk/	'Batek'	/b<ra>tɛk/	'Bateks'
/smɛj/	'Semai'	/s<ra>mɛj/	'Semais'
/tmwɛn/	'Temuan'	/tm<a>wɛn/	'Temuans'

Collective-like forms of three non-human nouns have also been recorded:

/hajɛ̄ʔ/	'house'	/h<ra>jɛ̄ʔ/	'houses'
/hapɔj/	'lean-to'	/h<ra>pɔj/	'lean-tos'
/ɟhūʔ/	'tree'	/ɟ<ra>hūʔ/	'gaps (in canopy/wood)'

Note that collective plural allomorphs associated with monosyllabic bases have not been documented.

From Schebesta (1928a:810811) and onwards, the /<ra>/<a>/ morpheme of Northern Aslian languages has been described as signalling plural number in human nouns. The present-day Jahai exponent is typically restricted to rather fixed constructions involving a human noun preceded by a determining attributive plural pronoun, e.g. /heʔ ɟ<ra>haj/ 'we Jahai', /gin b◄ra>boʔ/ 'you/them women'. It is never used in counting or quantification. Its function is not primarily to signal true plural number, it seems, but rather to collectivise several referents of the same class into a common group. See also §4.1.4.2. for verb-to-noun derivation of such collective human nouns. Collective plural forms do not feed further derivation, with the possible exception of relative (§4.12.2).

4.1.2 Diverse plural

Diverse plural (DP) is formed by means of total reduplication of the nominal lexeme. The resulting form signals diversity of members or kinds of a class. The process is probably borrowed from Malay and is mostly applied to Malay loan words, although indigenous nouns often receive the same treatment. A similar function of nominal reduplication is noted for Perak Malay by Zaharani (1991:116, 129).

Forms borrowed from Malay

/tmpət/	'place'	/tmpət-tmpət/	'various (kinds of) places'
/pagiʔ/	'morning'	/pagiʔ-pagiʔ/	'various (kinds of) mornings'
/kritəh/	'car'	/kritəh-kritəh/	'various (kinds of) cars'

Indigenous forms

/mnraʔ/	'people'	/mnraʔ-mnraʔ/	'various (kinds of) people'
/ɟlmɔl/[15]	'mountain'	/ɟlmɔl-ɟlmɔl/	'various (kinds of) mountains'
/kɛn/	'child'	/kɛn-kɛn/	'various (kinds of) children'

4.1.3 Unitisation

Nouns modified by a numeral, the quantifier /kɔm/ 'many' or the interrogative /mɛj siʔ/ 'how many' typically receive an affix involving the phoneme /n/ when the number of referents is in focus. A similar situation was described by Schebesta (1928a:811) for Jahai, and equivalent morphemes have been attested in Semai (Diffloth 1976a:236), Jah Hut (Diffloth 1976b:99–100) and Semelai (Kruspe 2004:218–219). This morpheme is intended to demarcate discrete units of the referent of the noun in order to make it more suitable for counting and quantification. It will here be labelled *unitiser* (UNIT), the resulting form will be called *unitised noun*, and the process will be referred to as *unitisation*. This terminology is inspired by that of Lucy (1992:73), who discusses unitisation in relation to numeral classifiers in Yucatec Maya.

---

15 This is possibly not an indigenous Jahai form but a borrowing of Temiar /ɟlmɔl/ 'mountain' (/jɛlmɔl/ in Benjamin's 1976a:111 transcription).

The morpheme has three different allomorphs which are determined by the structure of the base. With monosyllabic bases, a CC affix consisting of the prespecified /n/ and an underspecified consonant attaches as a prefix to the left edge of the CVC string of the base. The underspecified slot is filled by a copy of the final consonant of the base through the process of coda copy.

/sɛc/	'meat'	/nc-sɛc/	'[unit of] meat'
/teʔ/	'soil'	/nʔ-teʔ/	'[unit of] soil; place'
/tɔm/	'water'	/nm-tɔm/	'[unit of] water'
/can/	'leg'	/nn-can/	'[unit of] leg'

Affixation in sesquisyllabic bases involves the infixation of a similar CC affix at the left edge of the final CVC string of the base.

/ktɔʔ/	'daylight'	/k<nʔ>tɔʔ/	'[unit of] day'
/ɟʔes/	'root'	/ɟ<ns>ʔes/	'[unit of] root'
/ʔnaŋ/	'side'	/ʔ<nŋ>naŋ/	'[unit of] side'
/lwɛj/	'honey'	/l<nj>wɛj/	'[unit of] honey'

In disyllabic forms, the prespecified /n/ is infixed to become onset of the penultimate syllable:

/baboʔ/	'woman'	/b<n>aboʔ/	'[unit of] woman'
/kaɲcɔʔ/	'grandchild'	/k<n>aɲcɔʔ/	'[unit of] grandchild'
/tmpət/	'place'	/t<n>mpət/	'[unit of] place'
/ɟlmɔl/	'mountain'	/ɟ<n>lmɔl/	'[unit of] mountain'

The following examples illustrate the use of unitised nouns:

(13) lpəs tigaʔ k<nʔ>tɔʔ ja=ɟok ba=klap
after three day<UNIT> IRR=to.move GOAL=Kelap
'After three days I will move to Kelap.'

(14) ʔakɛt ŋɔk nɛj t<n>mpət
PROH to.sit one place<UNIT>
'Don't stay in one place.'

(15) mɛj siʔ ɟ<n>lmɔl
what number mountain<UNIT>
'How many mountains?'

(16) ʔoʔ ʔɛk nɛj nc-sɛc
3S to.give one UNIT-meat
'He gave away one piece of meat.'

(17) duwaʔ t<n>agiɲ blap
two wood<UNIT> nothing
'Two pieces of wood is nothing.'

Most nouns, including several borrowed from Malay and English, are susceptible to unitisation, and the process therefore appears to be fully productive. Nouns which refer to substances, which are conventionally considered to be 'mass nouns', are treated the same way as nouns which refer to inherently discrete (or bounded) entities such as physical objects, so-called 'count nouns'. But whereas unitisation of inherently discrete nouns

usually appears to simply *actualise* the natural 'discreteness' which already belongs to the concept of the noun in question, unitisation of non-discrete (or unbounded) entities such as substances rather *creates* the unit to be counted (see Croft 1994:162–163 and Bisang 1999:120–121 for discussions on creative and actualising individuation in relation to quantifiers and classifiers). This creative aspect of unitisation of non-discrete entities leaves room for a variety of context-dependent interpretations of unitised mass nouns. Hence, the construction /nɛj nm-tɔm/ 'one [unit of] water' (from /tɔm/ 'water') may refer to any amount of water.

Still, the morphologically identical treatment of count and mass nouns and the frequently utilised possibility of unitising inherently discrete nouns in counting and quantification indicate that *all* nouns may be characterised by a high degree of indeterminateness, as is common in Southeast Asia, and that all nouns in their underived form are basically unspecified as to unit.

A special interpretation is encountered with unitised forms of nouns which may be considered to be containers of some form. In such cases, the referent of the noun functions as a measurement and is referred to by its unitised form to indicate that its contents make up a full unit. Similar meanings have been identified in Semelai by Kruspe (2004:219).

/baniʔ/	'quiver'	/b<n>aniʔ/	'quiverful'
/bas/	'bus'	/ns-bas/	'busload'
/gunih/	'sack'	/g<n>unih/	'sackful'
/kapal/	'aircraft'	/k<n>apal/	'aircraftload'

Unitised nouns do not feed further derivation.

4.1.4 Verb-to-noun derivation

4.1.4.1 Nominalising /n/

Affixation of /n/ in verbs typically derives verbal nouns which denote the state of being or act/manner/way of doing whatever is denoted by the verb. The process is fully productive and may be applied to any verb, including borrowings from Malay. The affix will here be labelled *nominaliser* (NM). Nominalised verbs behave syntactically like ordinary nouns and become NP heads or modifying nouns of NP heads. The morphological processes involved are identical to those described for the unitisation of nouns (§4.1.3) and allomorphs of the /n/ affix are thus determined by the structure of the base. Examples include the following:

Monosyllabic:	/cɨp/	'to go'	/np-cɨp/	'act of going'
	/sam/	'to hunt'	/nm-sam/	'act of hunting'
	/cɔl/	'to tell'	/nl-cɔl/	'act of telling'
Sesquisyllabic:	/ɟhɨt/	'to smoke'	/ɟ<nt>hɨt/	'act of smoking'
	/tbɔh/	'to beat'	/t<nh>bɔh/	'act of beating'
	/ʔnaj/	'to bathe'	/ʔ<nj>naj/	'act of bathing'
Disyllabic:	/ckwɨk/	'to talk'	/c<n>kwɨk/	'act of talking'
	/kajil/	'to fish'	/k<n>ajil/	'act of fishing'
	/pikir/	'to think'	/p<n>ikir/	'act of thinking'

As noted, verbal nouns in Jahai typically name the general state or activity designated by the verb. *All* verbal nouns appear to be able to have this unmarked interpretation. In the terminology of Comrie and Thompson (1985:350), for example, such general verbal nouns are referred to as *action/state nominalisations*. The following sentences exemplify such readings:

(18) c<n>kwɨk ʔoʔ ton lajin-lajin, braʔ ja=ʔtʔet
to.speak<NM> 3S that CONT-to.be.different NEG IRR=to.know
'That way of speaking of his is very different. I don't understand [it].'

(19) lpəs nk-ɟok japɛ̃h wek
after NM-to.move 1P.EXCL to.go.back
'After moving we went back.'

(20) tmpət nk-ŋɔk japɛ̃h pdəh
place NM-to.sit 1P.EXCL to.be.near
'Our place of residence is nearby.'

(21) slamat nk-wek btʔɛt
safety NM-to.go.back to.be.good
'Have a good trip back!'

In addition to these general state and action/manner interpretations of nominalisations, more specialised and concrete meanings of the kind described by Benjamin (1976b:176–177) for Temiar, Diffloth (1976b:98–99) for Jah Hut, Kruspe (2004:222–223) for Semelai and Svantesson (1983:92–94) for Kammu, where nominalisations may denote different roles associated with the nominalised verb, are also present in Jahai. Thus, Jahai nominalisations frequently also denote the patient/theme (*object nominalisation*), instrument (*instrumental nominalisation*) or location (*locative nominalisation*) of an action. Most nominalised verbs have such specialised extensions but the process does not appear to be quite as productive as that of action/state nominalisation. Table 4.1 exemplifies the different meanings of nominalisations.

Table 4.1: Examples of meanings of nominalisations in Jahai

	ACTION	OBJECT	INSTRUMENT	LOCATION
/np-cɨp/ from /cɨp/ 'to go'	'act/manner of going'	–	'legs' 'car' 'wheelchair'	'route' 'path'
/nc-kɛc/ from /kɛc/ 'to cut'	'act/manner of cutting'	'thing cut' 'circumcised person'	'cutter' 'knife'	'place of cutting'
/nm-cɔm/ from /cɔm/ 'to burn'	'act/manner of burning'	'fuel'	'fire'	'swidden'
/k<n>ajil/ from /kajil/ 'to fish'	'act/manner of fishing'	'fishing catch'	'fishing rod'	'fishing place'
/h<n>aluh/ from /haluh/ 'to shoot with blowpipe'	'act/manner of shooting'	'game' 'quarry'	'blowpipe'	'hunting place'

The meanings of some concrete nominalisations may become very narrow and specific, such as the object interpretation of the nominalisation /nc-kɛc/ 'thing cut' (from /kɛc/ 'to cut'), which is sometimes used to mean 'circumcised person', 'Muslim'. Also, the instrumental interpretation of /k<n>ajil/ ('fishing-rod') has become so specific that other types of fishing gear, such as casting nets, cannot be referred to as /k<n>ajil/.

Nominalised forms do not feed further derivation. The syntactic behaviour of nominalisations is briefly introduced in §5.1.4.3.

4.1.4.2 Collective plural /<ra>/ and /<a>/

Most stative and several dynamic verbs may be nominalised by means of the collective plural morpheme (COLL) described for human nouns in §4.1.1. The resulting form is a collective noun denoting a group of people characterised by the state, or performing the action, designated by the verb. The process appears to be fairly productive. The pattern of allomorphic variation is identical to that described for human nouns in §4.1.1. Also, like the collectivised human nouns, collective plural nominalisations are typically restricted to rather fixed constructions with a preceding determining attributive plural pronoun.

/hɲɟut/	'to be heavy'	/hɲ<a>ɟut/	'heavy people'
/hgik/	'to be afraid'	/h<ra>gik/	'frightened people'
/crɔʔ/	'to be hungry'	/c<ra>rɔʔ/	'hungry people'
/kbis/	'to be dead/to die'	/k<ra>bis/	'dead people'
/ptis/	'to be sick'	/p<ra>tis/	'sick people'
/manɛh/	'to be old'	/m<ra>nɛh/	'old people'
/pcah/	'to break'	/p<ra>cah/	'breakers'
/tbɔh/	'to hit'	/t<ra>bɔh/	'fighters'

Collective plural nominalisations do not feed further derivation, with the possible exception of relative (§4.12.2).

4.1.4.3 Nominalising /m/

The presence of a nominalising morpheme involving the phoneme /m/ is not so much in evidence in Jahai as in some other Aslian languages, such as Jah Hut (Diffloth 1976b:98) and Semelai (Kruspe 2004:224–225), where it is described as a non-productive morpheme once used to derive agentive- or instrumental-like nominalisations. The two examples identified in Jahai can be given a related interpretation:

/wĩʔ/	'left'	/mʔ-wĩʔ/	'left-handed person'
/tem/	'right'	/mm-tem/	'right-handed person'

4.1.5 Numeral-to-noun derivation

4.1.5.1 Nominalising /n/

Affixation of the morpheme involving /n/ (NM) in numerals produces nouns which denote the state of being the number designated by the numeral. The process, the allomorphic patterns of which are identical to those described for unitisation of nouns (§4.1.3) and nominalisation of verbs (§4.1.4.1), is marginal and involves the numerals 1–8.

/nɛj/	'one'	/nj-nɛj/	'the state of being one'
/duwaʔ/	'two'	/d<n>uwaʔ/	'the state of being two'
/tigaʔ/	'three'	/t<n>igaʔ/	'the state of being three'
/ʔmpat/	'four'	/ʔ<n>mpat/	'the state of being four'
/limɛʔ/	'five'	/l<n>imɛʔ/	'the state of being five'
/nɛm/	'six'	/nm-nɛm/	'the state of being six'
/tuɟoh/	'seven'	/t<n>uɟoh/	'the state of being seven'
/lapan/	'eight'	/l<n>apan/	'the state of being eight'

4.1.5.2 Collectivising coda copy

Affixation of a morpheme involving coda copy (COLL) in numerals produces nouns which denote a group of the number of referents (usually humans) designated by the numeral. The process is marginal. Like the collectivised nouns (§4.1.1), collectivised numerals are usually restricted to fixed constructions with a preceding determining attributive plural pronoun.

/duwaʔ/	'two'	/d<ʔ>waʔ/	'[group of] two'
/tigaʔ/	'three'	/ti<ʔ>gaʔ/	'[group of] three'
/spuloh/	'ten'	/spu<h>loh/	'[group of] ten'
/sblas/	'eleven'	/sb<s>las/	'[group of] eleven'

4.2 Classifiers

Enumeration and other types of quantification often involve the use of a noun functioning as a classifier (CLF). The set of classifiers is rather small, and only two are of frequent occurrence: /kɛn/ 'child' for human nouns, and /kmɔʔ ~ kbiʔ/ 'fruit' for all non-human animate and many inanimate nouns.[16] These may be considered as making up a basic, primary two-way system of classification. In addition, there is a handful of sporadically occurring classifiers with more specific usage. Some of these are indigenous forms and others are Malay loans. It is possible that many of the indigenous forms are in fact calques of Malay classifiers. To some extent, the use of these peripheral classifiers appears to overlap that of the more basic non-human classifier /kmɔʔ ~ kbiʔ/, and idiolectal and contextual variation in usage is evident. Table 4.2 lists the full set of classifiers so far identified for Jahai.

As is evident from the description of unitised nouns in §4.1.3, counting/quantification does not require classifiers. Furthermore, if used, a classifier and its modifying numeral seldom form a phrasal unit with the noun they refer to. Instead, the numeral and the classifier typically make up a separate NP which is in some way detached from and syntactically opposed to the noun. Frequently, the classifier NP is separated from the noun by a pause, or it may be used anaphorically to replace a noun which has been introduced earlier in the discourse or an implicit noun not previously introduced overtly.

---

16 The form /kmɔʔ/ is in free variation with /kbiʔ/, which is possibly a loan of Temiar /kəbəəʔ/, which also means 'fruit' (Benjamin, pers. comm.).

Table 4.2: Classifiers in Jahai

	Classifier	Original meaning	Class characteristics	Examples
Indigenous forms	kɛn	'child'	human	all human nouns
	kmɔʔ ~ kbɨʔ	'fruit'	non-human	animals fruit various objects
	tom	'tree'	vegetation	trees plants grass
	haliʔ	'leaf'	small flat objects	leaves paper sheets notes
	tuŋkɔ̄l	'stone'	spherical/ cubical objects	stones
	mit	'eye'	small objects	seeds stones
Forms borrowed from Malay	bidaŋ (from *bidang*)	'broad/flat piece'	large flat objects	mats blankets fields
	bataŋ (from *batang*)	'tree-trunk', 'stick', 'shaft'	oblong objects	trees pencils cigarettes
	kpiŋ (from *keping*)	'portion' 'piece'	flat objects	plates helpings of food
	biɟiʔ (from *biji*)	'seed'	small objects	seeds stones
	buwah (from *buah*)	'fruit'	spherical/ cubical objects	houses
	ʔikɔr (from *ekor*)	'tail'	animal	all animal nouns
	prduʔ (from *perdu*)	'base of tree-trunk'	clusters of objects	clusters bunches banana stems

Classifier constructions can thus be largely seen as pragmatically motivated references to and emphases and specifications of the noun they refer to rather than syntactically obligatory components of quantified NPs. Examples are given below. For further discussion of the syntactic characteristics of classifier constructions, see §5.1.1.1.4.

(22) jɛʔ bʔbɔʔ tomɛn, duwaʔ k<nʔ>mɔʔ
1S to.carry.on.back snakehead two CLF<UNIT>
'I carried snakeheads. Two of them.'

(23) ɟa=jɛʔ bdil ʔamɛŋ k=tɔm maŋəh ʔon lɛh,
RT=1S to.shoot siamang LOC=river Mangga there EMP
duwaʔ k<nʔ>mɔʔ
two CLF<UNIT>
'Then I shot siamangs by Mangga river! Two of them.'

(24) d=jɛʔ, nɛj k<nʔ>mɔʔ
CONTR=1S one CLF<UNIT>
'One for me.'
[Uttered by a man who distributes quartered quarry among his group]

In the few recorded cases where the classifier and the numeral appear to form a phrasal unit with the noun, the order of constituents is numeral–classifier–noun. This is illustrated in the following example:

(25) jɛʔ bdil spuloh k<nʔ>mɔʔ kasaʔ
1S to.shoot ten CLF<UNIT> sambar.deer
'I shot ten sambar deer.'

Jahai classifiers, it seems, therefore have mainly discourse-related functions, and it is noteworthy that discourse-related features have been of great importance in explaining the function of classifiers in neighbouring Malay (see e.g. Hopper 1986), a long-standing and important source of influence on Jahai.

4.2.1 Unitisation

Like ordinary nouns, indigenous classifiers are frequently unitised by means of the /n/ morpheme (UNIT) described in §4.1.3. This obviously follows from their co-occurrence with numerals. The noun to which the classifier construction refers is never unitised, since the numeral in such cases co-occurs with the classifier, not the noun.

/nn-kɛn/	from	/kɛn/	CLF: human
/k<nʔ>mɔʔ/	from	/kmɔʔ/	CLF: non-human
/nm-tom/	from	/tom/	CLF: vegetation
/h<n>aliʔ/	from	/haliʔ/	CLF: small flat objects
/t<n>uŋkɔ̃l/	from	/tuŋkɔ̃l/	CLF: spherical/cubical objects
/nt-mit/	from	/mit/	CLF: small objects

Unitised forms have been identified occasionally for two of the classifiers borrowed from Malay:

/ʔ<n>ikɔr/	from	/ʔikɔr/	CLF: animal
/p<n>rduʔ/	from	/prduʔ/	CLF: clusters of objects

The fact that classifiers are unitised by means of a separate morpheme is interesting because it suggests that the classifier does not itself carry the semantic component of unit (which is one of the functions traditionally associated with classifiers) but only adds a classifying dimension to the construction. At the very least, any component of unit in the classifier would be made redundant by the use of the unitiser. The general and rather unspecific nature of the classifier system would thus suggest that classifiers are semantically more or less dispensable, just as they are syntactically dispensable, which would seem to support the notion that their common presence is mainly discourse-motivated.

4.3 Personal pronouns

4.3.1 Pronominal distinctions

Jahai personal pronouns are marked obligatorily for singular, dual and plural number. A first vs second vs third person distinction is made for singular and dual pronouns, whereas plural pronouns have a first vs non-first person distinction. Inclusion vs exclusion of the second person is distinguished in first person dual and plural. At least three degrees of familiarity/politeness are distinguished in second person singular: intimate, familiar and distant. The former two are used with close friends and familiar persons respectively, the latter with strangers, children and spouses. No gender distinctions are made (see Table 4.3).

Table 4.3: Personal pronouns in Jahai

<table>
<tr><th></th><th colspan="3">Singular</th><th colspan="2">Dual</th><th colspan="2">Plural</th></tr>
<tr><td rowspan="2">1</td><td colspan="3" rowspan="2">jɛʔ</td><td>Inclusive</td><td>Exclusive</td><td>Inclusive</td><td>Exclusive</td></tr>
<tr><td>hɛj</td><td>jɛh</td><td>heʔ</td><td>japɛ̄h~pɛ̄h</td></tr>
<tr><td rowspan="2">2</td><td>Intimate</td><td>Familiar</td><td>Distant</td><td colspan="2" rowspan="2">ɟɨh</td><td colspan="2" rowspan="2">gin</td></tr>
<tr><td>miʔ</td><td>mɔh</td><td>paj</td></tr>
<tr><td>3</td><td colspan="3">ʔoʔ</td><td colspan="2">wih</td></tr>
</table>

A special system of honorific pronoun use indicates in-law relationship of the speaker to the referent in the second and third person singular (see Table 4.4). This system is drawn from the second and third person dual and plural forms of the general system of personal pronouns.

Table 4.4: In-law pronouns in Jahai

<table>
<tr><th></th><th colspan="3">Singular</th></tr>
<tr><th></th><th>to father-in-law</th><th>to daughter/son-in-law</th><th>to sister/brother-in-law</th></tr>
<tr><td>2</td><td rowspan="2">gin</td><td>wih</td><td>ɟɨh</td></tr>
<tr><td>3</td><td>gin</td><td>wih</td></tr>
</table>

4.3.2 Allomorphs of personal pronouns

Pronouns occur in stressed and unstressed form. The stressed form is used in answer to questions of the type 'who is that?', in prepositional phrases, and as argument NPs. The unstressed form is used as a pre-verbal subject agreement marker (a so-called *subject particle*; see §5.1.1.1), a pre-nominal determiner in NPs (e.g. /gin kɛn/ 'them children') and sometimes postverbally in imperative constructions. As a modifying postnominal possessor in NPs, a pronoun may be either stressed or unstressed, depending on whether emphasis is placed on possessor or possessed: /wɔŋ 'jɛʔ/ '*my* child', /ʼwɔŋ jɛʔ/ 'my *child*'.

Unlike some other Aslian languages, Jahai is not analysed here as having a set of bound allomorphs of pronouns (pronominal clitics), although somewhat reduced forms are in free variation with some of the unstressed pronouns when these occur as preverbal subject agreement markers.[17] This applies only to pronouns with a final glottal stop /ʔ/, which is usually dropped in this position: [jɛ ~ jɛʔ] /jɛʔ/ '1S', [ʔo ~ ʔoʔ] /ʔoʔ/ '3S', [hɐ̨ ~ hɐ̨ʔ] /heʔ/ '1P INCL'. Such deletion of syllable-final glottal stops is a common phonetic reduction in connected speech (see §2.3.1.1). Otherwise final consonants are retained and display word-final characteristics. Notably, the typically word-final prestopped allophone of /n/ is retained in the unstressed form of /gin/ '2/3P'. The unstressed pronouns are treated here then as free forms and not as clitics on the basis that they behave segmentally like phonological words (see also §3.3.1 and §5.1.1.1).[18]

4.4 Demonstratives

4.4.1 Basic demonstratives

Jahai has a multiterm system of basic demonstratives involving eight distinctions encoding accessibility, exteriority and elevation of locations in relation to the speaker and the addressee. These eight terms are given here, but, for a more detailed semantic description, see §4.4.1.1–4.4.1.8.

/ʔə̃h/	'here'
/ʔon/	'there (you know)'
/ʔũn/	'there (you don't know)'
/ʔaniʔ/	'there (away)'
/ʔadeh/	'there (beyond me)'
/ʔɲiʔ/	'there (beyond you)'
/ʔitih/	'there (up)'
/ʔujih/	'there (down)'

These demonstratives may function as either heads or modifiers in adverbial NPs. As heads of NPs they usually occur in prepositional phrases and then combine with any of a set of four prepositional proclitics expressing location at (/ka=/), motion to (/ba=/), motion from (/can=/) and similarity to (/pn=/) the location designated by the demonstrative (see also §4.9). The following examples illustrate such constructions:

(26) ja=wek ba=ʔə̃h lɛh
IRR=to.go.back GOAL=here EMP
'I will return here!'

(27) can=ʔon jɛʔ ɟok ba=pulɔw tuɟoh
SOURCE=there 1S to.move GOAL=Pulau Tujuh
'From there I moved to Pulau Tujuh.'

The /k=/ allomorph of the locative prepositional proclitic is not found in immediate combination with basic demonstratives but with the corresponding nominal demonstrative,

---

17 Bound allomorphs of pronouns have been described for e.g. Temiar (Benjamin 1976b:158–59), Jah Hut (Diffloth 1976b:86–87) and Semelai (Kruspe 2004:88–89).

18 The first person plural exclusive pronoun /japɛ̃h/ is in free variation with a reduced variant /pɛ̃h/ in the preverbal subject agreement position.

e.g. /k=tə̃h/ '(at) here', literally 'at this'. However, it may be combined with basic demonstratives if the identification-marking proclitic /la=/ is inserted between it and the demonstrative: /k=la=ʔə̃h/ '(at) here' (see §4.12.3).

The following example illustrates a basic demonstrative in the form of a NP head which is not part of a PP.

(28) ʔoʔ b-tk-tek ʔadeh
3S PROG-IMPF-to.sleep there
'He was sleeping there [beyond me].'

Basic demonstratives are infrequent as modifiers of NP heads, as this position is usually filled by the corresponding nominal demonstrative (see §4.4.2). They are found in postnominal position. Examples include the following:

(29) kuciŋ ʔūn
cat there
'the cat over there'

(30) slaj heʔ ʔə̃h
swidden 1S here
'our swidden here'

The following sections describe the semantic characteristics of the demonstrative terms.

4.4.1.1 Speaker-anchored accessible /ʔə̃h/

This distinction is associated with referents conceived of as in some way accessible to the speaker, e.g. with regard to their proximity, perceptibility, reachability/approachability, possession and topicality in discourse. Speaker-proximal location is typical.

4.4.1.2 Addressee-anchored accessible /ʔon/

This distinction is associated with referents considered by the speaker to be 'cognitively accessible' to the addressee, i.e. referents which have the addressee's current or previous attention/knowledge. Proximity to addressee is common.

4.4.1.3 Speaker-anchored inaccessible /ʔaniʔ/

This distinction is associated with referents conceived of as inaccessible to the speaker, e.g. with regard to distance, imperceptibility, unreachability etc. Speaker-distal location is typical.

4.4.1.4 Addressee-anchored inaccessible /ʔūn/

This distinction is associated with the introduction of new referents, i.e. referents which do not have the addressee's current or previous attention/knowledge and therefore are 'cognitively inaccessible' to the addressee. There is no typical spatial patterning of referents; location is flexible.

4.4.1.5 Speaker-anchored exterior /ʔadeh/

This distinction is associated with referents located on the other side of the speaker from the addressee's position; distance is irrelevant.

4.4.1.6 Addressee-anchored exterior /ʔɲɨʔ/

This distinction is associated with referents located on the other side of the addressee from the speaker's position; distance is irrelevant.

4.4.1.7 Superjacent /ʔitɨh ~ ʔotɨh/

This distinction is associated with referents located above the speech situation, either in the immediate area of the speech situation (including vertically above) or with reference to landscape contour (uphill) or river profile (upstream). The allomorphs /ʔitɨh ~ ʔotɨh/ are in free variation and idiolectally determined.

4.4.1.8 Subjacent /ʔujih/

This distinction is associated with referents located below the speech situation, either in the immediate area of the speech situation (including vertically below) or with reference to landscape contour (downhill) or river profile (downstream).

4.4.2 Nominal demonstratives

The system of basic demonstratives described in §4.4.1 provides the basis for an identically categorised system of nominal demonstratives. The initial glottal stop /ʔ/ of the basic demonstratives is replaced by /t/ to form such nominal demonstratives.

/tə̄h/	'this'
/ton/	'that (you know)'
/tūn/	'that (you don't know)'
/taniʔ/	'that (away)'
/tadeh/	'that (beyond me)'
/tɲɨʔ/	'that (beyond you)'
/titɨh/	'that (up)'
/tujih/	'that (down)'

The primary function of nominal demonstratives is that of adnominal modifiers in the NP. They are always found in postnominal position. This is illustrated by the following examples:

(31) hajɛ̄ʔ tūn
house that
'that house'

(32) ɟhūʔ titɨh
tree that
'that [up] tree'

Occasionally nominal demonstratives occur pronominally as NP heads:

(33) ja=gɐj tə̃h
IRR=eat this
'[I] will eat this.'

The initial /t/ of the nominal demonstratives is likely to be related to the relative marking affix /t-/ described in §4.12.2, as both seem to serve to co-ordinate heads of NPs with modifiers. However, the /t/ of the nominal demonstratives, unlike the relative marker, has fused phonologically with its host roots and the resulting forms are therefore unanalysable.

4.4.3 Temporal demonstrative

Jahai has a temporal demonstrative /wɛj/ 'past' which functions as a postnominal modifier in the NP. It signals past status of the noun it modifies. The following examples illustrate its use.

(34) britis wɛj
British past
'those British [of the past]'

(35) ʔɛj jɛʔ wɛj
father 1S past
'my late father'

4.5 Numerals and quantifiers

4.5.1 Numerals

All Jahai numerals above the number one are Malay loans.[19] The basic numerals are the following:

1	nɛj	8	lapan
2	duwaʔ	9	smilan
3	tigaʔ	10	spuloh
4	ʔmpat ~ ʔmpət	11	sblas
5	limɛʔ	12	duwaʔ blas
6	nɛm	100	sratos
7	tuɟoh	1000	sribuh

Numerals typically represent prenominal modifiers of nouns or classifiers, which are then usually unitised (cf. §4.1.3 and §4.2.1):

(36) nɛj t<n>mkal
one man<UNIT>
'one man'

---

19 Diffloth (1976c:33) suggests that the only Mon-Khmer numeral in Jahai, /nɛj/ 'one', is a loan of Temiar /nɛj/ 'one' (or /nɛy/ in his transcription), the common Northern Aslian reflex being /naj/. However, Benjamin (1976a:113) gives the differing form /nej/ (or /ney/) for Temiar, and a sound shift of final syllable /a/ to /ɛ/ appears to be a regular process that distinguishes Jahai from other Northern Aslian languages, so Diffloth's suggestion must be regarded as uncertain.

(37) tiga? nn-can
three UNIT-foot
'three feet'

(38) spuloh nn-kɛn
ten UNIT-CLF
'ten people'

Numerals may also function as heads of NPs, as in the following example:

(39) ja=gɐj nɛj
IRR=to.eat one
'[I] will eat one.'

As heads, they are frequently modified by preposed determining attributive dual and plural pronouns, as in the following examples (cf. §4.3.2):

(40) wih duwa?
3D two
'them two'

(41) gin tiga?
2/3P three
'you/them three'

The numeral /nɛj/ 'one' may take the causative prefix /pi-/ (cf. §4.7.2.2) to signal 'causation of oneness', the resulting form /pi-nɛj/ meaning 'one each', 'one by one', 'one at a time', 'one after another' or the like. This is illustrated by the following example:

(42) ʔoʔ ʔɛk pi-nɛj pi-nɛj k<n?>mɔʔ
3S to.give CAUS-one CAUS-one CLF<UNIT>
'He gave away one by one.'

4.5.2 Quantifiers

Like numerals, quantifiers are usually prenominal modifiers of a noun or classifier. Only two indigenous quantifiers have been recorded: /kɔm/ 'many' and /pɛ̃w/ 'other'. The former frequently co-occurs with a unitised noun or classifier (see §4.1.3 and §4.2.1).

(43) kɔm nh-bɔh
many UNIT-fruit
'many fruits'

(44) pɛ̃w mnraʔ
other person
'other person'

Other quantifiers are Malay loans and include /smwɛʔ/ 'all' (from Malay *semua* 'all'), /tjap-tjap/ 'every' (from Malay *tiap-tiap* 'every') and /masiŋ-masiŋ/ 'each' (from Malay *masing-masing* 'each'). These are exemplified below.

(45) smwɛʔ ʔɔ̃t
all dog
'all dogs'

(46) tjap-tjap ʔɔ̃t
every dog
'every dog' (elic.)

(47) masiŋ-masiŋ ʔɔ̃t
each dog
'each dog' (elic.)

Quantifiers may also occur independently and then represent heads of NPs.

4.6 Interrogatives

The set of Jahai interrogative words (or WH words) display five basic distinctions of questioning, described in §4.6.1–4.6.5. They are represented by five roots, four of which appear (optionally or obligatorily) in conjunction with an interrogative prefix /ma-/, the exact function of which is unclear (and it is therefore not glossed separately). Interrogatives are typically clause-initial elements questioning arguments or adjuncts in the clause, or NP-initial elements questioning modifiers of NP heads. Occasionally they are also found in predicate position. The syntactic characteristics of WH words are described in §5.1.2.2 and §5.1.4.2.

4.6.1 Person-questioning /makɛn/

The interrogative /makɛn/ is used for the questioning of person (typically human beings) and corresponds to English 'who?' and 'whose?'. Its root /kɛn/ is the classifier denoting human beings. The interrogative prefix /ma-/ is obligatory.

(48) ka=puŋhuluh ba=ʔaniʔ, makɛn
SUBJ=headman GOAL=there who?
'Who is headman over there?'

(49) makɛn hajɛ̃ʔ
whose? house
'whose house?'

The reduplication /makɛn-makɛn/ represents an indefinite form meaning 'whoever', 'anybody'.

4.6.2 Time-questioning /mapuʔ/

The form /mapuʔ/ is a time-questioning interrogative and corresponds to English 'when?'. There is no known root */puʔ/.

(50) cn=mapuʔ
SOURCE=when?
'Since when?'

4.6.3 Item/situation/reason-questioning /mamɛj ~ mɛj/

The interrogative /mamɛj ~ mɛj/ is used for questioning item or situation, corresponding to English 'what?' and 'which?', or reason, then corresponding to English

'why?'. The interrogative prefix /ma-/ is optional and does not appear to carry any additional meaning.

(51) mamɛj k=hɲɟut
what? REL=to.be.heavy
'What [is it that] is heavy?'

(52) mɛj paj d?-de?
what? 2S.DIS IMPF-to.do
'What are you doing?'

(53) mamɛj mɔh ɟim
why? 2S.FAM to.cry
'Why do you cry?'

If /mɛj/ is combined with the word /si?/ 'number', the resulting construction questions number and corresponds to English 'how many?'.

(54) mɛj si? nn-kɛn
what? number UNIT-CLF
'How many people?'

The reduplication /mɛj-mɛj/ represents an indefinite form meaning 'whatever', 'whichever', 'anything'.

4.6.4 Manner-questioning /ma?acin ~ ?acin/

The form /ma?acin ~ ?acin/ is used for questioning manner and thus corresponds to English 'how?'. The interrogative prefix /ma-/ is optional and does not appear to carry any additional meaning.

(55) ma?acin mɔh de?
how? 2S.FAM to.do
'How do you do it?'

4.6.5 Location-questioning /lbah/

The interrogative /lbah/ is used to question location and corresponds to English 'where?'. Unlike the other interrogatives, it is never found in combination with the interrogative prefix /ma-/. It frequently appears in conjunction with the prepositional proclitic /b=/ 'to'.

(56) b=lbah mi? ja=ɟok
GOAL=where? 2S.INT IRR=to.move
'Where will you move?'

(57) lbah paj ŋɔk
where? 2S.DIS to.sit
'Where did you live?'

The reduplication /lbah-lbah/ represents an indefinite form meaning 'wherever', 'anywhere'.

4.7 Verbs

Verbs, which function as predicates of clauses, are defined syntactically for Jahai as words which may be negated, that is, they may be preceded by the negative marker /braʔ/ and receive an irrealis proclitic.

Like those of nouns (cf. §4.1), the lexeme forms of verbs are usually morphologically simplex; that is, they are represented by monomorphemic roots and do not contain traces of morphological processes that are synchronically non-productive. However, numerous exceptions to this pattern display fossilised morphology in the form of copying and thus resemble structurally the imperfective forms described in §4.7.1.1. Such exceptions have no identifiable synchronic roots. Verbs from different semantic domains are represented among these morphologically complex lexemes, although a great majority of them denote bodily (especially oral) actions. Some examples are given below:

/hchəc/	'to whistle'
/sksɛ̄k/	'to devour'
/ɟkɟik/	'to breathe'
/ckwik/	'to talk'
/pkpɔ̄k/	'to smack'
/lkluk/	'to laugh'
/tkjɔk/	'to pluck'
/luʔlɛ̄ʔ/	'to roll'
/bʔbɔʔ/	'to carry on one's back'
/tsdes/	'to stumble'
/lslɔs/	'to gnaw'
/kmɟ̃im/	'to taste'
/ɲnɲin/	'to hop'
/hɲɟaɲ/	'to stand up'
/pɲlɔɲ/	'to sing'
/pŋsɛŋ/	'to say'
/sŋsɛ̄ŋ/	'to walk with a limp'
/kŋjɛŋ/	'to flap wings'
/taŋwoŋ/	'to carry on one's shoulder'
/pŋjuŋ/	'to play a flute'
/krkɛr/	'to yell'

A small set of complex lexemes denote mental states or processes, e.g. the following:

/ʔtʔet/	'to know'
/ɦmhɔm/	'to like'
/jɲjɛɲ/	'to dream'

Occasional examples of stative verbs denoting properties also occur, including the following:

/btʔɛt/	'to be good'
/bclac/	'to be smooth'
/ɟɲɟɛŋ/	'to be wide'

A handful of verbs appear to be derived from nominal roots through an identical process of copying. This type of derivation is marginal and unproductive. Examples include the following:

/ktkit/	'to fart'	/kit/	'buttocks'
/ʔtʔɔ̃t/	'to stroke an animal'	/ʔɔ̃t/	'dog'
/smsɔm/	'to buzz around a nest'	/sɔm/	'nest'
/kŋliŋ/	'to make sound'	/kliŋ/	'sound'
/srsir/	'to move along the side of something'	/sir/	'side'

The lexeme form of a verb is semantically rather neutral and non-specific with regard to e.g. its temporal location (tense) and internal constituency (aspect and Aktionsart). Thus, for example, verbs in their lexeme form may be used to denote situations that are past and bounded as well as present and unbounded. They may also be used to denote future situations if these refer to recurrent, predictable phenomena. Furthermore, lexeme forms of verbs are used in imperative constructions. Some examples of such uses of lexeme forms of verbs are given in the following sentences:

(58) japɛ̃h cɨp ktɔʔ wɛj
1P.EXCL to.go day past
'We went yesterday.'

(59) jɛʔ hgik ba=taɟuʔ
1S to.fear GOAL=snake
'I'm afraid of snakes.'

(60) bolan limɛʔ ʔoʔ hĩc
month five 3S to.rain
'It will rain in May.'

(61) lɔj
to.run
'Run!'

The bare lexeme form of a verb is not used in isolation as a citation form. Instead, the preferred citation form for most verbs is the irrealis-marked lexeme form (cf. §4.7.4.1):

(62) ja=cɨp
IRR=to.go
'[I] will go.'

For other verbs, typically those denoting situations that are usually not associated with a first person subject, the preferred citation form consists of the lexeme combined with a preverbal subject-marking third person singular pronoun:

(63) ʔoʔ kapɨj
3S to.fly
'It flies.'

For some speakers, however, the desiderative-marked verb is the preferred citation form (cf. §4.7.4.2):

(64) ma=cɨp
DES=to.go
'to want to go'

A basic semantic distinction, relevant to several of the aspectual categories described in §4.7.1, needs to be made between *stative* and *dynamic* verbs. Stative verbs inherently denote situations that lack internal change. A subset of stative verbs is made up of adjective-like verbs that denote properties. The syntactic behaviour of these property verbs deviates to some extent from that of other verbs (see §5.1.1). Also, the citation forms of property verbs do not include an irrealis proclitic or preverbal pronoun (cf. above). They are treated as verbs on the basis that they are negatable. Hence Jahai is described here as not having a separate class of adjectives.[20] Some examples of property verbs follow:

/tmkɛt/	'to be cold'
/bkɨt/	'to be hot'
/cnhɔ̃t/	'to be short'
/btec/	'to be long'
/pcɛ̃ʔ/	'to be wet'
/laʔis/	'to be bad'
/pdəh/	'to be near'
/krɨŋ/	'to be dry'
/ghɛl/	'to be tired'
/bəw/	'to be big'
/ʔhəj/	'to be small'

Dynamic verbs, which inherently denote situations that involve some form of change, may be subdivided into several classes, and the main distinctions relevant to the following analysis include *punctual* verbs, which denote instantaneous situations, and *durative* verbs, which denote situations that last for a period of time.

Jahai exhibits a host of affixes and clitics which are added productively to verbs in order to specify their meaning, and the following sections describe the grammatical categories that are marked morphologically on the verb. These include aspect/Aktionsart (§4.7.1), causative (§4.7.2), and verbalisation of nouns and numerals (§4.7.3), which are expressed through derivational affixation, as well as modality (§4.7.4), expressed through postderivational procliticisation. The paradigm of derivational morphology associated with verbal lexemes is exemplified in Table 4.5. Many derived forms may, in turn, feed further derivation.

4.7.1 Aspect and Aktionsart

Jahai exhibits a rich set of derivational affixes which are added to verbs to describe the characteristics of the situation denoted by the verb. Such characteristics often pertain to the internal temporal constituency of the situation and may then be assigned to the grammatical category of aspect (Comrie 1976:3). Four such aspectual distinctions may be identified in Jahai, each of which is represented by a separate morpheme: *imperfective* (§4.7.1.1), *progressive* (§4.7.1.2), *iterative* (§4.7.1.3) and *continuative* (§4.7.1.4). The terminology and definitions of aspects employed here draw on those of Comrie (1976), Dahl (1985), Bybee (1985) and Bybee et al. (1994).

---

20 The treatment of property-signalling words as verbs rather than adjectives finds further support in the fact that they do not typically occur in their underived form as modifiers in the NP.

Table 4.5: Paradigm of verbal derivation in Jahai. (?) = unattested

Derivational morpheme		**Monosyllabic** /cip/ 'to go'	**Sesquisyllabic** /kɟeŋ/ 'to listen'	**Disyllabic** /gulɛm/ 'to carry'
Aspect/Aktionsart (§4.7.1)	Imperfective	cp-cip	k<ŋ>ɟeŋ	gu<m>lɛm
	Progressive[21]	b-cpcip	b-kŋɟeŋ	b-gulɛm
	Iterative	lp-cip	l-kɟeŋ	l-gulɛm
	Continuative	cip-cip	kɟeŋ-kɟeŋ	gulɛm-gulɛm
	Distributive	cip-cip	k<iŋ>ɟeŋ	g<im>lɛm
	Reciprocal	ca-cip	k<a>ɟeŋ	–
	Affix /m/	mp-cip	mŋ-ɟeŋ (?)	mm-lɛm (?)
Causative (§4.7.2)		pp-cip, pi-cip, pr-cip, tr-cip	k<ri>ɟeŋ	g<ri>lɛm
Nominalisation (§4.1.4.1)		np-cip	k<nŋ>ɟeŋ	g<n>ulɛm
Collective plural nomin. (§4.1.4.2)		–	k<ra>ɟeŋ	g<ra>lɛm (?)
Relative (§4.12.2)		t-cip (?)	t-kɟeŋ (?)	t-gulɛm (?)

Other morphemes express characteristics that mainly have to do with what a situation is like with respect to its inherent spatial or participant constituency or the manner in which it takes place. Such non-temporal characteristics will be referred to here unconventionally as Aktionsart, and the distinctions expressed morphologically in Jahai verbs include *distributive* (§4.7.1.5) and *reciprocal* (§4.7.1.6). An additional morpheme, /m/, whose function is unknown, is treated tentatively in this context (§4.7.1.7).

However, it is inappropriate in the case of Jahai to make too sharp a distinction between temporal aspect and non-temporal Aktionsart, because the two categories sometimes merge. Thus, the morpheme signalling progressive aspect frequently also conveys notions of Aktionsart-like characteristics like automaticity, randomness and aimlessness. Indeed, in some cases the main purpose of the progressive morpheme is to give such non-temporal meanings (see §4.7.1.2). Similarly, the continuative morpheme is often used to indicate non-temporal intensity in addition to its basically temporal meaning of extended action (see §4.7.1.4). In this work, therefore, aspect and Aktionsart are not treated as distinctly separate categories and they are discussed under a common heading.

4.7.1.1 Imperfective

Most verbs have a corresponding imperfective form (IMPF).[22] The imperfective morpheme is an affix made up only of underspecified consonants, the positions of which are filled through the morphological process of copy. Allomorphs are determined by base

21 With mono- and sesquisyllabic lexemes, the progressive morpheme is usually added to the derived imperfective form and not to the underived lexeme. See §4.7.1.2.

22 However, verbs whose lexeme forms exhibit morphological complexity in the form of copying (exemplified in §4.7) cannot feed imperfective derivation since they already display a structure identical to that of imperfectives and probably represent 'frozen' imperfectives. A similar situation is described for Semelai by Kruspe (2004: 108–109).

structure. The allomorph associated with monosyllabic bases is represented by a CC prefix consisting of copies of the consonants of the CVC string of the base. The allomorph associated with sesqui- and disyllabic bases consists of a C affix, infixed at the left edge of the final CVC string of the base and filled by a copy of the final consonant of the base.[23]

/sut/	'to sob'	/st-sut/	'to be sobbing'
/kɛc/	'to cut'	/kc-kɛc/	'to be cutting'
/weʔ/	'to exist'	/wʔ-weʔ/	'to be existing'
/ciɸ/	'to hiss'	/cɸ-ciɸ/	'to be hissing'
/pɨs/	'to sweep'	/ps-pɨs/	'to be sweeping'
/sam/	'to hunt'	/sm-sam/	'to be hunting'
/rɨn/	'to sit down'	/rn-rɨn/	'to be sitting down'
/teɲ/	'to plait'	/tɲ-teɲ/	'to be plaiting'
/pɛŋ/	'to chop'	/pŋ-pɛŋ/	'to be chopping'
/dor/	'to balance'	/dr-dor/	'to be balancing'
/ɟaw/	'to wash'	/ɟw-ɟaw/	'to be washing'
/lɔj/	'to run'	/lj-lɔj/	'to be running'
/pkɨp/	'to scorch'	/p<p>kɨp/	'to be scorching'
/kwac/	'to swim'	/k<c>wac/	'to be swimming'
/tʔoc/	'to ask'	/t<c>ʔoc/	'to be asking'
/brik/	'to brake'	/b<k>rik/	'to be braking'
/bliʔ/	'to buy'	/b<ʔ>liʔ/	'to be buying'
/rwis/	'to cut grass'	/r<s>wis/	'to be cutting grass'
/knɔm/	'to urinate'	/k<m>nɔm/	'to be urinating'
/smɛɲ/	'to ask'	/s<ɲ>mɛɲ/	'to be asking'
/kɟeŋ/	'to listen'	/k<ŋ>ɟeŋ/	'to be listening'
/bdɛl/	'to blowpipe'	/b<l>dɛl/	'to be blowpiping'
/cʔaj/	'to sing'	/c<j>ʔaj/	'to be singing'
/krlep/	'to forget'	/kr<p>lep/	'to be forgetting'
/ɟawap/	'to answer'	/ɟ<p>wap/	'to be answering'
/siruc/	'to slurp'	/si<c>ruc/	'to be slurping'
/bakɛs/	'to grow up'	/b<s>kɛs/	'to be growing up'
/turɛh/	'to tap poison'	/tu<h>rɛh/	'to be tapping poison'
/gulɛm/	'to carry'	/gu<m>lɛm/	'to be carrying'
/tigil/	'to go around'	/ti<l>gil/	'to be going around'

The term 'imperfective' is not unproblematic. It is conventionally defined as an aspect that views a situation from within and thereby makes explicit reference to its internal temporal structure (Comrie 1976:24–25; Bybee et al. 1994:125–127). This general definition includes as imperfective forms more specific aspects such as habitual, continuous, progressive and iterative.

Jahai imperfectives have the purpose of describing a situation as prevailing at a particular reference point, and they are used frequently in discourse for setting up background situations to which the main course of events is related. However, they do not appear to have any habitual reading, for example, and it might therefore be argued that they

---

23 For details of processes of copying, relevant to several derivational categories described in this chapter, see §3.2.1.

express an aspect more specific than the conventional imperfective and should therefore be given a different label, such as continuous or progressive (cf. the classification of aspects proposed by Comrie 1976:25). However, continuous/progressive readings are closely associated with another aspectual morpheme, /b-/, here referred to as progressive, which is frequently found in combination with imperfective forms (see §4.7.1.2). The imperfective appears to convey a more vague notion of 'ongoing' which is superordinate to the 'developing' and 'unfolding' nature of progressive forms. This more general meaning motivates the use of the term 'imperfective' in the present work. Still, the difference between imperfective and progressive in Jahai is very subtle and difficult to express formally. A similar stance is taken on aspectual distinctions in Temiar in the more recent works by Benjamin (1996).

The following examples illustrate imperfective aspect:

(65) ʔoʔ k<c>wac ɦajiʔ tɔm taniʔ
3S to.swim<IMPF> also river that
'He was also swimming in that river.'

(66) ʔoʔ tk-tek d=ʔɛj ʔoʔ ton
3S IMPF-to.sleep CONTR=father 3S that
'That father of his was sleeping.'

(67) gin bj-baj k=hajɛ̄ʔ lɛh
2/3P IMPF-dig LOC=house EMP
'They were digging by the house!'

(68) ʄa=jɛʔ cɨp lɛh d=jɛʔ ton, jɛʔ dj-duj
RT=1S to.go EMP CONTR=1S that 1S IMPF-to.hunt
'I [for my part] had already left. I was hunting.'

(69) ʔoʔ ks-kas lɛh ka=knʔac ʔoʔ ton
3S IMPF-to.pinch EMP SUBJ=father.in.law 3S that
'His father-in-law was pinching!'

Imperfective forms feed further derivational morphology in the form progressive (§4.7.1.2), iterative (§4.7.1.3), continuative (§4.7.1.4), nominalisation (§4.1.4.1) and relative (§4.12.2).

4.7.1.2 Progressive

Virtually every verb may be given a progressive meaning by means of a derivational affix involving the phoneme /b/ (PROG), originally borrowed from Malay. It usually occurs in combination with the imperfective form of the verb, especially in the case of mono- and sesquisyllabic bases. The progressive morpheme has two basic allomorphs determined by the structure of the base. With sesqui- and disyllabic bases the affix consists of a simple /b/ prefixed to the left edge of the base.

/cɨp/	'to go'	/b-cpcɨp/	'to be going'
/kɔt/	'to take'	/b-ktkɔt/	'to be taking'
/tek/	'to sleep'	/b-tktek/	'to be sleeping'
/weʔ/	'to exist'	/b-wʔweʔ/	'to be existing'
/caraʔ/	'to talk'	/b-caraʔ/	'to be talking'
/ʔimbus/	'to ambush'	/b-ʔimbus/	'to be ambushing'

/kdɨh/	'to say'	/b-kdɨh/	'to be saying'
/bləh/	'to enter'	/b-bləh/	'to be entering'
/ɟim/	'to cry'	/b-ɟmɟim/	'to be crying'
/dun/	'to cover'	/b-dndun/	'to be covering'
/samoŋ/	'to put together'	/b-samoŋ/	'to be putting together'
/bdɛl/	'to blowpipe'	/b-bldɛl/	'to be blowpiping'
/kanər/	'to carry on shoulder'	/b-kanər/	'to be carrying ...'
/baj/	'to dig'	/b-bjbaj/	'to be digging'

The second allomorph, which occurs rarely, is associated with monosyllabic bases and is made up of a CC prefix consisting of the prespecified /b/ and an underspecified consonant. The latter position is filled by a copy of the final consonant of the base through the process of coda copy.

/tek/	'to sleep'	/bk-tek/	'to be sleeping'
/ʔɛk/	'to give'	/bk-ʔɛk/	'to be giving'
/ɟok/	'to move'	/bk-ɟok/	'to be moving'
/ŋɔk/	'to sit'	/bk-ŋɔk/	'to be sitting'
/cah/	'to cut'	/bh-cah/	'to be cutting'
/ʔɛl/	'to look'	/bl-ʔɛl/	'to be looking'
/lɔj/	'to run'	/bj-lɔj/	'to be running'

The rare occurrence of this allomorph is due to the fact that the progressive morpheme is usually found in combination with imperfective forms, which are minimally disyllabic.

Progressive aspect is typically used in describing situations that prevail at a particular reference point, hence its frequent occurrence in combination with imperfective forms. As noted in §4.7.1.1, the progressive adds to the rather vague imperfective meaning of 'ongoing' a further notion of the situation being in progress and 'developing' or 'unfolding'. In many cases, translating progressive forms into English as 'to be V-ing along' or 'to be V-ing on' is the best way of catching this subtle difference in meaning. In the examples listed above, however, progressives are given the same English translation as imperfectives.

Note that stative verbs sometimes also take the progressive morpheme, implying that stative situations may also be thought of as being in progress. Following Comrie's definitions, such a distribution would occasion the use of the term 'continuous' rather than 'progressive' (Comrie 1976:12, 25). However, too little is known at this point about the use of progressive aspect with stative verbs, such as possible restrictions, for this more general term to be used. Using the term 'progressive' instead of 'continuous' also has the advantage of avoiding confusion of 'continuous' with 'continuative', a form of aspect described in §4.7.1.4.

The following examples illustrate the use of progressive aspect in combination with imperfective forms.

(70) jɛʔ sɜ̃h kɛj kawip, ʔoʔ b-pk-pok
1S to.encounter such bear 3S PROG-IMPF-to.stride
'I saw such a bear! It was striding along.'

(71) heʔ b-ɟk-ɟok ba=pɛ̃w hajɛ̃ʔ
1P.INCL PROG-IMPF-to.move GOAL=other house
'We were moving on to another house.'

(72) jɛʔ ja=b-cp-cɨp ba=ɟlmɔl
1S IRR=PROG-IMPF-to.go GOAL=mountain
'I will be walking along towards the mountain.'

(73) ʔoʔ b-dʔ-deʔ bakɔʔ ka=ʔũn
3S PROG-IMPF-to.make trap LOC=there
'He was setting a trap over there.'

(74) luŋan ʔoʔ b-tk-tek pn=ʔəh
binturong 3S PROG-IMPF-to.sleep EQU=here
'The binturong was sleeping like this.'

(75) pɛ̃h ja=b-b<l>dɛl
1P.EXCL IRR=PROG-to.blowpipe<IMPF>
'We will be blowpiping.'

In addition to its basically temporal meaning, progressive aspect frequently also conveys non-temporal notions of the situation being characterised by randomness, aimlessness, casualness, routine or, sometimes, habituality. What these non-temporal notions appear to have in common is an element of automaticity: the situation does not need a lot of input of energy in order to occur or progress. Thus, in the following examples, the progressive not only signals that the situation is in progress, but also that it is performed in an aimless, casual or routinely fashion:

(76) jɛʔ b-ch-cɨh slaj, japɛ̃h b-tanɛm padɛj
1S PROG-IMPF-to.cut swidden 1P.EXCL PROG-to.plant rice
'I was clearing a swidden. We were planting rice.'

(77) jɛʔ ja=b-rihat lɛh d=jɛʔ, ja=b-tk-tek
1S IRR=PROG-to.rest EMP CONTR=1S IRR=PROG-IMPF-to.sleep
'I [for my part] will be resting. I will be sleeping.'

A more habitual interpretation of the progressive is illustrated in the following example:

(78) japɛ̃h b-ŋk-ŋɔk k=slaj heʔ ʔəh wɛj
1P.EXCL PROG-IMPF-to.sit LOC=swidden 1P.INCL here past
'We used to live in this swidden of ours.'

If combined with punctual verbs in their non-imperfective, lexeme form, the progressive fully takes on these non-temporal notions and loses its temporal meaning altogether. The resulting form signals randomness or casualness, as in the following examples:

(79) lpəs ton japɛ̃h b-ɟumpaʔ paj
after that 1P.EXCL PROG-to.meet 2S.DIST
'After that we ran into you.'

(80) ʔoʔ bk-ʔɛk ba=jɛʔ
3S PROG-to.give GOAL=1S
'He gave [X] to me [in passing].'

Similar non-temporal meanings have been described for the *ber-* prefix of Malay dialects, the source of Jahai /b-/. For example, Asmah (1985:204–206) reports that one of the functions of *ber-* in Kedah peasant Malay is to signal 'action without specificity of purpose or target'. The full range of conceptually related meanings of Malay *ber-* is discussed by Benjamin (1993:371–85).

Occasionally, the element of automaticity in progressive forms appears to result in a reduction of the valency of the verb, producing passive-like forms. These exceptional cases are discussed in §5.4.2.

Progressive forms feed further derivational morphology only in the form of relative (§4.12.2).

4.7.1.3 Iterative

Dynamic verbs are given an iterative meaning by means of a derivational morpheme involving the phoneme /l/ (ITER). This has two allomorphs determined by base structure. The allomorph associated with monosyllabic bases is a CC prefix consisting of the prespecified /l/ and an underspecified consonant. The latter is filled by a copy of the final consonant of the base through the process of coda copy. With sesqui- and disyllabic bases, the affix consists of a simple /l/ prefixed to the left edge of the base.

/cɨp/	'to go'	/lp-cɨp/	'to go repeatedly'
/hɔk/	'to throw'	/lk-hɔk/	'to throw repeatedly'
/pɛl/	'to drip'	/ll-pɛl/	'to drip repeatedly'
/kul/	'to call'	/ll-kul/	'to call repeatedly'
/gɐj/	'to eat'	/lj-gɐj/	'to eat repeatedly'
/saʔot/	'to call someone'	/l-saʔot/	'to call someone repeatedly'
/kdɨh/	'to say'	/l-kdɨh/	'to say repeatedly'
/bdɛl/	'to shoot'	/l-bdɛl/	'to shoot repeatedly'

Defined in the widest possible way, the Jahai iterative signals temporal multiplicity of the event or process designated by the verb. Commonly, this entails a truly iterative interpretation — that is, the multiple repetition of a complete action on a single occasion (cf. Bybee et al. 1994:160, 317). This meaning is particularly apparent with punctual verbs but may also be applied to durative verbs. The following examples illustrate this typically iterative sense:

(81) ʔoʔ lt-sot ʔoʔ lt-but, ʔoʔ sot, ʔoʔ but, ʔoʔ sot, ʔoʔ but
3S ITER-to.cut 3S ITER-to.eat 3S to.cut 3S to.eat 3S to.cut 3S to.eat
'He cut [repeatedly] and ate [repeatedly]. He cut, he ate. He cut, he ate.'
[Of a man who carved up meat and ate the pieces successively]

(82) ʔoʔ ll-ʔɛl k=ɟam
3S ITER-to.look LOC=watch
'He looked at his watch now and then.'

However, the iterative is not limited to actions repeated on a single occasion but may also be used to indicate temporally separated and irregular repetition, which gives more of a frequentative or even habitual reading (cf. Bybee et al. 1994:127, 165, 317). This is illustrated in the following two examples:

(83) pagiʔ-pagiʔ japɛ̄h lr-pɨr
DP-morning 1P.EXCL ITER-to.court
'We courted in the mornings.'

(84) ha=mɔh lj-gɐj k=tə̄h
Q=2S.FAM ITER-to.eat LOC=this
'Do you usually eat here?'
[Uttered by a man who joined the author at a food stall in Grik]

With durative verbs, the iterative form is frequently also used to indicate that the process designated by the verb is not continuous but is taking place intermittently, that is by stages, by instalments, at intervals or the like:

/cɨp/	'to go'	/lp-cɨp/	'to go by stages'
/ʔɛc/	'to defecate'	/lc-ʔɛc/	'to defecate little by little'
/wek/	'to return'	/lk-wek/	'to return by stages'
/ʔɛ̃m/	'to drink'	/lm-ʔɛ̃m/	'to drink little by little'
/baj/	'to dig'	/lj-baj/	'to dig a little at a time'
/ckwɨk/	'to talk'	/l-ckwɨk/	'to talk now and then'
/kɟeŋ/	'to listen'	/l-kɟeŋ/	'to listen now and then'

The following examples illustrate the use of iterative aspect to signal such an intermittent process:

(85) hɛj lc-ʔɛc, bokan hɛj ja=ʔɛc btol
1D.INCL ITER-to.defecate NEG 1D.INCL IRR=to.defecate straight
'We defecated little by little. We didn't defecate in one go.'

(86) ɟa=jɛʔ lp-cɨp lɛh ba=ʔə̃h
RT=1S ITER-to.go EMP to=here
'Then I walked here [bit by bit].'

Iterative forms derived from monosyllabic bases feed further derivational morphology in the form of progressive (§4.7.1.2), continuative (§4.7.1.4), distributive (§4.7.1.5), nominalisation (§4.1.4.1) and relative (§4.12.2). Iterative forms derived from sesqui- and disyllabic bases possibly only feed further derivation in the form of relative.

Productive use of iterative /l/ in verbs has not been described for other Aslian languages. However, fossilised traces of an /l/ morpheme indicating intensity or great numbers have been identified in Jah Hut expressives (Diffloth 1976b:84). Also, many animal names in several Aslian languages contain traces of an infixed /l/ which is suggested by Diffloth (1976b:100–101) to be linked to the erratic, 'step by step' movements of these animals (see also Kruspe 2004:86). For Jahai examples, see §4.1.

4.7.1.4 Continuative

Verbs are given a continuative and intensive interpretation by means of total reduplication. This will be referred to here as continuative aspect (CONT). The process usually applies to lexeme forms of verbs:

/cɛp/	'to catch'	/cɛp-cɛp/	'to keep on catching'
/get/	'to cut'	/get-get/	'to keep on cutting'
/muc/	'to eat'	/muc-muc/	'to keep on eating'
/tek/	'to sleep'	/tek-tek/	'to keep on sleeping'
/ŋɔk/	'to sit'	/ŋɔk-ŋɔk/	'to keep on sitting'
/ɟampiʔ/	'to heal'	/ɟampiʔ-ɟampiʔ/	'to keep on healing'
/kwɛ̃s/	'to sweep'	/kwɛ̃s-kwɛ̃s/	'to keep on sweeping'
/bləh/	'to enter'	/bləh-bləh/	'to keep on entering'
/ɟim/	'to cry'	/ɟim-ɟim/	'to keep on crying'
/tanɛm/	'to plant'	/tanɛm-tanɛm/	'to keep on planting'

/pɲlɔɲ/	'to sing'	/pɲlɔɲ-pɲlɔɲ/	'to keep on singing'
/kɟeŋ/	'to listen'	/kɟeŋ-kɟeŋ/	'to keep on listening'
/bdɛl/	'to blowpipe'	/bdɛl-bdɛl/	'to keep on blowpiping'
/ʔɛl/	'to look'	/ʔɛl-ʔɛl/	'to keep on looking'
/gej/	'to eat'	/gej-gej/	'to keep on eating'
/pɔj/	'to dry'	/pɔj-pɔj/	'to keep on drying'

It may also involve reduplication of derived forms such as imperfectives, iteratives and distributives. However, progressive forms do not reduplicate to form continuatives.

/ɟiʔ/	'to burn'	/ɟʔ-ɟiʔ-ɟʔ-ɟiʔ/	'to be keeping on burning'
/kwẽs/	'to sweep'	/k<s>wẽs-k<s>wẽs/	'to be keeping on sweeping'
/cɔl/	'to tell'	/cil-cɔl-cil-cɔl/	'to keep on telling'

Continuative forms feed further derivational morphology only in the form of progressive (§4.7.1.2), as illustrated by the following examples:

/ɟumpaʔ/	'to meet'	/b-ɟumpaʔ-ɟumpaʔ/	'to be keeping on meeting'
/kdɨh/	'to say'	/b-kdɨh-kdɨh/	'to be keeping on saying'

Continuative aspect is used to signal that a situation is kept ongoing for an extended period of time, often with the additional notion that it is done to completion. It may also carry non-temporal meanings of intensity and commitment. The following examples illustrate the use of continuative aspect:

(87) ʔoʔ gej-gej nasiʔ ton
3S CONT-to.eat rice that
'He kept on eating that rice [until he was finished].'

(88) ʔoʔ ʔɛl-ʔɛl ka=ʔap cn=ʔaniʔ
3S CONT-to.look SUBJ=tiger from=there
'The tiger kept on watching from over there.'

(89) wih pɔj sɛc ton, sɛc kasaʔ ton, wih pɔj-pɔj kriŋ
3D to.dry meat that meat sambar.deer that 3D CONT-to.dry dry
'They dried that meat, that sambar meat. They kept on drying until it was dry.'

(90) gin k<s>wẽs-k<s>wẽs
2/3P CONT-to.sweep<IMPF>
'They were keeping on sweeping.'

(91) jɛʔ ŋɔk k=ton lɛh, jɛʔ ŋɔk-ŋɔk k=ton
1S stay LOC=that EMP 1S stay-CONT LOC=that
'I stayed there! I kept staying there.'

Reduplicated property verbs do not carry the temporal notion of continuation but signal only intensity:

/lajin/	'to be different'	/lajin-lajin/	'to be very different'
/ʔhəj/	'to be small'	/ʔhəj-ʔhəj/	'to be very small'

Continuative forms are related to, but should not be confused with, the multiple repetition of verbs that occurs commonly in Jahai as a discourse strategy to signal continuation and extension of an action. Continuative forms, which always involve a

single reduplication, are distinguished from such purely discourse-motivated repetition by stress patterns.

The Jahai continuative has meanings similar to those of reduplicated verbs in Malay (see e.g. Zaharani 1991:114–115 for Perak Malay), and the process of total reduplication is likely to be a borrowing from Malay (cf. §3.2.4 and §4.1.2).

4.7.1.5 Distributive

Most verbs may receive a derivational morpheme indicating that the state or action designated by the verb is characterised by non-temporal multiplicity; that is, it signals distribution of action/state over several locations and/or participants. This will be referred to here as *distributive Aktionsart* (DISTR). The morpheme is constructed by means of copy (according to the same patterns as those encountered with the imperfective forms described in §4.7.1.1) and a prespecified vowel /i/.

/cɨp/	'to go'	/cip-cɨp/	'to go [here and there]'
/lwec/	'to climb up'	/l<ic>wec/[24]	'to climb up [here and there]'
/ʔɛk/	'to give'	/ʔik-ʔɛk/	'to give [here and there]'
/ɟok/	'to move'	/ɟik-ɟok/	'to move [here and there]'
/ŋɔk/	'to sit'	/ŋik-ŋɔk/	'to sit [here and there]'
/sapuh/	'to sweep'	/s<ih>puh/	'to sweep [here and there]'
/boh/	'to put'	/bih-boh/	'to put [here and there]'
/tbɔh/	'to hit'	/t<ih>bɔh/	'to hit [here and there]'
/tanɛm/	'to plant'	/t<im>nɛm/	'to plant [here and there]'
/kɔm/	'to be many'	/kim-kɔm/	'to be many [here and there]'
/kriŋ/	'to be dry'	/k<iŋ>riŋ/	'to be dry [here and there]'
/kɟeŋ/	'to listen'	/k<iŋ>ɟeŋ/	'to listen [here and there]'
/bdil/	'to shoot'	/b<il>dil/	'to shoot [here and there]'
/ʔɛl/	'to look'	/ʔil-ʔɛl/	'to look [here and there]'
/jol/	'to throw'	/jil-jol/	'to throw [here and there]'
/cɔl/	'to tell'	/cil-cɔl/	'to tell [here and there]'
/sjər/	'to swim'	/s<ir>jər/	'to swim [here and there]'
/bəw/	'to be big'	/biw-bəw/	'to be big [here and there]'

Distributive Aktionsart may be used to denote an action carried out by, or a state characteristic of, more than one individual, location or object, typically more or less simultaneously. Indeed, distributive forms usually have a non-singular subject. However, it is not number of subject as such that is in focus (distributive is not obligatory with plural subjects), but rather the fact that the action or state is saliently distributed over several, often spatially distinct, subjects. The following examples illustrate this meaning:

(92) he? ja=ɟik-ɟok haden
1P.INCL IRR=DISTR-to.move tomorrow
'We will move [in different directions] tomorrow.'

---

[24] Note that all forms ending in a palatal are homophonic with the corresponding imperfective form, the epenthetic penultimate nucleus of which is realised as [i] before the palatal coda (see the rules of epenthesis described in §2.4.4.2): /l<c>wec/ [lic'wɛc'].

(93) wih rih-rɔh slaj
3D DISTR-to.clear swidden
'They cleared [different parts of] a swidden.'

(94) kritəh-kritəh japɛ̃h biw-bəw
DP-car 1P.EXCL DISTR-to.be.big
'Our cars are big.'

Distributive is also used to express that a singular subject is involved in an action, or characterised by a state, which is distributed over multiple locations or multiple participants acted upon. Thus, if combined with a singular subject, a distributive form like /ʔil-ʔɛl/ (from /ʔɛl/ 'to look') suggests that the subject either looks at something from different locations, or looks in different directions from the same location. The following examples illustrate distributive with a singular subject:

(95) mnraʔ ʔoʔ hit-hūt can=ɟlmɔl
person 3S DISTR-to.reveal.oneself SOURCE=mountain
'The person revealed himself [here and there] on the mountain.'

(96) ʔoʔ ŋik-ŋɔk pn=ʔə̃h
3S DISTR-to.sit EQU=here
'He sat down [here and there] like this.'

Distributive Aktionsart typically implies simultaneity, but note that in some cases (especially with singular subjects) spatial/participant multiplicity necessarily also entails temporal multiplicity. Again, however, focus is on non-temporal distribution. Focus on temporal multiplicity is reserved for the iterative aspect (§4.7.1.3), which may feed distributive derivation to signal that an action is distributed over several locations and/or participants at multiple points in time:

(97) ʔoʔ lil-ʔɛl can=tkih
3S ITER.DISTR-to.look SOURCE=backside
'He looked [around now and then] from behind.'

For some speakers, the distributive form of the verb is preferred in citation forms of verbs that denote actions that are inherently distributed in character. Examples include the following:

/hip-hap/ 'to disperse'
/sih-sə̃h/ 'to meet [each other]'
/ʔim-ʔəm/ 'to embrace [each other]'
/ʔ<iw>ŋiw/ 'to gaze around'

The use of the close front vowel /i/ as penultimate nucleus in distributive verb forms often creates a marked acoustic contrast to the final syllable vowel (except of course in those cases where the latter is represented by the same /i/ and perhaps also the adjacent /e/), and it might be suggested that this opposing variation in quality between nuclei is related iconically to the notion of spatial variation expressed by the verb form. A marginal parallel may be found in English forms like *chit-chat*, *flip-flop*, *zig-zag*, *mishmash* and *tick-tack*, where identical vowel alternation possibly reflects similar notions of spatial alternation. In his discussion of Semai expressives, Diffloth (1976d:254) refers to such alternation as 'antiphonic reduplication'.

Distributive forms feed further derivational morphology in the form of progressive (§4.7.1.2), iterative (§4.7.1.3), continuative (§4.7.1.4), nominalisation (§4.1.4.1) and possibly relative (§4.12.2).

4.7.1.6 Reciprocal

Mono- and sesquisyllabic verbs may receive a rarely occurring, marginal but possibly productive derivational morpheme involving the vowel /a/ (REC). This morpheme has two allomorphs, determined by base structure. With monosyllabic bases, a CV prefix is attached to the base, consisting of the prespecified /a/ and an underspecified consonant filled by a copy of the initial consonant of the base. In sesquisyllabic bases, the /a/ is infixed at the left edge of the final CVC string of the base.

Monosyllabic:	/cip/	'to go'	/ca-cip/	'to go together'
	/cɔl/	'to tell'	/ca-cɔl/	'to tell each other'
	/gej/	'to eat'	/ga-gej/	'to eat together'
Sesquisyllabic:	/smɛɲ/	'to ask'	/s<a>mɛɲ/	'to ask each other'
	/bdil/	'to shoot'	/b<a>dil/	'to shoot at each other'
	/ʔnaj/	'to bathe'	/ʔ<a>naj/	'to bathe each other'

The function of this morpheme is not known in detail, but its distribution tentatively suggests that it signals some form of 'reciprocal' or 'joint' action. Stative verbs appear not to take the /a/ affix. Identical processes of affixation of /a/ have been described for Central Aslian languages, notably Semai (Diffloth 1976a:238–239), Temiar (Benjamin 1976b:172–173; 1996) and Lanoh (Benjamin 1996). In Semai it is described as marking resultative, and in Temiar it signals the middle voice of the verb. It is possible that the /a/ in Jahai has a similar valency-related use, but, due to its rare occurrence, nothing is known of its syntactic properties and therefore no further conclusions will be drawn here.

Reciprocal forms feed further derivational morphology in the form of progressive (§4.7.1.2), iterative (§4.7.1.3), nominalisation (§4.1.4.1) and possibly relative (§4.12.2).

4.7.1.7 Affix /m/

A little-known process, described here tentatively and uncertainly in connection with aspect and Aktionsart, involves the rare affixation of /m/ to dynamic verbs (M). With monosyllabic bases, a CC prefix is added, consisting of the prespecified /m/ and an underspecified consonant. The latter position is filled by a copy of the final consonant of the base through the process of coda copy. With sesqui- and disyllabic bases, an identical CC prefix replaces base segments to the left of the final CVC string, a process which has not been noticed elsewhere. Notably, the underspecified vowel nucleus of the affix is not always realised according to the rules of epenthesis outlined in §2.4.4.2. Thus, if the copied consonant is glottal, the epenthetic vowel is often represented by [ə] instead of [a]. In two respects, then, affixation of /m/ deviates from related morphological processes.

The use of the /m/ affix is restricted to isolated, occasional narratives, but may be very frequent within those narratives, sometimes being attached to almost every verb. Its distribution has not yet provided any clues as to its exact function, and speakers, claiming /m/-affixed forms to be obsolete, are unable to elaborate on its meaning. Features of the narrative delivery, such as articulation and gestures, occasionally suggest that it signals that an action is carried out with particular dedication and zest, but this interpretation is far

from certain. The fact that speakers claim the affix to be obsolete possibly also suggests that it represents an archaic morpheme which is currently used to add a flavour of ancientness and importance to a story. Its deviant morphophonemic characteristics may support such an interpretation. Still, it appears to be largely productive within its restricted domain. The following forms have been recorded:

/mp-cɨp/	/cɨp/	'to go'
/mt-get/	/get/	'to cut'
/mk-wek/	/wek/	'to go back'
/mʔ-ɟiʔ/	/ɟiʔ/	'to burn'
/mʔ-ʔiʔ/	/cʔiʔ/	'to pour'
/mʔ-bɔʔ/	/bʔbɔʔ/	'to carry on one's back'
/ms-dɛs/	/dɛs/	'to move'
/ms-lɛs/	/plɛs/	'to smear poison on blowpipe dart'
/ms-gəs/	/gəs/	'to carve'
/ms-was/	/was/	'to split'
/mh-ʔəh/	/pʔəh/	'to heat'
/mh-boh/	/boh/	'to put'
/mh-cɔh/	/cɔh/	'to bite'
/mm-ʔɛ̃m/	/ʔɛ̃m/	'to drink'
/mŋ-ɟɨŋ/	/ɟɨŋ/	'to take'
/ml-pel/	/pel/	'to cook'
/ml-dɛl/	/bdɛl/	'to shoot'
/ml-ʔɛl/	/ʔɛl/	'to look'
/ml-pal/	/kipal/	[meaning unknown]
/ml-gɔl/	/gɔl/	'to carry'
/mr-ʔɔr/	/ʔɔr/	'to order'
/mj-gej/	/gej/	'to eat'
/mj-haj/	/haj/	'to follow'

The following examples illustrate its use:

(98) jɛʔ mh-cɔh ʔoʔ
1S M-to.bite 3S
'I bit him.'

(99) ɟa=jɛʔ ms-was
RT=1S M-to.split
'Then I split [X].'

/m/-affixed forms feed further derivational morphology in the form of progressive (§4.7.1.2), as shown by the following example: /b-mr-ʔɔr/ from /ʔɔr/ 'to order'.

4.7.2 Causatives

Jahai exhibits a variety of affixes which, when added to typically intransitive verb roots, turn the experiencing subject of the base verb into a direct object representing a causee, and enable the introduction of a new subject argument representing the causer (see §5.4.1). Affixes include /p-/, /pi-/, /pr-/, /tr-/ and /<ri>/ and variations of these, and the verb forms

which they derive are grouped together here as causatives. All of the affixes are glossed as CAUS.

The rich set of causativising affixes can to some extent be explained as allomorphy determined by either phonotactics or semantics; in some cases, however, affixes appear to occur in free variation. It has not been possible, for example, to determine whether there are clear semantic differences between /p-/, /pi-/, /pr-/ and /tr-/, all of which are associated with monosyllabic bases; some bases are able to take any of these four prefixes, which would seem to suggest some semantic distinction, e.g. /pp-cɨp/, /pi-cɨp/, /pr-cɨp/ and /tr-cɨp/ from /cɨp/ 'to go'. It is possible that /pr-/ and /tr-/ in such a case would represent more productive and semantically more general types of causativisation. On the whole, this varied system of causativising affixes seems to suggest that causative processes form a rather unstable and dynamic part of Jahai grammar.

4.7.2.1 Causative /p-/

Infrequently, causatives are formed through the prefixation of /p-/ to the base. With sesqui- and disyllabic bases, this process has been attested only occasionally, e.g. /p-hgik/ 'to frighten' from /hgik/ 'to be afraid', the main causativising strategy for such bases being infixation of /<ri>/ (see §4.7.2.5). With monosyllabic bases, the allomorph consists of a CC prefix composed of the prespecified /p/ and an underspecified consonant. The latter position is filled by a copy of the final consonant of the base. The following forms have been attested:

/cɨp/	'to go'	/pp-cɨp/	'to help a child walk'
/ʔɛk/	'to give'	/pk-ʔɛk/	'to make someone give'
/ɟɨʔ/	'to burn'	/pʔ-ɟɨʔ/	'to make something burn'
/cɔl/	'to tell'	/pl-cɔl/	'to teach someone to speak'
/gej/	'to eat'	/pj-gej/	'to feed'

4.7.2.2 Causative /pi-/

More frequently, causatives are formed through prefixation of a fully prespecified /pi-/. This process is usually limited to monosyllabic bases, but occasional examples have been found in combination with sesqui- and disyllabic bases. In such cases, the initial consonant of the base is always a glottal, e.g. /pi-ʔnaj/ 'to wash' from /ʔnaj/ 'to bathe'.

Causatives formed by means of /pi-/ often have rather specific and restricted meanings, and the process appears to be restricted to a limited set of roots, which suggests that this type of causativisation is not fully productive. Also, its valence-increasing properties are not always apparent, as the derived form sometimes does not carry a meaning very different from that of the base. Examples of /pi-/ in combination with monosyllabic bases include the following:

/cɨp/	'to go'	/pi-cɨp/	'to help a child walk'
/kap/	'to bite'	/pi-kap/	'to tear apart with one's teeth'
/muc/	'to eat'	/pi-muc/	'to feed'
/tek/	'to sleep'	/pi-tek/	'to lull'
/wek/	'to go back'	/pi-wek/	'to lift'
/gɨs/	'to descend'	/pi-gɨs/	'to pour'

/boh/	'to put'	/pi-boh/	'to put'
/ʔɛ̃m/	'to drink'	/pi-ʔɛ̃m/	'to suckle'
/tɨl/	'to touch'	/pi-tɨl/	'to place/put'
/dul/	'to hide oneself'	/pi-dul/	'to hide something'
/haj/	'to follow'	/pi-haj/	'to make someone follow'

The following examples illustrate the use of causative /pi-/:

(100) ʔoʔ pi-ʔɛ̃m d=wɔŋ ʔoʔ
3S CAUS-to.drink CONTR=child 3S
'She suckled her baby.'

(101) ɟa=ʔoʔ pi-gɨs tɔm
RT=3S CAUS-to.descend water
'Then he poured out the water.'

(102) ʔoʔ pi-wek haɦiʔ ʔoʔ tə̃h
3S CAUS-to.go.back tail 3S this
'He lifted its tail.'

(103) gin pi-ʔnaj pɛ̃w bulɔʔ, pɛ̃w bulɔʔ, pɛ̃w
2/3P CAUS-to.bathe other bamboo.tube other bamboo.tube other
bulɔʔ
bamboo.tube
'They washed one bamboo tube after another.'

Causatives derived by means of the /pi-/ affix feed further derivational morphology in the form of iterative (§4.7.1.3) and possibly imperfective (§4.7.1.1), progressive (§4.7.1.2), continuative (§4.7.1.4) and relative (§4.12.2).

4.7.2.3 Causative /pr-/

Another causative affix consists of a prespecified affix /pr-/, which like /pi-/ is prefixed to monosyllabic bases. Examples of /pr-/ include the following:

/cɨp/	'to go'	/pr-cɨp/	'to cause something to move'
/hit/	'to tremble'	/pr-hit/	'to cause someone to tremble'
/soc/	'to wash'	/pr-soc/	'to cause someone to wash'
/wek/	'to go back'	/pr-wek/	'to cause someone to go back'
/jol/	'to throw'	/pr-jol/	'to cause someone to throw'
/hir/	'to be frightened'	/pr-hir/	'to frighten'
/gej/	'to eat'	/pr-gej/	'to feed'
/lɔj/	'to run'	/pr-lɔj/	'to cause something to run away'

The following examples illustrate the use of causative /pr-/:

(104) ʔoʔ pr-gej ba=kneh ʔoʔ ton
3S CAUS-to.eat GOAL=wife 3S that
'He supported that wife of his.'

(105) wa=b-pr-hit gin b<ra>boʔ ʔon
IRR.3S=PROG-CAUS-tremble 2/3P woman<COLL> there
'He was going to make those women tremble [with fear].'

Causatives derived by means of the /pr-/ affix feed further derivational morphology in the form of imperfective (§4.7.1.1), progressive (§4.7.1.2), iterative (§4.7.1.3), continuative (§4.7.1.4), nominalisation (§4.1.4.1) and possibly relative (§4.12.2).

4.7.2.4 Causative /tr-/

Yet another causative affix which is prefixed to monosyllabic bases involves a prespecified /tr-/. Again, it is difficult to determine whether this causativising affix is semantically different from the others, but some of the translations seem to indicate that causatives derived with /tr-/ are more likely to involve causation through intentional, often oral instruction. A similar but more evident distinction has been noted in Semelai by Kruspe (2004:127–129), who proposes that /tar-/ (the Semelai equivalent of Jahai /tr-/) signals mediated causation. The following Jahai examples illustrate such instructional causation:

/cɨp/	'to go'	/tr-cɨp/	'to train someone to walk'
/gɨm/	'to deliberate'	/tr-gɨm/	'to summon to a deliberation'
/lɔj/	'to run'	/tr-lɔj/	'to tell someone to run away'

Causatives derived by means of the /tr-/ affix feed further derivational morphology in the form of imperfective (§4.7.1.1), progressive (§4.7.1.2), iterative (§4.7.1.3), continuative (§4.7.1.4), nominalisation (§4.1.4.1) and possibly relative (§4.12.2).

4.7.2.5 Causative /<ri>/

The causative allomorph /<ri>/ occurs as an infix in sesqui- and disyllabic bases and is wholly determined by base structure. It consists of a fully prespecified affix attached at the left edge of the final CVC string of the base. It will be regarded here as a phonotactically determined allomorph of all the causative morphemes described above. Examples include the following:

/ckɨp/	'to close one's eyes'	/c<ri>kɨp/	'to close someone's eyes'
/bkɨt/	'to be hot'	/b<ri>kɨt/	'to heat'
/hgɨk/	'to be afraid'	/h<ri>gɨk/	'to frighten'
/sɟuʔ/	'to be cold'	/s<ri>ɟuʔ/	'to cool something'
/kbɨs/	'to die/be dead'	/k<ri>bɨs/	'to kill'
/manɛh/	'to be old'	/m<ri>nɛh/	'to make something old'
/pɑah/	'to snap/break'	/p<ri>ɑah/	'to break something'
/btɔj/	'to be red'	/b<ri>tɔj/	'to make something red'

The following examples illustrate the /<ri>/ allomorph:

(106) ɟa=jɛʔ k◀ri>bɨs ʔoʔ
RT=1S to.die<CAUS> 3S
'Then I killed it.'

(107) heʔ ja=p<ri>ɑah tawɛs
1P.INCL IRR=to.break<CAUS> [type.of.tree]
'We will fell a *tawas* tree.'

Causatives derived by means of the /<ri>/ affix feed further derivational morphology in the form of imperfective (§4.7.1.1), progressive (§4.7.1.2), iterative (§4.7.1.3), continuative (§4.7.1.4) and possibly relative (§4.12.2).

4.7.3 Noun/numeral-to-verb derivation

4.7.3.1 Noun-to-verb derivation

Nouns may be turned into intransitive property verbs by means of the affix /b-/ (PROP), related to the progressive morpheme described in §4.7.1.2. The affix usually consists of a /b-/ prefixed to the nominal base, but with some bases this is in free variation with a full syllable allomorph /br-/. The process, which appears to be productive but not very frequent, results in verbs that denote utilisation, possession or containment of the noun in question. What these meanings have in common is a notion of 'property' (hence the glossing of the morpheme as PROP) which is likely to be closely linked to the notions of automaticity, routine and habituality that are associated with the /b-/ morpheme signalling progressive aspect in verbs. Similar denominal processes characterise the Malay prefix *ber-*, the source of Jahai /b-/. Zaharani (1991:87–88), for example, provides an almost identical semantic characterisation of *ber-* affixed to nominal roots in Perak Malay. Similar functions are also described for Semelai (Kruspe 2004:153–155).

/teʔ/	'earth/soil'	/b-teʔ ~ br-teʔ/	'to own/till land'
/hajɛ̃ʔ/	'house'	/b-hajɛ̃ʔ/	'to reside in a house'
/ʔikəʔ/	'fish'	/b-ʔikəʔ/	'to have fish'
/kritəh/	'car'	/b-kritəh/	'to be motorised'
/tɔm/	'water'	/b-tɔm/	'to have/use/contain water'
/hawɛ̃n/	'pig'	/b-hawɛ̃n/	'to have a pig/pork'
/twiɲ/	'headache'	/b-twiɲ/	'to have a headache'
/blaw/	'blowpipe'	/b-blaw/	'to have/use a blowpipe'

The following examples illustrate its use:

(108) ʔoʔ b-kneh
3S PROP-wife
'He was married.'

(109) ʔoʔ b-slaj ka=ʔũn
3S PROP-swidden LOC=there
'He tilled a swidden there.'

(110) braʔ wa=b-nasiʔ
NEG IRR.3S=PROP-rice
'He doesn't have any rice.'

Such noun-to-verb derivations possibly only feed further derivation in the form of relative (§4.12.2).

4.7.3.2 Numeral-to-verb derivation

Similarly, numerals borrowed from Malay may be verbalised by the same /b-/ affix (PROP). The resulting form means 'to be characterised by' the number designated by the

numeral. The only indigenous Jahai numeral /nɛj/ 'one' appears not to take this affix. Again, close parallels are found in Semelai (Kruspe 2004:152) and Perak Malay (Zaharani 1991:88–89).

/duwaʔ/	'two'	/b-duwaʔ/	'to be two'
/tigaʔ/	'three'	/b-tigaʔ/	'to be three'
/ʔmpat/	'four'	/b-ʔmpat/	'to be four'
/limɛʔ/	'five'	/b-limɛʔ/	'to be five'

4.7.4 Markers of modality

Jahai distinguishes at least three types of modality by means of proclitics attached to the verb: *irrealis*, *desiderative* and *hortative*. The domain of attachment of these proclitics is the clause, but their phonological host is always a verb; hence their treatment here in connection with verbal morphology. Following the definitions of Bybee (1985:165–169) and Bybee et al. (1994:176ff., 319ff.), irrealis and desiderative are mainly to be regarded as agent-oriented modalities; that is, they express internal or external conditions pertaining to the agent with respect to the predicate situation. Hortative, on the other hand, which is a type of directive or imperative, represents a speaker-oriented modality; that is, it expresses the speaker's imposition of conditions on the addressee.

The use of the term 'irrealis' calls for some clarification. Irrealis may be defined as expressing situations which are unreal or non-actual in some sense (Bybee et al. 1994:236–237). Strictly speaking, the three modal distinctions in Jahai may all be categorised as irrealis, since they all express non-actual situations. It may even be argued that the presence of modal proclitics distinguishes irrealis situations from realis ones. The reason for restricting the term irrealis to just one of the modal proclitics (/ja=/ ~ /wa=/, described in §4.7.4.1) is that this has a number of typical irrealis meanings that are difficult to bring together under a different label.

It should also be pointed out that it is difficult in the case of irrealis to make a sharp distinction between modality, tense and aspect. This is because the Jahai irrealis frequently conveys notions of both future and present as well as ongoing of a situation (see §4.7.4.1).

4.7.4.1 Irrealis

The irrealis morpheme has two allomorphs: /wa=/, which is used with subjects in third person singular (IRR.3S), and /ja=/, which is used with all other subjects (IRR). These portmanteau clitics replace the otherwise usually obligatory preverbal subject-marking pronouns, and, like these, are regarded as subject-agreement markers (§5.1.1.1). A disambiguating pronoun or full NP denoting the subject may precede the irrealis marker but is syntactically optional. In the absence of a disambiguating pronoun or full NP, the /ja=/ allomorph usually implies first person, typically in the singular.

Irrealis is used in a variety of contexts which are connected by the notion that the situation expressed by the clause is not a reality. Thus, irrealis may be used to indicate that a situation will happen at some point in the future (immediate or distant), either in relation to the time of speaking or to some other time of reference in the past. This tense-like, future sense of irrealis is also intimately linked to notions of intention and expectation. The following examples illustrate such future meanings:

(111) ʔnuj ja=cɨp, ja=sam ba=ʔə̄h, ja=bdil gaw, brɨj tə̄h
soon IRR=to.go IRR=to.hunt GOAL=here IRR=to.shoot pig evening this
'I will leave soon. I will hunt nearby. I will shoot a pig. Tonight.'

(112) heʔ ja=gej, mɔh pon, heʔ ja=gej skaliʔ
1P.INCL IRR=to.eat 2S.FAM too 1P.INCL IRR=to.eat together
'We will eat. You too. We will eat together.'

(113) jɛh duwaʔ haden ba=ʔūn, jɛh ja=b-dkdak
1D.EXCL two tomorrow GOAL=there 1D.EXCL IRR=PROP-fish.trap
'Tomorrow we will go there. We will trap fish.'

(114) wa=ll-cɔl sʔoʔ, ja=dɛs
IRR.3S=ITER-to.tell just IRR=to.move
'He will just go on telling stories. I will leave.'

(115) wa=pi-ʔɛ̆m wɔŋ ʔoʔ
IRR.3S=CAUS-to.drink child 3S
'She will suckle her baby.'

(116) mɛj wa=sam
what IRR.3S=to.hunt
'What will he hunt?'

Irrealis is also used to denote ongoing present situations where the predicate involves some type of goal that has not yet been reached and where the situation has therefore not been fully realised. The goal may be expressed by the inherent semantics of the verb and is then especially evident in verbs of achievement, as in the following example, where the verb /wek/ 'to go back ~ to return' refers to the whole process of going back and reaching one's starting-point:

(117) ja=wek
IRR=to.return
'I'm on my way back!'
[Frequently uttered by people who, walking through the village,
are about to reach their house after returning from the forest]

The goal may also be expressed overtly, e.g. as a PP, as in the following case, where the situation is not considered to be completed until one has reached the stated endpoint:

(118) ja=cɨp ba=hɨp
IRR=to.go GOAL=forest
'I'm going to the forest!'
[Frequently uttered by people who are leaving the village]

Such present readings of the irrealis take on almost aspect-like properties in that they express that a situation is ongoing and will be completed at some point.

Sometimes, irrealis is also used to indicate the desire or need of the subject to realise a situation, as in the following example:

(119) ja=knɔm
IRR=to.urinate
'I've got to do a wee-wee!'
[Uttered by a man who was looking in vain for a secluded place to pass water]

Finally, irrealis is obligatory in negated clauses. No distinction is made between negation of past, present and future situations. All display the same construction consisting minimally of the negative marker /braʔ/ (which is in free variation with the borrowed Malay negative marker /bokan/) and the irrealis-marked verb.

(120) c<n>kwɨk ʔoʔ ton lajin-lajin, braʔ ja=ʔtʔet
to.speak<NM> 3S that CONT-to.be.different NEG IRR=to.know
'That way of speaking of his is very different. I don't understand [it].'

(121) braʔ wa=muŋkɛr lagiʔ ka=ʔap ton
NEG IRR.3S=to.wake.up again SUBJ=tiger that
'That tiger never woke up again!'

(122) braʔ wa=lr-gɨr ka=ʔūn
NEG IRR.3S=ITER-to.roll LOC=there
'There was no thunder.'

(123) braʔ gin ja=wek
NEG 2/3P IRR=to.go.back
'They won't come back.'

4.7.4.2 Desiderative

The desiderative proclitic /ma=/ (DES), which expresses the desire, wish or intention of the subject to perform the action designated by the verb, is exceedingly rare and hardly ever occurs in spontaneous speech. For some speakers, however, the desiderative-marked verb is the preferred citation form.

/cɨp/	'to go'	/ma=cɨp/	'to want to go'
/lmah/	'to find'	/ma=lmah/	'to want to find'
/tigil/	'to go around'	/ma=tigil/	'to want to go around'

Like the irrealis proclitics, the desiderative appears to be mutually exclusive with the otherwise usually obligatory preverbal subject-marking pronoun.

The desiderative may be combined with all pronouns and hence any person. However, it usually occurs only with animate, typically human, subjects.

(124) japɛ̄h ma=pimɨc baraŋ japɛ̄h
1P.EXCL DES=to.take.back thing 1P.EXCL
'We want to take back what is ours.'

In the following sentence, the subject is represented by a natural phenomenon which is in some way animated:

(125) ktɔʔ ma=hĩc
sky DES=to.rain
'It will rain.' (lit. 'The sky wants to rain.')

A similar desiderative prefix /maʔ-/ is described for Kintaq by Asmah (1976:955–956), and a proclitic /ma=/ signals irrealis in Semelai (Kruspe 2004:89). Similarly, an affix /-m-/, associated with pronominal proclitics and expressing modal categories like intentional, imperative and purposive, has been described for Temiar by Benjamin (1976b:180–182; 1996).

4.7.4.3 Hortative

Three modals, /ha=/, /ca=/ and /ka=/, are procliticised to verbs to propose an addressee to carry out an action. Their domain of attachment is the imperative clause. They express meanings that range from mild commands, over incitements and encouragements, to humble suggestions. It has not yet been possible to discern any distributional differences between /ha=/, /ca=/ and /ka=/, and they are referred to here under the common label of hortative (HORT). They are distinct from true imperatives, which are used to issue direct commands and which are expressed with the verb lexeme only (cf. §4.7). Presumably, the hortative proclitic /ha=/ bears some semantic relation to the interrogative proclitic /ha=/, described in §4.10.3.

/cɨp/	'to go'	/ha=cɨp/	'please go'
/wek/	'to go back'	/ha=wek/	'please go/come back'
/kɟeŋ/	'to listen'	/ka=kɟeŋ/	'please listen'
/ʔɛl/	'to look'	/ca=ʔɛl/	'please look'

Examples include the following:

(126) ha=cɨp ba=ʔəh
HORT=to.go GOAL=here
'Please come here!'

(127) wek ha=ʔɛl
to.go.back HORT=to.look
'Come back and have a look!'

(128) ca=ʔɛl gamah ʔoʔ
HORT=to.look picture 3S
'Have a look at his photos!'

4.8 Expressives

Expressives, a category of words which forms a distinct word-class in many Austroasiatic languages, denote sensory perceptions of the speaker — visual, auditory, tactile, olfactory, gustatory, emotional or other — in relation to a particular phenomenon (Diffloth 1972, 1976d). They differ from most other words in that they are largely iconic rather than symbolic; that is, the relation between their meaning and form is not characterised by arbitrariness but likeness of some sort. They often display peculiar phonological and morphological features, and they function syntactically like sentence adjuncts. For previous descriptions of expressives in individual languages, see e.g. Diffloth (1976b:84–85) for Jah Hut, Diffloth (1976d) for Semai, Benjamin (1976b:177–178) for Temiar, Svantesson (1983:78-79, 115–125) for Kammu and Kruspe (2004:396–402) for Semelai.

The class of expressives does not appear to be as evident in Jahai as in other Aslian languages, or at least it takes a slightly different form. A distinction needs to be made on frequency and individual grounds between *onomatopoeic forms* (§4.8.1) and what will be referred to here as *expressive elaboration*. This latter manifestation of expressives is a highly marginal phenomenon associated with only a couple of speakers and as such is treated separately in Chapter 6.

4.8.1 Onomatopoeic forms

Onomatopoeic forms are words displaying acoustic iconicity; that is, their phonological form, drawn from the ordinary phonemic inventory, bears some resemblance to the acoustic characteristics of the phenomenon they denote. They usually serve as syntactically optional adjuncts that add life to a narrative. Frequently, however, they turn up in predicate position and then function syntactically like verbs. Some onomatopoeic forms in predicate position may be negated and also be subject to derivational processes normally associated with verbs, e.g. imperfective aspect (§4.7.1.1). By definition, such forms are to be regarded as verbs in this position. Onomatopoeic forms are numerous, frequent, semantically specific, highly conventionalised and shared by the whole speech community. Examples include the following:

/rop/	'(sound of a muntjac deer)'
/bat/	'(sound of a bird landing on a branch)'
/cik/	'(sound of a blowpipe dart hitting muscle of prey)'
/klɨk/	'(sound of a person walking)'
/klak/	'(sound of something falling)'
/pok/	'(sound of an animal falling to the ground)'
/chok/	'(sound of running water or waterfall)'
/kɔk/	'(sound of a hornbill)'
/hɔkɔ̃k/	'(sound of a leaf-monkey)'
/ʔɔk/	'(sound of boiling)'
/rɨɸ/	'(sound of flapping)'
/jaɸ/	'(sound of flying or leaping)'
/luɸ/	'(sound of dashing)'
/koɸ/	'(sound of a blowpipe dart hitting stomach of prey)'
/kɔɸ/	'(sound of a small object falling to the ground)'
/rɔɸ/	'(sound of a blowpipe dart hitting canopy)'
/plɔ̃ɸ/	'(sound of a blowpipe being fired)'
/plɛ̃s/	'(sound of a blowpipe being fired)'
/cɛ̃h/	'(sound of a blowpipe dart hitting a vine)'
/wɛ̃h/	'(sound of a blowpipe being fired)'
/krãŋ/	'(sound of carving)'
/pəw/	'(sound of a shotgun)'
/prəw/	'(sound of an animal fleeing from one tree to another)'

In some cases there is a correlation between size/speed and vowel quality, as illustrated by the following set of forms:

/tik/	'(sound of a small raindrop falling to the ground)'
/tək ~ tɨk/	'(sound of a big raindrop falling to the ground)'
/tak/	'(sound of a big raindrop falling quickly to the ground)'
/tuk ~ tɔk/	'(sound of a big raindrop falling slowly to the ground)'

Onomatopoeic forms which function as adjuncts are not subject to any morphological operations with the exception of iconic total reduplication, as in the following examples:

(129) ʔoʔ kŋliŋ ka=bhɔl, rop-rop-rop-rop
3S to.make.sound SUBJ=muntjac.deer [sound of a muntjac deer ...]
'The muntjac deer barked [sound].'

(130) ʔoʔ lɔj prəw-prəw-prəw ba=ʔaniʔ
3S to.run [sound of an animal fleeing ...] GOAL=there
'It fled [sound] over there.'

Sometimes different onomatopoeic forms are used in succession to convey whole sequences of events. Thus, in the following example, where the course of events involves a hornbill taking off, flying away and landing in a different spot, only the first event (the taking off) is described with an ordinary clause, whereas the following events (the flight and the landing) are described in a highly iconic fashion with only onomatopoeic forms.

(131) ʔoʔ kɟɔw wel ka=ʔaj ton, jaɸ-jaɸ-jaɸ-jaɸ
3S to.take.off again SUBJ=game that [sound of flying ...]
bat
[sound of landing]
'That bird took off again [sound of flying] [sound of landing].'

The following example illustrates an onomatopoeic form in predicate position (see also §5.1.1.2):

(132) luɸ ka=bhɔl
[sound of dashing] SUBJ=muntjac.deer
'The muntjac deer [sound of dashing]!'

4.9 Prepositions

As a rule, prepositions are obligatorily proclitics which attach to the initial constituent of a NP, including nouns, personal pronouns, demonstratives, interrogatives and numerals. They express the categories of *location*, *goal*, *source* and *contrast*. An additional preposition expressing the category of *equation* is proclitic to pronouns and demonstratives but is otherwise a free morpheme. Several of the prepositional proclitics also have two allomorphs, a full-syllabic form with a prespecified vowel nucleus /a/, and a corresponding reduced, half-syllabic form with an underspecified vowel nucleus. The two allomorphs are usually in complementary distribution which may be determined by assimilatory or phonotactic factors but, at least in one case, also by semantics.

Whereas most of the prepositions can be used to license the occurrence of a NP and are then predicative, one of them — /ka=/ — may also be used non-predicatively to mark the grammatical relation of subject. Another one — contrastive /d=/ — can only be used non-predicatively to make pragmatic contrasts. Brief mentioning of such non-predicative usage is included here, but for a more detailed discussion, see §5.2.

4.9.1 Location/instrument/subject /ka=/

The phonologically reduced form /k=/ of the prepositional proclitic /ka=/ expresses location in, at or by the referent of the NP (LOC):

(133) japɛ̃h ŋɔk k=klap
1P.EXCL to.sit LOC=Kelap
'We stayed at Kelap.'

(134) jɛʔ pʔəh k=buloʔ
1S to.cook LOC=bamboo.tube
'I cooked [it] in bamboo tubes.'

A more metaphorical reading is exemplified in the following sentence:

(135) ʔoʔ ʔɛl k=ʔawɛ̃ʔ ton
3S to.look LOC=type.of.tortoise that
'It looked at that tortoise.'

The fully syllabic form /ka=/ is used in the literally locative sense only with demonstratives and some monosyllabic nouns:

(136) jɛʔ b-pŋsɛŋ ka=ʔũn
1S PROG-to.talk LOC=there
'I was talking over there.'

(137) cɨp ka=sɛɲ
to.go LOC=front
'Walk in front!'

It is also used to mark the thematic role of instrument.

(138) jɛʔ cek ʔaj ka=taɟiʔ
1S to.stab game.animal INSTR=knife
'I stabbed the animal with a knife.' (elic.)

Otherwise /ka=/ is used non-predicatively to mark postverbal NPs that represent the syntactic subject of a clause (see §5.2.1.1.1).

(139) jɛʔ lkluk ka=jɛʔ
1S to.laugh SUBJ=1S
'I laughed.'

4.9.2 Goal /ba=/

The prepositional proclitic /ba=/ (GOAL) typically expresses concrete motion to or towards the referent of the NP:

(140) japɛ̃h ja=wek ba=hajɛ̃ʔ
1P.EXCL IRR=to.go.back GOAL=house
'We will go back to the house.'

It may also denote metaphorical motion to a recipient:

(141) ʔoʔ ʔɛk dwiʔ ba=jɛʔ
3S to.give money GOAL=1S
'He gave me money.'

The phonologically reduced form /b=/ is only found in combination with the demonstratives /ʔitih/ 'there, above' and /ʔujih/ 'there, below' and the interrogative /lbah/ 'where?'.

4.9.3 Source /can=/

Corresponding to English 'from', the prepositional proclitic /can=/ (SOURCE) expresses concrete or metaphorical motion away from the referent of the NP. The full allomorph /can=/ is normally in free variation with a phonologically reduced /cn=/ but is the preferred form with monosyllabic words, whereas reduced /cn=/ is the preferred form with disyllabic words with a pre-final vowel /a/.

(142) jɛʔ ɟok can=mɔh cɔs taniʔ
1S to.move SOURCE=headwaters Cos that
'I moved from those headwaters of the river Cos.'

(143) ʔoʔ ʔɛl-ʔɛl ka=ʔap cn=ʔaniʔ
3S to.look-CONT SUBJ=tiger SOURCE=there
'The tiger kept on watching from over there.'

(144) sɛc cn=kasaʔ
meat SOURCE=sambar.deer
'meat from a sambar deer' (elic.)

4.9.4 Contrast /d=/

The prepositional proclitic /d=/ (CONTR) is used to express that the referent of the NP is in some way contrasted with other explicit or implicit referents which are potential competitors for the same relation. In the following clause, it attaches to a postverbal subject NP to signal that the subject of the clause, as opposed to some other, implicit participant, is going to carry out the action (see also §5.2.1.1.2).

(145) hɛj ja=kajil d=hɛj
1D.INCL IRR=to.fish CONTR=1D.INCL
'We [as opposed to X] will go fishing.'

The following example illustrates similar contrasting, but this time of a direct object (see also §5.2.2.1).

(146) ɟa=ʔoʔ bdil lɛh d=hawɛ̃n
RT=3S to.shoot EMP CONTR=pig
'Then he shot the pig [as opposed to shooting X].'

If attached to a NP denoting a recipient, it signals that the referent receives some object for keeping, or is rendered a special favour intended for his or her benefit, as in the following example (see also §5.2.3).

(147) ʔɛk lɛh d=jɛʔ
to.give EMP CONTR=1S
'Give to me!'

It is occasionally also used in possessive-like constructions (see also §5.1.4.3).

4.9.5 Equation /pon ~ pn=/

The prepositional proclitic /pn=/ (EQU), used with pronouns and demonstratives, and its corresponding free form /pon/, used with nouns, express likeness to the referent of the NP. This is most likely a loan of Malay *pun* 'also', 'too'. Examples include the following:

(148) ɟa=jɛʔ tikah k=ɟhūʔ lɛh pn=ʔə̄h
RT=1S to.spring LOC=tree EMP EQU=here
'Then I sprang up in a tree like this!'

(149) ʔoʔ gɐj pon ʔap
3S to.eat EQU tiger
'He eats like a tiger.' (elic.)

4.10 Auxiliaries and adverbs

Jahai exhibits a number of proclitics, particle-like grammatical words and adverbs which modify verbs or clauses. Several of these can be arranged in the categories of temporal or modal auxiliaries (*relational tense* and *root possibility*), negation markers (*negative* and *prohibitive*) and interrogative particles (*interrogative*), all of which are found in preverbal position. These are described in §4.10.1–4.10.3. An emphatic particle is described in §4.10.4., adverbs are treated in §4.10.5, and additional adverbial elements are listed in §4.10.6. Many of these forms are of Malay origin.

4.10.1 Temporal and modal auxiliaries

4.10.1.1 Relational tense

The proclitic /ɟa=/ attaches to the first constituent of the obligatory part of the clause, referred to here as the *core* (cf. §3.3.2.3, §5.1.1), to signal *relational tense* (RT). This may signify that a situation is anterior to the time of reference (cf. the terminology of Bybee 1985:159–161 and Bybee et al. 1994:54, 61–63, 318). It is then readily translated into English as 'already' or 'ever'. The following examples illustrate its use:

(150) ɟa=jɛʔ ʔtʔet
RT=1S to.know
'I already know.'

(151) ha=ɟa=mɔh bdil ka=mutah
Q=RT=2S.FAM to.shoot INSTR=mortar
'Have you ever used a mortar?'

(152) ɟa=sɔc ka=snlɔc
RT=to.be.gone SUBJ=blowpipe.dart
'The darts are already used up!'

(153) hobiʔ ton ɟa=bɲɟiʔ
root.crop that RT=to.be.tall
'That root crop had already grown tall.'

It may also signify that a situation is posterior to the time of reference, and is then preferably translated into English as 'then'.

(154) ɟa=ʔoʔ lɔj ba=ʔaniʔ
RT=3S to.run GOAL=there
'Then it fled in that direction.'

(155) jɛʔ wk-wek, ɟa=jɛʔ gɛs k=slaj
1S IMPF-to.return RT=1S to.descend LOC=swidden
'I was going back. Then I went down through the swidden.'

If combined with irrealis forms, it signals that a situation is about to take place, as in the following example:

(156) ɟa=wa=sɔc tep tə̄h
RT=IRR.3S=to.be.gone cassette this
'This cassette is about to run out.'

The common notion of these meanings is one of temporal contrast.

The relational tense marker frequently merges with the third person singular pronoun /ʔoʔ/ to form a portmanteau morpheme /ɟoʔ/ (RT.3S).

(157) ɟoʔ cɨp
RT.3S to.go
'Then he left.'

The relational tense marker is likely to be derived from the Malay perfect marker *sudah* or *dah* 'done', 'finished'.

4.10.1.2 Root possibility

The modal auxiliary /boleh/ (from Malay *boleh* 'to be able to') signals that the agent is able to or is allowed to carry out the action designated by the verb. Following Bybee et al. (1994:178, 320), it will be referred to here as *root possibility* (RP). It occurs at the left edge of the core of a clause.

(158) boleh ja=gej
RP IRR=to.eat
'[I] can eat.'

(159) boleh jɛʔ cɨp
RP 1S to.go
'I was able to leave.'

4.10.2 Negation markers

4.10.2.1 Negative

Clause negation is expressed with a negative marker /braʔ/ (NEG), which occurs in free variation with the borrowed negative marker /bokan/ (from Malay *bukan* 'no', 'not') before the core of the clause or, occasionally, before the preverbal subject NP. As a negator at clause level it always co-occurs with an irrealis-marked verb (see §4.7.4.1 and §5.1.1.3).

(160) braʔ wa=btʔɛt
NEG IRR.3S=to.be.good
'It's not good.'

(161) bokan ja=b-ŋk-ŋɔk kampoŋ
NEG IRR=PROG-IMPF-to.sit village
'I wasn't living in a village.'

(162) bra? japɛ̄h ja=pktɔk
NEG 1P.EXCL IRR=to.cook
'We didn't cook.'

The negative marker may also be used to negate NPs and PPs. Furthermore, it sometimes occurs in predicate position, taking on properties associated with existential verbs (see also §5.1.1.2).

An additional indigenous negative marker /bŋit/, recorded by Schebesta (1928a:822), has been noted only sporadically in the present material. Its function and distribution are unclear.

4.10.2.2 Prohibitive

The prohibitive marker /?akɛt/ (PROH) is placed before a verb to express negative imperative.

(163) ?akɛt lkluk
PROH to.laugh
'Don't laugh!'

(164) ?akɛt rɔh bəw
PROH to.cut to.be.big
'Don't make a big clearing.'

(165) ?akɛt de? pn=?on
PROH to.make EQU=that
'Don't do like that!'

(166) ?akɛt lɔj, paj ton
PROH to.run 2S.DIS that
'You there! Don't run!'

It is also used as an isolated interjection to issue admonitions or warnings:

(167) ?akɛt
PROH
'Don't!'/'Watch out!'

4.10.3 Interrogative

Yes/no questions may be formed from any statement through the attachment of an interrogative proclitic /ha=/ (Q) to the first constituent of the clause.

(168) ha=ɟa=mɔh gej
Q=RT=2S.FAM to.eat
'Have you eaten?'

(169) ha=gin ja=ciweh
Q=2/3P IRR=to.climb.up
'Will you climb up?'

(170) ha=wa=wek ba=ʔũn
Q=IRR.3S=to.go.back GOAL=there
'Will he return there?'

(171) ha=mɔh ja=haj tkɨh jɛʔ
Q=2S.FAM IRR=to.follow backside 1S
'Will you follow me?'

4.10.4 Emphasis

A particle /lɛh/ (EMP), from Malay *lah*, may be placed after any word, phrase, clause or sentence for emphasis and affirmation. Emphatic readings are illustrated in the following examples:

(172) japɛ̄h ŋɔk k=tə̄h lɛh
1P.EXCL to.sit LOC=this EMP
'We stayed *here*!'

(173) pɛ̄h ŋɔk lɛh k=ton
1P.EXCL to.sit EMP LOC=that
'We *stayed* there!'

(174) braʔ wa=ʔŋɨc lɛh
NEG IRR.3S=to.burn EMP
'He didn't *burn* [X]!'

Affirmative use of /lɛh/ is exemplified in the following answer:

(175) ʔoʔ lɛh
3S EMP
'[It's] him all right.'

If combined with imperative forms, /lɛh/ adds additional emphasis to the exhortation:

(176) cɨp lɛh
to.go EMP
'Now, go!'

(177) ʔɛk lɛh d=jɛʔ
to.give EMP CONTR=1S
'Now, give to me!'

4.10.5 Adverbs

The adjective-like property verbs usually function as predicates but may also modify other verbs, in which case they are regarded here as adverbs. In this capacity they always follow the verb they modify and are frequently separated from it by the emphatic particle /lɛh/ described in §4.10.4.

(178) ʔakɛt rɔh bəw
PROH to.cut to.be.big
'Don't make a big clearing.'

(179) gej lɛh hakɨj
to.eat EMP to.be.slow
'Eat slowly!'

(180) jɛʔ cɨp bɲɟiʔ
1S to.go to.be.far
'I walked far.'

Schebesta (1928a:822) stated that adverbs are regularly formed by means of the optional prefix *le*. This prefix is likely to be related to the interposed emphatic particle identified here. It should be noted that a few property verbs may contain frozen traces of a prefix /l-/ which originally had the function of co-ordinating verbs with modifying property verbs. These include /lktwət/ 'to be quick', /lwɛs/ 'to be wide' and /laʔis/ 'to be bad'.

4.10.6 Other adverbial elements

Additional adverbials include the following set of indigenous, postverbal forms:

/haji?/	'also', 'too'
/ɟaga?/	'only'
/sʔoʔ/	'just', 'only'
/wel/	'again'
/ʔnuj/	'soon'

Also, the following forms borrowed from Malay occur frequently, usually clause-initially:

/lagiʔ/	'again', 'also', 'moreover', 'still', 'yet'
/baruʔ/	'again', 'anew', 'until'
/tros/	'at once', 'straightaway'
/dah/	'already'

4.11 Conjunctions

Jahai phrases and clauses may be related to each other by means of a set of five conjunctions. There are two indigenous, co-ordinating conjunctions: a disjunctive proclitic /ha=/ 'or' and an additive particle /ʔalɔʔ/ 'and'. These may be used to co-ordinate full clauses or phrases, typically noun phrases. The additive /ʔalɔʔ/ is simply placed between the two co-ordinated elements, whereas the disjunctive /ha=/ is attached to the first constituent of each co-ordinated element. The co-ordinating function of the disjunctive is closely linked to the main function of /ha=/ as an interrogative (cf. §4.10.3), and it is mostly used in interrogative contexts.

(181) ja=cɨp ba=grik ʔalɔʔ ja=bliʔ bras
IRR=to.go GOAL=Grik and IRR=to.buy rice
'I will go to Grik and I will buy rice.' (elic.)

(182) ha=mɔh ha=jɛʔ
or=2S.FAM or=1S
'you or me'

Subordinating conjunctions, of which three have been documented, are borrowings from Malay: a causal /sbap/ 'because' (from Malay *sebab* 'cause'), a temporal /lpəs/ 'after', 'when' (from Malay *lepas* 'after'), and a conditional /kaluh ~ kalɔw/ 'if' (from Malay *kalau* 'if'). These are placed at the beginning of the subordinate clause. Subordinate clauses may either precede or follow the main clause (cf. §5.1.1.4).

With the exception of the temporal /lpəs/, conjunctions are rare. Co-ordination normally involves the combination of phrases/clauses without conjunctions.

4.12 Co-ordinating morphemes in the NP

Jahai has two relative-marking morphemes, which have the purpose of subordinating heads of NPs with some of their modifiers: a proclitic /k=/ (§4.12.1) and a prefix /t-/ (§4.12.2). Although the difference between them is not always clear-cut and speakers claim they are interchangeable, they usually display quite distinct characteristics from both a syntactic and a semantic point of view. A third co-ordinating morpheme, the proclitic /l= ~ la=/, marks the identificational co-ordination of a NP with a deictic element (§4.12.3).

4.12.1 Relative marker /k=/

The relative marker /k=/ (REL) is attached proclitically to the initial constituent of a phrase or clause to signal relativisation. Thus, any such element may become a modifier in the NP. The relative clause/phrase always occurs at the right edge of the NP (see also §5.1.3 and §5.1.4.1). Some examples follow:

(183) ʔoʔ k=tmkal
3S REL=man
'the one who is a man'

[As opposed to another person, who is a woman; note that the relative phrase is intended to disambiguate the referent as no gender distinction is made in the third person singular pronoun]

(184) puŋhuluh k=hapaʔ
headman REL=to.die
'the headman who died'

(185) jaʔ ʔoʔ tə̄h k=lawaʔ ʔikan
grandmother 3S this REL=to.collect fish
'this grandmother of his, who caught fish'

(186) makɛn koleh k=pn=ʔə̄h
whose cup REL=like=here
'Whose cup [is it] that [looks] like this?'

(187) slaj k=wih rh-rɔh
swidden REL=3D IMPF-to.clear
'the swidden that they were clearing'

Sometimes, the relative marker /k=/ is replaced or supplemented by /jaŋ/, a borrowing of the Malay relative marker *yang*:

(188) pos jaŋ bəw
post REL to.be.big
'a military post that is big'

(189) hobiʔ jaŋ k=jɛʔ tanɛm ktɔʔ wɛj
root.crop REL REL=1S to.plant day past
'the root-crop that I planted yesterday' (elic.)

4.12.2 Relative marker /t-/

Unlike the relative marker /k=/, the marker /t-/ (also glossed as REL) does not appear to attach to phrases or clauses but only at word-level (hence it is treated here as an affix rather than a clitic), usually to verbs but occasionally also nouns and numerals. Its main purpose therefore appears to be to derive adjective-like modifiers of NP heads from individual verbs, which makes the resulting forms reminiscent of participles. However, its additional albeit occasional use with nouns and numerals warrants the treatment of it here as a relative marker operating at word level rather than a morpheme for deriving a separate class of participles. Still, given their derived character, it is convenient to translate many of these forms into English as participles or adjectives rather than relative clauses. Also, unlike relative elements introduced by the relative marker /k=/, /t-/-affixed forms are not attached at the right edge of the NP, but are found postnominally between the NP head and other modifiers, notably demonstratives (see §5.1.4.1). Examples of the relative marker /t-/ with verbs include the following:

(190) mnraʔ t-kul
person REL-to.call
'calling person'

(191) kɛn t-ʔaʄɔʔ
child REL-to.be.small
'small child'

(192) ʔoʔ t-bəw
3S REL-to.be.big
'the big one'
[Avoidance term denoting elephant]

(193) ʔoʔ t-dɛs
3S REL-to.move
'the moving one'

(194) ʔap t-b-tadɔʔ
tiger REL-PROG-to.wait
'waiting tiger'

The following phrases exemplify the relative marker /t-/ with nouns and numerals:

(195) heʔ t-mnraʔ
1P.INCL REL-person
'we humans'

(196) wih t-kʄih ka=ʔūn
3D REL-boy LOC=there
'those [two] boys'

(197) ʔoʔ t-nɛj
3S REL-one
'the single one'

4.12.3 Identification marker /l= ~ la=/

The identification marker /l= ~ la=/ (ID) is attached proclitically to the initial constituent of a NP to co-ordinate this NP with a preceding deictic element, typically a nominal demonstrative but occasionally also proper names. This signals that the NP affirmatively identifies or specifies the referent referred to by the deictic. The deictic element itself represents a NP, so the identification-marked NP is considered here to be a subordinate modifier similar to the relativised constituents illustrated in §4.12.1 (see also §5.1.3.1). The following phrases exemplify its use.

(198) tə̃h, l=ksɨj ʔoʔ ʔə̃h
this ID=husband 3S here
'This [is] her husband here.'

(199) tə̃h, l=nɛj ʔə̃h
this ID=one here
'this one here'

(200) ton, l=tahoh ʔon
that ID=Tahoh there
'That [is] Tahoh there.'

(201) tahanɛn, l=kneh ʔoʔ
Tahanan ID=wife 3S
'Tahanan [is] his wife.'

If a deictic head is absent, the identification-marked NP may make up an independent NP in its own right. Such constructions usually include a locative head noun, as in the following examples:

(202) l=krpiŋ ʔoʔ
ID=upper.side 3S
'its upper side'

(203) l=sɛɲ ʔoʔ
ID=front 3S
'its front'

The allomorph /la=/ occurs when identification-marked demonstratives and nouns form part of a prepositional phrase headed by the prepositional proclitic /k=/ 'LOC'. The set of nouns recorded in this position is limited and the construction gives the impression of being rather idiomatic. Some examples follow:

(204) k=la=ʔə̃h
LOC=ID=here
'here'

(205) k=la=ʔon
LOC=ID=there
'there'

(206) k=la=hɨp
LOC=ID=forest
'in the forest'

(207) k=la=hajɛ̃ʔ
LOC=ID=house
'in the house'

Given its affirmative and identifying function, it is likely that the identification marker bears some relationship to the phrase-final emphatic and affirmative particle /lɛh/, described in §4.10.4.

4.13 Summary

This chapter has been concerned with the identification and characterisation of Jahai word classes as well as the morphological categories associated with them. Nouns were shown to be a semantically well-defined word class displaying productive derivational morphology related to quantification. This situation is similar to the one described by Schebesta (1928a:810–812). The typologically unusual system of unitisation has parallels in some other Aslian languages. Furthermore, nouns may be formed by means of derivation of verbs and numerals, a productive pattern encountered in other Aslian languages as well. Other nominal classes of particular interest include pronouns and demonstratives, which display rich systems of distinctions.

With regard to the class of verbs, a significant feature is its system of derivation, with various morphemes expressing aspect/Aktionsart and causative. In terms of complexity, productivity and regularity, this system is to be placed on a par with the elaborate verbal paradigms described for other Aslian languages. As far as aspect/Aktionsart is concerned, Jahai appears to exhibit a somewhat richer set of distinctions than other Aslian languages described, the categories of iterative and distributive being unattested so far elsewhere in Aslian. A conspicuous feature is the presence of a number of grammatical morphemes borrowed from Malay, including the multifunctional /b-/ prefix, the continuative, and possibly some of the causative allomorphs, as well as several auxiliary and adverbial elements and conjunctions.

Jahai was also shown to exhibit a set of morphemes used to co-ordinate heads of NPs with some of their modifiers. Manifold processes of co-ordination within the NP may be a feature peculiar to Jahai in the Aslian context, as similar morphemes are not well-attested in other Aslian languages.

5 *Syntax*

This chapter outlines the syntactic characteristics of Jahai. It should be viewed as a brief and tentative introduction to the subject, as several aspects of Jahai syntax are yet to be thoroughly examined and only the most manifest patterns are presented. Beginning with a description of the structure of clauses and phrases (§5.1), the chapter goes on to cover the arguments of the clause (§5.2), valence and transitivity (§5.3), and valence-affecting operations (§5.4). Although not presented within a specific theoretical framework, the description has gained a lot of inspiration from the monostratal approach to syntactic formalisation proposed by Van Valin and LaPolla (1997). The account is based mostly on spontaneous rather than elicited data.

5.1 The structure of clauses and phrases

This section is concerned with the identification and order of syntactic constituents, investigating statements (§5.1.1), questions (§5.1.2), relative clauses (§5.1.3), noun phrases (§5.1.4) and prepositional phrases (§5.1.5).

5.1.1 Statements

Clauses may be broadly divided into two types. One type requires a preverbal subject-marking pronoun. Such clauses are referred to here as *common clauses* and are described in §5.1.1.1. The other type, primarily associated with a subset of stative verbs, does not require an obligatory preverbal subject-marking pronoun. Such clauses are described in §5.1.1.2. Negated constructions are dealt with in §5.1.1.3. Subordination is discussed briefly in §5.1.1.4.

5.1.1.1 Common clauses

5.1.1.1.1 The verb and its arguments

The basic order of elements in the common clause may be summarised as follows:

(SUBJ) SUBJ:AGR **V** (DIRECT OBJECT) (OBLIQUE ARGUMENT)

The predicating element of common clauses is always a verb, dynamic or stative. This is preceded by an obligatory subject marker, in non-irrealis clauses represented by an unstressed pronoun. Such pronouns will be referred to as *subject particles*. The verb and its subject particle form the minimal common clause, referred to here as the *core*, as illustrated in the following examples:

(1) ʔoʔ lɔj
3S to.run
'He ran.'

(2) ʔoʔ hm̄ic
3S to.rain
'It rains.'

(3) jɛʔ hgik
1S to.be.afraid
'I'm afraid.'

(4) hɛj pɛk
1D.INCL to.split
'We split.'

The subject particles will be viewed here as a kind of subject agreement on the verb, although this characterisation is not altogether unproblematic. Recall, for example, that subject particles are treated on phonological grounds as free forms rather than clitics (see §3.3.1 and §4.3), in spite of their inability to receive stress. This phonologically independent nature is somewhat at odds with the treatment of them as markers of agreement. Their absence from WH questions is also atypical of agreement (see §5.1.2.2). However, their syntactic behaviour is otherwise notably agreement-like, including their obligatory presence in clauses as well as citation forms of verbs (but absence with the adjective-like property verbs), their mutual exclusivity with the vaguely subject-encoding irrealis proclitics (see below), their fixed pre-verbal position and inability to be separated from the verb by other elements, and the optionality of a full phrase denoting the subject. Such features motivate the treatment of them here as agreement markers rather than full subject arguments. In reality, they are likely to represent an intermediate stage between subject argument and agreement.[25]

In irrealis constructions the subject particle is replaced by an irrealis proclitic, which on top of its mainly modal function also vaguely encodes person of subject. Recall that a third person singular form /wa=/ contrasts with a general form /ja=/ used with all other subjects (cf. §4.7.4.1). Like the subject particles, these portmanteau clitics are to be regarded as subject-agreement markers. Their vague and generalised nature with regard to person is in accordance with the subject suppression associated with irrealis. Similar suppression has been noted in Semelai by Kruspe (2004:161). In these constructions, it is the verb and the irrealis proclitic that form the *core* of the clause. Examples are given in (5) and (6).

(5) ja=sam
IRR=to.hunt
'[I] will hunt.'

---

25 I am grateful to Dr Arthur Holmer, Department of Linguistics and Phonetics, Lund University, for his insightful suggestions in this context.

(6) wa=cɨp
IRR.3S=to.go
'He will leave.'

A full noun phrase denoting the subject may optionally precede the subject agreement marker. The position of such NPs is always to the immediate left of the marker. Their referent typically represents known information and may therefore be considered as topic; this would explain the relative infrequency of such NPs, as the subject agreement marker may often be sufficient for the identification of the referent. Examples (7) and (8) illustrate clauses in which subject particles are preceded by such full NPs.

(7) puŋhuluh ʔoʔ lɔj
headman 3S to.run
'The headman fled.'

(8) ʔap ʔoʔ kbis
tiger 3S to.die
'The tiger died.'

Examples (9) and (10) illustrate clauses in which full subject NPs precede irrealis-marked verbs. In (10), the subject NP is represented by a disambiguating pronoun. Note that this now occupies the optional NP position and is therefore not a subject particle.

(9) wɔŋ wa=gej
child IRR.3S=to.eat
'The child will eat.'

(10) japɛ̃h ja=cɨp
1P.EXCL IRR=to.go
'We will leave.'

Other arguments are typically found in postverbal position and consist either of an unmarked noun phrase or a prepositional phrase. Direct objects are usually represented by NPs and are typically found to the immediate right of the verb. This is exemplified in (11) and (12).

(11) ja=bdil gaw
IRR=to.shoot pig
'[I] will shoot a pig.'

(12) ʔoʔ tampɛr jɛʔ
3S to.claw 1S
'It clawed me.'

Third arguments typically follow the direct object and are always represented by a PP. They are referred to here as oblique arguments. Example (13) illustrates a clause with both direct object and an oblique argument in the form of a recipient.

(13) ʔoʔ ʔɛk taɟiʔ ba=wɔŋ
3S to.give knife GOAL=child
'He gave the knife to the child.' (elic.)

The linear order of these clause elements is rather fixed, with the subject agreement marker (subject particle or irrealis proclitic) always occurring immediately to the left of the verb, and the direct object to the right of the verb. The unmarked location of the oblique

argument is to the right of the direct object, as illustrated in (13), but occasionally it precedes it, as illustrated in the following clause:

(14) jɛʔ ʔɛk ba=mɔh sorɛt
1S to.give GOAL=2S.FAM letter
'I gave you a letter.' (elic.)

No verb requires a direct object or oblique argument, and the most stable and only obligatory portion of the clause is therefore made up of the verb and its subject agreement marker. Note also that an oblique argument does not require the presence of a direct object. This is further discussed in §5.3.1.

The order of clause elements outlined here is considered here to be fundamental and to represent the basic word order of the language; hence Jahai may be characterised as an SVO language.

5.1.1.1.2 Permuted arguments

Two processes of argument permutation may alter the basic order of elements outlined in §5.1.1.1.1. One involves an optional postverbal prepositional phrase denoting the subject. In the unmarked case, this occurs to the immediate right of the verb, thus separating the verb from any subsequent arguments and adjuncts. Sometimes other elements may occur between the verb and the subject PP. It co-occurs with the coreferential obligatory preverbal subject marker. It is mutually exclusive with the preverbal subject NP — they typically cannot both occur in the same clause — and may be regarded as a permuted variant of this NP. It also appears to have a different pragmatic function, signalling that its referent is in focus. (However, the data contains one example of co-occurrence of a pronominal subject NP in preverbal position and PP in postverbal position (17), the significance of which is unclear.) Postverbal subject PPs may be headed by the prepositional proclitic /ka=/ 'SUBJ' or, less commonly, /d=/ 'CONTR' (see §5.2.1.1). The following clauses exemplify the pattern.

(15) ʔoʔ kɛc ka=kaɲcɔʔ ʔawɛj
3S to.cut SUBJ=grandchild rattan
'The grandchild cut the rattan.'

(16) ʔoʔ cɨp ka=ʔɛj jɛʔ wɛj ba=ʔəh
3S to.go SUBJ=father 1S past GOAL=here
'My late father came here.'

(17) wih ja=cɔm ka=wih ʔɔs
3D IRR=to.burn SUBJ=3D fire
'They will make a fire.'

(18) jɛʔ jiʔ d=jɛʔ
1S to.refuse CONTR=1S
'I refuse.'

(19) jɛʔ lkluk ka=jɛʔ
1S to.laugh SUBJ=1S
'I laughed.'

(20) ja=cɨp ba=ʔũn d=jɛʔ
IRR=to.go GOAL=there CONTR=1S
'I will go over there.'

(21) ʔoʔ b-k<ŋ>ɟeŋ lɛh ka=ʔoʔ cn=kjɔm tə̃h
3S PROG-to.listen<IMPF> EMP SUBJ=3S SOURCE=underside this
'It was listening from down here!'

The second type of permutation involves a noun phrase representing the direct object in clause-initial position; that is, to the left of the preverbal subject marker or the full subject NP if such an NP exists. As with subjects, preverbal and postverbal direct objects are mutually exclusive and cannot both occur in the same clause. Permutation of direct objects is relatively uncommon and the exact function of this operation is unclear. Examples are given in (22) and (23).

(22) blɨʔ ʔoʔ ʔoʔ gantoŋ
leg 3S 3S to.hang
'He hung up its leg.'

(23) ʔikəʔ jɛʔ gj-gej
fish 1S IMPF-to.eat
'I was eating fish.'

5.1.1.1.3 Adjuncts

Adjuncts may be defined as those elements of a clause which are not arguments of the verb. They typically include locative and temporal NPs and PPs, but also onomatopoeic forms and other adverbial elements. Adjunct elements are usually located to the right of the verb or any arguments following the verb, as in examples (24), (25) and (26), but they may occasionally turn up between the verb and the following argument, as in (27), or between postverbal arguments. Adjunct elements cannot occur between the verb and the preceding subject marker, or between the subject marker and a preceding subject NP, but they sometimes occur clause-initially, as in (28).

(24) ja=kajil ba=taseʔ
IRR=to.fish GOAL=lake
'[I] will go fishing in the lake.'

(25) jɛʔ sə̃h ʔap ktɔʔ wɛj
1S to.encounter tiger day past
'I saw a tiger yesterday.'

(26) jɛʔ bdɛl plɛ̃s
1S to.shoot [sound of blowpipe]
'I shot [sound]!'

(27) ʔoʔ paŋkoh pn=ʔon ʔawɛj manɔw
3S to.hold EQU=there vine *manau*
'He held the *manau* vine like this.'

(28) limɛʔ ktɔʔ japɛ̃h ŋɔk b=ʔitɨh
five day 1P.EXCL to.sit GOAL=there
'We stayed up there for five days.'

5.1.1.1.4 Detached phrases

Phrases which are separated from the clause by a pause are referred to here as *detached phrases*. This is in accordance with the terminology of Van Valin and LaPolla (1997). These may be described as being located outside the clause but inside the sentence. If such a phrase represents an argument of the verb in the clause from which it is detached, it must be coreferential with an argument within the clause. This clause-internal coreferential argument may be represented by a pronoun or a full phrase. Detached phrases may occur in Jahai both before and after the clause, and hence a distinction is made between left-detached phrases and right-detached phrases. As will be shown, these two positions assign phrases with different pragmatic functions. In the examples given, the pause separating the detached phrase from the clause is represented by a comma.

Left-detached phrases contain topicalised NPs or PPs that represent either an argument of the verb or an adjunct element. As noted, those which represent arguments of the verb must be coreferential with an element within the clause. If the detached phrase denotes the subject of the clause, this is typically coreferential with the obligatory preverbal subject-agreement marker. Such a construction is illustrated in example (29).

(29) ʔalɛh, ʔoʔ kul ba=biʔ ʔoʔ ton
girl 3S to.call GOAL=mother 3S that
'As for the girl, she called to her mother.'

In example (30), the left-detached phrase is represented by a noun denoting the direct object of the clause. This is coreferential with an NP within the clause in the form of a classifier construction.

(30) rampɔw, jɛʔ bdil duwaʔ kbiʔ
macaque 1S to.shoot two CLF
'As for macaques, I shot two of them.'

Left-detached phrases may also be represented by elements which are coreferential with a modifying referent of an argument within the clause. Example (31) illustrates a left-detached NP which is coreferential with the possessing modifier of the direct object of the clause. This modifier is represented in the clause by a pronoun.

(31) hobiʔ hɛj, kasaʔ ʔoʔ gej haliʔ ʔoʔ
root.crop 1D.INCL sambar.deer 3S to.eat leaf 3S
'As for our root-crop, the sambar deer ate its leaves.'

Like left-detached phrases, right-detached phrases contain NPs and PPs that represent either an argument of the verb or an adjunct element. Again, if such a right-detached phrase represents an argument of the verb, it must be coreferential with an argument within the clause. But whereas the function of left-detached phrases is one of topicalisation, right-detached phrases can be said to be specifying in character. This is illustrated in examples (32) and (33), where right-detached NPs form specifications of a more generic coreferential direct object in the clause.

(32) ja=lawaʔ tʔaʔ, tʔaʔ paku?
IRR=to.collect edible.plant edible.plant *paku*
'[I] will collect plants, *paku* plants.'

(33) hɛj ja=b-dkdak ʔikəʔ ba=ʔadeh, ʔikəʔ tŋraŋ
1D.INCL IRR=PROP-fish.trap fish GOAL=there fish *tengrang*
'We will trap fish back there, *tengrang* fish.'

Classifier constructions are a very common type of right-detached NP. These consist of a numeral and a classifier and are typically coreferential with the direct object of the clause. Again, the purpose of the right-detached phrase is one of specification, this time with regard to the number of the referent. The following examples illustrate such a construction.

(34) jɛʔ bdɛl ʔamɛŋ, duwaʔ k<nʔ>biʔ
1S to.shoot siamang two CLF<UNIT>
'I shot siamangs, two of them.'

(35) jɛʔ bʔbɔʔ luŋan, duwaʔ k<nʔ>mɔʔ
1S to.carry.on.back binturong two CLF<UNIT>
'I carried binturongs, two of them.'

As noted in §4.2, classifier NPs almost never include the noun itself. Instead, they are typically syntactically opposed coreferents of the noun, either in the form of right-detached phrases, as in (34) and (35), or, less commonly, arguments in the clause, in which case the coreferential noun may appear as a left-detached phrase, as was illustrated in example (30).

5.1.1.1.5 Summary

As the preceding sections reveal, the typical Jahai clause consists of a stable unit made up of a verb and an obligatory preverbal subject agreement marker (subject particle or irrealis proclitic). This unit, referred to as the *core* of the common clause, may be preceded by an optional NP denoting the subject, and it may be followed by one or two optional phrases representing a direct object and/or an oblique argument. The linear order of these two latter arguments is free, but usually the direct object precedes the oblique argument. Adjuncts are non-arguments which may occur in any postverbal position but usually to the right of the arguments. They may also occur clause-initially.

Two types of argument permutation may alter this basic clause structure. The first one involves the subject, the optional preverbal NP of which may be replaced by a postverbal PP. This is a very common type of operation. The second one involves the occurrence of a direct object in clause-initial position instead of the usual postverbal position.

Another type of operation is that of detached phrases. Thus, any referential element in the clause may occur sentence-initially or sentence-finally, separated from the clause by a pause. If this referent is an argument of the verb, it also has to be represented in some form in its usual position within the clause. In other words, such detached phrases must be coreferential with an argument in the clause. The assignment of a detached phrase to the left of the clause signals topicalisation, whereas a detached phrase to the right of the clause is associated with specification.

5.1.1.2 Clauses without subject particle

A structure different from that outlined in §5.1.1.1 applies to clauses whose predicating element is represented by a member of a subset of stative, intransitive verbs in the form of (1) most property verbs (e.g. /bəw/ 'to be big', /ʔhəj/ 'to be small' and /btʔɛt/ 'to be good'), and (2) the existential verbs /sɔc/ 'to be gone' and /weʔ/ 'to exist', as well as less frequent clauses whose predicating element is represented by non-verbal classes like onomatopoeic forms, the negative marker, PPs or WH elements (see §5.1.2.2). These clauses differ from the common clause pattern in that they do not require — but may occasionally take — a preverbal subject particle. The subject is optionally represented either by a preverbal NP or

a postverbal PP, corresponding to the mutually exclusive subject phrases described in §5.1.1.1.1 and §5.1.1.1.2. The core of such clauses is represented only by the verb. The following examples illustrate such clauses with property and existential verbs.

(36) lej mi? bəw
body 2S.INT to.be.big
'Your body is big.'

(37) ghɛl ka=jɛ?
to.be.tired SUBJ=1S
'I'm tired.'

(38) pinɛŋ we?
betelnut to.exist
'There are betelnuts.'

(39) we? ka=lwɛj
to.exist SUBJ=honey
'There is honey.'

(40) sɔc ka=gin gop
to.be.gone SUBJ=2/3P stranger
'Them[26] strangers are gone.'

Examples (41) and (42) illustrate the same pattern, but now with an onomatopoeic form as predicating element.

(41) luɸ ka=bhɔl
[sound] SUBJ=muntjac.deer
'The muntjac deer [sound of dashing].'

(42) pəw ka=bdil gin
[sound] SUBJ=gun 2/3P
'Their gun [sound of shotgun].'

Onomatopoeic forms in verbal position are negatable and sometimes also subject to derivative morphology associated with verbs. Thus, by definition, they are to be regarded as verbs in this position.

The same clause pattern applies to constructions where the negative marker /bra?/ forms a predicating element. Such constructions denote non-existence of the negated argument and take on clause-like properties in that they frequently appear as isolated, full statements. The negative marker is here possibly to be likened to an existential verb denoting the state of non-existence. It is not classified here as a verb, however, as it is not able to receive irrealis markers and itself be negated. The borrowed Malay negative marker /bokan/, which is otherwise in free variation with /bra?/, cannot be used in such constructions. The following examples illustrate the pattern:

(43) dwi? bra?
money NEG
'There is no money.'/'Money is non-existent.'

---

26 The colloquial/dialectal English form 'them' is preferred to standard English 'those' as a translation of the prenominal determining attributive plural pronoun as it corresponds better to the Jahai meaning. This is because the Jahai form lacks the locational deictic connotations which might be associated with 'those'. In Jahai NPs, locational deixis is instead expressed by postnominal determiners (see §5.1.4.1).

(44) braʔ ka=susuh
NEG SUBJ=milk
'There is no milk.'/'Milk is non-existent.'

In the case of irrealis constructions, however, clauses which do not require a subject particle behave according to the same pattern as the common clauses described in §5.1.1.1. Thus, the verb takes a subject-encoding irrealis proclitic. The core of the clause now corresponds to the verb and the irrealis proclitic. Examples are given below:

(45) ja=ghɛl
IRR=to.be.tired
'[I] will be tired.'

(46) ʔaj wa=sɔc
game.animal IRR.3S=to.be.gone
'The game will be gone.'

5.1.1.3 Negated constructions

Negation is signalled by the negative marker /braʔ/, which occurs in free variation with the borrowed negative marker /bokan/ (from Malay *bukan* 'no', 'not') before an irrealis-marked predicate, or, occasionally, before the optional subject NP. Its combination with irrealis is obligatory. Clauses containing verbs which require a subject particle (cf. §5.1.1.1) are negated in the same way as clauses containing verbs which do not (§5.1.1.2). The ability of property verbs to take the irrealis proclitic and thus to be negated motivates the treatment of them here as stative verbs rather than adjectives.

In the unmarked case, a negative construction represents a propositional negation of the clause, signalling that the entire proposition expressed by the clause is negated. Some examples follow.

(47) jɛʔ braʔ ja=ʔtʔet
1S NEG IRR=to.know
'I don't know.'

(48) braʔ heʔ ja=kɟeŋ kliŋ kritəh
NEG 1P.INCL IRR=to.hear sound car
'We didn't hear the sound of cars.'

(49) kampoŋ braʔ wa=weʔ
village NEG IRR.3S=to.exist
'There was no village.'

(50) braʔ ja=ʔɛ̃m tɛh
NEG IRR=to.drink tea
'[I] didn't drink tea.'

(51) braʔ wa=bləh ka=pɛ̃ʔ paj
NEG IRR.3S=to.enter SUBJ=older.sibling 2S.DIS
'Your brother didn't go in.'

In cases where the scope of negation is not on the whole proposition but only part of the clause, e.g. on one of its arguments, the use of the negative marker and the irrealis proclitic is identical to that of propositional negation. Such negation is disambiguated from

propositional negation by means of the contrastive prepositional proclitic /d=/ (cf. §4.9.4, §5.2), which is attached to the negated argument. This is exemplified in (52) and (53) below. In (52) it is the subject which is being negated. This is signalled by a postverbal subject PP, headed by the contrastive preposition. In (53) it is the direct object which is negated and represented by a PP headed by the same preposition.

(52) bokan wa=gej d=ʔoʔ
NEG IRR.3S=to.eat CONTR=3S
'He [as opposed to X] didn't eat.'

(53) braʔ japɛ̃h ja=pktɔk d=gin kdek
NEG 1P.EXCL IRR=to.cook CONTR=2/3P squirrel
'We didn't get to cook them squirrels [as opposed to cooking X].'

The position typically associated with the negative marker, that is the left periphery of the core of the clause, is the same as that associated with the relational tense proclitic /ɟa=/ (§4.10.1.1) and the modal auxiliary signalling root possibility /boleh/ (§4.10.1.2).

5.1.1.4 Subordination

Subordination at clause level is not so much in evidence in Jahai (for subordination at NP level, however, see §5.1.3). Such co-ordination is instead rendered by placing two main clauses in series. Subordinator-like forms, of which there are three, are borrowings from Malay (see also §4.11): causal /sbap/ 'because', temporal /lpəs/ 'after'/'when', and conditional /kaluh ~ kalɔw/ 'if'. These may introduce either the initial or the final clause of such series.

(54) ja=gej sbap jɛʔ crɔʔ
IRR-to.eat because 1S to.be.hungry
'I will eat because I'm hungry.' (elic.)

(55) lpəs jɛʔ gej jɛʔ cɨp
after 1S to.eat 1S to.go
'I left when I had eaten.'

(56) kalɔw jɛʔ crɔʔ ja=gej
if 1S to.be.hungry IRR=to.eat
'I will eat if I'm hungry.' (elic.)

However, NPs headed by action/state nominalisations sometimes behave like contracted clauses and complement main clauses in a subordination-like manner. This is exemplified in §5.1.4.3.

5.1.2 Questions

5.1.2.1 Yes/no questions

As noted in §4.10.3, yes/no questions may be formed from any statement through the attachment of an interrogative proclitic /ha=/ to the first constituent of the clause.

(57) ha=mɔh ʔtʔet
Q=2S.FAM to.know
'Do you know?'

(58) ha=gop ʔoʔ bdil kasaʔ
Q=stranger 3S to.shoot sambar.deer
'Did the stranger shoot a sambar deer?' (elic.)

(59) ha=wa=wek ba=ʔũn
Q=IRR.3S=to.go.back GOAL=there
'Will he return there?'

(60) ha=weʔ ka=ʔaj ba=ʔaniʔ
Q=to.exist SUBJ=game.animal GOAL=there
'Is there game over there?'

5.1.2.2 WH questions

WH questions are typically formed by placing the WH element in clause-initial position. This applies to all WH words. The questioned position of the corresponding statement is empty. The pattern is illustrated in the following examples:

(61) makɛn but kmɔʔ
who? to.eat fruit
'Who ate the fruit?' (elic.)

(62) mɛj paj dʔ-deʔ
what? 2S.DIS IMPF-to.do
'What are you doing?'

(63) lbah ʔoʔ ŋɔk
where? 3S to.sit
'Where does he live?'

(64) b=lbah mɔh cip
GOAL=where 2S.FAM to.go
'Where did you go?'

(65) mapuʔ mɔh ja=wek
when? 2S.FAM IRR=to.go.back
'When will you go back?'

Person-questioning /makɛn/ 'who?' and item-questioning /mɛj/ 'what/which?' typically replace subjects, direct objects or oblique arguments, time-questioning /mapuʔ/ 'when?' and location-questioning /lbah/ 'where?' typically replace temporal and locative adjuncts respectively, manner-questioning /maʔacin/ 'how?' replaces manner-related adverbial adjuncts, and reason-questioning /mɛj/ 'why?' replaces causal clauses. The WH word fully replaces the questioned position; in the case of questioned subjects, for example, note that the subject agreement marker is absent (see example (61)).

It is also possible to form WH questions by attaching the relative-marking proclitic /k=/ to the first constituent of a clause preceded by the WH word. These constructions appear to represent relativisation of the clause to the WH element, and they look like NPs with relative clauses (cf. §5.1.3). Indeed, the WH words that appear in such constructions typically question the nominal categories person and item in the form of subjects, direct objects and, less evidently, oblique arguments. The constructions are infrequent and the examples given here have all been elicited.

(66) makɛn k=but kmɔʔ
who? REL=to.eat fruit
'Who [is it that] ate the fruit?' (elic.)

(67) mamɛj k=but kmɔʔ
what? REL=to.eat fruit
'What [is it that] ate the fruit?' (elic.)

(68) mamɛj k=ʔoʔ but
what? REL=3S to.eat
'What [is it that] he ate?' (elic.)

(69) makɛn k=paj ʔɛk tmakɔw
who? REL=2S.DIS to.give tobacco
'Whom [is it that] you gave tobacco?' (elic.)

A third type of WH question involves a WH word/phrase in predicating position. Such questions have a structure similar to that of clauses with predicates that do not require a subject particle, outlined in §5.1.1.2. This is illustrated in the following examples:

(70) puŋhuluh ba=ʔaniʔ makɛn
headman GOAL=there who?
'Who is headman over there?'

(71) b=lbah ka=siput
GOAL=where? SUBJ=snail
'Where is the snail?'

(72) maʔacin ka=mɔh ton
how? SUBJ=2S.FAM that
'How about you there?'

(73) mɛj siʔ nn-kɛn ka=gop
what? number UNIT-CLF SUBJ=stranger
'How many strangers [were there]?'

There are no apparent restrictions as to which WH words may be used in this position, but location-specifying /lbah/ 'where?' is particularly frequent.

5.1.3 Relative clauses

Relativised elements occur at the right-hand periphery of NPs (see §5.1.4.1) and are introduced by means of the relative-marking proclitic /k=/, which is attached to the first constituent of the relativised element. Hence, relativised elements in Jahai are always externally headed. Not only clauses may be relativised, but also NPs and, occasionally, PPs. Such relativised elements are possibly to be likened to contracted or implicit clauses of some form. Some examples follow:

(74) mnraʔ k=cɨp ba=hɨp
people REL=to.go GOAL=forest
'the people who went to the forest' (elic.)

(75) slaj k=wih rh-rɔh
swidden REL=3D IMPF-to.clear
'the swidden that they were clearing'

(76) ʔoʔ k=tmkal
3S REL=man
'the one who is a man'

(77) koleh k=pn=ʔə̃h
cup REL=EQU=here
'the cup that [looks] like this'

The following discussion concerns only relative clauses proper, and not relativised NPs or PPs.

Relative clauses differ from ordinary clauses only in that the relativised position is empty in the relative clause (cf. the WH questions described in §5.1.2.2). Relativised positions appear to be largely restricted to subject and direct object. Examples are given below. In the case of relativised subjects, note that the subject agreement marker is absent.

Relativised subjects:

(78) bdil k=pcah
gun REL=to.break
'the gun that broke' (elic.)

(79) hawẽn k=gɐj hobiʔ
pig REL=to.eat root.crop
'the pig that ate the root-crop' (elic.)

(80) tmkal k=bdil kasaʔ
man REL=shoot sambar.deer
'the man who shot the sambar deer' (elic.)

Relativised direct objects:

(81) hobiʔ k=jɛʔ tanɛm ktɔʔ wɛj
root.crop REL=1S to.plant day past
'the root-crop that I planted yesterday' (elic.)

(82) tmakɔw k=ʔoʔ ʔɔm ka=ʔoʔ
tobacco REL=3S to.want SUBJ=3S
'the tobacco that he wants' (elic.)

(83) tɛh k=tmkal ʔoʔ ʔɛk
tea REL=man 3S to.give
'the tea that the man gave away' (elic.)

Relativised oblique arguments and other positions appear to be avoided, although recipients are occasionally accepted.

(84) tmkal k=jɛʔ ʔɛk tmakɔw
man REL=1S to.give tobacco
'the man whom I gave tobacco' (elic.)

In terms of which positions may be replaced, this situation is parallel to that described for WH questions formed by means of the relative marker (see §5.1.2.2). However, restrictions are greater for relativisations than for clausal WH questions, the former being restricted to mainly subjects and direct objects whereas the latter also involves other arguments as well as adjuncts.

Occasional examples suggest that relative clauses may occur without a head and that they may therefore make up a NP on their own. There is also one example which suggests that such independent relative clauses may take the subject-marking proclitic /ka=/:

(85) maʔacin ka=k=jɛʔ cɔl ka=ʔũn
how? SUBJ=REL=1S to.tell LOC=there
'How about what I just said?'

5.1.3.1 Subordination by means of the identification marker

A second type of NP subordination involves the identification marker /l= ~ la=/ (see §4.12.3), which is attached proclitically to the initial constituent of a NP to subordinate this NP to a preceding coreferential deictic element, typically a demonstrative pronoun. Such constructions always occur as independent statements, never forming part of clauses.

(86) tə̃h, l=bawac
this ID=pig.tailed.macaque
'This [is] a pig-tailed macaque.'

5.1.4 Noun phrases

5.1.4.1 The basic order of NP elements

The Jahai noun phrase consists typically of a referent in the form of a head noun (or nominal compound) and possible modifiers of that noun. Such modifiers include pronominal determiners and quantifiers (including numerals), which precede the head noun, and nominal modifiers (nominal and pronominal possessors), verbal modifiers (the relativised participle-like forms, see §4.12.2), deictic determiners (such as locative and temporal demonstratives, demonstrative pronouns or locative PPs) and relative clauses/ phrases (see §5.1.3), which follow the head noun. WH modifiers, which precede the head noun and display special characteristics, are treated in §5.1.4.2. The order of modifiers in relation to their head noun may be summarised as follows:

(DET:PRO) (QNT) **N** (POSS:N/PRO) (V) (DET:DEIC) (REL)

Thus, the head is preceded by modifiers associated with person and quantity, whereas it is followed by modifiers primarily signalling quality and locality. The following constructed example illustrates the maximal NP. The head is boldfaced.

(87) gin tigaʔ **wɔŋ** jɛʔ t-bəw ton k=wek can=hip
2/3P three child 1S REL-to.be.big that REL=to.return SOURCE=forest
'them three big children of mine there, who returned from the forest'

The following examples represent authentic NPs:

(88) wɔŋ kɟih jɛʔ tə̃h
child boy 1S this
'this son of mine'

(89) gin bit
2/3P ant
'them ants'

(90) hali? ʔo? ton
leaf 3S that
'that leaf of its'

(91) bɛr ʔo? wɛj
younger.sibling 3S past
'his late brother'

(92) mnra? can=tkam
people SOURCE=Tekam
'people from Tekam'

(93) kdek ka=ʔũn
squirrel LOC=there
'the squirrel over there'

Prenominal modifiers, unlike postnominal ones, are frequently associated with affixal derivation of the head noun. Thus, a quantifier usually co-occurs with a unitised noun (cf. §4.1.3), and a determining attributive non-singular pronoun typically co-occurs with the collective plural/nominalising morpheme if the head noun is human (§4.1.1, §4.1.4.2). This is exemplified in (94) and (95).

(94) duwa? h<n>rkit
two night<UNIT>
'two nights'

(95) heʔ tm<a>kal tə̄h
1P.INCL man<COLL> this
'we men here'

In the absence of a nominal head, the head position is interpreted here as belonging to another NP element, e.g. a pronoun (96), classifier (97), quantifier/numeral (98) or demonstrative (99), which then serves as a referring element on its own.

(96) gin t-kɔm ʔadeh
2/3P REL-to.be.many there
'the many ones back there'

(97) tiga? kbi?
three CLF
'three heads'

(98) hɛj duwa?
1D.INCL two
'we two'

(99) tə̄h
this
'this one'

Interrogatives (WH words) may also in themselves serve as NP heads, as shown by several of the examples in §5.1.2.2.

With head nouns representing single concrete referents, the postnominal position usually associated with possessor modifiers is sometimes filled by a third person singular pronoun /ʔoʔ/, which cannot be interpreted as designating a possessor but which instead is

coreferential with the head itself. Such readings are exemplified below. Note that the postnominal third person pronoun is referentially ambiguous and can also be used to denote a possessor.

(100) ʔap ʔoʔ
tiger 3S
'the tiger'

(101) mawẽʔ ʔoʔ ton
gibbon 3S that
'that gibbon'

(102) pẽw ʔoʔ ton
other 3S that
'that other one'

The significance of this reiteration of the referent of the head is unclear. It bears some functional similarity to the suffix *-nya* in Malay, which may signal both possession and definiteness.

Another feature associated with the possessor position of NPs is that NP heads representing kinship terms and words denoting body parts or other parts-of-a-whole almost invariably display a possessing modifier, usually in the form of a pronoun. This is exemplified in (103) and (104).

(103) kaɲcɔʔ ʔoʔ ton
grandchild 3S that
'that grandchild of his'

(104) ʄʔeŋ bliʔ ʔoʔ tə̄h
bone thigh 3S this
'this thighbone of its'

5.1.4.2 WH modifiers of NPs

WH modifiers of NPs include possessor-questioning /makɛn/ 'whose?', item-questioning /mɛj/ 'what/which?' and number-questioning /mɛj siʔ/ 'how many?' (see §4.6). These are located in NP-initial position and correspond to other NP modifiers, which they question and replace. More specifically, possessor-questioning /makɛn/ 'whose?' replaces nominal or pronominal possessor modifiers; item-questioning /mɛj/ 'what/which?' replaces deictic modifiers, verbal modifiers, as well as relative clauses; and number-questioning /mɛj siʔ/ 'how many?' replaces the quantifier modifier. In the following examples, proper (but constructed) answers to the WH-modified NPs have been included in order to illustrate the position of the questioned element (cf. also the order of modifiers given in §5.1.4.1):

(105) makɛn ʔɔ̃t
whose? dog
'whose dog?'

ʔɔ̃t jɛʔ
dog 1S
'my dog!'

(106) mɛj hajẽʔ
which? house
'which house?'

hajẽʔ tũn
house that
'that house!'

(107) mɛj si? k<n?>tɔ? tiga? k<n?>tɔ?
what? number day<UNIT> three day<UNIT>
'how many days?' 'three days!'

This NP-initial location of WH modifiers of NPs, as well as their mutual exclusivity with the corresponding questioned modifiers, is analogous to the situation encountered in clausal WH questions (§5.1.2.2), in which the WH element is found in clause-initial position and the position of the questioned element is empty.

5.1.4.3 Nominalisations

Within the NP, nominalisations (cf. §4.1.4) generally behave like ordinary nouns, forming a NP head and taking the same modifiers. This applies to the action/state and concrete nominalisations formed by means of the /n/ nominaliser as well as the human nominalisations formed by means of the collective plural morpheme. Some examples of such NPs follow:

(108) nj-gej paj
NM-to.eat 2S.DIS
'your food'

(109) c<n>kwɨk ʔoʔ ton
to.speak<NM> 3S that
'that way of speaking of his'

(110) nk-wek c an=slaj crjkaj
NM-to.return SOURCE=swidden *Cerikai*
'the return from *Cerikai* swidden'

(111) gin g<ra>hɛl taniʔ
2/3P to.be.tired<COLL> that
'them tired ones over there'

As part of heads in the form of compounds, nominalisations may represent either the head or the modifying element of the compound:

(112) t<n>anɛm hobiʔ
to.plant<NM> root.crop
'tuber-planting'

(113) tmpət nk-ŋɔk
place NM-to.sit
'living-quarters'/'place of residence'

With heads in the form of action nominalisations, nominal and pronominal modifiers in possessor position infrequently take the form of a PP headed by the contrastive proclitic /d=/, as exemplified in (114).

(114) nt-gat d=japɛ̃h
NM-to.waylay CONTR=1P.EXCL
'waylaying of ours'

This construction has not been documented with other types of NP heads in the present material. However, Schebesta (1928a:814) speculates about the possibility of an optional genitive marker *de* and provides the following example (represented in Schebesta's original

orthography): *kilad de karɛi* 'Karei's lightning-flash'. In the present orthography this would correspond to the following phrase:

(115) kilat d=karɛj
flash.of.lightning CONTR=Karei
'Karei's flash of lightning'

It is possible that the contrastive preposition is used in this position to emphasise the possessor in some way (cf. the use of the contrastive preposition to mark arguments in the clause, described in §5.2).[27] However, it is not clear why it appears to be restricted in the present material to NPs headed by an action nominalisation.

The syntactic behaviour of NPs headed by nominalisations is similar to that of ordinary NPs. Thus, they may function as arguments and adjuncts in the clause, and they can be headed by prepositional proclitics. This is exemplified in the following clauses:

(116) c<n>kwɨk ʔoʔ btʔɛt
to.speak<NM> 3S to.be.good
'His way of speaking is good.'

(117) jɛʔ ʔaŋket k<n>ajil
1S to.take to.fish<NM>
'I took the fishing-rod.'

(118) jɛʔ sam ba=nm-sam ʔaj
1S to.hunt GOAL=NM-to.hunt game.animal
'I hunted at the [place of] game-hunting.'

However, action/state nominalisations also occur in NPs which function as clausal complements and which may be described as contracted clauses. Such complements are sometimes introduced by subordinators and other elements typically associated with clauses and do appear to represent a form of subordination. Examples are given below.

(119) ʔakɛt cɨp ba=hɨp, ɟa=np-cɨp ʔap lɛh
PROH to.go GOAL=forest RT=NM-to.go tiger EMP
'Don't go to the forest. The tiger is already on the move!'
[lit. 'Don't go to the forest; already tiger-movement!']

(120) lpəs nk-ɟok japɛ̃h wek
after NM-to.move 1P.EXCL to.go.back
'After moving we went back.'

The characteristics of such constructions are not known in detail.

5.1.5 Prepositional phrases

Jahai adpositions usually occur as proclitics attached to the first constituent of the NP[28] and are thus prepositions. All such constructions are structurally identical and will be referred to generically as prepositional phrases. Their structure may be summarised as follows:

---

27 Semelai exhibits possessive markers /də= ~ dɔ/ (Kruspe 2004:92), and Rischel (1995:140–141, 145) describes possession-related usage of the words /di/ and /dɔ/ in Minor Mlabri.

28 The only exception to this is the *equative* preposition, a free-form variant of which is used with nouns (see §4.9.5).

PREP=[NP]

In terms of function, however, prepositions may be broadly divided into two types. One type is predicative and licenses the occurrence of its NP in a clause. PPs headed by such prepositions represent either oblique arguments or adjuncts. Such PPs are illustrated in (121) and (122).

(121) k=hajɛ̄ʔ
LOC=house
'in the house'

(122) ba=wɔŋ jɛʔ
GOAL=child 1S
'to my child'

The second type of preposition does not license the occurrence of its NP. This non-predicative type of preposition is first and foremost associated with PPs denoting the postverbal subject (see §5.1.1.1.2 and §5.2.1.1) and is then essentially a case marker. An example is given in (123).

(123) ka=wɔŋ ʔoʔ
SUBJ=child 3S
'his child ...'

A non-predicative function is also associated with the contrastive preposition /d=/, which may be characterised as a pragmatic marker.

The function of prepositions in relation to arguments is further discussed in §5.2.

5.2 The arguments

The following sections describe arguments in the clause. These are organised into three categories: the subject, the direct object, and oblique arguments. This division is motivated by the syntactic behaviour of arguments in terms of their position and phrasal identity as outlined in §5.1. The subject (§5.2.1) is typically associated with a preverbal position and is usually represented by an agreement marker. The direct object (§5.2.2) is typically a NP in postverbal position. Oblique arguments (§5.2.3) are typically postverbal PPs and represent a range of argument relations. A summary of prepositional usage with arguments is given in §5.2.4.

5.2.1 The subject

As has been shown, the syntactic subject of the clause is represented in common clauses by an obligatory preverbal agreement marker in the form of an unstressed but otherwise phonologically free personal pronoun (the subject particle) or, in irrealis constructions, by an irrealis proclitic, which vaguely encodes person of subject. The subject particle and the irrealis proclitic are mutually exclusive. This subject agreement marker may display cross-reference with an optional full phrase denoting the subject. If occurring to the left of the verb, this phrase is represented as a NP; postverbally it is represented as a PP.

As was noted in §5.1.1.2, a different pattern applies to clauses whose predicate is represented by a member of a subset of stative verbs (including existential verbs and most property verbs), an onomatopoeic form, a WH word or a negative marker. Here, the subject particle is typically absent and the subject is represented only by the optional

preverbal NP or postverbal PP. In irrealis constructions, however, it is also represented by the agreement-marking irrealis proclitic, as in common clauses.

Subject is considered here to be a purely syntactic relation which is largely independent of semantic relations. This is suggested e.g. by the fact that the subject agreement marker, along with its possible cross-reference phrases, may denote participants representing diverse semantic roles. The most salient thematic roles so far identified for the subject are exemplified in the following sentences.

AGENT:

a. jɛʔ tbɔh paj
1S to.hit 2S.DIS
'I hit you.'

b. gin jɨj bras
2/3P to.carry rice
'They carried rice.'

EXPERIENCER:

a. ʔoʔ ʔɛl-ʔɛl ka=ʔap cn=ʔaniʔ
3S CONT-to.look SUBJ=tiger SOURCE=there.DIST
'The tiger kept on watching from over there.'

b. ha=miʔ ʔtʔet jaʔ ʔoʔ
Q=2S.INT to.know grandmother 3S
'Do you know his grandmother?'

INSTRUMENT:

a. taɟiʔ ʔoʔ get
knife 3S to.cut
'The knife made a cut.' (elic.)

b. ʔoʔ kac ka=cnrɔs ʔoʔ
3S to.scratch SUBJ=claw 3S
'Its claw made a scratch.'

FORCE:

a. ɟa=ʔoʔ tpis ka=tɔm
RT=3S to.wash.away SUBJ=river
'Then the river washed [it] away.'

b. ʔɔs ʔoʔ gɨj
fire 3S to.produce.smoke
'The fire produced smoke.' (elic.)

PATIENT:

a. ʔap ʔoʔ kbis
tiger 3S to.die
'The tiger died.'

b. ɟa=sɔc ka=snlɔc
RT=to.be.gone SUBJ=blowpipe.dart
'The darts are already used up!'

THEME:

a. ʔoʔ ɟlɔɸ ka=pluruh klɛŋ hɛ̃ŋ ʔoʔ
3S to.whiz SUBJ=bullet inside mouth 3S
'The bullet went into its mouth.'

b. batuʔ ʔoʔ rɛs ba=tɔm
stone 3S to.fall GOAL=water
'The stone fell into the water.' (elic.)

Thus, the syntactic subject may be represented by arguments which are either animate or inanimate, volitional or non-volitional, agent-like or patient-like, and so on, and the syntactic relation of subject therefore appears to neutralise the semantic relations. This relative independence of the subject from semantic relations can also be seen in a few exceptional cases where the progressive aspect morpheme added to verbs has a valence-reducing effect, whereby a theme or patient which would be given object status in an active sentence is turned into subject in what looks like a passive construction. This is further discussed in §5.4.2.

Just like the syntactic relation of subject can be shown to neutralise semantic relations, it appears to also operate independently of pragmatic relations. This is most apparent in the coding properties of the mutually exclusive preverbal subject NP, associated with topic, and postverbal PP, associated with focus. Although these two reflect different pragmatic functions, their coding properties are identical in that they are both cross-referenced by the subject agreement marker, suggesting that they represent the same syntactic relation. It is not possible therefore to place a pragmatic relation, say topic, on a par with a particular syntactic relation. In this case, the syntactic relation of subject neutralises the pragmatic relations of topic and focus.

The pattern of subject assignment is also independent of the valence of the verb. In other words, the location and marking of the subject is identical in intransitive and transitive constructions. This is exemplified in the following clauses involving postverbal subject PPs:

(124) jɛʔ tek ka=jɛʔ
1S to.sleep SUBJ=1S
'I slept.'

(125) ʔoʔ sjiŋ lɛh ka=ʔɛj jɛʔ wɛj slaj
3S to.burn EMP SUBJ=father 1S past swidden
'My old father burned a swidden.'

The preverbal pronoun is here cross-referenced by the postverbal PP, which in both cases is headed by the preposition /ka=/. No syntactic distinction is made between the subject of intransitive and transitive clauses, so Jahai, it seems, does not display the ergative-like patterns described for some other Aslian languages, including Jah Hut (Diffloth 1976b:92–94) and Semelai (Kruspe 2004:106), or the unusual three-way system of alignment ascribed to Temiar by Benjamin (1996).

The lack of a subject particle in clauses with some stative predicates calls for some elaboration. The stative predicates in question are typically property or existential verbs, and the essentially non-volitional and patient-like character of their single argument could be argued to have something to do with the fact that it is not represented by a subject particle. This may well be the case, but it should be noted that the argument in every other way behaves like other subjects. Thus, it turns up as a full phrase in the usual subject slots, taking a subject-marking prepositional proclitic in postverbal position, and in irrealis constructions (which include negated clauses) it is encoded in the irrealis proclitic. Therefore, the presence vs. absence of a subject particle cannot be conditioned by different syntactic identities of the argument.

5.2.1.1 The postverbal subject PP

As noted, a full phrase denoting the subject takes the form of a prepositional phrase in postverbal position. Such PPs signal that the subject is in focus. Two prepositional proclitics may be used, /ka=/ and /d=/, the distribution of which appears to be determined by syntax and pragmatics respectively.

5.2.1.1.1 Subject-marking /ka=/

Postverbal subject PPs are typically headed by the prepositional proclitic /ka=/, glossed in such cases as 'SUBJ'. The function of this PP is to disambiguate the referent of a subject from other potential subject referents in the discourse. For example, it is normally used in answer to 'who?' questions. Thus, it appears to be associated with subject focus. Examples are given in (126) and (127).

(126) ɟa=ʔoʔ lɔj wel ka=pɛ̃w ba=ʔaniʔ
RT=3S to.run again SUBJ=other GOAL=there
'Then *the other one* fled over there again.'

(127) jɛʔ kul ka=jɛʔ
1S to.call SUBJ=1S
'*I* called.'

But whereas the presence of the PP has pragmatic motivation, the form of its prepositional proclitic is syntactically determined in that it cross-references the PP with the subject agreement marker, and it does so irrespective of the semantic relation of its referent; hence the treatment of it here as a syntactic marker and the glossing of it as 'SUBJ'.

The situation is complicated by the fact that the semantic relation of instrument is similarly marked with the proclitic /ka=/. Such PPs are not cross-referenced by the subject agreement marker; hence their status is distinct from that of subject. Occasionally this leads to ambiguity between subject and instrument, as in the following two overtly identical clauses:

(128) ʔoʔ get ka=taɟiʔ
3S to.cut SUBJ=knife
'The knife made a cut.' (elic.)

(129) ʔoʔ get ka=taɟiʔ
3S to.cut INSTR=knife
'He made a cut with the knife.' (elic.)

The distinct status of subject versus instrument is also evidenced by the fact that both PPs may occur in the same clause, as shown in (130). The first of the two PPs is cross-referenced by the subject particle.

(130) ʔoʔ get ka=tmkal ka=taɟiʔ
3S to.cut SUBJ=man INSTR=knife
'The man made a cut with the knife.' (elic.)

In sum, argument-marking /ka=/ is primarily associated with the syntactic relation of subject, as PPs headed by it are typically cross-referenced by the subject agreement marker. So, whereas the presence of the PP is pragmatically motivated, the form of the preposition

itself is syntactically determined. Other use of argument-marking /ka=/ is restricted to the semantic relation of instrument. Thus, although homonymous, subject-marking /ka=/ is treated here as profoundly distinct from the instrumental-marking proclitic. It is possible, however, that both originate in the similarly homonymous locative proclitic.[29]

5.2.1.1.2 Contrastive /d=/

The subject may also be represented by a postverbal PP headed by the contrastive proclitic /d=/, glossed as 'CONTR'. As is the case with subject PPs headed by /ka=/, the presence of those headed by /d=/ is pragmatically determined as they are associated with focus. Unlike /ka=/, however, /d=/ is not specifically associated with the syntactic relation of subject. As will be further shown in §5.2.2.1 and §5.2.3, contrastive /d=/ can also head other postverbal arguments of the predicate and therefore appears to operate independently of syntactic and semantic relations. As its name implies, its function is instead to contrast an argument with other explicit or implicit participants which are potential competitors for the same relation. Therefore it is not only the presence of the PP that is pragmatically determined, but also the form of the preposition itself. Some examples of subject PPs headed by /d=/ follow.

(131) ʔoʔ lɔj d=ʔoʔ
3S to.run CONTR=3S
'*He* [as opposed to X] ran.'

(132) japɛ̄h wek lɛh d=japɛ̄h
1P.EXCL to.go.back EMP CONTR=1P.EXCL
'*We* [as opposed to X] went back.'

(133) ɟa=ʔoʔ cip lɛh d=ʔap tə̄h
RT=3S to.go EMP CONTR=tiger this
'Then *this tiger* [as opposed to X] left.'

(134) jɛʔ gej d=jɛʔ duwaʔ kpiŋ
1S to.eat CONTR=1S two CLF
'*I* [as opposed to X] had two helpings.'

(135) tadɔʔ k=tə̄h d=miʔ, ja=cip ja=ʔɛl ba=ʔaniʔ
to.wait LOC=this CONTR=2S.INT IRR=to.go IRR=to.look GOAL=there
'*You* wait here. I will go and have a look over there.'

So whereas subject PPs headed by /ka=/ are associated with a subject in focus, those headed by /d=/ are associated with a contrasted argument in focus which happens to be the subject.

5.2.2 The direct object

As was shown in §5.1, the direct object of a clause is typically represented by a postverbal NP; it is the only argument which is not required to be represented by a PP in postverbal position. Alternatively, the direct object NP may occur to the left of the preverbal subject marker or the full subject NP if such an NP exists. The ability to occur as

---

[29] The semantic motivation for such an origin is unclear, however, although it could be suggested that a notion of 'location of action/state' would explain the connection.

a NP in postverbal position is what distinguishes the direct object syntactically from other arguments.

Some direct object-like arguments represented by an unmodified noun are non-referential (so-called *inherent arguments*), in which case they are considered an intrinsic part of the meaning of the predicate and therefore do not represent a true argument. The postverbal noun in the following example could be interpreted in this way:

(136) ja=sam ʔaj
IRR=to.hunt game
'[I] will hunt game.'

Such non-referential arguments, which are difficult to disambiguate from referential ones, will not be considered further here.

The most salient thematic roles so far identified for the direct object include those of PATIENT, THEME, PERCEPT and BENEFICIARY. These are exemplified below.

PATIENT:

a. ja=gej bap jɛʔ
IRR=to.eat food 1S
'[I] will eat my food.'

b. ʔoʔ tampɛr jɛʔ
3S to.claw 1S
'It clawed me.'

THEME:

a. wa=jɨj ʔɔs k=taniʔ
IRR.3S=to.carry.in.hand fire LOC=there
'He will carry the fire over there.'

b. jɛʔ ɟɪŋ ka=jɛʔ lɛh wel bdil
1S to.take SUBJ=1S EMP again gun
'I took the gun again.'

PERCEPT:

a. miʔ ʔɛl hajɛ̃ʔ jɛʔ
2S.INT to.see house 1S
'You see my house.'

b. jɛʔ jiʔ ʔoʔ
1S to.dislike 3S
'I dislike him.'

BENEFICIARY:

a. japɛ̃h b-pimpin gin
1P.EXCL PROG-to.guide 2/3P
'We were guiding them.'

b. ja=pr-gej wɔŋ jɛʔ
IRR=CAUS-to.eat child 1S
'[I] will support my child.'/'[I] will feed my child.'

Other, less prominent roles would include e.g. LOCATION and DESIRE. When compared to those of the subject, the thematic roles associated with the direct object display less semantic variation. They all seem to fit into a generalised semantic role that Van Valin and LaPolla (1997:141) refer to as the *undergoer* macrorole. Thus, there is a clearer connection between syntactic and semantic relations in the case of direct object than in the case of subject.

5.2.2.1 Oblique direct objects

The overwhelming majority of direct objects are represented by NPs and, as noted, the ability to take on this phrasal identity in postverbal position is what distinguishes direct objects syntactically from other arguments, which are always represented by PPs postverbally. Sometimes, however, arguments which behave semantically like unequivocal direct objects appear as PPs. In many cases, these involve the pragmatically determined contrastive proclitic /d=/, which also has the ability to attach to other postverbal arguments to signal that an argument is contrasted with other explicit or implicit participants which are potential competitors for the same relation (cf. the contrasting of subjects in §5.2.1.1.2 and the summary in §5.2.4). Such direct object PPs are exemplified in (137), (138) and (139).

(137) ɟa=ʔoʔ bdil lɛh d=hawɛ̄n
RT=3S to.shoot EMP CONTR=pig
'Then he shot *the pig* [as opposed to shooting X].'

(138) japɛ̄h kɟeŋ d=mawɛ̄ʔ ʔoʔ ton
1P.EXCL to.hear CONTR=gibbon 3S that
'We heard *that gibbon* [as opposed to hearing X].'

(139) braʔ japɛ̄h ja=pktɔk d=gin kdek
NEG 1P.EXCL IRR=to.cook CONTR=2/3P squirrel
'We didn't get to cook *them squirrels* [as opposed to cooking X].'

Furthermore, some verbs seemingly allow direct objects to appear either as the usual NPs or as PPs headed by the locative /k=/ or goal /ba=/ prepositional proclitics. The infrequency of such PPs prevents a more thorough analysis, but Benjamin's (1996) account of the goal-marking preposition *ma-* 'to'/'towards' in neighbouring Temiar offers an illuminating parallel. In Temiar, the goal marker may be added to the transitive object of a verb of action 'to mark it as the Goal towards which action is directed, rather than as a Patient which undergoes the action and is affected by it' (Benjamin 1996:51). This implies that the action is attempted, partial or benefactive rather than leading to a complete change of state, and that emphasis is placed on the directing outwards of the actor's action rather than its result. Clear English parallels may be found in sentences like *He shot the pig* versus *He shot at the pig*, or *She ate the apple* versus *She ate away at the apple*. Presumably such readings are to be given to the Jahai examples below (140b–d), which also suggest that a further distinction is made in Jahai between the goal to which an action is directed, signalled by /ba=/, and the location at which an action takes place, signalled by /k=/.

(140) a. ja=gej bap jɛʔ
IRR=to.eat food 1S
'[I] will eat my food.'

b. pɛ̄h gej ba=tahɛʔ
1P.EXCL to.eat GOAL=type.of.salty.tuber
'We ate away at the tubers.'

c. gej sʔoʔ k=nasiʔ tūn
to.eat just LOC=rice that
'Just eat away at that rice.'

d. gin get ba=barɔŋ
2/3P to.cut GOAL=tapir
'They cut away at the tapir.'

As noted by Benjamin for Temiar, these constructions are not as completely transitive as those involving an ordinary direct object. Structurally they behave like the true oblique arguments described in §5.2.3.

5.2.3 Oblique arguments

Arguments other than those representing the syntactic relations of subject and direct object are invariably oblique; that is, they are always marked with prepositions. Clause-internally they are found only in postverbal position. These features characterise them syntactically. Such arguments are referred to here generically as *oblique arguments*.

A number of prepositional proclitics have the ability to head oblique arguments, including /k= ~ ka=/ (which usually introduces arguments representing the thematic roles of LOCATION and INSTRUMENT), /ba=/ (GOAL and RECIPIENT) and /can=/ (SOURCE). The contrastive preposition /d=/ sometimes replaces /ba=/ to mark a RECIPIENT. Typical examples are given below.

LOCATION: ʔaj ʔisiʔ k=buloh
game.animal to.insert LOC=bamboo.tube
'Put the meat in the bamboo tube.'

INSTRUMENT: mamɛj mɔh ja=dʔ-deʔ ka=ʔat ton
what? 2S.FAM IRR=IMPF-to.do INSTR=stick that
'What will you be doing with that stick?'

GOAL: gin ja=hantər ba=kdah
2/3P IRR=to.send GOAL=Kedah
'They were going to send [X] to Kedah.'

RECIPIENT: ɟa=jɛʔ ʔɛk ba=wɔŋ kɟɨh jɛʔ ton
RT=1S to.give GOAL=child boy 1S that
'Then I gave [X] to that son of mine.'

SOURCE: ʔoʔ hantər krbɔw can=taj lɛn
3S to.send buffalo SOURCE=Thailand
'He sent buffaloes from Thailand.'

However, these semantic relationships hold only as a general outline of the most typical meanings of prepositions, as there is no absolute correlation between prepositions and thematic roles. Instead, many thematic roles exhibit a marked flexibility in terms of which prepositions they are associated with. The choice of different prepositions for arguments representing one and the same thematic role reflects subtle semantic distinctions pertaining to e.g. motion and intention. This is illustrated in (141) by RECIPIENTS, which like GOALS are generally introduced by goal-marking /ba=/, but which may also be introduced by contrastive /d=/ or location-marking /k=/, each giving a slightly different meaning to the notion of receiving.

(141) a. ʔɛk ba=wɔŋ kɟih jɛʔ
to.give GOAL=child boy 1S
'Give [X] to my son!'
[with focus on change of location]

b. ɟa=jɛʔ ʔɛk d=wɔŋ kɟih jɛʔ tə̃h
RT=1S to.give CONTR=child boy 1S this
'Then I gave [X] to this son of mine.'
[with focus on change of possessor]

c. ʔɛk k=wɔŋ kɟih jɛʔ
to.give LOC=child boy 1S
'Give [X] to my son!'
[with focus on location of recipient]

Similar flexibility may be illustrated by arguments representing the thematic role of LOCATION, which can be introduced by the location-marking /k=/ or the goal-marking /ba=/.

(142) a. ʔoʔ boh k=hajẽʔ
3S to.put LOC=house
'He put [X] in the house.' (elic.)
[with focus on location]

b. ʔoʔ boh ba=hajẽʔ
3S to.put GOAL=house
'He put [X] in the house.' (elic.)
[with focus on change of location]

Note that the semantic relation of RECIPIENT is not treated as a separate relation distinct from other oblique arguments. Thus, no separate syntactic relation of indirect object is posited. This is because RECIPIENTS generally do not behave differently from GOALS, e.g. they require a licensing preposition which, as we have seen, is variable and which does not mark a specific syntactic relation. However, RECIPIENTS do hold a special status among oblique arguments in that they appear to be the only such argument which can take the contrastive proclitic /d=/. The significance of this ability is unclear.

5.2.4 The functions of prepositions: a summary

As is evident from the discussion in the previous sections, prepositions take on a variety of types and functions. Three categories will be summarised here. The first category corresponds to predicative prepositions. These semantically determined prepositions license the occurrence of a NP in a clause, and the resulting PP functions either as an adjunct or as an oblique argument. The prepositions /k= ~ ka=/ 'LOC', /ba=/ 'GOAL', /can=/ 'SOURCE' and /ka=/ 'INSTR' may be used to mark oblique arguments, whereas /k= ~ ka=/ 'LOC', /ba=/ 'GOAL' and /pn= ~ pon/ 'EQU' may occur as adjunct prepositions.

A second category corresponds to argument-marking non-predicative prepositions. These syntactically determined prepositions, which do not license the occurrence of a NP, are essentially case markers. The resulting PP functions as a specific syntactic argument. Only one such preposition is posited for Jahai, subject-marking /ka=/, which heads subject PPs in postverbal position. Homonymous to the prepositions marking instrument and,

partly, location, subject-marking /ka=/ may possibly originate in the system of predicative prepositions, from where it has developed into a syntactic marker.

The third category of prepositions, represented only by contrastive /d=/, poses a problem in that it cannot be clearly associated with fixed semantic or syntactic relations. Thus, it may introduce the syntactic arguments of both subject and direct object, as well as the oblique argument of recipient. Recall also that it occurs infrequently as a possessive marker on possessors (§5.1.4.3). However, with none of these relations does it represent the primary type of marking; it either replaces a standard marker (as in the case of subject and recipient) or it attaches where there is typically no marker at all (as in the case of direct object and possessor). It therefore has no 'home base' in terms of syntactic or semantic relations and 'floats' more or less freely on top of these. As noted, however, its occurrence can be largely seen as pragmatically motivated, because wherever it attaches it appears to single out its argument and contrast it with other explicit or implicit participants competing for the same relation. This is particularly evident in the case of the syntactic relations of subject and direct object (cf. §5.2.1 and §5.2.2), but also in the case of recipients is it used primarily to contrast its argument, albeit perhaps with greater semantic consequences (a contrasted RECIPIENT becomes BENEFICIARY; see e.g. example 141b). Like subject-marking /ka=/, contrastive /d=/ is non-predicative in that it does not license the occurrence of its NP (although in the case of RECIPIENTS it replaces prepositions that do). At the same time, however, it differs from both subject-marking /ka=/ and most predicative prepositions in that it is not associated with a specific argument. It thus appears to represent a distinct type of pragmatically assigned prepositions. It is likely that further investigation will reveal additional aspects of the nature of contrastive /d=/, wherefore the present characterisation is to be regarded as tentative.

5.3 Valence and transitivity

This section provides a brief discussion of the problems of valence and transitivity of Jahai verbs and their consequences for the argument structure of the clause. It is an area which has not been charted in detail, so the intention here is simply to convey the basic patterns that emerge from the spontaneous text material.

5.3.1 The optional nature of arguments

As has been noted, the only syntactically obligatory portion of most Jahai clauses is the verb and the subject agreement marker, represented by the subject particle or the subject-encoding irrealis proclitic. While this is true for the most part, the claim has to be modified somewhat. First, recall that a subset of stative verbs (as well as certain other predicating elements) do not require the subject particle, and, although they are frequently accompanied by a subject in the form of a NP or a PP, such verbs may occur in isolation. Notably, their citation form does not include the subject-marking element. If negated, however, they take the irrealis proclitic just like other verbs, and this is the reason for the treatment of them here as verbs rather than a separate class of adjectives. Second, the subject particle occurs in statements and questions, but no second person marking is involved in imperative constructions. Third, the subject particle may be left out of any clause if this clause forms part of a continuous sequence of clauses denoting the actions of one and the same subject, which has been introduced in the first clause of the sequence. This is illustrated in (143).

(143) hɛj pɛk can=ʔəh, boh, pɛk can=ʔəh, boh ba=ʔəh
1D to.chop SOURCE=here to.put to.chop SOURCE=here to.put GOAL=here
'We chopped [off a piece] from here [and] put [it down]. Chopped from here [and] put [it] here.'

Last, some speakers occasionally drop the subject particle if the predicate is followed by a subject PP, as illustrated in (144). The resulting clause is reminiscent of those involving the stative verbs mentioned above. However, it is not possible to omit the pronoun if the subject is represented by a preverbal NP.

(144) lɔj ka=ʔaj prəw ba=ʔaniʔ
to.run SUBJ=game.animal [sound] GOAL=there.DIST
'The animal fled [sound] over there.'

In spite of these variations, the predominant pattern that emerges is that a subject agreement marker is necessary for most clauses to be grammatical and complete. It is noteworthy that citation forms of verbs other than the stative ones just mentioned include a subject particle or irrealis proclitic (cf. §4.7).

However, overt subject arguments — in the form of preveral NPs or postverbal PPs — are always syntactically optional. The subject agreement marker is frequently sufficient enough for the identification of the subject referent. Hence Jahai may be characterised as a pro-drop language. Similarly, no verb requires an overt direct object for a clause to be grammatical and complete. This is in spite of the fact that clauses lack direct object agreement. In fact, typical transitive verbs are more often used in an overtly intransitive manner. Some examples are given below.

(145) ʔoʔ ɟiŋ lɛh ka=kaɲcɔʔ ʔoʔ ton, ʔoʔ boh ba=hɨp
3S to.take EMP SUBJ=grandchild 3S that 3S to.put GOAL=forest
'That grandchild of his took [X] and placed [X] outside.'
[The phrase /ba=hɨp/ 'to the forest' is synonymous to 'outside']

(146) ɟa=gin bʔbɔʔ
RT=2/3P to.carry.on.back
'Then they carried [X].'

(147) ʔɛk ba=wɔŋ kɟɨh jɛʔ
to.give GOAL=child boy 1S
'Give [X] to my son!'

(148) ʔoʔ prisɛʔ ka=knʔac ʔon
3S to.keep SUBJ=father-in-law there
'That father-in-law kept [X].'

But although these clauses are considered to be fully grammatical and complete, it is clear that they are typically used only when the referent of the absent direct object in some respect represents 'known information'. Either it may have been introduced overtly earlier in the discourse, or it may constitute common knowledge of the speaker and addressee which is not introduced explicitly. A previously introduced referent need not be sequentially connected to the object-less clause. This distinguishes omitted direct objects from the omitted subject agreement markers described above, which implied overt agreement at the beginning of a tightly connected sequence of clauses (cf. example (143) above). As to implicit common knowledge, this may in some cases be encoded in the semantics of the verb. For example, in (146) above, the unmarked direct object referent of

the verb /bʔbɔʔ/ 'to carry on one's back' is some type of quarry, and the verb itself therefore encodes enough information for the identification of the object. Only when this object needs to be disambiguated or specified in some way is it necessary to introduce it overtly.

Similarly, potentially ditransitive verbs do not require an overt direct object or oblique argument if these are known from the discourse context. They may occur with subject agreement only, or with subject agreement and either the direct object or the oblique argument, or with all three. This is exemplified below with clauses involving the three-place verb /ʔɛk/ 'to give'. Note that clauses exhibiting both direct object and an oblique RECIPIENT overtly, exemplified in (152) and (153), are rather uncommon in spontaneous speech.

(149) jɛʔ ʔɛk
1S to.give
'I gave [X to X].'

(150) gin ʔɛk guləh
2/3P to.give sugar
'They gave sugar [to X].'

(151) ɟa=jɛʔ ʔɛk k=wɔŋ
RT=1S to.give LOC=child
'Then I gave [X] to the child.'

(152) ʔɛk d=jɛʔ ʔɔs
to.give CONTR=1S fire
'Give me the lighter!'

(153) ʔoʔ ʔɛk taɟiʔ ba=wɔŋ
3S to.give knife GOAL=child
'He gave the knife to the child.' (elic.)

5.3.2 Transitivity and imperfectivity

Kruspe (2004:111–115) describes for Semelai a clear association between imperfective aspect and lowered transitivity, as does Diffloth (1976b:96) for Jah Hut and Benjamin (1976b:171–172; 1996) for Temiar. There may be a tendency towards a similar association in the present Jahai material (see §4.7.1.1 for a description of Jahai imperfectives), although the situation is obscured by the fact that direct objects are frequently dropped and that transitive verbs thus frequently behave intransitively also in their root form, as shown in §5.3.1. Thus, it is difficult to identify a clear association between verbs in the imperfective form and a lack of direct objects, given that direct objects are syntactically optional anyway.

As also noted by Kruspe (2004:114) for Semelai, direct objects that turn up in combination with verbs in the imperfective can often be interpreted as inherent arguments — that is, non-referential arguments which form an intrinsic part of the meaning of an activity verb (cf. §5.2.2) and which therefore do not represent a true transitive object. Possible such readings are illustrated in the following examples, where an unmodified direct object noun follows immediately after the verb.

(154) gin ja=p<l>ʔɔl hobiʔ
2/3P IRR=to.roast<IMPF> tuber
'They will be roasting tubers.'

(155) heʔ ja=b-gj-gej kmɔʔ dorɛn
1P.INCL IRR=PROG-IMPF-to.eat fruit durian
'We will be eating durian fruit.'

(156) ʔoʔ gu<m>lɛm ʔat
3S to.carry.on.shoulders<IMPF> stick
'He was carrying a stick.'

A non-referential interpretation is possible also in clauses where the argument is not found to the immediate right of the verb, as when separated from the verb by some element, or fronted to a clause-initial position. This is illustrated below.

(157) japēh ch-cih lɛh slaj
1P.EXCL IMPF-to.cut EMP swidden
'We were clearing a swidden.'

(158) ʔikəʔ jɛʔ gj-gej
fish 1S IMPF-to.eat
'I was eating fish.'

No clear examples have been identified in which a verb in the imperfective is combined with a direct object modified by determiners, relative clauses or the like, which would seem to indicate that such arguments are required to be non-referential.

5.4 Valence-affecting operations

Valence-affecting operations so far documented in Jahai include valence-increasing causativisation (§5.4.1) and a marginal pattern of valence reduction associated with the progressive morpheme (§5.4.2).

5.4.1 The causatives

As noted in §4.7.2, Jahai has a number of causative affixes, the semantic differences of which are unclear. What they all have in common, however, is that they increase the valence of the verb and thereby allow for the introduction of an additional argument, resulting in changes in the grammatical relations. Thus, if an intransitive base verb is causativised, the subject of the base verb is turned into a direct object representing a causee, and a new argument representing the causer becomes subject. The causer always has to be animate and volitional. This is illustrated in (159).

(159) a. ʔoʔ kbis
3S to.die
'It died.'

b. jɛʔ k<ri>bis ʔoʔ
1S to.die<CAUS> 3S
'I killed it.'

Transitive and ditransitive verbs are rarely causativised. If they are, the same argument pattern applies as in the case of intransitive verbs: the subject of the base verb is turned into a direct object representing a causee, and a new argument representing the causer becomes subject. The remaining arguments, including the direct object of the base verb, have not been found to be overt in such constructions.

As other transitive verbs, causativised verbs do not require an overt direct object (cf. §5.3.1). This is illustrated in (160) and (161).

(160) jɛʔ pr-hir lɛh
1S CAUS-to.be.afraid EMP
'I scared [X]!'

(161) ɟa=jɛʔ pi-dol k=la=hip
RT=1S CAUS-to.hide LOC=ID=forest
'Then I hid [X] in the forest.'

5.4.2 The progressive as passive

The progressive morpheme described in §4.7.1.2 is typically a strictly aspect/Aktionsart marker affixed to verbs in the imperfective to signal that a situation is in progress and 'developing' or 'unfolding', often with an additional non-temporal notion of automaticity. As noted, its use indicates that the situation does not need a lot of input of energy in order to occur or progress. In combination with punctual verbs in their lexeme (non-imperfective) form, the progressive fully takes on such non-temporal notions and loses its temporal meaning altogether. Usually, such forms do not exhibit an argument structure different from that of other forms of the verb.

In a few exceptional cases, however, they attract subjects whose referents would normally be represented as direct objects with the verb in question. In example (162), involving the progressive form of the verb /ʔɛk/ 'to give', the subject-encoding irrealis proclitics are cross-referenced by postverbal subject PPs representing the thematic relation of THEME, that is the object that is given.

(162) wa=bk-ʔɛk ka=gtah, wa=bk-ʔɛk ka=klapah sawit
IRR.3S=PROG-to.give SUBJ=rubber IRR.3S=PROG-to.give SUBJ=oil.palm
'Rubber will be given. Oil palms will be given.'

Thus, an original subject argument representing an AGENT is lost and replaced by the THEME in a valence-reducing operation that looks like a proper passive construction. The notion of automaticity encoded in the progressive morpheme appears to play down and suppress the AGENT role to a point where it is lost and its syntactic slot is filled by the THEME. As shown by example (163), however, the same verb form need not always be given a passive reading.

(163) ʔoʔ bk-ʔɛk ba=jɛʔ
3S PROG-to.give GOAL=1S
'He gave [X] to me [in passing].'

The reason for this discrepancy is not known, and the scarcity of examples precludes further conclusions. It is not unlikely that the valence-reducing effect of the progressive is a recent development (perhaps under influence from Standard Malay, the *ber-* prefix of which is used to create middle voice forms) and that we are witnessing a passive in the

making. Kruspe (2004:117–118) describes for Semelai a *b(r)-* prefix which is productively and regularly employed as a valence-reducing marker of middle voice.

5.5 Summary

This chapter has introduced briefly the main features of Jahai syntax. Tentative and incomplete, the analysis will surely be subject to future revisions. Nevertheless, it is possible to draw some preliminary conclusions. Thus, the order of constituents appears rather straightforward and largely conforms to the patterns described for other Aslian languages. The pattern of agreement also finds similarities in closely related languages. The marking of postverbal arguments, with subjects of both intransitive and transitive constructions contrasting with the direct object, appears to reflect an accusative system of grammatical relations. This is in contrast to the ergative-like patterns identified in some other Aslian languages.

Several features are unclear and in need of further investigation, including for example the syntactic behaviour of action/state nominalisations, the properties of the pragmatically determined preposition /d=/, as well as valence-affecting operations.

6 *Expressive elaboration*

This chapter provides a brief outline of a marginal linguistic phenomenon in Jahai referred to here as 'expressive elaboration'. An introduction to the phenomenon and its status is given in §6.1, and a description follows of its phonological (§6.2), morphological (§6.3), semantic (§6.4) and syntactic (§6.5) characteristics. The analysis should be regarded as tentative.

6.1 Introduction

Many onomatopoeic forms (see §4.8.1) and a set of roots of primarily stative verbs may be subject to what will be referred to here as 'expressive elaboration'. This involves a number of regular morphological operations and results in adjuncts at sentence level which express the sensory perceptions of the speaker (visual, auditory or tactile) in relation to the phenomenon described by the clause. In this respect, expressively elaborated forms behave in much the same way as expressives do in other Austroasiatic languages, as described e.g. by Diffloth (1972, 1976d).

However, expressive elaboration, unlike the onomatopoeic forms described in §4.8.1, gives the impression of being a highly marginal phenomenon in Jahai. It is very infrequent and not a feature of the whole speech community. Almost all of the elaborated forms analysed below were elicited from a single male speaker, originally from a different Jahai group sometimes referred to as Mengkah. Other speakers assumed a bewildered attitude towards them or rejected them outright, instead suggesting the corresponding verbal or onomatopoeic base as the correct form. All speakers agree on the existence of most such bases, as well as the ordinary verbal morphology associated with them.

Spontaneous use of expressive elaboration was observed rarely in only one other individual, another male speaker who also originates from a different group and who grew up in a Temiar-speaking environment. It may be of some significance that both of these speakers have spent long periods away from the Jahai community in close contact with speakers of other Aslian languages like Temiar and Semai, where expressives form an important word class (see Benjamin 1976b:177–178; Diffloth 1976d). It could therefore even be suggested that their habit of applying expressive elaboration to Jahai words is borrowed and that the phenomenon is foreign to Jahai.[30] No systematic comparison has

30 Interestingly, Gérard Diffloth (pers. comm.) notes that expressives have not previously been documented in Northern Aslian languages.

been made with expressives in neighbouring languages, so no conclusions will be drawn as to this possibility. At any rate, idiolectal variation clearly plays an important role in usage.

A possible reason for the marginal position of expressive elaboration may be the complex aspectual/Aktionsart system of Jahai verbs, with its unusual categories of iterative (§4.7.1.3) and distributive (§4.7.1.5), which is likely to extend into semantic domains associated with expressives in many other Mon-Khmer languages. In other words, the 'expressiveness' of the system of verbal derivation exists at the expense of the class of expressives.

Expressive elaboration displays structural peculiarities which are not characteristic of the Jahai linguistic system as a whole, particularly with regard to the vowel inventory and word structure. The marginal, non-universal character of expressive elaboration is the reason why such deviant features are not included in the general analysis of Jahai but warrant separate treatment.

Given the limited data and restricted usage, the following description should not be considered an exhaustive analysis of expressive elaboration. Further elicitation is likely to provide considerably more information about the characteristics of these processes.

6.2 Phonological peculiarities

On the whole, expressive elaboration involves the same phonemes as other words and are characterised by the same phonetic realisations. Interestingly, however, a nasal phonemic diphthong /ũɔ̃/, which is not attested elsewhere, has been found in two instances of expressive elaboration:

/litlũɔ̃t/	'the appearance of leeches moving'
/riŋrũɔ̃ŋ/	'the appearance of a bird turning its head'

Also, phonemically nasal counterparts of the mid vowels /e/ and /o/, which are not part of the ordinary phonemic inventory (cf. §2.2), turn up occasionally in expressive elaboration. Furthermore, phonemically nasal vowels sometimes form nuclei in pre-final syllables. This is otherwise not allowed (cf. §2.4.3).

/gliʔwẽʔ/[31]	'the appearance of the fluttering wings of a butterfly'
/kpõtpẽt/	'the feeling of waking up to the sound of munching'

Another characteristic is the common use of otherwise infrequent phonemes. For example, the phonemically nasal vowels, which are ordinarily rather infrequent, are commonplace in expressive elaboration. Similarly, the voiceless bilabial fricative /ɸ/, the most infrequent consonant phoneme, is disproportionately common in expressive elaboration. Finally, expressively elaborated forms are often tetrasyllabic, a clear violation of ordinary word structure, which permits a maximum of three syllables (cf. §2.4.2).

/klawẽhwɔ̃h/	'the appearance of several crooked objects'
/prakɔɸkɔɸ/	'the sound of a squirrel dropping empty nutshells'
/placiɸcɔ̃ɸ/	'the appearance of impressions in skin'

31 No root */gwẽʔ/ has been attested. Otherwise it would be possible to postulate a phonemically nasal counterpart of the vowel /e/ for the general phonemic inventory of Jahai.

6.3 Morphological characteristics

Expressive elaboration involves morphemes which are partly similar to those of other word classes, especially verbs. Thus, expressive elaboration may include the use of the collectivising morpheme /<ra> ~ <a>/ (from nominal morphology, cf. §4.1.1 and §4.1.4.2), infixed /<la>/ (which is likely to have some connection with the iterative /l/ of verbal morphology, cf. §4.7.1.3), infixed /<na>/ (possibly linked to the nominalising/unitising /n/, cf. §4.1.3 and §4.1.4.1), as well as reduplication with vowel alternation (according to a pattern similar to that found in the distributive form of verbs, cf. §4.7.1.5). Of these, the equivalents of collective and distributive appear to be the most significant, as they are present in most of the recorded examples of expressive elaboration according to a very regular pattern.

This similarity to primarily verbal morphology sometimes renders it difficult to make a sharp distinction on morphological grounds alone between verbs and examples of expressive elaboration. However, when forms are tetrasyllabic we can safely categorise them as examples of expressive elaboration, since tetrasyllabicity is otherwise disallowed (§2.4.2). The following exemplification is restricted to such tetrasyllabic forms.

Expressively elaborated tetrasyllabic forms assume what appears to be a regular and highly standardised shape consisting of a final CVC syllable (a monosyllabic root or a final syllable of a sesqui- or disyllabic root) preceded from right to left by a penultimate reduplication of the final syllable (with or without vowel alternation), an antepenultimate syllable made up of one of the infixes /<ra>/, /<la>/ or /<na>/ (by far the most common being the collective plural morpheme /<ra>/), and a pre-antepenultimate half syllable. In the case of sesqui- and disyllabic roots, this half syllable is represented by the initial consonant of the root. In the case of monosyllabic roots, it is represented by another, non-predictable consonant, usually a stop. The pattern is exemplified in the following forms:

Root: /ckɨp/ 'to close one's eyes'
/c<ra><kip>kɨp/ '(appearance of many people closing their eyes)'

Root: /cnhə̃t/ 'to be short'
/cn<a><hit>hə̃t/ '(appearance of many short objects)

Root: /hɲʄut/ 'to be heavy'
/hɲ<a><ʄit>ʄut/ '(feeling of carrying many heavy objects)'

Root: /tək/ '(sound of a big raindrop falling to the ground)'
/k<ra><tik>tək/ '(sound of many big raindrops falling to the ground)'

Root: /rɨɸ/ '(sound of flapping)'
/k<na><rɨɸ>rɨɸ/ '(sound of something flapping repeatedly, like a bird's wings, or a piece of cloth being shaken, or a song ringing in one's head)'

Root: /kɔɸ/ '(sound of a small object falling to the ground)'
/r<la><kɔɸ>kɔɸ/ '(sound of empty nutshells being thrown to the ground by a squirrel, one after another)'

Root: /rɔɸ/ '(sound of a blowpipe dart hitting canopy)'
/k<na><rɔɸ>rɔɸ/ '(sound of blowpipe darts hitting canopy, one after another)'

Root: /gc̈ih/ 'to be dark'
/g<ra><cih>c̈ih/ '(appearance of many black objects)'

Root: /cgil/ 'to be uneven'
/c<ra><gil>gil/ '(feeling of sitting on many uneven surfaces)'

6.4 Semantic characteristics

A distributive reading, represented by reduplication with possibly iconic vowel alternation (referred to as antiphonic reduplication by Diffloth 1976d:254; cf. §4.7.1.5), is present in an overwhelming majority of the expressive forms recorded. This involves notions of spatial distribution and irregularity. Sometimes, however, the reduplication of the final syllable is total, which appears to represent iteration.

The infixed /<ra> ~ <a>/ is likely to carry collective meanings similar to the ones it conveys in nominal morphology. The meanings of /<la>/ and /<na>/ are less clear, but possibly they have some connection to iteration and unitisation respectively.

There is frequently a direct correlation between the size of an entity referred to and vowel quality in the final syllable. Consistently, central and back vowels like /ɨ, ə, a, o, ɔ/ or their nasal counterparts are associated with bigness, whereas the high front vowel /i/ or its nasal counterpart signals smallness. Intermediate size is represented by the low front vowel /ɛ/ or its nasal counterpart. The following forms exemplify such correlation:

/kramitmɨ̃t/	'(appearance of large impressions in skin)'
/kramitmɛ̃t/	'(appearance of small impressions in skin)'
/kramitmĩt/	'(appearance of tiny impressions in skin)'
/cnahithə̃t/	'(appearance of many short objects)'
/cnahithɛ̃t/	'(appearance of many very short objects)'
/cnahithĩt/	'(appearance of many very short, hardly visible objects)'
/klaʔicʔũc/	'(appearance of large larvae crawling)'
/klaʔicʔɛ̃c/	'(appearance of small larvae crawling)'
/klaʔicʔĩc/	'(appearance of tiny larvae crawling)'

6.5 Syntactic characteristics

Only sporadically have instances of expressive elaboration been recorded in spontaneous conversation, and therefore no firm conclusions will be drawn here about its syntactic behaviour. As described for other Aslian languages, however, they appear to function as a form of adjunct at sentence level, which follows the clause it modifies and is separated from it by a pause. It sums up the perceptions of the speaker and complements or specifies the information expressed in the clause. Such detached elements with specifying functions appear to be associated with a syntactic slot referred to here as a right-detached phrase (see §5.1.1.1.4). An example is given in (1).

(1) ca=ʔɛl ton, klawɛ̃hwɔ̃h-klawɛ̃hwɔ̃h
HORT=to.look that [appearance of several crooked objects ...]
'Look at those! [expressive]'
[Uttered by a man inspecting the warped shafts of a blowpipe]

6.6 Summary

Displaying a number of peculiar structural features not characteristic of the Jahai linguistic system as a whole, expressive elaboration gives the impression of being a highly marginal (and perhaps even foreign) phenomenon in Jahai. It is suggested that this marginal status is at least partly related to the complexity of the aspect/Aktionsart system of verbs, some categories of which are likely to extend into semantic domains associated with expressives in other Mon-Khmer languages. However, a more extensive analysis of expressive usage in other idiolectal and dialectal varieties of Jahai is necessary in order to evaluate the tentative results presented here.

7 *Concluding remarks*

The present work has provided a characterisation of the phonology, morphology and syntax of contemporary Jahai. Certain issues have been dealt with in detail; others have been only briefly introduced. Indeed, much of the analysis is to be regarded as tentative and a basis for further research. Nevertheless, several conclusions may be drawn about the nature of the Jahai linguistic system. Thus, by and large, Jahai behaves in several important respects like other Aslian languages. In most cases, this typical Aslian behaviour was expected; the 3 x 3 vowel system, the distinctive vowel nasality, the peculiar realisations of word-final nasal consonants, the lack of contrastive tone, the polysyllabic lexemes, the rich systems of pronominal and demonstrative distinctions, the great number of Malay loanwords, and the patterns of word order and agreement are all expected features shared by most of Jahai's closest relatives.

However, at least one characteristic encountered in Jahai was more or less unexpected. The morphology of Northern Aslian languages has hitherto been tentatively considered to be less productive and more fossilised than that of languages belonging to other branches of Aslian, notably Central Aslian, frequently described as exhibiting some of the most regular and productive systems of verbal morphology in Southeast Asia. The present work indicates that Northern Aslian languages display equally impressive systems of derivation, if not more so. As has been shown, Jahai productively employs a range of intricate processes of word-formation to create distinctions pertaining to aspect, Aktionsart, nominalisation and causativisation of verbs, as well quantification and verbalisation of nouns. However, it is possible that this productivity and regularity of the morphological system is particular to Jahai and not characteristic of the whole Northern Aslian subgroup. If Benjamin (1976a:77) is correct in suggesting that Jahai has moved into its present area of distribution fairly recently from the south, as indicated by intra-Aslian lexical borrowing, it could be suggested that its morphological complexity results rather from early contacts with Central Aslian languages. Still, parts of the Jahai system of verbal derivation do not have known equivalents elsewhere in Aslian. Additional work on other Northern Aslian languages is necessary in order to solve this issue.

As has been shown recurrently throughout this work, a pervasive feature of the Jahai linguistic system is its readiness to incorporate and adapt foreign elements. While this receptivity is perhaps superficially not surprising, its magnitude and significance should not be underestimated. Thus, at least one-fifth of the lexical items collected for the present study is of Malay origin. Furthermore, foreign segmental material, as well as complete affixes, have been incorporated into the Jahai system of word formation, and at least two

processes of affixation, referred to here as outer affixation and total reduplication, were suggested to have been borrowed from Malay. Moreover, the majority of auxiliary and adverbial elements and conjunctions identified are of Malay origin, and some of the prepositional proclitics also look conspicuously Malay. Also, much of the system of classifiers may be a calque of the Malay system. In addition, occasional lexical and morphological elements are likely to have been borrowed from Central Aslian languages like Temiar. Apparently, Jahai speakers have long been inclined to pick up useful linguistic features from their neighbours and make them their own. By analogy with the Jahai way of life, the Jahai language may with good reason be characterised as highly utilitarian, adaptable and opportunistic.

This receptivity is in keeping with the dynamic linguistic situation of the Northern Aslian speech communities. In an age when governments worry about the future of languages of whole nations, the 1000 or so Jahai proudly — and non-literately — maintain their linguistic identity in an environment where multilingualism and constant linguistic contact and change are the norm. For the Jahai as well as their language, flexibility and adaptation are tantamount to survival.

With cultural and linguistic diversity on the retreat, the opportunities of studying traditional, small-scale speech communities are becoming increasingly rare. The language of the Jahai, spoken by people who belong to the tiny and constantly dwindling part of humanity that still pursues a hunter-gatherer existence, provides a precious source of information about communication systems in a traditional, non-literate society, and, essentially, about what it means to be human.

Appendix I: Jahai rhyming glossary

The following glossary constitutes the full list of 1730 synchronic Jahai words collected by the author for the present study. Items represent lexeme forms of words, many of which are roots or may at least be regarded as synchronically monomorphemic (see §3.1). In some cases the forms given here are derivations of existing roots. As a rule, however, derived forms are not included. The lexeme form of a word is usually tantamount to the preferred citation form; however, note that the agreement marker on verbs, which is part of the citation forms of most verbs, is not included here (cf. §4.7). Citation forms of names of birds, fish, snakes, trees and vines commonly include the generic names of these classes (/kawɔ̃t/, /ʔikəʔ/, /taɟuʔ/, /tom/ and /ʔawɛj/ respectively) followed by a proper, specifying name, e.g. /ʔikəʔ bawuŋ/, literally 'fish *baung*'. However, unless there is no ambiguity involved, only the more specific part of the names has been listed here.

Employing the International Phonetic Alphabet (IPA), words are given in both phonemic and phonetic form, in order to fully illustrate the phonological analysis presented in Chapter 2. The phonetic transcriptions generally convey the exact pronunciation, although some standardisation has been made, especially in cases of free variation. For example, the alveolar rhotic /r/ is consistently transcribed as a trill [r] except in cases where it is preceded by [n], in which case it is usually realised as an approximant preceded by a stop transition and transcribed accordingly: [^{d}ɹ]. In reality, however, the trill is in free variation with the approximant [ɹ] in all positions and a flap [ɾ] in syllable-initial position (see §2.3.1.4). In cases where there are similar but in some respect phonemically different varieties of the same word, the alternate forms are listed together and separated by ~, as in /ɗip ~ tip/ 'to be flat'. For a detailed discussion on the phonetic realisation of phonemes, see Chapter 2.

Items are listed phonemically in rhyming format, which means that words are analysed backwards and therefore arranged according to their final phoneme. Phonemes, in turn, are arranged according to the following relative order of modes of articulation: vowels, stops, fricatives, nasals, laterals, rhotics and approximants. Furthermore, for each mode of articulation phonemes are ordered according to place of articulation, with 'front' phonemes first and 'back' phonemes last. Vowels are further ordered from high to low. Finally, where relevant, voiceless phonemes precede voiced ones. This gives the following order of phoneme presentation: /i, e, ɛ, ɨ, ə, a, u, o, ɔ, p, b, t, d, c, ɟ, k, g, ʔ, ɸ, s, h, m, n, ɲ, ŋ, l, r, w, j/. Nasal vowels are not ordered separately from oral vowels unless they occur in minimal pairs, in which case oral vowels precede nasal ones. The rhyming format may be confusing at first to readers accustomed to traditional alphabetic and initial ordering, but it has clear advantages in the case of Mon-Khmer languages. This is because Mon-Khmer languages seldom have suffixes, and the end of the word is therefore usually part of the

root and not affected by morphophonemic processes. Also, the final syllable is the most informative part of a word in that it always receives stress and contains the greatest phonemic variation (Diffloth 1976b:102).

English translations express the meaning of the Jahai words as exactly as possible, and explanatory descriptions are provided within square brackets wherever direct translation is not possible. Latin names are given for identifiable species and genera of plants and animals; unidentified species are described in as much detail as possible, e.g. [a type of hornbill]. Proper names are divided up into *toponyms* (placenames) and *ethnonyms* (ethnic groups). English synonyms are separated by a solidus (/) and different meanings by a semicolon (;).

Forms thus far identified as definite or likely loans from or via Malay, including words of English origin, are marked with an asterisk (*).

Phonemic form	**Phonetic form**	**English translation**
/-p/		
fip ~ tip	'fip˺ ~ 'tip˺	to be flat
ctip ~ ctip	cᶜə'tip˺ ~ cᶜə'tip˺	to assemble; to grow together
sĩp	'sĩp˺	fruit without a seed
kawip	ka'wip˺	sun bear (*Helarctos malayanus*)
tep*	'tep˺	cassette
pdep	pə'dẹp˺	to peep
kʔep	kə'ʔẹp˺	[a type of centipede]
klep	kə'lẹp˺	[a type of tuber]
krlep	kər'lẹp˺	to forget
titɛp	ti'tɛp˺	opposite side
ktɛp	kə'tɛp˺	to carry in one's mouth
pdɛp	pə'dɛp˺	to stare
cɛp	'cᶜɛp˺	to catch
kakɛp	ka'kɛp˺	to remember
ksɛp	kə'sɛp˺	to adorn oneself with leaves for good luck
mɛp*	'mɛ̃p˺	map
ɲɛp	'ɲɛ̃p˺	to be raw
lɛp	'lɛp˺	to fold; to plait
halɛp	ha'lɛp˺	to raft
piplɛ̃p	pip˺'lɛ̃p˺	to blaze; to twinkle
cprɛp	cᶜəp˺'rɛp˺	babbler (*Malacopteron*)
jɛp	'jɛp˺	to count
pipjɛp	pip˺'jɛp˺	to teach
sntip	sɜn'tip˺	to pound
cip	'cᶜip˺	to go; to move
sɟip	sə'ɟᶻip˺	to flow (of tears)
pkip	pə'kip˺	to scorch
crkip can	cᶜɜr'kip˺ 'cᶜaᵈn	heel
krʔip	kər'ʔip˺	to crouch
hip	'hip˺	wood/forest; outside
raŋip	ra'ŋip˺	to crunch

grlɨp	gərˈlɨpˀ	to lay a floor
təp	ˈtəpˀ	to carry on one's back
krʔəp	kərˈʔəpˀ	to burp
harəp*	haˈrəpˀ	to believe; to have faith
bap	ˈbapˀ	rice; food
sbap*	səˈbapˀ	because; but
ttap*	təˈtapˀ	to stay in one place; to settle down
stap	səˈtapˀ	to split
kap	ˈkapˀ	to bite
tiŋkap*	tiŋˈkapˀ	window
laŋkap*	laŋˈkapˀ	[a type of palm (*Arenga obtusifolia*)]
hagap	haˈgapˀ	Sumatran rhinoceros (*Dicerorhinus sumatrensis*)
ʔap	ˈʔapˀ	[generic name for large felines (*Panthera*)]; to encounter a large feline
ʔap tmtum	ˈʔapˀ təmˈtuᵇmˀ	black panther (*Panthera pardus*)
ʔap ʔawɛj	ˈʔapˀ ʔaˈwɛ̃jˀ	leopard (*Panthera pardus*)
maʔap*	maˈʔapˀ	pardon
hap ~ hiphap	ˈhapˀ ~ hipˀˈhapˀ	to spread in different directions
blap	bəˈlapˀ	nothing
klap	kəˈlapˀ	[toponym: Kelap]
slap	səˈlapˀ	to die
ɟawap*	ɟʑaˈwapˀ	to answer
hawap	haˈwapˀ	[a type of large freshwater gastropod]
sajap*	saˈjapˀ	wing
tjap-tjap*	tijapˀ-tiˈjapˀ	every
sagup	saˈgupˀ	cloud; fog
jup	ˈjupˀ	to be flat
cukop*	cᶜuˈkopˀ	to be full
gop	ˈgopˀ	stranger (derogatory)
ʔop	ˈʔopˀ	to stop
sop	ˈsopˀ	lung
rop	ˈropˀ	to walk; [sound of walking; sound of muntjac deer]
hatɔ̃p	haˈtɔ̃pˀ	to be light (of weight)
dɔp	ˈdɔpˀ	to sneak; to whisper
kɔp	ˈkɔpˀ	to move hut
khɔp	kəˈhɔpˀ	to gobble down
ɲɔp	ˈɲɔ̃pˀ	to chew
karɔp	kaˈrɔpˀ	dead bamboo

/-t/

ptpit	pətˀpitˀ	[a type of small animal]
cpĩt	cᶜəˈpĩtˀ	to squeeze
bit	ˈbitˀ	[a type of ant]

kfĩt	kə'fĩt˺	to sting
kacĩt	ka'c^{ɕ}ĩt˺	to rattle
ʄit	'ʄ^{ʑ}it˺	to collect/gather
kit	'kit˺	bottom; buttocks
kit ktɔʔ	'kit˺ kə'tɔʔ	sun
kit tɔm	'kit˺ 'tɔ^{b}m	mouth of river
ktkit	kət˥'kit˺	to fart
ksĩt	kə'sĩt˺	birthmark
hit	'hit˺	to rattle; to tremble
mit	'mĩt˺	eye; CLF: small objects
mit taʄiʔ	'mĩt˺ ta'ʄ^{ʑ}iʔ	blade of knife
mit ktɔʔ	'mĩt˺ kə'tɔʔ	sun
mit knajil	'mĩt˺ kəna'jil	fishing hook
mit ʔhəj	'mĩt˺ ʔə'həj	seed
hnit	hə'nĩt˺	[a type of fruit]
bŋit	mə̃'ŋĩt˺	NEGATIVE MARKER
rŋit	rə'ŋĩt˺	sandfly
plit	pə'lit˺	to fade away; to extinguish by itself
ktlĩt	kət˥'lĩt˺	to feel the sharpness of e.g. a knife
krtwĩt	kərɜt˥'wĩt˺	wrinkles
wtwĩt	wət˥'wĩt˺	[a type of red bird]
ʔmpet	ʔəm'pẹt˺	[a type of cobra]
kbet	kə'bẹt˺	old woman
kbet clah	kə'bẹt˺ c^{ɕ}ə'lah	old bachelor
puket*	pu'kẹt˺	pocket
ʔaŋket*	ʔaŋ'kẹt˺	to get; to take
get	'gẹt˺	to cut
lget	lə'gẹt˺	mountain pass
ʔtʔet	ʔət˥'ʔẹt˺	to know
pahet	pa'hẹt˺	[a type of fruit]
plet	pə'lẹt˺	to be dense/heavy
wet	'wẹt˺	to flow (of river); to run (of e.g. trail)
lumpɛt*	lum'pɛt˺	to jump
bɛt	'bɛt˺	to lash; to slither
lbɛt	lə'bɛt˺	to be heavy (of rain)
rbɛt	rə'bɛt˺	to tie
kaltɛt	kal'tɛt˺	to flip with one's finger
plaɲcɛt	pəlaɲ'c^{ɕ}ɛt˺	mouse deer (*Tragulus*)
kɛt	'kɛt˺	to cut
ʔikɛt*	ʔi'kɛt˺	to lash
ʔakɛt	ʔa'kɛt˺	PROHIBITIVE
tmkɛt	təm'kɛt˺	to be cold (of weather, objects etc.)
hlaŋkɛt	həlaŋ'kɛt˺	[a type of ant]
haʔɛ̃t	hã'ʔɛ̃t˺	bad smell/stench
btʔɛt	bət˥'ʔɛt˺	to be beautiful/fine/good

ʔtʔɛt	ʔət˥ʔɛt˩	to stretch oneself
sɛt	ˈsɛt˩	to pour
sɛtsɛt	sɛt˥sɛt˩	sunbird (*Nectarinia*; *Anthreptes*)
stsɛt	sɜt˥sɛt˩	to pick one's teeth
laɲsɛt*	laɲˈsɛt˩	langsat (*Lansium domesticum*)
rɲsɛ̃t	rəɲˈsɛ̃t˩	to sob
bhɛt	bəˈhɛt˩	to be sweet
kmɛt	kəˈmɛt˩	[a type of frog]; vulva
grinɛt*	gəriˈnɛt˩	grenade
ʔiŋɛt*	ʔiˈŋɛt˩	to believe; to think; to remember
bulɛt	buˈlɛt˩	to slither
bnolɛt	mãnoˈlɛt˩	many
hrtlɛt	həˈrɜt˥lɛt˩	to be tired
parɛt	paˈrɛt˩	[a type of large insect]
barɛt	baˈrɛt˩	Malay tapir (*Tapirus indicus*)
ɟarɛ̃t	ɟᶻaˈrɛ̃t˩	[a type of small grasshopper]
sorɛt*	soˈrɛt˩	letter
stwɛt	sɜt˥wɛt˩	[a type of fruit]
hwɛ̃t	hə̃ˈwɛ̃t˩	to stain
jɛt	ˈjɛt˩	to ogle
sitjɛ̃t	sit˥jɛ̃t˩	to blow (of wind)
pɨt	ˈpɨt˩	to blow (of person and wind); to extinguish fire
ktɨt	kəˈtɨt˩	egg
ddɨt˩	dəˈdɨt˩	[a type of bird]
ɟɨt	ˈɟᶻɨt˩	to wipe
bkɨt	bəˈkɨt˩	to be light; to be warm; to be hot
bŋkɨt	mãŋˈkɨt˩	sweat
hrkɨt	həˈrkɨt˩	evening; night
bʔɨt	bəˈʔɨt˩	to be sweet
sɨt	ˈsɨt˩	honeycomb
sɨ̃t ~ ʔotsɨ̃t	ˈsɨ̃t˩ ~ ʔot˥sɨ̃t˩	to rub oneself
bhɨt	bəˈhɨt˩	bamboo
ɟhɨt	ɟᶻəˈhɨt˩	to smoke; to suck
blhɨt	bəlˈhɨt˩	to be tasteless
kmɨt	kəˈmɨ̃t˩	gall-bladder; knee-cap
tnɨt	təˈnɨ̃t˩	lower lip
ktlɨt	kət˥lɨt˩	glow-worm
hmlɨ̃t	həmˈlɨ̃t˩	mythical cannibals
ptpət	pət˥pət˩	to place one's hand on something
tmpət*	təmˈpət˩	place
sitkət	sit˥kət˩	to laugh
ckət	cᶜəˈkət˩	to kill/slash
ragət	raˈgət˩	to carry on one's back
hə̃t	ˈhə̃t˩	to sting (of mosquito)

cnhət	cᶜɜn'hət˺	to be short
ttlət	tɜt˺'lət˺	to stare
klət	kə'lət˺	to swallow
rət	'rət˺	to tie
wawət	wa'wət˺	rat
lktwət	ləkət˺'wət˺	to be fast; quickly
jət	'jət˺	to watch/see
jət	'jət˺	to hurt
ʔmpat ~ ʔmpət*	ʔəm'pat˺ ~ ʔəm'pət˺	four
bat	'bat˺	to sit (of bird); to roost; [sound of bird landing on a branch]
cat*	'cᶜat˺	to paint
sikat*	si'kat˺	comb
hakat	ha'kat˺	to rub; to scratch
pukat*	pu'kat˺	large net
gat	'gat˺	to waylay
ʔat	'ʔat˺	stick
sat	'sat˺	to leave poisoned blowpipe dart to dry
rihat*	ri'hat˺	rest; to rest
phat*	pə'hat˺	to hammer
ɟahat*	ɟᶻa'hat˺	to be bad
hamat*	ha'mat˺	to roll something
slamat*	səla'mat˺	safety
kilat*	ki'lat˺	lightning
calat	cᶜa'lat˺	[a type of lizard]
barat*	ba'rat˺	west
karat*	ka'rat˺	dirt
kawat*	ka'wat˺	to march
kwat*	kə'wat˺	to be strong
swat	sə'wat˺	to build a hut
siput*	si'put˺	snail
rumput*	rum'put˺	grass
but	'but˺	to eat
dut	'dut˺	navel
hɲɟut	həɲ'ɟᶻut˺	to be heavy
ʔũt	'ʔũt˺	to argue
sut	'sut˺	to sob
prasut	pəra'sut˺	to lose hair
hũt	'hũt˺	to reveal oneself
ɟaŋut*	ɟᶻa'ŋũt˺	beard
ʔŋut	ʔə'ŋũt˺	throat
parut*	pa'rut˺	scar
surut*	su'rut˺	to push
pot	'pot˺	to suck
klutbot	kəlut˺'bot˺	[a type of larvae]

ʔikot*	ʔi'kot˺	to follow
saʔot	sa'ʔot˺	to call someone
ʄtʔot	ʄᶻət˺'ʔot˺	[toponym]
sot	'sot˺	to carve/cut; straight
kasot*	ka'sot˺	shoe
slot	sə'lot˺	to drown
lawot*	la'wot˺	sea
kajot	ka'jot˺	to be pregnant
dbɔt	də'bɔt˺	to climb a tree (of bear)
cbɔt	cᶜə'bɔt˺	to devour
dɔt	'dɔt˺	vulva
kɔt	'kɔt˺	to take
hakɔ̃t	ha'kɔ̃t˺	[a type of tuber]
ʔɔ̃t	'ʔɔ̃t˺	dog
ʔtʔɔ̃t	ʔət˺'ʔɔ̃t˺	to stroke an animal
hɔ̃t	'hɔ̃t˺	to perceive taste
krtlɔ̃t	kəɾət˺'lɔ̃t˺	kidney
smutlɔ̃t	səmut˺'lɔ̃t˺	brain
wɔ̃t	'w̃ɔ̃t˺	to bend something down (e.g. a tree)
kawɔ̃t	kã'w̃ɔ̃t˺	bird (generic)
kawɔ̃t batuʔ*	kã'w̃ɔ̃t˺ ba'tuʔ	broadbill; pitta (*Psarisomus*; *Pitta*)
rajɔ̃t	rã'ɟɔ̃t˺	[a type of small animal]

/-c/

hĩc	'hĩc˺	to rain
btec	bə'tẹ̈ⁱc˺	to be long
lwec	lə'wẹ̈ⁱc˺	to ascend; to climb up
rmpɛc	rəm'pɛⁱc˺	hardwood tree (*Turtur tigrinus*)
kbɛc	kə'bɛⁱc˺	to spit slowly to the ground
kɛc	'kɛⁱc˺	to cut
gɛ̃c	'gɛ̃ⁱc˺	to scratch (of thorns)
ʔɛc	'ʔɛⁱc˺	belly; excrement; to defecate
sɛc	'sɛⁱc˺	meat
knɛc	kə'nɛ̃ⁱc˺	comb
ŋɛc	'ŋɛ̃ⁱc˺	to tickle
lɛc	'lɛⁱc˺	to miss target; to be wrong
taʔɨc	tã'ʔɨⁱc˺	[a type of large bird]
krʔɨc	kər'ʔɨⁱc˺	to heat
trhɨc	tɜr'hɨⁱc˺	[a type of small bird]
sɨc	'sɨⁱc˺	to sting
grsɨc	gər'sɨⁱc˺	to feel
pimɨc	pi'mɨ̃ⁱc˺	to take back
ʔŋɨc	ʔə'ŋɨ̃ⁱc˺	to burn
ʔalɨc	ʔa'lɨⁱc˺	to pass
harɨc	ha'rɨⁱc˺	[magic word uttered to stop rain]

tə̃c	'tə̃^{i}c˺	to follow
hchəc	hic˺'hə^{i}c˺	to whistle
cməc	c^{c}ə'mə̃^{i}c˺	boil; wound
plə̃c	pə'lə̃^{i}c˺	to disappear
klcbac	kəlic˺ba^{i}c˺	[a type of millipede]
kac	'ka^{i}c˺	to scratch
knʔac	kən'ʔa^{i}c˺	father-in-law
bclac	bic˺la^{i}c˺	to be smooth; to be slippery
srac	sə'ra^{i}c˺	[magic word uttered to stop rain]
bawac	ba'wa^{i}c˺	pig-tailed macaque (*Macaca nemestrina*)
kwac	kə'wa^{i}c˺	to swim
huc	'hu^{i}c˺	to drink
muc	'mũ^{i}c˺	to eat fish or meat
kmuc	kə'mũ^{i}c˺	ghost
siruc	si'ru^{i}c˺	to slurp
grũc	gə'rũ^{i}c˺	slender-toed gecko (*Cyrtodactylus*)
poc	'po^{i}c˺	[a type of terrapin]
scboc	sic˺bo^{i}c˺	to lick one's lips
tʔoc	tə'ʔo^{i}c˺	to ask for something
soc	'so^{i}c˺	to wash
hoc	'ho^{i}c˺	[a type of fruit]
pɔc	'pɔ^{i}c˺	to collect small objects
ʔmpɔc	ʔəm'pɔ^{i}c˺	salt
bɔc	'bɔ^{i}c˺	to tell a lie
tɔc	'tɔ^{i}c˺	to tickle by poking one's finger
cɔc ~ cəc	'c^{c}ɔ^{i}c˺ ~ 'c^{c}ə^{i}c˺	scar
kɔc	'kɔ^{i}c˺	to gnaw
lkɔc	lə'kɔ^{i}c˺	fontanel
tʔɔc	tə'ʔɔ^{i}c˺	to climb
sɔc	'sɔ^{i}c˺	to be finished; to be gone; to stop
ɲɔc	'ɲɔ̃^{i}c˺	to move snout (of monkey)
lɔc	'lɔ^{i}c˺	bow
snlɔc	sɜn'lɔ^{i}c˺	blowpipe dart
canwɔc	c^{c}an'wɔ^{i}c˺	collared scops-owl (*Otus lempiji*)

/-k/

tik	'tik˺	[sound of raindrop]
mudik*	mu'dik˺	to return
cik	'c^{c}ik˺	[sound of blowpipe dart hitting muscle of prey]
hgik	hə'gik˺	to fear
lanik	la'nik˺	lie/haunt (of animal)
palīk	pa'līk˺	[a type of small animal]
clīk	c^{c}ə'līk˺	to keep a good lookout sideways while walking

ɲiɲik	ɲiˈrikˀ	bee-eater (*Merops*)
brik*	bəˈrikˀ	to brake
grik ~ grit	gəˈrikˀ ~ gəˈritˀ	[toponym: Gerik]
tek	ˈtẹkˀ	to lie down; to marry; to sleep
lntek	lɜnˈtẹkˀ	tongue
kdek	kəˈdẹkˀ	squirrel
kdek ʔabuʔ	kəˈdẹkˀ ʔaˈbuʔ	Provost's squirrel (*Callosciurus prevosti*)
kdek creh	kəˈdekˀ cᶜəˈrẹh	plantain squirrel (*Callosciurus notatus*)
cek	ˈcᶜẹkˀ	to stab; to throw a spear
plek	pəˈlẹkˀ	to fall (of tree)
klek	kəˈlẹkˀ	quill of porcupine
wek	ˈwẹkˀ	to go back
piwek	piˈwẹkˀ	to lift
wikwek	wikˀˈwẹkˀ	to turn around; to move around
pɛk	ˈpɛkˀ	to split; to chop
pɛ̃k	ˈpɛ̃kˀ	to prick (of thorns)
lapɛ̃k	laˈpɛ̃kˀ	mud
bɛk*	ˈbɛkˀ	bag
ckbɛk	cᶜəkˀˈbɛkˀ	to spear
hrkbɛk ~ hrkbãk	hərəkˀˈbɛkˀ ~ hərəkˀˈbãkˀ	to sink to the ground (of animal shot)
tɛk	ˈtɛkˀ	[ethnonym: other group of Jahai]; [sound of snails (?)]
batɛk	baˈtɛkˀ	[ethnonym: Batek]
baltɛk	balˈtɛkˀ	friend
kktɛ̃k	kəkˀˈtɛ̃kˀ	to click (with the tip of the tongue towards the palate)
hntɛk*	hənˈtɛkˀ	to hit; to pound
kdɛk	kəˈdɛkˀ	to be bitter
ʔɲcɛk	ʔəɲˈcᶜɛkˀ	[a type of tree]
ʔɛk	ˈʔɛkˀ	to give
sksɛ̃k	səkˀˈsɛ̃kˀ	to devour quickly
hɛ̃k	ˈhɛ̃kˀ	to snap and fall (of e.g. tree-branch)
hkhɛ̃k	həkˀˈhɛ̃kˀ	to be ragged/torn (of cloth)
ʔanɛk	ʔaˈnɛkˀ	girl
lanɛk*	laˈnɛkˀ	[a type of porcupine]
tulɛk*	tuˈlɛkˀ	to push someone or something
blɛk	bəˈlɛkˀ	to lick
snrɛk	sɜnᵈˈɹɛkˀ	to go out
tawɛ̃k	tãˈwɛ̃kˀ	butterfly (generic)
lajɛ̃k	lãˈʃɛ̃kˀ	to tread
cik	ˈcᶜikˀ	to wade across a river
ʄkʄik	ʄᶻəkˈʄᶻikˀ	to breathe
bkik	bəˈkikˀ	pigeon (generic)
sik	ˈsikˀ	to be caught
hḭ̃k	ˈhḭ̃kˀ	to breathe with difficulty

rhɨk	rəˈhɨkˀ	red
lɨk	ˈlɨkˀ	to swallow (of snake)
klɨk	kəˈlɨkˀ	[sound of a person walking]
sklɨk	səkˀˈlɨkˀ	to bind
rɨk	ˈrɨkˀ	to stab
ʔawɨk	ʔaˈwɯkˀ	[a type of small tree-dwelling mammal]
ckwɨk	cᶜəkˀˈwɯkˀ	to crackle; to bellow (of e.g. macaque, gaur, rhinoceros); to talk/chat
pkpə̃k	pəkˀˈpə̃kˀ	to smack
bə̃k	ˈbə̃kˀ	to untie
kktə̃k	kəkˀˈtə̃kˀ	to click (with the tip of the tongue towards the alveolar ridge)
carək*	cᶜaˈrəkˀ	stream-bed
pak	ˈpakˀ	to clap
napak	naˈpakˀ	wild boar (*Sus scrofa*)
clapak	cᶜəlaˈpakˀ	boar/male pig
dkdak	dəkˀˈdakˀ	fish-trap
pʔak	pəˈʔakˀ	to cook
krhak	kərˈhakˀ	to clear one's nasal cavity
lak	ˈlakˀ	to have sore eyes
calak	cᶜaˈlakˀ	lizard (generic)
klak	kəˈlakˀ	[sound of something falling]
krak	kəˈrakˀ	[a type of plant/fruit]
kujak	kuˈjakˀ	to butcher; to tear into pieces
kapuk	kaˈpukˀ	flying squirrel (*Petaurista*)
tbuk*	təˈbukˀ	to peck (of bird)
duk	ˈdukˀ	to pounce upon
taduk	taˈdukˀ	[a type of tree]
dkduk	dəkˀˈdukˀ	chest
pucuk* lɔc	puˈcᶜukˀ ˈlɔⁱcˀ	arrowhead
tkuk	təˈkukˀ	[a type of hornbill]
ʔuk	ˈʔukˀ	to blow
suk	ˈsukˀ	to light up
skŋuk	səkˀˈŋūkˀ	[a type of frog]
haluk	haˈlukˀ	[a type of lizard]
ʄuluk	ʄᶻuˈlukˀ	to keep a fire burning
lkluk	ləkˀˈlukˀ	to laugh
rkruk	rəkˀˈrukˀ	to go along a watercourse
trjuk	tɜrˈjukˀ	[ethnonym: different group of Jahai]
pok	ˈpokˀ	to break open; to stride; [sound of an animal falling to the ground]
cok	ˈcᶜokˀ	to tear apart
ckcok	cᶜəkˀˈcᶜokˀ	Diard's trogon (*Harpactes diardii*)
ʄok	ˈʄᶻokˀ	to move from one place to another
krkok	kərˈkokˀ	casque of hornbill
chok	cᶜəˈhokˀ	[sound of running water or waterfall]

planok*	pəla'nõk˺	mouse deer (*Tragulus*)
pɔk	'pɔk˺	round object
lumpɔk	lum'pɔk˺	[a type of tree]
cmalpɔk	cᶜəmal'pɔk˺	[a type of millipede]
btɔk	bə'tɔk˺	malkoha; cuckoo (*Phaenicophaeus*; *Cuculus*); papaya (*Carica papaya*)
kluktɔk	kəluk˺'tɔk˺	[a type of large terrestrial gastropod]
pktɔk	pək˺'tɔk˺	to cook meat in its own juice
dɔk	'dɔk˺	ipoh poison
cɔk	'cᶜɔk˺	bag
ku'cɔ̃k	ku'cɔ̃k˺	[a type of bird]
ckcɔk	cᶜək˺'cᶜɔk˺	banded palm civet (*Hemigalus derbyanus*)
kɔk	'kɔk˺	[sound of a hornbill]
crikɔ̃k	cᶜəri'kɔ̃k˺	to jabber/chatter (of monkey)
hakɔk	ha'kɔk˺	to throw
hokɔ̃k	ho'kɔ̃k˺	to burn off fur/feathers from a killed animal
hɔkɔ̃k	hɔ'kɔ̃k˺	[sound of a leaf-monkey]
kikkɔk	kik˺'kɔk˺	to sit with one's knees pulled up
tŋkɔk	təŋ'kɔk˺	nape of the neck
ʔɔk	'ʔɔk˺	[sound of boiling]
ʄuʔɔk	ʄᶻu'ʔɔk˺	tiger (*Panthera tigris*)
sɔk	'sɔk˺	hair
sɔk kawɔ̃t	'sɔk˺ kã'wɔ̃t˺	feather; plumage; down
sɔk kuj	'sɔk˺ 'kuj̃	hair of the head
sksɔ̃k	sək˺'sɔ̃k˺	to suck out (e.g. a snail from its shell)
rksɔk	rək˺'sɔk˺	keelback (generic) (*Xenochrophis*, *Amphiesma*)
hɔk	'hɔk˺	to throw
bhɔk	bə'hɔk˺	to vomit
ŋɔk	'ŋɔ̃k˺	to sit; to stay; to reside
sŋɔk	sə'ŋɔ̃k˺	[a type of bird]
brŋɔ̃k	bərˈŋɔ̃k˺	[a type of toad]
malɔk*	ma'lɔk˺	to be embarrassed
wɔk	'wɔk˺	to dig tubers
wawɔk	wa'wɔk˺	to have a sad expression on one's face
jɔk	'jɔk˺	to undress; to stretch someone's arm
tjɔ̃k ~ tkjɔ̃k	ti'jɔ̃k˺ ~ tək˺'jɔ̃k˺	to point
tkjɔk	tək˺'jɔk˺	to pull; to pluck

/-ʔ/

ʄampiʔ*	ʄᶻam'piʔ	to heal
labiʔ*	la'biʔ	[a type of soft-shelled turtle]
hobiʔ*	ho'biʔ	root-crop; tuber (generic)
hatĩʔ	ha'tĩʔ	tail
ptiʔ*	pə'tiʔ	box

bldi?*	bələ'di?	bucket
ci?	'c^{c}i?	to flare up; louse
?acĩ?	?a'c^{c}ĩ?	dog
cuci?*	c^{c}u'c^{c}i?	to clean
kuɲci?*	kuɲ'c^{c}i?	key; to mend
biʄi?*	bi'ʄ^{ʒ}i?	seed; CLF: small objects
taʄi?*	ta'ʄ^{ʒ}i?	knife
bɲʄi?	məɲ'ʄ^{ʒ}i?	to be distant (horizontally)
pagi?*	pa'gi?	morning
lagi?*	la'gi?	again; also; moreover; still; yet
si?	'si?	number
mɛj si?	'mɛj 'si?	how many?
?isi?*	?i'si?	to put inside; to insert
nasi?*	na'si?	cooked rice
bsi?*	bə'si?	iron
bhi?	bə'hi?	to be full/to have had enough to eat
mi?	'mĩ?	PRONOUN 2S.INT
tani?	ta'nĩ?	that (away)
?ani?	?a'nĩ?	there (away)
skali?*	səka'li?	all; together
hali?	ha'li?	leaf; CLF: small flat objects
hali? gadiŋ	ha'li? ga'di^{g}ŋ	[a type of leaf]
bli?*	bə'li?	to buy
ʄli?	ʄ^{ʒ}ə'li?	[toponym: Jeli]
ʄari?*	ʄ^{ʒ}a'ri?	finger
mnri?	mən'^{d}ɹi?	[ethnonym: Menriq]
wĩ?	'wĩ?	left (side)
dwi?*	də'wi?	money
m?wĩ?	ma?'wĩ?	left-handed person
ji?	'ji?	to cry (of infant); to dislike; to refuse; to reject
kaji?	ka'ji?	bat
kajĩ?	kã'jĩ?	little finger
haji?	ha'ji?	also
te?	'tẹ?	earth; ground; soil
?ite?*	?i'tẹ?	duck; water-fowl
ste?*	sə'tẹ?	to shoot with slingshot
de?	'dẹ?	to make; to do
tase?*	ta'sẹ?	lake
he?	'hẹ?	PRONOUN 1P.INCL
male?	ma'lẹ?	brown hornbill (*Ptilolaemus tickelli*)
cole?	c^{c}o'lẹ?	cannibal
ʄle?	ʄ^{ʒ}ə'lẹ?	thorn
ti?le?	ti?'lẹ?	to point (with one's finger, hand or lips); to show

tareʔ*	ta'rẹʔ	to pull
weʔ	'wẹʔ	to exist; to be found; to be present; to occur; to be available
pẽʔ	'pẽʔ	older sibling
pcẽʔ	pə'cᶜẽʔ	to be wet
joh waŋkẽʔ	jowaŋ'kẽʔ	tree-branch
sisɛʔ	si'sɛʔ	to dance
prisɛʔ	pəri'sɛʔ	to keep
gasɛʔ	ga'sɛʔ	[a type of tree]
brasɛʔ	bəra'sɛʔ	to remain; to stay behind
tahɛʔ	ta'hɛʔ	[a type of salty tuber]
cuhẽʔ	cᶜũ'hẽʔ	to flow; to spread
mɛʔ	'mẽʔ	female child
limɛʔ*	li'mẽʔ	five
gamɛʔ	ga'mẽʔ	to halt/stop
samɛʔ*	sa'mẽʔ	with; likewise
cmɛʔ	cᶜə'mẽʔ	to be sharp
cinɛʔ*	cᶜinẽʔ	[ethnonym: Chinese]
trinɛʔ	təri'nẽʔ	height
banɛʔ	ba'nẽʔ	to give birth
bnɛʔ	mə̃'nẽʔ	size
knɛʔ*	kə'nẽʔ	to score a bull's eye
taɲɛʔ*	ta'ɲɛʔ	to ask
puɲɛʔ*	puɲɛʔ	to have
ɲʔɲɛʔ	ɲāʔ'ɲẽʔ	to kill an animal
siŋɛʔ	si'ŋẽʔ	[toponym: Singor]
buŋɛʔ*	bu'ŋẽʔ	flower
brilɛʔ	bəri'lɛʔ	to pull off (e.g. a mat from the ground); to roll (of animal); to turn around
liʔlɛʔ	liʔlɛʔ	to call for somebody; to rub against (of e.g. cat)
luʔlẽʔ	luʔlẽʔ	to roll (of animal)
gʔlɛʔ	gaʔlɛʔ	to say no; to decline
slɛʔ	sə'lɛʔ	[a type of snake]
ciwẽʔ	cᶜĩ'wẽʔ	to open one's eyes
ʔawẽʔ	ʔā'w̃ẽʔ	[a type of tortoise]
mawẽʔ	mā'w̃ẽʔ	gibbon (*Hylobates*)
smwɛʔ*	səm'wɛʔ	all
jɛʔ	'jɛʔ	PRONOUN 1S
hajẽʔ	hā'j̃ẽʔ	house; hut
hapɨʔ	ha'pəʔ	bag
bɨʔ	'bəʔ	mother
bɨʔ masəj	'bəʔ ma'səj	sow
kbɨʔ	kə'bəʔ	fruit; CLF: animals, several objects
kbɨʔ ʔntɛp	kə'bəʔ ʔən'tɛpˀ	testicles
knbɨʔ	kən'bəʔ	[a type of rattan]

tiʔ	'təʔ	to pound
tiʔ siriŋ	'təʔ si'riᵍŋ	[a type of tree-dwelling mammal]
ptiʔ	pə'təʔ	forehead
ktiʔ	kə'təʔ	skin
tʔtiʔ	taʔ'təʔ	old; grandparent
ʄiʔ	'ʄᶻəʔ	to burn; to make fire; to roast
ʄaʄiʔ ~ ʄaʄaʔ	ʄᶻa'ʄᶻəʔ ~ ʄᶻa'ʄᶻaʔ	old woman
bɲʄiʔ	mə̃ɲ'ʄᶻəʔ	to be distant (vertically); to be high; to be deep
kiʔ	'kəʔ	to vomit
piʔ̃iʔ	pĩʔ̃əʔ	to put a child to sleep
haʔiʔ	ha'ʔəʔ	yes; right; correct
cʔiʔ	cᶜəʔəʔ	to pour fluid
bhiʔ	bə'həʔ	to hug
mamiʔ	mã'məʔ	[a type of colourful larvae]
baniʔ	mã'nəʔ	quiver for blowpipe darts
tɲiʔ	tə'ɲəʔ	that (beyond you)
ʔɲiʔ	ʔə'ɲəʔ	there (beyond you)
bliʔ	bə'ləʔ	upper leg
piʄəʔ	pi'ʄᶻəʔ	to fly up
ʔikəʔ	ʔi'kəʔ	fish
jaŋkəʔ	jaŋ'kəʔ	jaw/mandible
sipaʔ*	si'paʔ	to kick
hapaʔ	ha'paʔ	to die
lʔpaʔ	laʔ'paʔ	[ethnonym: unidentified easterly group of Semang]
nampaʔ*	nam'paʔ	to see
ʄumpaʔ*	ʄᶻum'paʔ	to meet
cubaʔ*	cᶜu'baʔ	to taste
kbaʔ	kə'baʔ	[a type of tuber]
rbaʔ	rə'baʔ	rotten wood
taʔ	'taʔ	ancestor; ancestral spirit; father's father
mataʔ	ma'taʔ	spear
lataʔ*	la'taʔ	waterfall
kntaʔ ~ gntaʔ	kən'taʔ ~ gən'taʔ	[ethnonym: Kintaq]
badaʔ*	ba'daʔ	Sumatran rhinoceros (*Dicerorhinus sumatrensis*)
ladaʔ*	la'daʔ	pestle
kudaʔ*	ku'daʔ	horse
ldaʔ	lə'daʔ	armpit
bacaʔ*	ba'cᶜaʔ	to talk
ciʔcaʔ	cᶜiʔ'cᶜaʔ	gecko
cʔcaʔ	cᶜaʔ'cᶜaʔ	common kingfisher (*Alcedo atthis*)
saʄaʔ*	sa'ʄᶻaʔ	only
pʄaʔ	pə'ʄᶻaʔ	to hide oneself
krʄaʔ*	kər'ʄᶻaʔ	to work

luka?*	lu'ka?	to hit a target; to wound
paŋka?*	paŋ'ka?	to hit; to throw
ɟaŋka?*	ɟᶻaŋ'ka?	to go
slaŋka?*	səlaŋ'ka?	collar-bone
bŋka?	məŋ'ka?	molar tooth
tiga?*	ti'ga?	three
ɟaga?	ɟᶻa'ga?	only
dahaga?*	daha'ga?	to be thirsty
raga?*	ra'ga?	basket
tga?*	tə'ga?	to flex muscle
t?a?	ta'?a?	vegetables
t?a? pɔk	ta'?a? 'pɔk˜	[a type of mushroom]
pasa?	pa'sa?	to live
kasa?	ka'sa?	sambar deer (*Cervus unicolor*)
brasa?	bəra'sa?	many
piha?	pi'ha?	to be separate
tuha?*	tu'ha?	to ripen
cnha?	cɜn'ha?	to jest
ma?	'ma?	mother
lma?*	lə'ma?	intestines
la?	'la?	penis
hala?	ha'la?	to fly
tula?*	tu'la?	to push someone to go somewhere
mula?*	mu'la?	to begin
bla?	bə'la?	to be alone
kira?*	ki'ra?	to count
pihira?	pihi'ra?	to rear animals
cara?	cᶜa'ra?	to talk
pra?	pə'ra?	to live/reside
bra?	bə'ra?	NEGATIVE MARKER
pnra?	pən'ᵈɹa?	[a type of sweet-smelling flower]
mnra?	mən'ᵈɹa?	human being; people; person
?awa?	?a'wa?	elder
lawa?	la'wa?	to collect/gather
duwa?*	du'wa?	two
duwa? puloh*	du'wa? pu'loh	twenty
brwa?	bər'wa?	wind
ja?	'ja?	grandmother
caja?	cᶜa'ja?	to believe
laja?	la'ja?	leech
sraja?*	səra'ja?	[a type of tree (*Shorea*)]
buja?*	bu'ja?	crocodile (*Tomistoma*, *Crocodilus*)
hrja?	hər'ja?	[a type of tuber]
mapu?	ma'pu?	when?
?ampu?*	?am'pu?	to convey

kabuʔ	kaˈbuʔ	[a type of tuber]
ʔabuʔ*	ʔaˈbuʔ	dust
ʔabuʔ* ʔɔs	ʔaˈbuʔ ˈʔɔs	ashes
labuʔ*	laˈbuʔ	bottle
batuʔ*	baˈtuʔ	stone
pintuʔ*	pinˈtuʔ	door
bantuʔ*	banˈtuʔ	to catch something that is falling
hantuʔ*	hanˈtuʔ	ghost
braduʔ	bəraˈduʔ	to rest
prduʔ*	pərˈduʔ	CLF: cluster
cũʔ	ˈcᶜũʔ	to be blind
baɟuʔ*	baˈɟᶻuʔ	clothes
taɟuʔ	taˈɟᶻuʔ	snake (generic)
sɟuʔ*	səˈɟᶻuʔ	to be cold (of weather, objects etc.)
pakuʔ*	paˈkuʔ	*Filex* fern
takuʔ ~ ʔakuʔ	taˈkuʔ ~ ʔaˈkuʔ	to steal
blakũʔ	bəlaˈkũʔ	to be brown
bukuʔ*	buˈkuʔ	book
pusuʔ*	puˈsuʔ	termite mound
hũʔ	ˈhũʔ	to make hooting sound
prahuʔ*	pəraˈhuʔ	boat
ɟhũʔ	ɟᶻəˈhũʔ	tree
bamuʔ*	mãˈmũʔ	bamboo
somuʔ	soˈmũʔ	tapir's snout; horn of rhinoceros (?)
gmuʔ*	gəˈmũʔ	fat/grease
lmuʔ*	ləˈmũʔ	cattle
sanuʔ	saˈnũʔ	ghost
ɲaɲuʔ	ɲãˈɲũʔ	to die
baruʔ*	baˈruʔ	again; anew; until
baboʔ	baˈboʔ	female; woman
taboʔ	taˈboʔ	big digit (thumb or big toe)
loboʔ*	loˈboʔ	plunge-pool beneath waterfall
toʔ	ˈtoʔ	ancestor; grandparent; [ethnonym: To']
blatoʔ	bəlaˈtoʔ	crimson-winged woodpecker (*Picus puniceus*)
ɟoʔ	ˈɟᶻoʔ	[toponym]
kokoʔ	koˈkoʔ	to crow (of cock)
pgoʔ	pəˈgoʔ	to cook
ʔoʔ	ˈʔoʔ	PRONOUN 3S
sʔoʔ	səˈʔoʔ	a little; just
soʔ	ˈsoʔ	cubit (measurement)
sʔsoʔ	saʔˈsoʔ	blood vessel
tumoʔ*	tuˈmõʔ	to fight
tanoʔ*	taˈnõʔ	fin; horn
ʔiʔŋoʔ	ʔiʔˈŋõʔ	to lean one's head

buloʔ*	buˈloʔ	bamboo; bamboo tube
siʔroʔ	siʔˈroʔ	to decline
lawoʔ*	laˈwoʔ	fish/meat
joʔ	ˈjoʔ	outer shaft of blowpipe; to carry on one's shoulders
pɔʔ	ˈpɔʔ	mountain-top
kapɔʔ*	kaˈpɔʔ	axe; cheek
bʔbɔʔ	baʔˈbɔʔ	to carry on one's back
ktɔʔ	kəˈtɔʔ	day; sky
ktɔʔ tə̃h	kəˈtɔʔ ˈtə̃h	today
də̃ʔ	ˈdə̃ʔ	[toponym: Dok]
tadɔʔ	taˈdɔʔ	to wait
cundɔʔ*	cᶜunˈdɔʔ	to rest one's chin on something
cə̃ʔ	ˈcᶜə̃ʔ	to sting
kaɲcɔʔ	kaɲˈcᶜɔʔ	grandchild
ʔaʄɔʔ	ʔaˈʄᶻɔʔ	a little; to be small
pɲʄɔʔ	pəɲˈʄᶻɔʔ	blanket; fabric
ciˈkə̃ʔ	cᶜiˈkə̃ʔ	long-tailed shrike (*Lanius schach*)
bakɔʔ	baˈkɔʔ	trap/snare
takɔʔ	taˈkɔʔ	cup; glass
makɔʔ	maˈkɔʔ	egg
maŋkə̃ʔ	mãŋˈkə̃ʔ	to conceive; to be pregnant
tŋkɔʔ	təŋˈkɔʔ	[a type of bird]
klkɔʔ	kəlˈkɔʔ	nail
sɔʔ	ˈsɔʔ	to burn
ʔɲsɔʔ	ʔəɲˈsɔʔ	rotten (of wood)
hɔʔ	ˈhɔʔ	jar
mɔʔ	ˈmə̃ʔ	younger sister of parent
timɔʔ	tiˈmɔʔ	hard surface (stone, paved road, brick wall etc.)
camɔʔ	caˈmɔʔ	day after tomorrow
kmɔʔ	kəˈmɔʔ	fruit; pumpkin; CLF: animals, several objects
taŋɔʔ	taˈŋɔʔ	[a type of short millipede]
ʄŋɔʔ*	ʄᶻəˈŋə̃ʔ	to look down
lɔʔ	ˈlɔʔ	[toponym]
ʄalɔʔ*	ʄᶻaˈlɔʔ	casting-net
kalə̃ʔ	kaˈlə̃ʔ	[a type of large freshwater gastropod]
ʔalɔʔ	ʔaˈlɔʔ	and; with
ʄlɔʔ	ʄᶻəˈlɔʔ	hole
sarɔʔ	saˈrɔʔ	corpse
crɔʔ	cᶜəˈrɔʔ	to be hungry
krɔʔ	kəˈrɔʔ	back (of person)
karwɔʔ	karˈwɔʔ	[a type of terrapin]
piʔjɔʔ	piʔˈjɔʔ	to roar (of tiger)

/-ɸ/

ciɸ	'cᶜiɸ	to hiss (of e.g. bamboo in fire)
kiciɸ	ki'cᶜiɸ	to crackle/explode from heat
hĩɸ ~ hẽɸ	'hĩɸ ~ 'hẽɸ	to whistle shrilly (with one's fingers in one's mouth); to blow a fire
lĩɸ ~ lɸlĩɸ	'lĩɸ ~ ləp'lĩɸ	fontanelle
ciɾ̃iɸ	ci'ɾ̃iɸ	to extinguish; to fade away; to put out fire with water
priɸ ~ piɸriɸ	pə'riɸ ~ pip'riɸ	to flicker
cẽɸ	'cᶜẽɸ	to fan fire
cnɛɸ	cᶜə'nẽɸ	tail feathers
ɲɨɸ	'ɲɾ̃ɨɸ	to blow smoke
ɾɨɸ	'ɾɨɸ	[sound of flapping]
jɨɸ	'jɨɸ	to shake something (e.g. a piece of paper)
jaɸ	'jaɸ	[sound of flying or leaping]
tuɸ	'tuɸ	to spit
ʔuɸ	'ʔuɸ	to blow (e.g. a fire)
suɸ	'suɸ	to attack
ɲuɸ	'ɲũɸ	to haunt
luɸ ~ ɾuɸ	'luɸ ~ 'ɾuɸ	[sound of dashing]
cruɸ	cə'ruɸ	to descend
koɸ	'koɸ	[sound of blowpipe dart hitting stomach of prey]
ktɔɸ	kə'tɔɸ	to spit/splutter horizontally
cɔɸ	'cᶜɔɸ	to cause a blister (of fire)
kɔɸ	'kɔɸ	[sound of small object falling to the ground]
bsɔɸ	bə'sɔɸ	to be quick
hɔɸ	'hɔɸ	to prick a hole in something (e.g. a leaf)
lɔɸ	'lɔɸ	to leak
tilɔɸ	ti'lɔɸ	to dig
plɔ̃ɸ	pə'lɔ̃ɸ	[sound of a blowpipe being fired]
ɟlɔɸ	ɟᶻə'lɔɸ	to whiz
rɔɸ	'rɔɸ	[sound of a blowpipe dart hitting canopy]

/-s/

lipis*	li'pis	to be thin
tpis	tə'pis	to blow away; to wash away
cbis*	cᶜə'bis	to cleave
kbis	kə'bis	to die
tĩs bataŋ*	'tĩs ba'taᵍŋ	[a type of insect]
britis*	bəri'tis	[ethnonym: British]
ptis	pə'tis	to hurt; pain; sickness
ktis*	kə'tis	to break/snap
cscis	cis'cᶜis	to scratch oneself
kis	'kis	to dig

kĩs	'kĩs	ghost
ciʔĩs	cĩ'ʔĩs	to ask for something
hĩs	'hĩs	to blow one's nose; to snuffle
ɟnis*	ɟᶻə'nis	kind/sort
maŋis*	mā'ŋĩs	mangosteen (*Garcinia manggostana*)
klaŋis	kəla'ŋĩs	heart
tŋis	tə'ŋĩs	[a type of fish]
piris	pi'ris	to go across a watercourse
hiris*	hi'ris	to cut
wĩs	'wĩs	to split/to go different ways
bliwĩs	bəli'wĩs	to get out of bed
cawis	cᶜa'wis	[a type of small animal]
rwis	rə'wis	to cut grass
tjis	tə'jis	to bounce back; to kick something to pieces
tsdes	tis'dẹs	to stumble
ges	'gẹs	fragrance; odour; smell; stench
ɟʔes	ɟᶻə'ʔẹs	root
hampɛs	ham'pɛs	to live/reside
kmpẽs	kəm'pẽs	[a type of tree]
tbɛs*	tə'bɛs	to cut
lintɛs*	lin'tɛs	to go across
dɛs	'dɛs	to go and get
csɛcɛs	cᶜis'cᶜɛs	to tear
kɛs	'kɛs	to put together
bakɛs	ba'kɛs	to grow up; to be adult
gɛs	'gɛs	to descend; to go downhill
ʔɛs	'ʔɛs	ice
mɛs*	'mẽs	gold
plɛs	pə'lɛs	to smear poison on blowpipe dart
plẽs	pə'lẽs	[sound of blowpipe being fired]
rɛs	'rɛs	to fall
wẽs	'wẽs	frontal tuber (on forehead)
ciwɛs	cᶜi'wɛs	[toponym]
tawɛs	ta'wɛs	[a type of tree]
kwẽs	kə'wẽs	to sweep with one's fingers
lwɛs	lə'wɛs	to be wide
pjɛs*	pi'jɛs	R.P.S. (Rancangan Pengumpulan Semula; site of regroupment program)
pɨs	'pɨs	to sweep/wipe
tpɨs	tə'pɨs	valley
ɟkɨs	ɟᶻə'kɨs	[a type of porcupine]
taŋkɨs	taŋ'kɨs	?
gɨs	'gɨs	to descend/to climb down/to go down; to apply make-up
prgɨs	pər'gɨs	[a type of tree]

ʔis	'ʔis	to die
laʔis	la'ʔis	to be bad
nis	'nɨ̃s	mat
lajis	la'jis	[a type of palm leaf used for thatching]
lpəs*	lə'pəs	after; to leave
kə̃s	'kə̃s	to kill (?)
gəs	'gəs	to carve
lʔəs	lə'ʔəs	fatty tissue
kipas*	ki'pas	to wag tail
bas*	'bas	bus
tas	'tas	to cleave
kas	'kas	to pinch somebody
gas ~ gɛ̃s	'gas ~ 'gɛ̃s	[a type of skin disease]
hagas*	ha'gas	mosquito
ʔas	'ʔas	to believe; to have faith
has	'has	to steam (intransitive)
lanas*	la'nas	pineapple (*Ananassa*)
tronas*	təro'nas	[toponym: Petronas petrol station at Banding]
balas*	ba'las	to answer
sblas*	səbə'las	eleven
tlas	tə'las	[a type of tree]
bras*	bə'ras	husked rice
kras*	kə'ras	to be tough (of meat)
was	'was	to split; fork/junction
was tɔm	'was 'tɔ^{b}m	river confluence
puwas*	pu'was	to be satisfied
ʔuswas	ʔus'was	to meet (of rivers)
cjas	ci'jas	hand
ʔimbus*	ʔim'bus	to ambush
tũs ~ tmus	'tũs ~ tə'mũs	to grub; to grunt (of pig)
tstũs	tis'tũs	to lose hair
gus	'gus	to be together
nus	'nũs	upper lip
tanus	ta'nũs	boar's snout
barus*	ba'rus	camphor
kurus*	ku'rus	thin
pos*	'pos	post
sratos*	səra'tos	hundred
dos	'dos	to owe (?)
gos	'gos	belch
phos	pə'hos	to blow (of person)
hamos	ha'mos	to blow (of person)
ros	'ros	liver
tros*	tə'ros	at once; straightaway; straight

tpɔs	tə'pɔs	[a type of fruit]
cɔs	'cᶜɔs	[toponym: Banun]
kɔ̄s	'kɔ̄s	to scratch
gɔs	'gɔs	to live/exist
ʔɔs	'ʔɔs	fire
lanʔɔs	lan'ʔɔs	door
cnɔs ~ cnus	cᶜə'nɔs ~ cᶜə'nūs	ring
lslɔs	lis'lɔs	to gnaw
cnrɔs	cᶜɜnᵈɹɔs	nail; phalange

/-h/

pipih	pi'pih	[a type of rattan]
ctih	cᶜə'tih	to be easy
titih*	ti'tih	bridge
didih*	di'dih	to boil (intransitive)
cih	'cᶜih	to shoo away
gcĩh	gə'cᶜĩh	to be black; to be dark
haʔĩh	ha'ʔĩh	yes; right; correct
sih	'sih	to pound
gunih*	gu'nĩh	gunny bag; sack
malih*	ma'lih	[a type of palm-tree]
tlĩh	tə'lĩh	to fall to the ground
wih	'wih	PRONOUN 3D
tujih	tu'jih	that (down)
ʔujih	ʔu'jih	there (down)
lbeh	lə'bẹ̣h	honey
puteh*	pu'tẹ̣h	to be white
thteh	tah'tẹ̣h	oriental pied hornbill (*Anthracoceros albirostris*)
tadeh	ta'dẹ̣h	that (beyond me)
ʔadeh	ʔa'dẹ̣h	there (beyond me)
pihdeh	pih'dẹ̣h	to summon; to call for
seh	'sẹ̣h	to be good
kneh	kə'nẹ̃h	wife
haleh	ha'lẹ̣h	to be hungry
boleh*	bo'lẹ̣h	ROOT POSSIBILITY; to be able to
koleh	ko'lẹ̣h	cup
ciweh	ci'wẹ̣h	to climb up
ɟɔweh	ɟɔ'wẹ̣h	[a type of bird]
rajeh	ra'jẹ̣h	to beat (e.g. a pillow)
klipɛh	kəli'pɛh	to flay/skin
japɛ̄h ~ pɛ̄h	ja'pɛ̄h ~ 'pɛ̄h	PRONOUN 1P.EXCL
bɛh	'bɛh	younger brother of parent
dibɛh	di'bɛh	wallet
tɛh*	'tɛh	tea

cɛ̃h	'cᶜɛ̃h	[sound of a blowpipe dart hitting a vine]
trɟɛh	tɜrˈɟᶻɛh	to fly up (of bird)
manɛh	mã'nɛ̃h	to be long (of time); to be old
lɛh*	'lɛh	to go together; EMPHATIC PARTICLE
ʔalɛh	ʔa'lɛh	girl; virgin
bnalɛh	mɜ̃na'lɛh	to breastfeed
kralɛh	kəra'lɛh	giant squirrel (*Ratufa affinis*) (?)
plɛh	pə'lɛh	[ethnonym: Temiar]
turɛh	tu'rɛh	to tap poison from ipoh tree
wɛh	'wɛh	to cut off/to pick
wɛ̃h	'w̃ɛ̃h	[sound of blowpipe]
jɛh	'jɛh	PRONOUN 1D.EXCL
pihpɨh	pih'pəh	to wake up somebody
titɨh	ti'təh	that (up)
ʔitɨh	ʔi'təh	there (up)
kdɨh	kə'dəh	to say
cɨh	'cᶜəh	to fell/to cut down trees
ɟɨh	'ɟᶻəh	PRONOUN 2D
kɟɨh	kə'ɟᶻəh	boy
tkɨh	tə'kəh	behind; backside
rigɨh	ri'gəh	to break
pʔɨh	pə'ʔəh	fluid
sɨh	'səh	[a type of tuber (sweet potato?)]
rwɨh	rə'wəh	[toponym]
klapəh	kəla'pəh	shoulder
srupəh*	səru'pəh	same
tɜ̃h	'tɜ̃h	this
plitəh*	pəli'təh	small fire; lamp
kritəh*	kəri'təh	car
pdəh*	pə'dəh	to be near
kəh	'kəh	to lie down
pokəh	po'kəh	[a type of gecko]
ʔɜ̃h	'ʔɜ̃h	here
pʔəh	pə'ʔəh	to cook/heat in a fire
sɜ̃h ~ sihsɜ̃h	'sɜ̃h ~ sih'sɜ̃h	to meet
həh	'həh	to whistle with one's mouth open
maŋəh	mã'ŋɜ̃h	[toponym: Mangga]
guləh*	gu'ləh	sugar
bləh	bə'ləh	to enter
klapah sawit*	kəla'pah sawit	oil palm
japɜ̃h	ja'pɜ̃h	[ethnonym: group of Jahai]
cpah	cᶜə'pah	wood shavings from carving or scraping
kbah	kə'bah	to glare
lbah	lə'bah	where?
lbah lwɛj	lə'bah lə'wɛ̃j	honeybee

mutah*	mu'tah	mortar
gtah*	gə'tah	rubber
kdah*	kə'dah	[toponym: Kedah]
cah	'cah	[a type of tree]; to cut
ʔacah*	ʔa'cᶜah	to scold
pcah*	pə'cᶜah	to break; to move from one place to another
gaɟah*	ga'ɟᶻah	elephant (*Elephas maximus*)
tikah	ti'kah	to travel by air; to spring
prihkah	pərih'kah	to chatter (of macaque, people etc.)
laŋkah*	laŋ'kah	to step over
bŋkah ~ mŋkah (?)	mãŋ'kah	[toponym]
gagah	ga'gah	to walk
pʔah	pa'ʔah	to kneel; to sit by fire
sisah*	si'sah	remains
gamah*	ga'mah	photograph
lmah	lə'mah	to find; to meet
rnah*	rə'nãh	to be low (of height, voice etc.)
kalah*	ka'lah	to lose (e.g. a fight)
hmalah	həma'lah	young pig
blah*	bə'lah	to chop; to wane (of moon)
clah	cᶜə'lah	childless person
llah	lə'lah	oriole (*Oriolus*)
rah	'rah	to vomit
darah*	da'rah	blood
carah	cᶜa'rah	barbet (*Megalaima*)
marah*	ma'rah	to be angry; to forbid
buwah*	bu'wah	CLF: spherical/cubical objects
sapuh*	sa'puh	to clean/to sweep/to wipe
sribuh*	səri'buh	thousand
sltuh	sɜl'tuh	to attack
duduh	du'duh	[a type of bird]
suduh*	su'duh	to scrape
ɟuh	'ɟᶻuh	to blow away smoke
khkuh	kah'kuh	[a type of hornbill]
susuh*	su'suh	milk
bunuh*	bu'nuh	to kill
ɟnuh*	ɟᶻə'nuh	too many/much; enough
ɟnuh ~ ɟnūh (?)	ɟᶻə'nūh	[a type of small bird]
kaluh*	ka'luh	if
haluh	ha'luh	to shoot
puŋhuluh*	puŋhu'luh	headman
ruh	'ruh	[a type of wasp]
pluruh*	pəlu'ruh	bullet
biruh*	bi'ruh	blue
krjuh	kər'juh	to slither

boh	'boh	to place/put
ɟatoh	ɟᶻa'toh	to move; to wander
ctoh	cᶜə'toh	phlegm; snot
coh	'cᶜoh	to refuse
tuɟoh*	tu'ɟᶻoh	seven
paŋkoh*	paŋ'koh	to hold
saŋkoh	saŋ'koh	wreathed hornbill (*Rhyticeros undulatus*)
goh	'goh	manchild
pgoh	pə'goh	to boil (transitive)
ʔoh	'ʔoh	to cough
soh	'soh	game animal
ɟnoh	ɟᶻə'nõh	[a type of tree]
spuloh*	səpu'loh	ten
ruroh	ru'roh	to shed leaves
woh ~ wɔh	'woh ~ 'wɔh	to call (of bird)
joh	'joh	crown of tree; tree branch
pɔh	'pɔh	to fan; to hit with a flat hand
tapõh	ta'põh	pelvis
bɔh	'bɔh	fruit (generic)
brubɔh	bəru'bɔh	yellow-vented bulbul (*Pycnonotus goiavier*)
tbɔh*	tə'bɔh	to beat/to hit
mantɔh	mãn'tɔh	Malay person
lantõh	lan'tõh	porcupine (*Hystrix*, *Atherura*)
klihdõh	kəlih'dõh	to meander (of river)
cɔh	'cᶜɔh	to bite (of snake, bird etc.)
chcɔh	cᶜah'cᶜɔh	to mince
kalcõh	kal'cᶜõh	to crack
ɟɔh	'ɟᶻɔh	to ascend; to go uphill
nhɟɔh	nah'ɟᶻɔh	height
kɔh	'kɔh	[a type of tortoise]
taŋgõh	taŋ'gõh	to ring
kuhʔõh	kūh'ʔõh	[a type of turtle]
mɔh	'mõh	nose; PRONOUN 2S.FAM
mɔh tɔm	'mõh 'tɔᵇm	headwater/source of river
knmɔh	kən'mõh	name
kanɔh	ka'nɔh	to grow (of plants)
lanɔh	la'nõh	[ethnonym: Lanoh]
prɔh	pə'rɔh	[a type of tree]
srɔh	sə'rɔh	to give birth
rɔh	'rɔh	to cut; to clear
rawɔh	ra'wɔh	[a type of fruit tree]
jɔh	'jɔh	to drop
kajɔh*	ka'jɔh	to paddle

/-m/

cĩm	'cᶜĩm	to sizzle (of meat being roasted)
ɟim	'ɟᶻi^{b}m	to cry/weep
bnim	mə'nĩm	[a type of tree]
smsĩm	səm'sĩm	silver-eared mesia (*Leiothrix argentauris*)
harim	ha'ri^{b}m	quill of porcupine's tail
kmjɪm	kəm'jɪm	to taste
sbem	sə'bẹ^{b}m	to be full
tem	'tẹ^{b}m	right (side)
mmtem	məm'tẹ^{b}m	right-handed person
lmtem	ləm'tẹ^{b}m	to arrive; to become
tmdem	təm'dẹ^{b}m	to go into the jungle and stay there overnight in search of food
musem*	mu'sẹ^{b}m	season
slem*	sə'lẹ^{b}m	coin
ɟrem*	ɟᶻə'rẹ^{b}m	rapid
krem	kə'rẹ^{b}m	to play an instrument
dẽm*	'dẽm	lake
lkɛm	lə'kɛ^{b}m	brain
ʔẽm	'ʔẽm	breast; to drink
ʔẽm kaji?	'ʔẽm ka'ji?	[a type of thorn]
piʔẽm	pĩʔẽm	to breastfeed
masẽm*	mã'sẽm	to be sour
krhɛm ~ krhĩm	kər'hɛ^{b}m ~ kər'hĩm	to clear one's throat
nɛm*	'nẽm	six
tanɛm*	ta'nẽm	to plant
ɲɛm	'ɲẽm	to eat
gulɛm	gu'lɛ^{b}m	to carry on a stick
karɛm*	ka'rɛ^{b}m	to sink
sabɨm	sa'bɨ^{b}m	[ethnonym: Sabüm]
patɨm	pa'tɨ^{b}m	to pound
ktɨm	kə'tɨ^{b}m	to hit with side of fist
ckɨm	cᶜə'kɨ^{b}m	lump; chip
ckĩm	cᶜə'kĩm	peacock-pheasant (*Polyplectron*)
gɨm	'gɨ^{b}m	to deliberate
lgɨm	lə'gɨ^{b}m	valley floor
bhĩm	bə'hĩm	blood
krhɨm	tər'hɨ^{b}m	[a type of large beetle]
kmɨm	kə'mĩm	to smoke with the cigarette inside one's mouth
tmɨm	tər'nɨ^{b}m	to shoot
lɨm	'lɨ^{b}m	to be big; [toponym: Long]
pəm	'pə^{b}m	to sit (of bird)
katəm*	ka'tə^{b}m	to cut rattan
tmtəm	təm'tə^{b}m	to buzz (of insect)

kəm	ˈkəᵇm	to rub one's hands
kmkə̃m	kəmˈkə̃m	ankle
ʔəm ~ ʔimʔəm	ˈʔəᵇm ~ ʔimˈʔəᵇm	to hug
məm	ˈmə̃m	mother's milk
gaɾəm*	gaˈɾəᵇm	salt
pam*	ˈpaᵇm	to pump
lampam*	lamˈpaᵇm	[a type of carp (*Puntius*)]
kbam	kəˈbaᵇm	to smoke with the cigarette the wrong way round
padam*	paˈdaᵇm	to close one's eyes
ɟam*	ˈɟᶻaᵇm	hour; wristwatch
tkam	təˈkaᵇm	[toponym: Tekam]
sam	ˈsaᵇm	to go hunting
paham*	paˈhaᵇm	to understand
mam	ˈmãm	to drink milk from mother's breast
tamam	taˈmaᵇm	[a type of small animal]
lam	ˈlaᵇm	to squeeze
wam	ˈwaᵇm	[a type of rattan]
sjam	siˈjaᵇm	Thailand
hajam*	haˈjaᵇm	poultry
piɲjam*	piɲˈjaᵇm	to borrow
ɾjam	pəɾˈjaᵇm	time; when
pum	ˈpuᵇm	to lie face down
kutum	kuˈtuᵇm	flower bud
cdum	cᶜəˈduᵇm	to carry in one's arms
rdũm ~ rdup	ɾəˈdũm ~ ɾəˈdupˀ	to be cloudy
kɾnum	kəɾˈnũm	[a type of tree]
blum	bəˈluᵇm	[toponym: Belum]
jarum*	jaˈɾuᵇm	needle
tom	ˈtoᵇm	tree-base; CLF: vegetation
tom ktɔʔ	ˈtoᵇm kəˈtɔʔ	east
ptom	pəˈtoᵇm	day before yesterday
lantom	lanˈtoᵇm	to sting
mɲɟlom*	məɲɟᶻəloᵇm	[toponym: Mendelum]
hajom	haˈjoᵇm	young green tree python (*Chondropython viridis*)
tɔm	ˈtɔᵇm	fluid; river; water
tɔm mit	ˈtɔᵇm ˈmĩtˀ	tear
cɔm	ˈcᶜɔᵇm	to burn something; to make fire
kɔm	ˈkɔᵇm	many
ʔɔm	ˈʔɔᵇm	to want
ɟʔɔm	ɟᶻəˈʔɔᵇm	to be dirty
krʔɔm	kəɾˈʔɔᵇm	to cook in ashes or fire
sɔm	ˈsɔᵇm	bird's nest
smsɔm	səmˈsɔᵇm	to buzz around a nest (of birds, bees)

hmhɔm	həm'hɔ^{b}m	to like
knɔm	kə'nɔ̃m	urinary bladder; urine; to urinate
rlɔm	rə'lɔ^{b}m	[toponym]
kjɔm	kə'jɔ^{b}m	below; lower side; underside

/-n/

pimpin*	pim'pi^{d}n	to guide
bin	'bi^{d}n	to sleep on one's stomach
tin	'ti^{d}n	to kick open
ʔacin	ʔa'c^{c}i^{d}n	how?
gin	'gi^{d}n	PRONOUN 2/3P
risin*	ri'si^{d}n	ration
masin*	ma'si^{d}n	to be salty
brsin*	bər'si^{d}n	to sneeze
ɟamin*	ɟ^{ʑ}a'mĩn	to guarantee
mnmin	mən'mĩn	to play (games)
ɲnɲin	ɲən'ɲĩn	to hop on something
ɟilĩn	ɟ^{ʑ}i'lĩn	to jump over a high obstacle
tanlĩn	tan'lĩn	[a type of civet]
lajin*	la'ji^{d}n	to be different
ben	'bẹ^{d}n	to tie around
haden	ha'dẹ^{d}n	tomorrow; near future; soon
ʔɲcen	ʔəɲ'c^{c}ẹ^{d}n	to cook
ken	'kẹ^{d}n	to hear
bnken ~ mnken (?)	mən'kẹ^{d}n	to astonish
muŋken*	mūŋ'kẹ^{d}n	maybe
kawen	ka'wẹ^{d}n	to plant
blawen	bəla'wẹ^{d}n	mango
pɛ̃n	'pɛ̃n	to hit with an object
bɛn	'bɛ^{d}n	to wrap a blanket around one's body
tabɛn	ta'bɛ^{d}n	[a type of fruit]
krntɛn	kərən'tɛ^{d}n	wrinkles
kɛn	'kɛ^{d}n	child; CLF: human
sriken	səri'kɛ^{d}n	[a type of leaf]
makɛn	ma'kɛ^{d}n	who?; whose?
gɛn	'gɛ^{d}n	[a type of tree]
bagɛn	ba'gɛ^{d}n	[a type of monitor lizard (*Varanus*)]
tomɛn*	to'mɛ̃n	snakehead (*Channa micropeltes*)
tahanɛn	taha'nɛ̃n	[a type of rattan]
prnɛn	pər'nɛ̃n	to refuse to give
lɛn	'lɛ^{d}n	loincloth; to wear a loincloth
dorɛn*	do'rɛ^{d}n	durian (*Durio zibethinus*)
hawɛ̃n	hã'wɛ̃n	wild boar (*Sus scrofa*); pig
tmwɛn*	təm'wɛ̃n	[ethnonym: Temuan]
kɨn	'kɨ^{d}n	hole

gɨn	'gɨ^{d}n	to pour
rɨn	'rɨ^{d}n	to sit down
hmirɨn	həmi'rɨ^{d}n	to extinguish (by itself)
wɨn	'wʉ^{d}n	to crawl
jɨn	'jɨ^{d}n	to hop on something to test its supporting capacity
j̃ɨn	'j̃ɨn	to stop
dən	'də^{d}n	to run short/to run out
hnhən	hən'hə^{d}n	to devour
papan*	pa'pa^{d}n	plank; wall; [a type of snake]
lapan*	la'pa^{d}n	eight
sampan*	sam'pa^{d}n	boat
dan	'da^{d}n	to take a hit
can	'c^{c}a^{d}n	foot
ɟan	'ɟ^{ʑ}a^{d}n	to throw something to the ground
ʔikan*	ʔi'ka^{d}n	fish
bokan*	bo'ka^{d}n	NEGATIVE MARKER
bagan	ba'ga^{d}n	[a type of fruit]
tagan	ta'ga^{d}n	adult green tree python (*Chondropython viridis*)
tasan	ta'sa^{d}n	flying squirrel (*Petaurista*)
tahan*	ta'ha^{d}n	to endure
piŋan*	pi'ŋa^{d}n	helping; serving (of food)
luŋan	lu'ŋa^{d}n	binturong (*Arctitis binturong*)
smilan*	səmi'la^{d}n	nine
bolan*	bo'la^{d}n	moon
blan	bə'la^{d}n	[toponym]
hiran	hi'ra^{d}n	to be surprised
daran	da'ra^{d}n	[a type of tuber]
lawan*	la'wa^{d}n	to kill
tũn	'tũn	that (you don't know)
dun	'du^{d}n	to cover; to fill
ʔũn	'ʔũn	there (you don't know)
banun*	mã'nũn	[toponym: Banun]
gnun	gə'nũn	bamboo
krlun	kər'lu^{d}n	to retreat
run	'ru^{d}n	to fall to the ground (of e.g. fruit, leaf)
pon*	'po^{d}n	like; also
ɟpon*	ɟ^{ʑ}ə'po^{d}n	[ethnonym: Japanese]
sabon*	sa'bo^{d}n	soap
ton	'to^{d}n	that (you know)
ton ~ tnton	'to^{d}n ~ tɜn'to^{d}n	to climb up (e.g. a rope); to balance
ʔon	'ʔo^{d}n	there (you know)
panʔon	pan'ʔo^{d}n	[a type of tree]
lon	'lo^{d}n	to push

talon	ta'lo^{d}n	reticulated python (*Python reticulatus*)
tawon*	ta'wo^{d}n	year
bɔn	'bɔ^{d}n	[a type of fruit]
dɔn	'dɔ^{d}n	great grandparent

/-ɲ/

tawiɲ	ta'wi^{ɟ}ɲ	[a type of small spider]
twiɲ	tə'wi^{ɟ}ɲ	headache
peɲ	'pe̤^{iɟ}ɲ	daylight; sunshine
hɾɲpeɲ	həriɲ'pe̤^{iɟ}ɲ	goose bumps
teɲ	'te̤^{iɟ}ɲ	to plait; to make a net
pɛɲ	'pɛ^{iɟ}ɲ	to shoot (with bow or slingshot)
blʔɛɲ	bəl'ʔɛ^{iɟ}ɲ	green
sɛɲ	'sɛ^{iɟ}ɲ	before; front
sɛɲ ktɔʔ	'sɛ^{iɟ}ɲ kə'tɔʔ	west
hɛ̃ɲ	'hɛ̃ɲ	tooth
smɛɲ	sə'mɛ̃ɲ	to ask for; to request
trɛɲ	tə'rɛ^{iɟ}ɲ	to climb
ɾɲɾɛɲ	riɲ'rɛ^{iɟ}ɲ	to cause pain
jɲjɛɲ	jiɲ'jɛ^{iɟ}ɲ	to dream
tpɨɲ	tə'pɨ^{iɟ}ɲ	part of honeycomb (?)
tagɨɲ	ta'gɨ^{iɟ}ɲ	wood (material); firewood
tgɨɲ	tə'gɨ^{iɟ}ɲ	to tear something apart with one's teeth
hmɨɲ	hə'mɨ^{iɟ}ɲ	[magic word uttered to redress broken taboo]
daɾɨɲ	da'ɾɨ^{iɟ}ɲ	termite; small worm causing tooth decay
hɲɟaɲ	hiɲ'ɟ^{ʒ}a^{iɟ}ɲ	to rise; to stand
mamuɲ	mā'mūɲ	to beg
boɲ	'bo^{iɟ}ɲ	[a type of tree]
goɲ	'go^{iɟ}ɲ	[a type of outgrowth on trees]
raŋgoɲ	raŋ'go^{iɟ}ɲ	mouth harp
trpɔɲ	tɜr'pɔ^{iɟ}ɲ	[a type of tree]
gɔ̃ɲ	'gɔ̃ɲ	to refuse to give something
ʔɔ̃ɲ	'ʔɔ̃ɲ	to smell something
pɲlɔɲ	piɲ'lɔ^{iɟ}ɲ	to sing

/-ŋ/

kpiŋ*	kə'pi^{g}ŋ	portion; piece; CLF: flat objects
pŋpĩŋ	pəŋ'pĩŋ	Philippine glossy starling (*Aplonis panayensis*)
krpiŋ	kər'pi^{g}ŋ	above; top; upper side
cuntĩŋ	cun'tĩŋ	temple
gadiŋ*	ga'di^{g}ŋ	boar's tusk; elephant's tusk
taciŋ*	ta'c^{c}i^{g}ŋ	worm
cŋcĩŋ	c^{c}əŋ'c^{c}ĩŋ	biorbital area; spectacle

ʄĩŋ	ˈʄᶻĩŋ	to take
ʔasiŋ*	ʔaˈsi^{g}ŋ	to be different; other
masiŋ-masiŋ*	masi^{g}ŋ-maˈsi^{g}ŋ	each
pusiŋ*	puˈsi^{g}ŋ	to move in circles
kɲsiŋ	kəɲˈsi^{g}ŋ	banded palm civet (*Hemigalus derbyanus*)
cliŋhĩŋ	cᶜəliŋˈhĩŋ	to turn head (of bird)
kamiŋ*	kaˈmi^{g}ŋ	goat
baniŋ*	mãnˈi^{g}ŋ	[toponym: Banding]
kuniŋ*	kuˈnĩŋ	yellow
pniŋ*	pəˈnĩŋ	headache; confusion
baliŋ	baˈli^{g}ŋ	tiger (*Panthera tigris*)
maliŋ*	maˈli^{g}ŋ	to steal
bliŋ	bəˈli^{g}ŋ	upper arm
kliŋ	kəˈli^{g}ŋ	language; sound
ciŋlĩŋ	ciŋˈlĩŋ	to look over one's shoulder
gruŋliŋ	gəruŋˈli^{g}ŋ	to roll down
kɳliŋ	kəɳˈli^{g}ŋ	to make sound: e.g. howl/croak/bark
brlĩŋ	bərˈlĩŋ	long-tailed parakeet (*Psittacula longicauda*)
bariŋ	baˈri^{g}ŋ	[toponym: Baring]
kriŋ*	kəˈri^{g}ŋ	to be dry
cnñŋ	cɜnᵈɹĩŋ	to hold one's wrist
cawĩŋ	cãˈw̃ĩŋ	eyebrow
riŋwĩŋ	riŋˈwĩŋ	to dangle one's arms when walking
kalwiŋ	kalˈwi^{g}ŋ	[a type of fruit]
sajĩŋ	sãˈjĩŋ	friend
hɔrjĩŋ	hɔrˈjĩŋ	[a type of small animal]
tbeŋ	təˈbe̤^{g}ŋ	bamboo slat
ʔnteŋ	ʔənˈte̤^{g}ŋ	ear
kuceŋ*	kuˈcᶜe̤^{g}ŋ	cat
kʄeŋ	kəˈʄᶻe̤^{g}ŋ	to hear
keŋ	ˈke̤^{g}ŋ	to pull
paŋkeŋ*	paŋˈke̤^{g}ŋ	chair
ʄʔeŋ	ʄᶻəˈʔe̤^{g}ŋ	bone; skeleton
plʔeŋ	pəlˈʔe̤^{g}ŋ	fluid
barʔeŋ	barˈʔe̤^{g}ŋ	[a type of tortoise]
kɳseŋ	kəŋˈse̤^{g}ŋ	to dance
lheŋ	ləˈhe̤^{g}ŋ	phlegm; saliva
ʄleŋ	ʄᶻəˈle̤^{g}ŋ	to put inside
ʄreŋ	ʄᶻəˈre̤^{g}ŋ	soul
pɛŋ	ˈpɛ^{g}ŋ	to chop vegetables
tampɛŋ	tamˈpɛ^{g}ŋ	to ascend
batɛŋ*	baˈtɛ^{g}ŋ	[a type of rattan]
bintɛŋ*	mĩnˈtɛ^{g}ŋ	star
tuŋtẽŋ	tũŋˈtẽŋ	to slither
ʄɛŋ	ˈʄᶻɛ^{g}ŋ	whiskers

ɟŋɟɛŋ	ɟᶻəŋˈɟᶻɛ^{g}ŋ	to be wide
pgɛŋ*	pəˈgɛ^{g}ŋ	to hold
gluŋʔɛ̃ŋ	gəlũŋˈʔɛ̃ŋ	to be in pain; to scratch (of monkey hit by blowpipe dart)
pŋsɛŋ	pəŋˈsɛ^{g}ŋ	to say; to speak; to talk; to tell
sŋsɛ̃ŋ	sə̃ŋˈsɛ̃ŋ	to walk or run with a limp
ʔameŋ	ʔaˈmɛ̃ŋ	siamang (*Hylobates syndactylus*)
pineŋ*	piˈnɛ̃ŋ	betelnut
trineŋ*	təriˈnɛ̃ŋ	training
ʔaneŋ	ʔaˈnɛ̃ŋ	[a type of tree]
bneŋ*	mə̃ˈnɛ̃ŋ	thread
sneŋ*	səˈnɛ̃ŋ	to be easy
mneŋ*	mə̃ˈnɛ̃ŋ	to win
bileŋ*	biˈlɛ^{g}ŋ	to count
toleŋ	toˈlɛ^{g}ŋ	[a type of tuber]
kleŋ	kəˈlɛ^{g}ŋ	inside
klɛ̃ŋ	kəˈlɛ̃ŋ	bird of prey
klɛ̃ŋ hip	kəˈlɛ̃ŋ ˈhip˺	crested serpent-eagle (*Spilornis cheela*)
klɛ̃ŋ siwah	kəˈlɛ̃ŋ siˈwah	brahminy kite (*Haliastur indus*)
tŋleŋ	təŋˈlɛ^{g}ŋ	to see at a distance
preŋ	pəˈrɛ^{g}ŋ	egret (*Egretta*); [a type of cobra (*Naja*)]
greŋ	gəˈrɛ^{g}ŋ	[a type of monitor lizard (*Varanus*)]
ʔnreŋ	ʔənᵈɹɛ^{g}ŋ	[a type of fruit tree]
weŋ ~ wiŋweŋ	ˈwɛ^{g}ŋ ~ wiŋˈwɛ^{g}ŋ	to lean one's head
laweŋ	laˈwɛ^{g}ŋ	[a type of python (*Python*)]
ɟɔwɛ̃ŋ	ɟᶻɔ̃ˈw̃ɛ̃ŋ	scorpion (generic)
bwɛ̃ŋ	bə̃ˈw̃ɛ̃ŋ	great hornbill (*Buceros bicornis*)
cŋwɛ̃ŋ	cᶜə̃ŋˈw̃ɛ̃ŋ	to be awake
kŋjeŋ	kəŋˈjɛ^{g}ŋ	to flap wings
dɨŋ	ˈdɨ^{g}ŋ	big permanent house
cŋciŋ	cᶜəŋˈcᶜɨ^{g}ŋ	to stretch something
nŋkɨŋ	nəŋˈkɨ^{g}ŋ	small hut
ʔiŋ	ˈʔi^{g}ŋ	[a type of wasp]
cʔiŋ	cᶜəˈʔi^{g}ŋ	[a type of tuber]
bliŋ	bəˈli^{g}ŋ	to think of someone
rnajɨŋ	rənaˈjɨ^{g}ŋ	to be blunt (of knife)
cjɨŋ	ciˈjɨ^{g}ŋ	[toponym: Chiong]
sjɨŋ	siˈjɨ^{g}ŋ	to burn
krpəŋ	kərˈpə^{g}ŋ	[a type of insect]
tabəŋ	taˈbə^{g}ŋ	leaf monkey (*Semnopithecus*)
labəŋ	laˈbə^{g}ŋ	skull
prŋgəŋ	pərəŋˈgə^{g}ŋ	pharynx
rəŋ	ˈrə^{g}ŋ	[a type of small animal]
srampaŋ*	səramˈpa^{g}ŋ	fish spear
cliŋpaŋ	cᶜəliŋˈpa^{g}ŋ	to turn head (of bird)

cabaŋ*	cᶜa'baᵍŋ	branch
cabaŋ* tɔm	cᶜa'baᵍŋ 'tɔᵇm	tributary
kabaŋ	ka'baᵍŋ	[a type of fruit]
kobaŋ*	ko'baᵍŋ	mud
bataŋ*	ba'taᵍŋ	CLF: oblong objects
gantaŋ*	gan'taᵍŋ	potato
tŋtaŋ	təŋ'taᵍŋ	greater racket-tailed drongo (*Dicrurus paradiseus*)
bidaŋ*	bi'daᵍŋ	CLF: large flat objects
padaŋ*	pa'daᵍŋ	field
cadaŋ*	cᶜa'daᵍŋ	to be ready to do something
sladaŋ*	səla'daᵍŋ	wild ox/gaur (*Bos gaurus*)
paɲcaŋ ~ paɲcoŋ	paɲ'cᶜaᵍŋ ~ paɲ'cᶜoᵍŋ	to light up; to shine
ciŋcaŋ*	cᶜiŋ'cᶜaᵍŋ	to mince
kŋkaŋ	kəŋ'kaᵍŋ	Asian horned toad (*Megophrys monticola*)
pisaŋ*	pi'saᵍŋ	banana
lasaŋ	la'saᵍŋ	taste
maŋ	'mãŋ	to prevent something from falling
simaŋ*	si'maᵍŋ	to make magic (?)
bramaŋ	bəra'maᵍŋ	[a type of tree]
kmaŋ*	kə'maᵍŋ	to swell
ʔnaŋ	ʔə'naᵍŋ	side
paŋaŋ*	pa'ŋaᵍŋ	to roast
hŋaŋ*	hə'ŋaᵍŋ	rhinoceros hornbill (*Buceros rhinoceros*)
hilaŋ*	hi'laᵍŋ	to spin; to whirl
balaŋ*	ba'laᵍŋ	bottle
knalaŋ	kəna'laᵍŋ	[a type of tree]
blalaŋ*	bəla'laᵍŋ	[a type of large grasshopper]
baraŋ*	ba'raᵍŋ	thing
karaŋ	ka'raᵍŋ	[a type of tree-growing fern]
saraŋ*	sa'raᵍŋ	roof
praŋ*	pə'raᵍŋ	to shoot
krãŋ	kə'rãŋ	[sound of carving]
siŋraŋ	siŋ'raᵍŋ	to chatter
tŋraŋ	təŋ'raᵍŋ	[a type of fish]
swaŋ	sə'waᵍŋ	to go around
tŋwaŋ	təŋ'waᵍŋ	blue coral snake (*Maticora bivirgata*)
klwaŋ	kəl'waᵍŋ	flying fox
jaŋ	jãŋ	RELATIVE MARKER
ʔuŋjaŋ ~ ʔuŋjiŋ	ʔuŋ'jaᵍŋ ~ ʔuŋ'jiᵍŋ	to shake something
lŋpuŋ	ləŋ'puᵍŋ	to beat (of heart)
tuŋ	'tuᵍŋ	to fell; to bring down
katuŋ	ka'tuᵍŋ	[a type of tree]
cantũŋ	cᶜan'tũŋ	to stamp one's feet; to play bamboo tubes

gantuŋ*	gan'tu^{g}ŋ	to hang
tŋtũŋ	tə̃ŋ'tũŋ	[a type of large poisonous spider]
tuduŋ*	tu'du^{g}ŋ	to cover one's eyes from the sun
gŋguŋ	gəŋ'gu^{g}ŋ	[a type of civet]
kʔuŋ	kə'ʔu^{g}ŋ	[a type of tree]
lsuŋ	lə'su^{g}ŋ	mortar; [toponym]
gahuŋ*	ga'hu^{g}ŋ	cave
siɲuŋ	si'ɲũŋ	to move ears before attacking (of predator)
balũŋ hajam*	ba'lũŋ ha'ja^{b}m	[a type of fruit]
taluŋ	ta'lu^{g}ŋ	[a type of millipede]
tuluŋ*	tu'lu^{g}ŋ	to help
juluŋ	ju'lu^{g}ŋ	flat-headed cat (*Felis planiceps*)
bawuŋ*	ba'wu^{g}ŋ	[a type of catfish (*Mystus planiceps*)]
liŋwuŋ	liŋ'wu^{g}ŋ	to move in circles
pajuŋ*	pa'ju^{g}ŋ	umbrella
cjuŋ*	ci'ju^{g}ŋ	myna (*Acridotheres*)
sjuŋ*	si'ju^{g}ŋ	fang
pŋjuŋ	pəŋ'ju^{g}ŋ	to play flute
kampoŋ*	kam'po^{g}ŋ	village
lampoŋ*	lam'po^{g}ŋ	to float
criŋpoŋ	c^{ɕ}əriŋ'po^{g}ŋ	to roll oneself
tŋtoŋ	təŋ'to^{g}ŋ	collared kingfisher (*Halcyon chloris*)
kaltoŋ	kal'to^{g}ŋ	knee
tdoŋ	tə'do^{g}ŋ	[a type of snake]
koŋ	'ko^{g}ŋ	honey (?)
riŋkoŋ	riŋ'ko^{g}ŋ	[a type of frog]
krʔoŋ	kər'ʔo^{g}ŋ	to trumpet (of elephant)
ɟhoŋ	ɟ^{ʑ}ə'ho^{g}ŋ	to bark
haŋhoŋ	haŋ'ho^{g}ŋ	to blow (of person); to whistle
samoŋ*	sa'mo^{g}ŋ	to put something together
gnoŋ	gə'no^{g}ŋ	wooden material
rloŋ	rə'lo^{g}ŋ	[a type of snake]
taroŋ	ta'ro^{g}ŋ	[a type of lizard]
woŋ	'wo^{g}ŋ	to build a house
kawoŋ	ka'wo^{g}ŋ	great argus (*Argusianus argus*)
taŋwoŋ	taŋ'wo^{g}ŋ	to carry on one's shoulder
lajoŋ	la'jo^{g}ŋ	[a type of tree]
pɔŋ	'pɔ^{g}ŋ	to tap poison from ipoh tree
ʔmpɔŋ	ʔəm'pɔ^{g}ŋ	hole
ʔmpɔŋ kit	ʔəm'pɔ^{g}ŋ 'kit̚	anus
sŋpɔŋ	səŋ'pɔ^{g}ŋ	leafbird (*Chloropsis*)
katõŋ	ka'tõŋ	claw; nail
katõŋ cjas	ka'tõŋ ci'jas	finger nail
katõŋ can	ka'tõŋ 'c^{ɕ}a^{d}n	toe nail
cõŋ	'c^{ɕ}õŋ	to hit from above (e.g. of falling fruit)

kaɲcoŋ	kaɲˈcᶜɔᵍŋ	[a type of insect]
ɟɳɟɔŋ	ɟᶻəŋˈɟᶻɔᵍŋ	to smell (of rotten wood)
kɔŋ	ˈkɔᵍŋ	to plait
ʔõŋ	ˈʔõŋ	water
pʔɔŋ	pəˈʔɔᵍŋ	to heat poisoned blowpipe dart in fire
pahõŋ	pãˈhõŋ	thrush (*Zoothera*)
rŋhõŋ ʔɔs	rəŋˈhõŋ ˈʔɔs	charcoal
lamɔŋ	laˈmɔᵍŋ	[a type of fruit]
tanɔŋ	taˈnõŋ	neuropteran (generic)
ranɔŋ	raˈnõŋ	[a type of snake]
cnɔŋ	cᶜəˈnõŋ	beak/bill/nib
ʔilõŋ	ʔiˈlõŋ	masked civet (*Paguma larvata*)
balɔŋ	baˈlɔᵍŋ	[a type of tree]
dalɔŋ	daˈlɔᵍŋ	[a type of lizard]
clɔŋ	cᶜəˈlɔᵍŋ	wild dog (*Cuon alpinus*)
glɔŋ	gəˈlɔᵍŋ	wide path; way
snlɔŋ	sɜnˈlɔᵍŋ	to remain in one place to make a fire for a returning hunting party
barɔŋ	baˈrɔᵍŋ	Malay tapir (*Tapirus indicus*)
manrɔŋ	manᵈˈɹɔᵍŋ	skink (*Emoia*)
wɔŋ	ˈwɔᵍŋ	child; offspring
kwɔŋ	kəˈwɔᵍŋ	dove
brwõŋ	bərˈw̃õŋ	[a type of frog]
jõŋ	ˈjõŋ	[a type of salty tuber]
pjõŋ	pĩˈjõŋ	[a type of turtle]
kajõŋ	kãˈjõŋ	elbow
brjɔŋ	bərˈjɔᵍŋ	to burn

/-l/

kpil	kəˈpil	[a type of turtle]
tampĩl	tamˈpĩl	slow lori (*Nycticebus coucang*)
knabil	kənaˈbil	[a type of large spider]
til	ˈtil	to be hot
cantil	cᶜanˈtil	to lose one's footing
bdil*	bəˈdil	gun; to shoot with gun
klɗĩl	kəlˈdĩl	[a type of snake]
cil	ˈcᶜil	to blink
wakil	waˈkil	younger sibling
saŋkĩl	saŋˈkĩl	to peck (of bird)
cilkil	cilˈkil	to look around
kulkil	kulˈkil	to fall
tigil	tiˈgil	to go along a mountain-side
sil	ˈsil	[a type of tortoise]
raɲsil	raɲˈsil	[a type of tree]
khil	kəˈhil	to pound; to hit somebody with one's elbow

karil	ka'ril	paradise tree snake (*Chrysopelea paradisi*); [a type of high grass by the river]
krɨ̃l	kə'rɨ̃l	ankle-joint; wrist
klwɨ̃l	kəl'wɨ̃l	to come
kajil*	ka'jil	to fish
pel	'pe̤l	to cook
tel	'te̤l	track
mhel	mə'he̤l	to astonish (?)
wel	'we̤l	again
pɛl	'pɛl	to drip
tabɛl	ta'bɛl	to be thick
cntɛl	$c^{ɕ}$ɜn'tɛl	seat bone
bdɛl*	bə'dɛl	to shoot with blowpipe
gɛl	'gɛl	waist
ʔɛl	'ʔɛl	to look; to see
rasɛl	ra'sɛl	[a type of small animal]
psɛl	pə'sɛl	to sprinkle something along one's path
ksɛl	kə'sɛl	to fill
hɛl	'hɛl	to pant
ghɛl	gə'hɛl	to be tired
mhɛl*	mə'hɛl	to be expensive
taŋɛl*	ta'ŋɛl	to fall
klɛl	kə'lɛl	to rub (e.g. ointment on skin)
hnrɛl	hən^{d}ɹɛl	[a type of small animal]
wɛ̃l	'wɛ̃l	to wax (of moon)
crilwɛl	$c^{ɕ}$əril'wɛl	to curl
kjɛl	ki'jɛl	[toponym: KL (Kuala Lumpur)]
karjɛl	kar'jɛl	[a type of krait (*Bungarus*)]
tampɨl	tam'pɨl	[a type of large ant]
bɨl	'bɨl	to cover the ground with mats
tɨl	'tɨl	to pound
pitɨl	pi'tɨl	to touch
tltɨl	təl'tɨl	to pull blowpipe apart
dlɗɨl	dɜl'ɗɨl	heel
cɨl	'$c^{ɕ}$ɨl	to tattoo
kɨl	'kɨl	to rest
gɨl	'gɨl	tualang tree (*Koompassia excelsa*)
ɟlʔɨl	ɟ^{ʑ}ɜl'ʔɨl	to descend/to go down
jɨl	'jɨl	to finish
ʔipəl*	ʔi'pəl	apple
kmarbəl	kəmar'bəl	[a type of insect]
kəl	'kəl	to tie
taŋkəl	taŋ'kəl	to carry on one's back
məl	'mə̃l	to roll leaf (for smoking)
gimməl	gi'məl	to slither

mlməl	məl'mə̃l	to wind/swaddle
baləl	ba'ləl	small stone
kapal*	ka'pal	aircraft
smpal	səm'pal	blue-eared kingfisher (*Alcedo meninting*)
gibal	gi'bal	to fall
bdal*	bə'dal	to throw
kal	'kal	to lick
tmkal	təm'kal	man; male
haŋkal	haŋ'kal	to cover
rŋkal	rəŋ'kal	[a type of fruit]
kpral*	kəpə'ral	corporal
buwal*	bu'wal	to chat/to converse/to talk
ɟuwal*	ɟᶻu'wal	to sell
kwal	kə'wal	white-rumped shama (*Copsychus malabaricus*)
kilwal	kil'wal	to sit with one's legs crossed
trbul*	tɜr'bul	[a type of carp]
tul	'tul	summit/top
kul	'kul	to call; to summon
caŋkul*	cᶜaŋ'kul	to hoe
gul	'gul	pestle
timul*	ti'mul	to float
tuŋul	tu'ŋul	fireplace; stove
sjul	si'jul	cobra (*Naja*)
brul	bə'rul	[a type of small bird]
critol	cᶜəri'tol	to carry on one's neck
btol*	bə'tol	true; to be straight; to aim with blowpipe or gun
dol ~ pidol	'dol ~ pi'dol	to hide something
kol	'kol	to lie down
klkol	kəl'kol	to kneel
gol	'gol	to rest with one's arms behind one's neck
prlʔol ~ prlʔil	pərəl'ʔol ~ pərəl'ʔil	to pluck
bisol*	bi'sol	boil; wound
ʔɲsol	ʔəɲ'sol	to be embarrassed; to feel shame
brol	bə'rol	[a type of tree]
jol	'jol	to tell someone to leave; to throw
butəl*	bu'təl	bottle
dadə̃l	da'də̃l	[a type of reed snake (*Macrocalamus*)]
cəl	'cᶜəl	to tell
hɟəl	hə'ɟᶻəl	to raise something to an upright position
snkə̃l	sən'kə̃l	circumcision
tuŋkə̃l	tuŋ'kə̃l	stone; CLF: spherical/cubical objects
gəl	'gəl	to carry on one's back or shoulder
pʔəl	pə'ʔəl	to bake in fire; to roast

bhɔl	bə'hɔl	muntjac (barking) deer (*Muntiacus muntjac*)
ɟlmɔl	ɟᶻɜl'mɔ̃l	mountain
brawɔl	bəra'wɔl	thick-billed pigeon (*Treron curvirostra*)
hiljɔl	hil'jɔl	to hop

/-ɾ/

prpɨ̂r ~ prpɨr	pəɾ'pɨ̂r: ~ pəɾ'pɨr:	to flicker
pikir*	pi'kir:	to think/to cogitate
sir	'sir:	side
srsir	səɾ'sir:	to go along the side of something
hir	'hir:	to be frightened
prhir	pəɾ'hir:	to frighten
samir	sa'mir:	water strider (*Gerridae*)
cɲɟir ~ cɲɟər	cɜɾ'ɲɟ̂r: ~ cɜɾ'ɲɟ̃ər:	to scent
ger taɟi?	'gɐ̣r: ta'ɟᶻi?	knife handle
her	'hɐ̣r:	to pull
trwer	tɜɾ'wɐ̣r:	to look up into the canopy to spot game
tampɛr*	tam'pɛr:	to claw; to scratch
ʔampɛr	ʔam'pɛr:	[a type of tree]
bɛr	'bɛr:	younger sibling
bɛ̃r	'bɛ̃r:	to tell a lie
prbɛr	pəɾ'bɛr:	lower arm
jadɛr	ja'dɛr:	[a type of tree]
paɲcɛr*	paɲ'cᶜɛr:	to spit horizontally
muŋkɛr	muŋ'kɛr:	to wake (up)
krkɛr	kəɾ'kɛr:	to yell
phɛr	pə'hɛr:	to walk slowly
tmɛr	tə'mɛr:	[ethnonym: Temiar]
ɟawɛr	ɟᶻa'wɛr:	Pacific swallow (*Hirundo tahitica*)
kwɛr*	kə'wɛr:	to sweep
pɨr	'pɨr:	to secretly call for someone at night
crbɨr	cᶜɜɾ'bɨr:	to come out (of leaves)
tɨr	'tɨr:	to play instrument
gɨr ~ gar	'gɨr: ~ 'gar:	to roll (of thunder); to vibrate
higɨr	hi'gɨr:	to sink
sagɨr	sa'gɨr:	[a type of tree]
grgɨr	gəɾ'gɨr:	to shiver
blaʔɨr	bəla'ʔɨr:	[a type of green viper (*Trimeresurus*)]
sɨr	'sɨr:	to slide forward on one's behind
bhɨr	bə'hɨr:	to be overgrown
hawɨr ~ hawɔr	ha'wɯr: ~ ha'wɔr:	buttock; tail
tapər	ta'pər:	bat
tər	'tər:	to pluck a bird
sitər ~ hitər	si'tər: ~ hi'tər:	to look for something

hantər*	han'tər:	to send
ʔakər	ʔa'kər:	[a type of snake]
kanər*	ka'nər:	to carry on one's shoulder, e.g. an animal
sajər	sa'jər:	flock/herd
sjər	si'jər:	to swim
prjər	pər'jər:	to accompany
tapar* cjas	ta'par: ci'jas	palm of hand
star	sə'tar:	[a type of fruit]
gntar*	gən'tar:	to shake; to vibrate
cicar	ci'c^{c}ar:	[a type of tree]
ʔaɟar*	ʔa'ɟ^{ʒ}ar:	to learn
kɟar	kə'ɟ^{ʒ}ar:	[toponym: Kejar]
tukar*	tu'kar:	to exchange
har	'har:	small path/trail
kmar*	kə'mar:	twins
laŋar*	la'ŋar:	to crash
war	'war:	plaintive cuckoo (*Cacomantis merulinus*)
gugur*	gu'gur:	to lose hair
pirʔur	pir'ʔur:	to growl (of stomach)
brhur	bər'hur:	[a type of snake]
simur*	si'mur:	east
smur*	sə'mur:	to spray water
dor	'dor:	to balance
dapor* ʔɛ̃m	da'por: 'ʔɛ̃m	nipple
gor	'gor:	lower leg
hor	'hor:	striped tit-babbler (*Macronous gularis*); to have a running nose
smlor	səm'lor:	[toponym: Semelor]
pɔr	'pɔr:	tiger shrike (*Lanius tigrinus*)
kubɔr*	ku'bɔr:	grave
slantɔr	səlan'tɔr:	[a type of snake]
dɔr	'dɔr:	to spread something
tacɔ̃r	ta'c^{c}ɔ̃r:	treeshrew (*Tupia*)
ʔikɔr*	ʔi'kɔr:	CLF: animals
tgɔr*	tə'gɔr:	to allow (?)
ʔɔr	'ʔɔr:	to order; to request
sɔr	'sɔr:	[a type of cockroach]
bsɔr	bə'sɔr:	[a type of tree]
hɔr	'hɔr:	to drill or carve a hole
ɲɔr	'ɲɔ̃r:	coconut palm
sŋɔr	sə'ŋɔ̃r:	silence; to be silent
lɔ̃r	'lɔ̃r:	to hiss
wɔr	'wɔr:	to go around

/-w/

bgiw	bə'giw	wind
knsiw	kən'siw	[ethnonym: Kensiw]
hwhĩw	həw'hĩwə̃	crested wood-partridge (*Rollulus rouloul*)
ʔiwŋiw ~ ʔiwɪŋəw	ʔiw'ŋĩwə̃ ~ ʔiw'ŋəə̃w	to gaze; to look
putew	pu'tẹw	tiger (*Panthera tigris*)
pɛ̃w	'pɛ̃w̃	different; other
ʄritɛw	ʄᶻəri'tɛw	to squat
sɛ̃w	'sɛ̃w̃	[a type of small mammal]
hɛ̃w	'hɛ̃w̃	to eat (vegetables, fruit)
mɛw	'mɛ̃w̃	[a type of civet]
ʔamɛw	ʔa'mɛ̃w̃	cat
jɛw	'jɛw	to have a runny nose
pəw	'pəw	[sound of a shotgun]
bəw	'bəw	to be big
ʄwʄəw	ʄᶻuw'ʄᶻəw	Achilles tendon
gəw	'gəw	to go along the crest of a ridge
bʔəw	bə'ʔəw	whip snake (*Ahaetulla*)
həw	'həw	[a type of tuber (*Discorea pentaphylla*?)]
məw	'mə̃w̃	to make sound (of monkey)
prəw	pə'rəw ~ p'rəw	[sound of an animal fleeing from one tree to another]
paw	'paw	side of body
rabaw	ra'baw	[toponym: Raba]
hntaw	hən'taw	*petai* fruit
daw	'daw	to reach (of e.g. tree branch)
cdaw	cᶜə'daw	rainbow
kakaw	ka'kaw	[a type of monitor lizard (*Varanus*)]
bukaw	bu'kaw	flower (generic)
gaw	'gaw	wild boar (*Sus scrofa*)
mŋsaw	məŋ'saw	daughter/son-in-law
blaw	bə'law	blowpipe
wãw	'w̃ãw̃	[a type of civet]
kawaw	ka'waw	bird (generic)
rjaw	ri'jaw	[a type of wasp]
snow	sə'now	[a type of cobra]
rampɔw	ram'pɔw	long-tailed macaque (*Macaca fascicularis*)
krbɔw*	kər'bɔw	buffalo
litɔw	li'tɔw	boy; young bachelor
kacɔw*	ka'cᶜɔw	to work
hiʄɔw*	hi'ʄᶻɔw	to be green
kʄɔw	kə'ʄᶻɔw	to take off; to fly up
sikɔw	si'kɔw	to roar; to meow (of leopard, cat)
bakɔw	ba'kɔw	[a type of tree]
tmakɔw*	təma'kɔw	tobacco

manɔw*	mã'nɔ̃w̃	*manau* rattan
kalɔw*	ka'lɔw	if
salɔw	sa'lɔw	to split
halɔw*	ha'lɔw	to scare
pulɔw*	pu'lɔw	island
pulɔw tuɟoh*	pu'lɔw tu'ɟ^{z}oh	[toponym: Pulau Tujuh]
barɔw	ba'rɔw	oriental white-eye (*Zosterops palpebrosus*)
gurɔw*	gu'rɔw	to joke/to jest
smrɔw	səm'rɔw	to heal
cɔwrɔw	c^{c}ɔw'rɔw	[a type of bird]
jɔw	'jɔw	to take

/-j/

lamij	la'mĩj̃	sister/brother-in-law
gej	'gẹj̃	to eat
prgej	pər'gẹj̃	to feed
ʔjʔej ʔɔs	ʔijʔẹj 'ʔɔs	smoke
ʔjʔej tɔm	ʔijʔẹj 'tɔ^{b}m	steam
sej	'sẹj̃	to raise something
lej	'lẹj̃	body
pantɛj*	pan'tɛj̃	sand
lantɛj*	lan'tɛj̃	bed; floor
krtɛj	kər'tɛj̃	[toponym]
dɛj	'dɛj̃	to ascend to a mountain pass
padɛj*	pa'dɛj̃	rice in the husk; unharvested rice; [a type of small green snake]
kdɛj*	kə'dɛj̃	shop
kɛj	'kɛj̃	such; [a type of tuber (yam?)]
ckɛj	c^{c}ə'kɛj̃	to be big
caŋkɛ̃j	c^{c}aŋ'kɛ̃j̃	[a type of frog]
sraŋkej	səraŋ'kɛj̃	[a type of large spider]
ʔɛj	'ʔɛj̃	father
sɛj	'sɛj̃	to shovel
hɛj	'hɛj̃	PRONOUN 1D.INCL
mɛj	'mɛj̃	what?
mɛj ~ mɛ̃j (?)	'mɛ̃j̃	to comb
kmɛj	kə'mɛj̃	[a type of larvae]
smɛj	sə'mɛj̃	[ethnonym: Semai]
nɛj	'nɛ̃j̃	one
manɛj	mã'nɛ̃j̃	otter (*Lutra sumatrana*)
bralɛj ~ blalɛj*	bəra'lɛj̃ ~ bəla'lɛj̃	elephant's trunk
gulɛj	gu'lɛj̃	to whirl
karɛj	ka'rɛj̃	thunder; an evil human-like mythical being
bcrɛj*	bəc^{c}ə'rɛj̃	to repudiate one's wife
sorɛj	so'rɛj̃	[a type of tortoise]

morɛj* batuʔ*	moˈrɛj baˈtuʔ	magpie robin (*Copsychus saularis*)
wɛj	ˈwɛ̃j	past
gawɛj*	gaˈwɛ̃j	to play; to work
ʔawɛj	ʔaˈwɛ̃j	rattan; root; rope; string; vine
lawɛ̃j	laˈwɛ̃j	leech
lwɛj	ləˈwɛ̃j	honey
ɟrwɛj	ɟᶻɜrˈwɛ̃j	[a type of tree]
jɛj	ˈjɛ̃j	fly
kapɨj	kaˈpɨ̃j	to fly
tɨj	ˈtɨ̃j	thunder; river source
sdɨj	səˈdɨ̃j	vomit
hljdɨj	həlijˈdɨ̃j	flat ground
hakɨj	haˈkɨ̃j	to be slow
gɨj	ˈgɨ̃j	to produce/blow smoke
sɨj	ˈsɨ̃j	to abound (of game)
ksɨj	kəˈsɨ̃j	husband
chɨj	cᶜəˈhɨ̃j	moon
brɨj	bəˈrɨ̃j	late afternoon
jɨj	ˈjɨ̃j	to carry in one's hand; to drive (a car); to transport
pəj	ˈpə̃j	to produce/blow smoke
mantəj	mãnˈtə̃j	pangolin (*Manis*)
ɟəj	ˈɟᶻə̃j	cave
ʔhəj	ʔəˈhə̃j	to be small
hjhəj	hijˈhə̃j	to yawn
paj	ˈpa̰j	new; PRONOUN 2S.DIS
baj	ˈba̰j	to dig
lambaj*	lamˈba̰j	to wave
ɟaj	ˈɟᶻa̰j	banana
kɟaj	kəˈɟᶻa̰j	[a type of tree]
kaj	ˈka̰j	to heal
pakaj*	paˈka̰j	to eat a lot
crjkaj	cᶜɜrijˈka̰j	[toponym]
gaj	ˈga̰j	to turn/roll over in one's sleep
tuŋgaj	tuŋˈga̰j	headman
ʔaj	ˈʔa̰j	game animal
cʔaj	cᶜəˈʔa̰j	to sing
pljsaj	pəlijˈsa̰j	antenna (of fish)
haj	ˈha̰j	to follow
ɟahaj ~ ɟhaj (?)	ɟᶻaˈha̰j	[ethnonym: Jahai]
rmaj	rəˈma̰j	[a type of fruit]
ʔnaj	ʔəˈna̰j	to bathe
blaj	bəˈla̰j	to look upwards
slaj	səˈla̰j	clearing; swidden
caraj	caˈra̰j	gliding frog (*Rhacophorus*)

manraj	man^{d}ɹaj̃	[a type of tuber]
waj	'waj̃	to cut
rwaj	rə'waj̃	soul
duj	'duj̃	to go hunting/gathering late afternoon
kuj	'kuj̃	head
kuruhuj	kuru'huj̃	[a type of owl]
ʔnuj	ʔə'nuj̃	soon
taŋuj	ta'ŋuj̃	rambutan (*Nephelium lappaceum*)
pruj	pə'ruj̃	to sprinkle food to animals
toj	'toj̃	older brother of parent
toj biʔ	toj 'bəʔ	older sister of parent
ʔoj	'ʔoj̃	to open; to pull off (e.g. clothes)
huroj	hu'roj̃	to become full
pɔj	'pɔj̃	to dry in the sun
ha'pɔj	ha'pɔj̃	hut
smpɔj mit	səm'pɔj 'mĩt̚	eyelid
bjbɔj	bij'bɔj̃	[a type of insect]
btɔj	bə'tɔj̃	to be red
kadɔj	ka'dɔj̃	[a type of long-snouted insect]
takɔj	ta'kɔj̃	sail-fin lizard (*Hydrosaurus*)
rgɔj	rə'gɔj̃	to be red
ʔɔj	'ʔɔj̃	INTERJECTION: hey!
thɔ̃j	tə'hɔ̃j̃	to be red
cnhɔ̃j	c^{c}ɜn'hɔ̃j̃	noon
nɔj	'nɔj̃	to copulate
laŋɔj	la'ŋɔ̃j̃	shadow
lɔj	'lɔj̃	to run; to flee; to hurry
prlɔj	pər'lɔj̃	star
rɔj	'rɔj̃	to leave
wɔj	'wɔj̃	banded kingfisher (*Lacedo pulchella*)
twɔ̃j	tə̃'wɔ̃j̃	evening/night
kwɔj	kə'wɔj̃	[a type of tuber]

Appendix II: English–Jahai finder list

The following finder list presents Jahai translations of alphabetically listed simplified English meanings of the Jahai lexemes given in the glossary (see Appendix I). It is intended as a search tool; for more detailed information the reader is referred to Appendix I. In cases where an English form has more than one Jahai translation, the Jahai forms are listed in rhyming order (see Appendix I). As in the glossary, Jahai forms thus far identified as definite or likely loans from or via Malay, including English words, are marked with an asterisk (*).

able, be	boleh*
abound (of game)	sɨj
accompany	prjər
Achilles tendon	ɟwɟəw
adorn	ksɛp
after	lpəs*
afternoon	brɨj
again	wel
aim (with blowpipe or gun) (v.)	btol*
aircraft	kapal*
all	smwɛʔ*
allow (?)	tgɔr*
alone	blaʔ
also	hajiʔ
ambush (v.)	ʔimbus*
ancestor	toʔ
and	ʔalɔʔ
angry	marah*
animal (game)	soh, ʔaj
animal, type of	ptpit
animal, type of	rajɔ̃t
animal, type of	palĭk
animal, type of	ʔawɨk
animal, type of	tiʔ siriŋ
animal, type of	cawis

animal, type of	tamam
animal, type of	hɔrjĩŋ
animal, type of	rəŋ
animal, type of	rasɛl
animal, type of	hnrɛl
animal, type of	sɛ̃w
ankle	kmkə̃m
ankle-joint	krɨ̃l
answer (v.)	ɟawap*, balas*
ant, type of	bit
ant, type of	hlaŋkɛt
ant, type of	tampɨl
antenna (of fish)	pljsaj
anus	ʔmpɔŋ kit
apple	ʔipəl*
argue	ʔūt
arm, lower	prbɛr
arm, upper	bliŋ
armpit	lda?
arrive	lmtem
arrowhead	pucuk* lɔc
ascend	ɟɔh, tampɛŋ, dɛj
ashes	ʔabuʔ* ʔɔs
Asian horned toad (*Megophrys monticola*)	kŋkaŋ
ask	tʔoc, taɲɛʔ*, ciʔɨs
assemble	ctip ~ ctɨp
astonish	bnken ~ mnken (?), mhel
attack (v.)	suɸ, sltuh
awake	cŋwɛ̃ŋ
axe	kapɔʔ*
babbler (*Malacopteron*)	cprɛp
bachelor, old	kbet clah
back (body part)	krɔʔ
backside	tkɨh
bad	ɟahat*, laʔɨs
bag	bɛk*, cɔk, hapiʔ
balance (v.)	ton ~ tnton, dor
bamboo	karɔp, bhɨt, buloʔ*, bamuʔ*, gnun
bamboo slat	tbeŋ
banana	pisaŋ*, ɟaj
banded kingfisher (*Lacedo pulchella*)	wɔj
banded palm civet (*Hemigalus derbyanus*)	ckcɔk, kɲsiŋ
barbet (*Megalaima*)	carah
bark (v.)	ɟhoŋ
basket	ragaʔ*

bat	kajiʔ, tapər
Batek (ethnonym)	batɛk
bathe	ʔnaj
beak	cnɔŋ
bear (*Helarctos malayanus*) (n.)	kawip
beard	ɟaŋut*
beat (v. tr.)	rajeh
beat (v., intr., of heart)	lŋpuŋ
because	sbap*
bee-eater (*Merops*)	ririk
beetle, type of	trhim
beg	mamuŋ
begin	mulaʔ*
belch (n.)	gos
believe	harəp*, cajaʔ, ʔas
belly	ʔɛc
bend (v.)	wɔ̃t
betelnut	pinɛŋ*
big	lim, bəw, ckɛj
big toe	tabo?
bind (v.)	sklik
binturong (bearcat; *Arctitis binturong*)	luŋan
bird	kawɔ̃t, kawaw
bird, type of	wtwit
bird, type of	ddit'
bird, type of	taʔic
bird, type of	trhic
bird, type of	ku'cɔ̃k
bird, type of	sŋɔk
bird, type of	tŋkɔʔ
bird, type of	ɟɔweh
bird, type of	duduh
bird, type of	ɟnuh ~ ɟnũh (?)
bird, type of	brul
bird, type of	cɔwrɔw
bird's nest	sɔm
bird of prey	klɛ̃ŋ
birthmark	ksĩt
bite (v.)	kap, cɔh
bitter	kdɛk
black panther (*Panthera pardus*)	ʔap tmtum
blade (of knife)	mit taɟiʔ
blanket	pɲɟɔʔ
blind	cũʔ
blink (v.)	cil

blood	darah*, bhĩm
blood vessel	sʔsoʔ
blow (of person)	pɨt, ʔuk, ʔuɸ, phos, hamos, ɟuh, haŋhoŋ
blow (of smoke)	ɲiɸ, gɨj, pəj
blow (of wind)	sitjɛ̃t
blowpipe	blaw
blowpipe dart	snlɔc
blowpipe, shaft of	joʔ
blue	biruh*
blue coral snake (*Maticora bivirgata*)	tŋwaŋ
blue-eared kingfisher (*Alcedo meninting*)	smpal
blunt	majiŋ
boat	prahuʔ*, sampan*
body	lej
boil (n.)	cməc, bisol*
boil (v., intr.)	didih*
boil (v., tr.)	pgoh
bone	ɟʔeŋ
book	bukuʔ*
borrow	piŋjam*
bottle (n.)	labuʔ*, balaŋ*, butɔl*
bottom (body part)	kit
bow (n.)	lɔc
box (n.)	ptiʔ*
boy	kɟɨh, litɔw
brahminy kite (*Haliastur indus*)	klɛ̃ŋ siwah
brain	smutlɔ̃t, lkɛm
brake (v.)	brik*
branch (n.)	cabaŋ*
break (v., intr)	ktis*, pcah*
break (v., tr.)	rigɨh, pok
breast	ʔɛ̃m
breastfeed	bnalɛh, piʔɛ̃m
breathe	ɟkɟɨk, hĩk
bridge (n.)	titih*
British (ethnonym)	britis*
broadbill (*Psarisomus*)	kawɔ̃t batuʔ*
brother of parent, older	toj
brother of parent, younger	bɛh
brother-in-law	lamij
brown	blakūʔ
brown hornbill (*Ptilolaemus tickelli*)	maleʔ
bucket	bldiʔ*
buffalo	krbɔw*

build (a house) woŋ
build (a hut) swat
bullet pluruh*
burn (v.) ʔŋic, hokɔ̃k, ɟiʔ, sɔʔ, sjiŋ, brjɔŋ, cɔm
burp (v.) krʔəp
bus bas*
butcher (v.) kujak
butterfly tawɛ̃k
buttock hawɨr ~ hawɔr
buy bliʔ*
buzz (v.) tmtəm, smsɔm
call (v.) saʔot, liʔlɛʔ, kul, pɨr
call (v., of bird) woh ~ wɔh
camphor barus*
cannibal coleʔ, hmlɨ̃t
car kritəh*
carp, type of ʁbul*
carp, type of (*Puntius*) lampam*
carry (in one's arms) cdum
carry (in one's hand) jɨj
carry (in one's mouth) ktɛp
carry (on one's back) təp, ragət, bʔbɔʔ, taŋkəl, gɔl
carry (on one's neck) critol
carry (on one's shoulder) joʔ, taŋwoŋ, kanər*
carry (on stick) gulɛm
carve gəs, sot
casque (of hornbill) krkok
cassette tep*
casting-net ɟalɔʔ*
cat kuceŋ*, ʔamɛw
catch (v.) cɛp, bantuʔ*, maŋ
catfish, type of (*Mystus planiceps*) bawuŋ*
cattle lmuʔ*
caught sɨk
cause a blister (of fire) cɔɸ
cause pain ɲɲrɛɲ
cave gahuŋ*, ɟəj
centipede, type of kʔep
chair paŋkeŋ*
charcoal rŋhɔ̃ŋ ʔɔs
chat (v.) buwal*
chatter (v.) siŋraŋ, prihkah, crikɔ̃k
cheek kapɔʔ
chest dkduk
chew ɲɔp

child	kɛn, wɔŋ
childless person	clah
Chinese (ethnonym)	cinɛʔ*
chop (v.)	pɛŋ, blah*
circumcision	snkɔ̃l
civet, type of	tanl̃in
civet, type of	gŋguŋ
civet, type of	mɛw
civet, type of	wãw
clap (v.)	pak
claw (n.)	klkɔʔ, cnrɔs, katɔ̃ŋ
claw (v.)	tampɛr*
clean (v.)	cuciʔ*
clear (one's nasal cavity)	krhak
clear (one's throat)	krhɛm ~ krhm̃ɨm
clear (v.)	rɔh
cleave	cbis*, tas
CLF: animals	ʔikɔr*
CLF: cluster	prduʔ*
CLF: large flat objects	bidaŋ*
CLF: oblong objects	bataŋ*
CLF: spherical/cubical objects	buwah*, tuŋkɔ̃l
click (v.)	kktɔ̃k, kktɛ̃k
climb (down)	gɨs
climb (v.)	dbɔt, tʔɔc, trɛɲ
climb up	lwec, ciweh
close one's eyes	padam*
clothes	baɟuʔ*
cloud	sagup
cloudy	rdũm ~ rdup
cobra, type of	sjul
cobra, type of	ʔmpet
cobra, type of	snow
cobra, type of	prɛŋ
cockroach, type of	sɔr
coconut palm	ɲɔr
coin	slem*
cold	tmkɛt, sɟuʔ*
collarbone	slaŋkaʔ*
collared kingfisher (*Halcyon chloris*)	tŋtoŋ
collared scops-owl (*Otus lempiji*)	canwɔc
collect (v.)	pɔc, ɟit, lawaʔ
comb (n.)	sikat*, knɛc
comb (v.)	mɛj ~ mɛ̃j (?)
come	klwl̃il

come out (of leaves)	crbɨr
common kingfisher (*Alcedo atthis*)	cʔcaʔ
convey	ʔampuʔ*
cook (v.)	pʔak, pktɔk, pgoʔ, pʔəh, krʔɔm, ʔɲcen, pel
copulate	nɔj
corporal	kpral*
corpse	sarɔʔ
cough (v.)	ʔoh
count (v.)	jɛp, kiraʔ*, bilɛŋ*
cover (v.)	dun, bɨl, haŋkal
cover one's eyes	tuduŋ*
crack (v.)	kalcɔ̃h
crash (v.)	laŋar*
crawl (v.)	wɨn
crested serpent-eagle (*Spilornis cheela*)	klɛ̃ŋ hip
crested wood-partridge (*Rollulus rouloul*)	hwhĩw
crimson-winged woodpecker (*Picus puniceus*)	blatoʔ
crocodile (*Tomistoma, Crocodilus*)	bujaʔ*
crouch (v.)	krʔɨp
crow (v.)	kokoʔ
crunch (v.)	raŋɨp
cubit	soʔ
cuckoo (*Cuculus*)	btɔk
cup	koleh, takɔʔ
curl (v.)	crilwɛl
cut (v.)	get, kɛt, kɛc, hiris*, tbɛs*, cah, waj, rwis, katəm*
dance (v.)	sisɛʔ, kŋseŋ
dangle (one's arms)	riŋwĩŋ
dark	gcĩh
daughter-in-law	mŋsaw
day	ktɔʔ
day after tomorrow	camɔʔ
day before yesterday	ptom
daylight	peŋ
decline (v.)	gʔlɛʔ, siʔroʔ
deep	bŋɟiʔ
defecate	ʔɛc
deliberate (v.)	gɨm
dense	plet
descend	cruɸ, ɟlʔil, gɛs
devour	cbɔt, sksɛ̃k, hnhən
Diard's trogon (*Harpactes diardii*)	ckcok
die	slap, hapaʔ, ɲaɲuʔ, kbis, ʔis

different	lajin*, ʔasiŋ*
dig (v.)	wɔk, tilɔɸ, kis, baj
dirt	karat*
dirty	ɟʔɔm
disappear	plə̃c
do	deʔ
dog	ʔɔ̃t, ʔacîʔ
door	pintuʔ*, lanʔɔs
dove	kwɔŋ
dragonfly	tanɔŋ
dream (v.)	jɲjɛɲ
drill (v.)	hɔr
drink (milk from mother's breast)	mam
drink (v.)	ʔɛ̃m, huc
drip (v.)	pɛl
drop (v.)	jɔh
drown (v. intr.)	slot
dry	kriŋ*
dry (v. tr.)	sat, pɔj
duck	ʔiteʔ*
durian (*Durio zibethinus*)	dorɛn*
dust	ʔabuʔ*
each	masiŋ-masiŋ*
ear	ʔnteŋ
earth	teʔ
east	tom ktɔʔ, simur*
easy	ctih, snɛŋ*
eat	but, muc, ɲɛm, hɛ̃w, gej, pakaj*
egg	ktɨt, makɔʔ
egret (*Egretta*)	prɛŋ
eight	lapan*
elbow	kajɔ̃ŋ
elder	ʔawaʔ, tʔtiʔ
elephant (*Elephas maximus*)	gaɟah*
eleven	sblas*
embarrassed	malɔk*, ʔɲsol
EMPHATIC PARTICLE	lɛh*
endure	tahan*
enough	ɟnuh*
enter	bləh
every	tjap-tjap*
exchange (v.)	tukar*
excrement	ʔɛc
exist	weʔ
expensive	mhɛl*

explode	kiciɸ
extinguish (v. intr.)	hmirɨn
extinguish (v. tr.)	plit, cirĩɸ
eye	mit
eyebrow	cawĩŋ
eyelid	smpɔj mit
fall (v.)	plek, rɛs, tl̃ih, run, kulkil, taŋɛl*, gibal
fan (v.)	cɛ̃ɸ, pɔh
fang	sjuŋ*
far	bɲɟi?
fart (v.)	ktkit
fast	lktwət, bsɔɸ
fat (n.)	gmu?*, l?əs
father	?ɛj
father's father	ta?
father-in-law	kn?ac
fear (v.)	hgik
feather	sɔk kawɔ̃t
feed	prgej
feel	grsɨc, ktlĩt
fell (v.)	cih, tuŋ
fern (*Filex*)	paku?*
fern, type of	karaŋ
field	padaŋ*
fight (v.)	tumo?*
fill (v., intr.)	ksɛl, huroj
finger	ɟari?*
fingernail	katɔ̃ŋ cjas
finish (v.)	jɨl
finished	sɔc
fire	?ɔs
fireplace	tuŋul
fish (n.)	?ikə?, ?ikan*
fish (v.)	kajil*
fish, type of	tŋis
fish, type of	tŋraŋ
fishing hook	mit knajil
fish spear	srampaŋ*
fish trap	dkdak
five	limɛ?*
flap (wings)	kŋjɛŋ
flat	tĩp ~ tɨp, jup
flat ground	hljdɨj
flat-headed cat (*Felis planiceps*)	juluŋ
flee (v.)	lɔj

flex (v.)	tgaʔ*
flicker (v.)	priɸ ~ piɸriɸ, prpɨr ~ prpɨr
flip (v.)	kaltɛt
float	lampoŋ*, timul*
floor	lantɛj*
flow (v.)	wet, sɟɨp, cuhɛ̃ʔ
flower	buŋɛʔ*, bukaw
flower bud	kutum
flower, type of	pnraʔ
fluid	pʔɨh, plʔeŋ
fly (n.)	jɛj
fly (v.)	halaʔ, kapɨj
fly up	piɟəʔ, trɟɛh, kɟɔw
flying fox	klwaŋ
flying squirrel (*Petaurista*)	kapuk, tasan
fold (v.)	lɛp
follow (v.)	ʔikot*, təc, haj
fontanelle	lkɔc, lĩɸ ~ lɸlĩɸ
food	bap
foot	can
forehead	ptɨʔ
forest	hɨp
forget	krlep
fork (of river) (v.)	was
four	ʔmpat ~ ʔmpət*
friend	baltɛk, sajĩŋ
frighten	prhir
frightened	hir
frog, type of	kmɛt
frog, type of	skŋuk
frog, type of	riŋkoŋ
frog, type of	brwə̃ŋ
frog, type of	caŋkɛ̃j
front	sɛɲ
frontal tuber	wɛ̃s
fruit	sĩp, bɔh, kbɨʔ, kmɔʔ
fruit, type of	hnit
fruit, type of	pahet
fruit, type of	stwɛt
fruit, type of	hoc
fruit, type of	tpɔs
fruit, type of	tabɛn
fruit, type of	bagan
fruit, type of	bɔn
fruit, type of	kalwiŋ

fruit, type of	kabaŋ
fruit, type of	balūŋ hajam*
fruit, type of	lamɔŋ
fruit, type of	rŋkal
fruit, type of	star
fruit, type of	rmaj
fruit tree, type of	rawɔh
fruit tree, type of	ʔnrɛŋ
full	cukop*, bhiʔ, sbem
gall-bladder	kmɨt
gastropod, type of	hawap
gastropod, type of	kluktɔk
gastropod, type of	kalɔ̄ʔ
gaur (*Bos gaurus*)	sladaŋ*
gaze (v.)	ʔiwŋiw ~ ʔiwŋəw
gecko, type of	ciʔcaʔ
gecko, type of	pokəh
get	dɛs
get out (of bed)	bliwīs
ghost	kmuc, hantuʔ*, sanuʔ, kīs
giant squirrel (*Ratufa affinis*) (?)	kralɛh
gibbon (*Hylobates*)	mawɛ̄ʔ
girl	ʔanɛk, mɛʔ, ʔalɛh
give (v.)	ʔɛk
give birth	banɛʔ, srɔh
glare (v.)	kbah
gliding frog (*Rhacophorus*)	caraj
glow-worm	ktlɨt
gnaw	kɔc, lslɔs
go (across)	piris, lintɛs*
go (along crest of a ridge)	gəw
go (along mountain-side)	tigil
go (along watercourse)	rkruk
go (alongside)	srsir
go (around)	swaŋ, wɔr
go (back)	wek
go (out)	snrɛk
go (together)	lɛh
go (v.)	cɨp, ɟaŋkaʔ*
goat	kamiŋ*
gobble (v.)	khɔp
gold	mɛs*
good	btʔɛt, seh
goose bumps	hrŋpeŋ
grandchild	kaɲcɔʔ

grandmother	jaʔ
grass	rumput*
grass, type of	karil
grasshopper, type of	ɟarɛ̃t
grasshopper, type of	blalaŋ*
grave	kubɔr*
great argus (*Argusianus argus*)	kawoŋ
great grandparent	dɔn
great hornbill (*Buceros bicornis*)	bwɛ̃ŋ
greater racket-tailed drongo (*Dicrurus paradiseus*)	tŋtaŋ
green	blʔɛɲ, hiɟɔw*
green tree python (*Chondropython viridis*)	hajom, tagan
grenade	grinɛt*
grow (of animates)	bakɛs
grow (of plants)	kanɔh
growl (of stomach)	pirʔur
grunt (v.)	tũs ~ tmus
guarantee (v.)	ɟamin*
guide (v.)	pimpin*
gun	bdil*
hair	sɔk
halt (v.)	gamɛʔ
hammer (v.)	phat*
hand	cjas
hang	gantuŋ*
hard surface	timɔʔ
hardwood tree (*Turtur tigrinus*)	rmpɛc
haunt (v.)	ɲuɸ
have	puɲɛʔ*
have a running nose	jɛw
have a sad expression on one's face	wawɔk
have sore eyes	lak
head (n.)	kuj
headache	twiɲ, pniŋ*
head-hair	sɔk kuj
headman	puŋhuluh*, tuŋgaj
headwater	mɔh tɔm, tɨj
heal (v., tr.)	ɟampiʔ*, simaŋ*, smrɔw, kaj
hear	ken, kɟeŋ
heart	klaŋis
heat (v.)	krʔic
heat poisoned blowpipe dart in fire	pʔɔŋ
heavy	hɲɟut
heavy (of rain)	lbɛt

heel — crkɨp can, dlɗil
height — trinɛʔ
help (v.) — tuluŋ*
helping — piŋan*
herd — sajər
here — ʔəh
hey! — ʔɔj
hide (v. intr.) — pɟaʔ
hide (v. tr.) — dol ~ pidol
high — bɲɟiʔ
hiss (v.) — ciɸ, lɔ̃r
hit (v.) — lukaʔ*, tbɔh*, ktɨm, pɛ̃n, cɔ̃ŋ, khil
hoe (v.) — caŋkul*
hold — paŋkoh*, pgɛŋ*
hold (one's wrist) — cnr̃ŋ
hole — ɟlɔʔ, kɨn, ʔmpɔŋ
honey — lbeh, koŋ, lwɛj
honeybee — lbah lwɛj
honeycomb — sɨt, tpɨɲ
hop (v.) — ɲnɲɨn, jɨn, hiljɔl
horn — tanoʔ*
hornbill, type of — tkuk
hornbill, type of — khkuh
horse — kudaʔ*
hot — bkɨt, til
hour — ɟam*
house — diŋ
how many? — mɛj siʔ
how? — ʔacin
hug (v.) — bhɨʔ, ʔəm ~ ʔimʔəm
human — mnraʔ
hundred — sratos*
hungry — crɔʔ, haleh
hunt (v.) — sam, duj
hurry (v.) — lɔj
hurt (v., intr.) — jɔ̃t, ptis
husband — ksɨj
hut — hajɛ̃ʔ, nŋkɨŋ, ha'pɔj
ice — ʔɛs
if — kaluh*, kalɔw*
insect, type of — parɛt
insect, type of — tĩs bataŋ*
insect, type of — krpəŋ
insect, type of — kaɲcɔŋ
insect, type of — kmarbəl

insect, type of	bjbɔj
insect, type of	kadɔj
inside	klɛŋ
intestines	lmaʔ*
ipoh poison	dɔk
iron	bsiʔ*
island	pulɔw*
Jahai (ethnonym)	ɟahaj
Japanese (ethnonym)	ɟpon*
jar	hɔʔ
jaw	jaŋkəʔ
jest (v.)	cnhaʔ, gurɔw*
jump (v.)	lumpɛt*, ɟilĩn
just	sʔoʔ
keelback (*Xenochrophis*, *Amphiesma*)	rksɔk
keep (fire burning)	ɟuluk
keep (v.)	prisɛʔ
Kensiw (ethnonym)	knsiw
key	kuɲciʔ*
kick (v.)	sipaʔ*, tjis, tin
kidney	krtlɔ̃t
kill (v.)	ckət, ɲʔɲɛʔ, kə̃s, bunuh*, lawan*
kind (n.)	ɟnis*
Kintaq (ethnonym)	knaʔ ~ gnaʔ
knee	kaltoŋ
knee-cap	kmɨt
kneel	pʔah, klkol
knife	taɟiʔ*
knife handle	ger taɟiʔ
know	ʔtʔet
krait, type of (*Bungarus*)	karjɛl
lake	taseʔ*, dɛ̃m*
lamp	plitəh*
langsat (*Lansium domesticum*)	laɲsɛt*
language	kliŋ
Lanoh (ethnonym)	lanɔh
larvae, type of	klutbot
larvae, type of	mamiʔ
larvae, type of	kmɛj
lash (v.)	bɛt, ʔikɛt*
laugh	sitkət, lkluk
lay (a floor)	grlɨp
leaf	haliʔ
leaf monkey (*Semnopithecus*)	tabəŋ
leaf, type of	haliʔ gadiŋ

leaf, type of — srikɛn
leafbird (*Chloropsis*) — sŋpɔŋ
leak (v.) — lɔɸ
lean (one's head) — ʔiʔŋoʔ, wɛŋ ~ wiŋwɛŋ
learn — ʔaɟar*
leave (v.) — rɔj
leech — lajaʔ, lawɛ̃j
left (side) — wîʔ
left-handed person — mʔwîʔ
leg, lower — gor
leg, upper — bliʔ
leopard (*Panthera pardus*) — ʔap ʔawɛj
letter — soret*
lick (v.) — scboc, blɛk, kal
lie (of animal) (n.) — lanik
lie down — tek, kəh, pum, bin, kol
lift (v.) — piwek
light (of weight) — hatɔ̃p
light (v.) — suk, paɲcaŋ ~ paɲcoŋ
lightning — kilat*
like (prep.) — pon*
like (v.) — hmhɔm
limp (v.) — sŋsɛ̃ŋ
lip, lower — tnit
lip, upper — nus
little — ʔaɟɔʔ, ʔhəj
little finger — kajîʔ
live — pasaʔ, gɔs, praʔ, hampɛs
liver — ros
lizard — calak
lizard, type of — calat
lizard, type of — haluk
lizard, type of — taroŋ
lizard, type of — dalɔŋ
loincloth — lɛn
long — btec
long-tailed macaque (*Macaca fascicularis*) — rampɔw
long-tailed parakeet (*Psittacula longicauda*) — brlĩŋ
long-tailed shrike (*Lanius schach*) — ci'kɔ̃ʔ
look (v.) — clĩk, ɟŋɔʔ*, ciŋlĩŋ, cilkil, ʔɛl, trwer, sitər ~ hitər, blaj
lose — kalah*
lose (hair) — prasut, tstũs, gugur*
lose (one's footing) — cantil
louse — ciʔ

low	rnah*
lump	ckɨm
lung	sop
magpie robin (*Copsychus saularis*)	morɛj* batuʔ*
make (v.)	deʔ
make sound	hūʔ, kŋliŋ, məw
Malay	mantɔh
malkoha (*Phaenicophaeus*)	btɔk
man	tmkal
manchild	goh
mango	blawen
mangosteen (*Garcinia manggostana*)	maŋis*
many	bnolɛt, brasaʔ, kɔm
map	mɛp*
march (v.)	kawat*
masked civet (*Paguma larvata*)	ʔilɔ̄ŋ
mat	nɨs
maybe	muŋken*
meander (v.)	klihdɔ̄h
meat	sɛc, lawoʔ*
meet	ɟumpaʔ*, ʔuswas, sɔ̄h ~ sihsɔ̄h, lmah
mend (v.)	kuŋciʔ*
Menriq (ethnonym)	mnriʔ
milk	susuh*, məm
millipede, type of	klɛbac
millipede, type of	cmalpɔk
millipede, type of	taŋɔʔ
millipede, type of	taluŋ
mince (v.)	chɛɔh, ciŋcaŋ*
miss (v.)	lɛc
molar tooth	bŋkaʔ
money	dwiʔ*
monitor lizard, type of (*Varanus*)	bagɛn
monitor lizard, type of (*Varanus*)	grɛŋ
monitor lizard, type of (*Varanus*)	kakaw
moon	bolan*, chɨj
morning	pagiʔ*
mortar (household utensil)	lsuŋ
mortar (weapon)	mutah*
mosquito	hagas*
mother	bɨʔ, maʔ
mountain	ɟlməl
mountain pass	lget
mountain-top	pəʔ, tul
mouse deer (*Tragulus*)	plaŋcɛt, planok*

mouth harp	raŋgoɲ
move (ears)	siɲuŋ
move (in circles)	pusiŋ*, liŋwuŋ
move (snout)	ɲɔc
move (v.)	kɔp, ɟok, ɟatoh
mud	lapɛ̃k, kobaŋ*
muntjac (barking) deer (*Muntiacus muntjac*)	bhɔl
mushroom, type of	tʔaʔ pɔk
myna (*Acridotheres*)	cjuŋ*
nail	klkɔʔ, cnrɔs, katɔ̃ŋ
name (n.)	knmɔh
navel	dut
near	pdəh*
neck, nape of	tŋkɔk*
needle	jarum*
NEGATIVE MARKER	bŋit, braʔ, bokan*
net (n.)	pukat*
new	paj
night	twɔ̃j, hrkit
nine	smilan*
nipple	dapor* ʔɛ̃m
noon	cnhɔ̃j
nose	mɔh
nothing	blap
number (n.)	siʔ
ogle (v.)	jɛt
oil palm	klapah sawit*
old	manɛh
one	nɛj
only	saɟaʔ*, ɟagaʔ
open (one's eyes)	ciwɛ̃ʔ
open (v.)	ʔoj
opposite side	titɛp
order (v.)	ʔɔr
oriental pied hornbill (*Anthracoceros albirostris*)	thteh
oriental white-eye (*Zosterops palpebrosus*)	barɔw
oriole (*Oriolus*)	llah
other	pɛ̃w
otter (*Lutra sumatrana*)	manɛj
outgrowth (on tree)	goɲ
outside	hɨp
overgrown	bhɨr
overnight (in forest) (v.)	tmdem
owe (?)	dos
owl, type of	kuruhuj

Pacific swallow (*Hirundo tahitica*)	ɟawɛr
paddle (v.)	kajɔh*
pain (n.)	ptis
paint (v.)	cat*
palm (of hand)	tapar* cjas
palm leaf, type of	lajɨs
palm, type of	malih*
palm, type of (*Arenga obtusifolia*)	laŋkap*
pangolin (*Manis*)	mantəj
pant (v.)	hɛl
papaya (*Carica papaya*)	btɔk
paradise tree snake (*Chrysopelea paradisi*)	karil
pardon	maʔap*
pass (v.)	ʔalɨc
past	wɛj
path	glɔŋ, har
peacock-pheasant (*Polyplectron*)	ckɨ̃m
peck (of bird) (v.)	tbuk*, saŋkɨ̃l
peep (v.)	pdep
pelvis	tapɔ̃h
penis	laʔ
people	mnraʔ
pestle	ladaʔ*, gul
petai fruit	hntaw
phalange	cnrɔs
pharynx	prŋgəŋ
Philippine glossy starling (*Aplonis panayensis*)	pŋpɨ̃ŋ
photograph	gamah*
pick (one's teeth)	stsɛt
pick (v.)	wɛh
pig	napak, clapak, hmalah, hawɛ̃n, gaw
pigeon	bkɨk
pig-tailed macaque (*Macaca nemestrina*)	bawac
pinch (v.)	kas
pineapple (*Ananassa*)	lanas*
pitta (*Pitta*)	kawɔ̃t batuʔ*
place (n.)	tmpət*
place (one's hand)	ptpət
plaintive cuckoo (*Cacomantis merulinus*)	war
plait (v.)	teŋ, kɔŋ
plank	papan*
plant (v.)	tanɛm*, kawen
plant, type of	krak
plantain squirrel (*Callosciurus notatus*)	kdek creh
play (games)	mnmin, gawɛj*

play (instrument)	krem, cantũŋ, pŋjuŋ, tir
pluck (v.)	tkjɔk, prlʔol ~ prlʔil, tər
pocket	puket*
point (v.)	tjɔ̃k ~ tkjɔ̃k, tiʔleʔ
pool (n.)	loboʔ*
porcupine (*Hystrix, Atherura*)	lantɔ̃h
porcupine, type of	lanɛk*
porcupine, type of	ɟkis
portion	kpiŋ*
post (military) (n.)	pos*
potato	gantaŋ*
poultry	hajam*
pounce upon	duk
pound (v.)	sntip, hntɛk*, tiʔ, sih, patim, til
pour	sɛt, cʔiʔ, gin
pregnant	kajot, maŋkɔ̃ʔ
prick (v.)	pɛ̃k, hɔɸ
PROHIBITIVE	ʔakɛt
PRONOUN 1D.EXCL	jɛh
PRONOUN 1D.INCL	hɛj
PRONOUN 1P.EXCL	japɛ̃h ~ pɛ̃h
PRONOUN 1P.INCL	heʔ
PRONOUN 1S	jɛʔ
PRONOUN 2/3P	gin
PRONOUN 2D	ɟih
PRONOUN 2S.DIS	paj
PRONOUN 2S.FAM	mɔh
PRONOUN 2S.INT	miʔ
PRONOUN 3D	wih
PRONOUN 3S	ʔoʔ
Provost's squirrel (*Callosciurus prevosti*)	kdek ʔabuʔ
pull (blowpipe apart)	tltil
pull (v.)	tareʔ*, keŋ, her
pump (n., v.)	pam*
push	surut*, tulɛk*, tulaʔ*, lon
put	boh
put (inside)	ɟleŋ, ʔisiʔ*
put (to sleep)	piʔiʔ
put (together)	samoŋ*, kɛs
python, type of (*Python*)	lawɛŋ
quill (n.)	klek, harim
quiver	baniʔ
raft (v.)	halɛp
ragged	hkhɛ̃k
rain (v.)	hĩc

rainbow	cdaw
raise (v.)	hɟɔl, sej
rambutan (*Nephelium lappaceum*)	taŋuj
rapid (n.)	ɟrem*
rat	wawət
ration	risin*
rattan	ʔawɛj
rattan, type of	manɔw*
rattan, type of	knbiʔ
rattan, type of	pipih
rattan, type of	wam
rattan, type of	tahanɛn
rattan, type of	batɛŋ*
rattle (v.)	kacĩt
raw	ɲɛp
reach (v.)	daw
ready	cadaŋ*
rear (v.)	pihiraʔ
red	rhik, btɔj, rgɔj, thɔ̃j
reed snake, type of (*Macrocalamus*)	dadɔ̃l
refuse (v.)	coh, jiʔ, prnɛn, gɔ̃ɲ
RELATIVE MARKER	jaŋ
remain (v.)	brasɛʔ, snlɔŋ
remains	sisah*
remember	kakɛp, ʔiŋɛt*
repudiate	bcrɛj*
request (v.)	smɛɲ
rest (v.)	rihat*, braduʔ, cundɔʔ*, kil, gol
reticulated python (*Python reticulatus*)	talon
retreat (v.)	krlun
return (v.)	mudik*
reveal	hūt
rhinoceros (*Dicerorhinus sumatrensis*)	hagap, badaʔ*
rhinoceros hornbill (*Buceros rhinoceros*)	hŋaŋ*
rice (cooked)	nasiʔ*
rice (husked)	bras*
rice (in the husk)	padɛj*
right (side)	tem
right-handed person	mmtem
ring (n.)	cnɔs ~ cnus
ring (v.)	taŋgɔ̃h
ripen	tuhaʔ*
river confluence	was tɔm
river-mouth	kit tɔm
roar (v.)	piʔjɔʔ, sikɔw

roast (v.)	paŋaŋ*, pʔɔl
roll (of thunder) (v. intr.)	gɨr ~ gar
roll (v. intr.)	luʔlɛ̃ʔ, gruŋliŋ, criŋpoŋ
roll (v. tr.)	hamat*, məl
roof	saraŋ*
root	ɟʔes
root-crop	hobiʔ*
rope	ʔawɛj
rotten (of wood)	ʔɲsɔʔ
round object	pɔk
rub (v.)	sɨ̃t ~ ʔotsɨ̃t, kəm, klɛl
rubber	gtah*
run (v.)	lɔj
run short	dən
Sabüm (ethnonym)	sabɨm
sack (n.)	gunih*
safety	slamat*
sail-fin lizard (*Hydrosaurus*)	takɔj
saliva	lheŋ
salt	ʔmpɔc, garəm*
salty	masin*
sambar deer (*Cervus unicolor*)	kasaʔ
same	srupəh*
sand	pantɛj*
sandfly	ŋit
satisfied	puwas*
say	kdɨh
scar	parut*, cɔc ~ cəc
scare (v.)	halɔw*
scent (v.)	crŋir ~ crŋər
scold (v.)	ʔacah*
scorch (v.)	pkɨp
score (v.)	knɛʔ*
scorpion	ɟɔwɛ̃ŋ
scrape (v.)	suduh*
scratch	hakat, gɛ̃c, kac, cscis, kɔ̃s, gluŋʔɛ̃ŋ
sea	lawot*
season	musem*
seat bone	cntɛl
see	jət, nampaʔ*, tŋleŋ, ʔɛl
seed	biɟiʔ*, mit ʔhəj
sell	ɟuwal*
Semai (ethnonym)	smɛj
send	hantər*
separate	pihaʔ

settle	ttap*
seven	tuɟoh*
shadow	laŋɔj
shake (v., tr.)	jɨɸ, ʔuŋjaŋ ~ ʔuŋjiŋ, gntar*
sharp	cmɛʔ
shaving (from carving)	cpah
shed (v.)	ruroh
shiver (v.)	grgɨr
shoe	kasot*
shoo (v.)	cih
shoot (v.)	steʔ*, haluh, tmɨm, pɛɲ, praŋ*, bdil*, bdɛl*
shop (n.)	kdɛj*
short	cnhə̃t
shoulder	klapəh
shovel (v.)	sɛj
siamang (*Hylobates syndactylus*)	ʔamɛŋ
sibling, older	pɛ̃ʔ
sibling, younger	wakil, bɛr
side	ʔnaŋ, sir
side (body part)	paw
silence	sŋɔr
silent	sŋɔr
silver-eared mesia (*Leiothrix argentauris*)	smsɨ̃m
sing	pɲloɲ, cʔaj
sink (v.)	hrkbɛk ~hrkbāk, karɛm*, higɨr
sister of parent, older	toj bɨʔ
sister of parent, younger	mɔʔ
sister-in-law	lamij
sit	kikkɔk, ŋɔk, rɨn, kilwal
sit (of bird)	pəm, bat
six	nɛm*
size	bnɛʔ
sizzle	cɨ̃m
skin (n.)	ktɨʔ
skin (v.)	klipɛh
skin disease, type of	gas ~ gɛ̃s
skink (*Emoia*)	manrɔŋ
skull	labəŋ
sky	ktɔʔ
sleep (v.)	tek
slender-toed gecko (*Cyrtodactylus*)	grũc
slide	sɨr
slither	bulɛt, krjuh, tuŋtɛ̃ŋ, gimәl
slow	hakɨj

slow lori (*Nycticebus coucang*)	tampĩl
slurp (v.)	siruc
smack (v.)	pkpə̃k
small	ʔaɟɔʔ, ʔhəj
smear (v.)	plɛs
smell (n.)	ges
smell (v.)	ʔɔ̃ɲ, ɟɳɟɔŋ
smoke (n.)	ʔjʔej ʔɔs
smoke (v.)	ɟhɨt, kmɨm, kbam
smooth	bclac
snail	siput*
snake	taɟuʔ
snake, type of	papan
snake, type of	padɛj*
snake, type of	slɛʔ
snake, type of	tdoŋ
snake, type of	rloŋ
snake, type of	ranɔŋ
snake, type of	klə̃ĩl
snake, type of	ʔakər
snake, type of	brhur
snake, type of	slantɔr
snakehead (*Channa micropeltes*)	tomɛn*
snap (v.)	hɛ̃k
sneak (v.)	dɔp
sneeze (v.)	brsin*
snot	ctoh
snout (of boar)	tanus
snout (of tapir)	somuʔ
snuffle (v.)	hĩs
soap (n.)	sabon*
sob (v.)	ɲɪsɛ̃t, sut
son-in-law	mɲsaw
soon	ʔnuj
soul	ɟreŋ, rwaj
sound	kliŋ
sound (animal fleeing from one tree to another)	prəw
sound (blowpipe dart hitting a vine)	cɛ̃h
sound (blowpipe dart hitting canopy)	rɔɸ
sound (blowpipe dart hitting muscle of prey)	cik
sound (blowpipe dart hitting stomach of prey)	koɸ
sound (blowpipe)	plɔ̃ɸ
sound (blowpipe)	plɛ̃s
sound (blowpipe)	wɛ̃h
sound (boiling)	ʔɔk

sound (carving)	krāŋ
sound (dashing)	luɸ ~ ruɸ
sound (falling)	klak
sound (flapping)	riɸ
sound (flying or leaping)	jaɸ
sound (hornbill)	kɔk
sound (leaf-monkey)	hɔkɔ̄k
sound (muntjac deer)	rop
sound (raindrop)	tik
sound (running water or waterfall)	chok
sound (shotgun)	pəw
sound (small object falling to the ground)	kɔɸ
sound (walking)	rop, klik
sour	masɛ̃m*
sow (n.)	biʔ masəj
speak	pŋsɛŋ
spear (n.)	mataʔ
spear (v.)	ckbɛk
spectacle	cŋcĩŋ
spider, type of	tawiɲ
spider, type of	tŋtūŋ
spider, type of	knabil
spider, type of	sraŋkɛj
spit (v.)	kbɛc, tuɸ, ktɔɸ, paɲcɛr*
split (v.)	stap, pɛk, wis, was, salɔw
spray (v.)	smur*
spread (v.)	hap ~ hiphap, dɔr
spring (v.)	tikah
sprinkle (v.)	psɛl, pruj
squat (v.)	ɟritɛw
squeeze (v.)	cpĩt, lam
squirrel	kdek
stab (v.)	cek, rik
stain (v.)	hwɛ̃t
stand (v.)	hŋɟaɲ
star	bintɛŋ*, prlɔj
stare (v.)	pdɛp, ttlət
steal	takuʔ ~ ʔakuʔ, maliŋ*
steam (n.)	ʔjʔej tɔm
steam (v., intr.)	has
stench	haʔɛ̃t
step over	laŋkah*
stick (n.)	ʔat
sting (v.)	kɰ̃t, hāt, sic, cɔ̄ʔ, lantom
stomach	ʔɛc

stone	batuʔ*, balәl
stop (v.)	ʔop, ɟɨn
straight	tros*
stranger	gop
stream-bed	carәk*
stretch (v. intr.)	ʔtʔɛt
stretch (v. tr.)	cŋciŋ
striped tit-babbler (*Macronous gularis*)	hor
stroke (v.)	ʔtʔɔ̃t
strong	kwat*
stumble	tsdes
such	kɛj
suck (v.)	ɟhit, pot, sksɔ̃k
sugar	gulәh*
summon	pihdeh
sun	kit ktɔʔ, mit ktɔʔ
sunbird (*Nectarinia*; *Anthreptes*)	sɛtsɛt
surprised	hiran
swallow (v.)	klɔ̃t, lik
sweat (n.)	bŋkɨt
sweep (v.)	kwɛ̃s, pɨs, sapuh*, kwɛr*
sweet	bhɛt, bʔɨt
swell (v.)	kmaŋ*
swidden	slaj
swim (v.)	kwac, sjәr
tail	haʃiʔ
tail feathers	cnɛɸ
take (a hit)	dan
take (back)	pimɨc
take (v.)	kɔt, ɟɨŋ, ʔaŋket*, jɔw
talk (v.)	ckwɨk, bacaʔ*, caraʔ
tap (poison) (v.)	turɛh, pɔŋ
tapir (*Tapirus indicus*)	barɛt,barɔŋ
taste (n.)	lasaŋ
taste (v.)	hɔ̃t, cubaʔ*, kmjɨ̃m
tasteless	blhɨt
tattoo	cɨl
tea	tɛh*
teach	pipjɛp
tear (n.)	tɔm mit
tear (v.)	cok, csɛɛs, tgɨɲ
tell	cɔl
tell a lie	bɔc, bɛ̃r
Temiar (ethnonym)	plɛh, tmɛr
temple (body part)	cunʃiŋ

Temuan (ethnonym)	tmwɛn*
ten	spuloh*
termite	dariŋ
termite mound	pusu?*
terrapin, type of	poc
terrapin, type of	karwɔ?
testicles	kbi? ?ntɛp
Thailand	sjam
that (away)	tani?
that (beyond me)	tadeh
that (beyond you)	tɲi?
that (down)	tujih
that (up)	titih
that (you don't know)	tũn
that (you know)	ton
there (away)	?ani?
there (beyond me)	?adeh
there (beyond you)	?ɲi?
there (down)	?ujih
there (up)	?itih
there (you don't know)	?ũn
there (you know)	?on
thick	tabɛl
thick-billed pigeon (*Treron curvirostra*)	brawɔl
thin	lipis*, kurus*
thing	baraŋ*
think	bliŋ, pikir*
thirsty	dahaga?*
this	tə̃h
thorn	ɟle?
thorn, type of	?ɛ̃m kaji?
thousand	sribuh*
thread (n.)	bnɛŋ*
three	tiga?*
throat	?ŋut
throw (v.)	hakɔk, hɔk, paŋka?*, ɟan, bdal*, jol
thrush (*Zoothera*)	pahə̃ŋ
thumb	tabo?
thunder	tij
thunder-spirit	karɛj
tickle (v.)	ŋɛc, tɔc
tie (v.)	rbɛt, rət, ben, kəl
tiger (*Panthera tigris*)	?ap, ɟu?ɔk, baliŋ, putew
tiger shrike (*Lanius tigrinus*)	pɔr
time	prjam

tired	ghɛl, hrtlɛt
toad, type of	brŋɔ̃k
tobacco	tmakɔw*
today	ktɔʔ tɔ̃h
toe nail	katɔ̃ŋ can
together	skaliʔ*, gus
tomorrow	haden
tongue	lntek
tooth	hɛ̃ɲ
tortoise, type of	ʔawɛ̃ʔ
tortoise, type of	kɔh
tortoise, type of	barʔeŋ
tortoise, type of	sil
tortoise, type of	sorɛj
touch (v.)	pitɨl
tough (of meat)	kras*
track (n.)	tel
training	trinɛŋ*
trap (n.)	bakɔʔ
tread (v.)	lajɛ̃k
tree	ɟhũʔ
tree, type of	ʔɲcɛk
tree, type of	taduk
tree, type of	lumpɔk
tree, type of	gasɛʔ
tree, type of	kmpɛ̃s
tree, type of	tawɛs
tree, type of	prgɨs
tree, type of	tlas
tree, type of	ɟnoh
tree, type of	prɔh
tree, type of	bnim
tree, type of	krnum
tree, type of	gɛn
tree, type of	panʔon
tree, type of	boɲ
tree, type of	trpɔɲ
tree, type of	ʔanɛŋ
tree, type of	bramaŋ
tree, type of	knalaŋ
tree, type of	katuŋ
tree, type of	kʔuŋ
tree, type of	lajoŋ
tree, type of	balɔŋ
tree, type of	raɲsil

tree, type of	brol
tree, type of	ʔampɛr
tree, type of	jadɛr
tree, type of	sagɨr
tree, type of	cicar
tree, type of	bsɔr
tree, type of	bakɔw
tree, type of	ɟrwɛj
tree, type of	kɟaj
tree, type of	cah
tree, type of (*Shorea*)	sraja?*
tree-base	tom
tree branch	joh, joh waŋkɛ̃?
treeshrew (*Tupia*)	tacɔ̃r
tremble (v.)	hit
tributary	cabaŋ* tɔm
true	btol*
trumpet (of elephant) (v.)	krʔoŋ
trunk (of elephant)	bralɛj ~ blalɛj*
tualang tree (*Koompassia excelsa*)	gɨl
tuber	hobi?*
tuber, type of	kɛj
tuber, type of	klep
tuber, type of	hakɔ̃t
tuber, type of	tahɛ?
tuber, type of	kba?
tuber, type of	hrja?
tuber, type of	kabu?
tuber, type of	daran
tuber, type of	sɨh
tuber, type of	tolɛŋ
tuber, type of	cʔɨŋ
tuber, type of	jɔ̃ŋ
tuber, type of	həw
tuber, type of	manraj
tuber, type of	kwɔj
turn (v. tr. and intr.)	wikwek, brilɛ?, gaj
turn head (of bird)	cliŋhĩŋ, cliŋpaŋ
turtle, type of	labi?*
turtle, type of	kuhʔɔ̃h
turtle, type of	pjɔ̃ŋ
turtle, type of	kpil
tusk	gadiŋ*
twenty	duwa? puloh*
twinkle (v.)	piplɛ̃p

twins	kmar*
two	duwaʔ*
umbrella	pajuŋ*
underside	kjɔm
understand	paham*
undress	jɔk
untie	bə̄k
until	baruʔ*
upper side	krpiŋ
urinary bladder	knɔm
urinate	knɔm
urine	knɔm
valley	tpis, lgim
vegetable	tʔaʔ
village	kampoŋ*
vine	ʔawɛj
viper, type of (*Trimeresurus*)	blaʔir
vomit (n.)	sdɨj
vomit (v.)	bhɔk, kɨʔ, rah
vulva	dɔt
wade (across)	cɨk
wag (v.)	kipas*
waist (body part)	gɛl
wait (v.)	tadɔʔ
wake (up) (v. intr.)	muŋkɛr
wake (v. tr.)	pihpih
walk (v.)	rop, gagah, phɛr
wallet	dibɛh
wane (v.)	blah*
want (v.)	ʔɔm
wash (v.)	soc
wash away	tpis
wasp, type of	ruh
wasp, type of	ʔiŋ
wasp, type of	rjaw
water	tɔm, ʔə̄ŋ
water strider (*Gerridae*)	samir
waterfall	lataʔ*
wave (v.)	lambaj*
wax (v.)	wɛ̃l
waylay	gat
weep (v.)	ɟim
west	barat*, sɛɲ ktɔʔ
wet	pcɛ̃ʔ
what?	mɛj

when?	mapuʔ
where?	lbah
whip snake (*Ahaetulla*)	bʔəw
whirl (v.)	hilaŋ*, gulɛj
whiskers	ɟɛŋ
whistle (v.)	hchəc, hĩɸ ~ hẽɸ, həh
white	puteh*
white-rumped shama (*Copsychus malabaricus*)	kwal
whiz (v.)	ɟlɔɸ
who?	makɛn
whose?	makɛn
wide	lwɛs, ɟŋɟɛŋ
wife	kneh
wild boar (*Sus scrofa*)	napak, clapak, hmalah, hawɛ̃n, gaw
wild dog (*Cuon alpinus*)	clɔŋ
win (v.)	mnɛŋ*
wind	brwaʔ, bgiw
wind (v.)	mlməl
window	tiŋkap*
wing	sajap*
wipe (v.)	ɟɨt
with	samɛʔ*
woman	baboʔ
woman, old	kbet, ɟaɟiʔ ~ ɟaɟaʔ
wood (material)	tagɨɲ, gnoŋ
wood (rotten)	rbaʔ
work (v.)	krɟaʔ*, kacɔw*, gaˈwɛj*
worm	taciŋ*
wrap (v.)	bɛn
wreathed hornbill (*Rhyticeros undulatus*)	saŋkoh
wrinkle (n.)	krtwĩt, krntɛn
wrist	krĩl
wrong	lɛc
yawn (v.)	hjhəj
year	tawon*
yell (v.)	krkɛr
yellow	kuniŋ*
yellow-vented bulbul (*Pycnonotus goiavier*)	brubɔh
yes	haʔɨʔ, haʔĩh
yet	lagiʔ*

References

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

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

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

—— 1976, The verb in Kentakbong. In Jenner, Thompson and Starosta, eds 1976, part 2, 951–970.

—— 1985, Language and the world-view of the Malay peasants. In Mohd. Taib Osman, ed. *Malaysian world-view*, 184–252. Singapore: Institute of Southeast Asian Studies.

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

—— 1992, Mon-Aslian contacts. *Bulletin of the School of Oriental and African Studies* 55:532–537.

Benjamin, Geoffrey, 1976a, Austroasiatic subgroupings and prehistory in the Malay Peninsula. In Jenner, Thompson and Starosta, eds 1976, part 1, 37–128.

—— 1976b, An outline of Temiar grammar. In Jenner, Thompson and Starosta, eds 1976, part 1, 129–188.

—— 1985a, In the long term: three themes in Malayan cultural ecology. In Karl L. Hutterer, A. Terry Rambo and George Lovelace, eds *Cultural values and human ecology in Southeast Asia*, 219–278. University of Michigan: Center for South and Southeast Asian Studies.

—— 1985b, On pronouncing and writing Orang Asli languages: a guide for the perplexed, part 1. *Orang Asli studies newsletter* 4:4–16. Hanover, NH: Department of Anthropology.

—— 1986, On pronouncing and writing Orang Asli languages: a guide for the perplexed, part 2. *Orang Asli studies newsletter* 5:3–28. Hanover, NH: Department of Anthropology.

—— 1987, 'Ethnohistorical perspectives on Kelantan's prehistory'. In Nik Hassan Shuhaimi bin Nik Abdul Rahman, ed. *Kelantan zaman awal: kajian arkeologi dan sejarah di Malaysia*, 108–153. Kota Bharu: Perpaduan Muzium Negeri Kelantan.

—— 1989, Achievements and gaps in Orang Asli research. *Akademika* 35:7–46. Universiti Kebangsaan Malaysia.

—— 1993, Grammar and polity: the cultural and political background to Standard Malay. In William A. Foley, ed. *The role of theory in language description*, 341–392. Berlin/New York: Mouton de Gruyter.

—— 1996, The anthropology of grammar: self and other in Temiar. Unpublished typescript, 95pp.

—— 2001, Orang Asli languages: from heritage to death?. In Razha Rashid and Wazir Jahan Karim, eds *Minority cultures of Peninsular Malaysia: survivals of indigenous heritage*, 101–122. Penang: Malaysian Academy of Social Sciences (AKASS).

—— 2004, Aslian languages: characteristics and usage. In Asmah Haji Omar, ed. *The encyclopedia of Malaysia,* Volume 12: *Languages and literatures*, 46–49. Kuala Lumpur: Archipelago Press.

—— In press, The current situation of the Aslian languages. To appear in Hein Steinhauer and James T. Collins, eds *Endangered languages of Southeast Asia.* London: School of Oriental and African Studies (in the series *Papers in Language Documentation and Description*).

Bisang, Walter, 1999, Classifiers in East and Southeast Asian languages: counting and beyond. In Jadranka Gvozdanović, ed. *Numeral types and changes worldwide*, 113–185. Berlin/New York: Mouton de Gruyter.

Bishop, Nancy M., 1992, A preliminary analysis of Maniq oral monophthongs. *Journal of Language and Linguistics*, vol. 10, No. 2.

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

—— 1996b, Who's who in Kensiw? Terms of reference and address in Kensiw. *Mon-Khmer Studies* 26:245–253.

—— 2001, Sociolinguistic aspects of Northern Aslian languages: a case study of Kensiw (Kensiu). In Razha Rashid and Wazir Jahan Karim, eds *Minority cultures of Peninsular Malaysia: survivals of indigenous heritage*, 83–99. Penang: Malaysian Academy of Social Sciences (AKASS).

Bishop, Nancy M. and Mary M. Peterson, 1993, *Maniq language survey report.* Bangkok: Thammasat University.

—— 1994, Kensiw glossary. *Mon-Khmer Studies* 23:163–195.

—— 1995, A selective Aslian bibliography. *Mon-Khmer Studies* 24:161–169.

Blagden, Charles Otto, 1902, A Malayan element in some of the languages of southern Indo-China. *Journal of the Straits Branch, Royal Asiatic Society* 38:1–27.

—— 1906a, Language. In Walter William Skeat and Charles Otto Blagden *Pagan races of the Malay Peninsula*, vol. 2, 379–472. London: Macmillan and Co.

—— 1906b, Comparative vocabulary of Aboriginal dialects. In Skeat and Blagden, 1906 vol. 2, 481–775.

Bloch, Jules, 1952, L'Andaman. In Marcel Cohen, ed. *Les langues du monde*, 512–521. Paris: Société linguistique de Paris.

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

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

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

—— 2000, Unitizer and nominalizer: the /n/ affix in Jahai. Paper presented at the Fifth International Symposium on Languages and Linguistics: Pan-Asiatic Linguistics, Ho Chi Minh City, Vietnam, November 16–17, 2000.

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

—— 2001b, Loanword phonology in Jahai. *Working Papers* 48:5–14. Lund University: Department of Linguistics.

—— 2002, A grammar of Jahai. Unpublished doctoral dissertation, Department of Linguistics, Lund University.

—— 2003, Attention, accessibility, and the addressee: the case of the Jahai demonstrative *ton*. *Pragmatics*, 13(3/4):363–379.

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

—— 2004b, Landscape terms and toponyms in Jahai: a field report. *Working Papers* 51, 17–29. Lund Univeristy: Department of Linguistics.

Bybee, Joan, 1985, *Morphology: a study of the relation between meaning and form*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bybee, Joan, Revere Perkins and William Pagliuca, 1994, *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Chicago/London: The University of Chicago Press.

Carey, Iskander, 1961, *Tengleq kui seroq: a study of the Temiar language with an ethnographical summary*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

—— 1970, A Mendrik vocabulary. *Federation Museums Journal New Series* 15:185–191.

Clements, George N. and Samuel J. Keyser, 1983, *CV-phonology*. Cambridge, Mass.: MIT Press.

Clifford, Hugh, 1891, Some notes on the Sakai dialects of the Malay Peninsula. *Journal of the Straits Branch, Royal Asiatic Society* 24:13–29.

Collings, H.D., 1949, 'A Temoq word list and notes' and 'Aboriginal notes'. *Bulletin of the Raffles Museum*, Ser. B, 4:69–94.

Comrie, Bernard, 1976, *Aspect: an introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Comrie, Bernard and Sandra A. Thompson, 1985, Lexical nominalization. In Timothy Shopen, ed. *Language typology and syntactic description: grammatical categories and the lexicon*, 349–398. Cambridge: Cambridge University Press.

Coope, A.E., 1993, *Malay–English English–Malay dictionary*. New York: Hippocrene Books, Inc. (paperback edition)

Cox, Merel J., Peter Paul van Dijk, Jarujin Nabhitabhata and Kumthorn Thirakupt, 1998, *A photographic guide to snakes and other reptiles of Peninsular Malaysia, Singapore and Thailand*. London: New Holland.

Crawfurd, John, 1820, *History of the Indian Archipelago*. 3 vols. Edinburgh.

Croft, William, 1994, Semantic universals in classifier systems. *Word* 45(2):145–171.

Dahl, Östen, 1985, *Tense and aspect systems*. Oxford: Blackwell.

Daly, D.D., 1880, Vocabulary of Perak and Kenering 'Semang'. *Journal of the Straits Branch of the Royal Asiatic Society*, No. 5:125–156.

Diffloth, Gérard, 1968, Proto-Semai phonology. *Federation Museums Journal* 13:65–73.

—— 1972, Ambiguïté morphologique en semai. In J.M.C. Thomas and L. Bernot, eds *Langues et techniques, nature et société, 1: Approche linguistique*, 91–93. Paris: Klincksieck.

—— 1974, Body-moves in Semai and in French. *Papers from the tenth regional meeting*, 128–138. Chicago Linguistic Society.

—— 1975, Les langues mon-khmer de Malaisie: classification historique et innovations. *Asie du sud-est et monde insulinde* 6 (4):1–19.

—— 1976a, Minor-syllable vocalism in Senoic languages. In Jenner, Thompson and Starosta, eds 1976, part 1:229–247.

—— 1976b, Jah-Hut: an Austroasiatic language of Malaysia. In Nguyen Dang Liem, ed. *South-east Asian linguistic studies* vol. 2:73–118. Canberra: Pacific Linguistics.

—— 1976c, Mon-Khmer numerals in Aslian languages. *Linguistics* 174:31–37.

—— 1976d, Expressives in Semai. In Jenner, Thompson and Starosta, eds 1976, part 1:249–264.

—— 1977, Towards a history of Mon-Khmer: Proto-Semai vowels. *Southeast Asian studies* 14:463–495.

—— 1979, Aslian languages and Southeast Asian prehistory. *Federation Museums journal (new series)* 24:2–16.

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

Egerod, Søren, 1965, An English–Atayal vocabulary. *Acta Orientalia* 29.

Endicott, Kirk M., 1990, Batek history and dialect group dynamics. Unpublished paper presented at the 89th Annual Meeting of the American Anthropological Association.

Evans, Ivor H.N, 1915, Some Semang vocabularies obtained in Pahang and Perak. *Journal of the Federated Malay States Museums* 6:115–125.

—— 1927, *Papers on the ethnology and archaeology of the Malay Peninsula*. Cambridge: Cambridge University Press.

Gafos, Adamantios I., 1998, Correspondence in Temiar: no need for long-distance spreading here. In Shobhana L. Chellian and Willem J. de Reuse, eds *Papers from the Fifth Annual Meeting of the Southeast Asian Linguistics Society*, 115–131. Tempe: Arizona State University.

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

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

Hewett, R.D., 1880, Vocabulary of Ulu Selama Semang. *Journal of the Straits Branch of the Royal Asiatic Society*, No. 5:125–156.

Hopper, Paul J., 1986, Some discourse functions of classifiers in Malay. In Colette Craig, ed. *Noun classes and categorization*, 309–325. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Huffman, Franklin E., 1986, *Bibliography and index of Mainland Southeast Asian languages and linguistics.* New Haven/London: Yale University Press.

Jenner, P.N., L.C. Thompson and S. Starosta, eds, 1976, *Austroasiatic studies.* 2 vols Honolulu: The University Press of Hawaii.

Klavans, Judith L., 1982, *Some problems in a theory of clitics.* Bloomington: Indiana University Linguistic Club.

—— 1985, The independence of syntax and phonology in cliticization. *Language* 61:95–120.

Kruspe, Nicole D., 1999, Semelai. Unpublished PhD thesis. University of Melbourne.

—— 2004, *A grammar of Semelai.* Cambridge: Cambridge University Press.

Ladefoged, Peter and Ian Maddieson, 1996, *The sounds of the world's languages.* Oxford: Blackwell Publishers.

Lindell, Kristina, 1974, A vocabulary of the Yuan dialect of the Kammu language. *Acta Orientalia* 36:191–207.

Lucy, John A., 1992, *Grammatical categories and cognition: a case study of the linguistic relativity hypothesis.* Cambridge: Cambridge University Press.

Lye Tuck-Po, ed., 2001, *Orang Asli of Peninsular Malaysia: a comprehensive and annotated bibliography.* CSEAS Research Report Series No. 88. Kyoto: Center for Southeast Asian Studies, Kyoto University.

McCarthy, John J., 1982, Prosodic templates, morphemic templates, and morphemic tiers. In H. van der Hulst and N. Smith, eds *The structure of phonological representations* I. Dordrecht: Foris.

Marantz, Alec, 1982, Re reduplication. *Linguistic Inquiry* 13:435–482.

Matisoff, James A., 1973, Tonogenesis in Southeast Asia. In Larry M. Hyman, ed. *Consonant types and tone.* Southern California Occasional Papers in Linguistics, No. 1:71–95. Los Angeles: University of Southern California.

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

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

Nik Safiah Karim and Ton Bte Ibrahim, 1979, Semoq Beri: some preliminary remarks. *Journal of the Malaysian Branch, the Royal Asiatic Society* 47:123–129.

Ogilvie, Charles, 1949, Che Wông word list and notes. *Bulletin of the Raffles Museum Series B* 4:11–43.

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

Peterson, Mary M., 1993, Lost in space: discovering spatial locatives in Kensiw. Paper presented at the Faculty of Liberal Arts' 30th Anniversary Seminar. Bangkok: Thammasat University.

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

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

Reichling, A.J.B.N., E.M. Uhlenbeck and W. Sidney Allen, 1977, Questionnaire. *Lingua descriptive studies*, 11–71. Amsterdam: North-Holland Publishing Company.

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

Savage, H.E., 1926, A Pangan vocabulary. *Journal of the Malayan Branch, Royal Asiatic Society*, 147–153.

Schebesta, Paul, 1928a, Grammatical sketch of the Jahai dialect, spoken by a Negrito tribe of Ulu Perak and Ulu Kelantan, Malay Peninsula. (Translated and re-worked by C.O. Blagden). *Bulletin of the School of Oriental Studies* 1928:803–826.

—— 1928b, *Among the forest dwarfs of Malaya*. Kuala Lumpur: Oxford University Press.

—— 1931, Grammatical sketch of the Ple-Temer language. (Translated and re-worked by C.O. Blagden). *Journal of the Royal Asiatic Society* 1931:641–652.

—— 1952, *Die Negrito Asiens: Geschichte, Geographie, Umwelt, Demographie und Antropologie der Negrito*. Vienna: St. Gabriel Verlag, Studia Instituti Antropos.

—— 1954, *Ethnographie der Negrito, I. Wirtschaft und Soziologie*. Vienna: St. Gabriel Verlag, Studia Instituti Antropos.

—— 1957, *Ethnographie der Negrito, III. Religion und Mythologie*. Vienna: St. Gabriel Verlag, Studia Instituti Antropos.

Schmidt, Wilhelm, 1901, Die Sprachen der Sakei und Semang auf Malacca und ihr Verhältnis zu den Mon-Khmer-Sprachen. *Bijdragen tot de Taal-, Land- en Volkenkunde* 8:401–583.

—— 1903, The Sakai and Semang languages in the Malay Peninsula and their relation to the Mon-Khmer languages. *Journal of the Straits Branch of the Royal Asiatic Society* 39:38–45.

Shorto, Harry L., Judith M. Jacob and E.H.S. Simmonds, 1963, *Bibliographies of Mon-Khmer and Tai linguistics*. London: Oxford University Press.

Skeat, Walter William and Charles Otto Blagden, 1906, *Pagan races of the Malay Peninsula*, 2 vols. London: Macmillan and Co. (Reprint: 1966, London: Frank Cass)

Spencer, Andrew, 1991, *Morphological theory: an introduction to word structure in generative grammar*. Oxford, UK/Cambridge, US: Blackwell.

Strange, Morten, 1998, *Tropical birds of Malaysia and Singapore*. Hong Kong: Periplus Editions.

Svantesson, Jan-Olof, 1983, *Kammu phonology and morphology*. Lund: CWK Gleerup.

Swettenham, F.A., 1880, Vocabulary of Ijok Semang. *Journal of the Straits Branch of the Royal Asiatic Society*, No. 5:125–156.

Ter Mors, Christine, 1983, Affix to X. Broselow & McCarthy versus Marantz: a reinterpretation. *The Linguistic Review* 3:275–298.

Trombetti, Alfredo, 1923, *Elementi di glottologia.* Volume pubblicato a spese della cassa di risparmio in Bologna. Bologna.

van der Sluys, Cornelia M.I., 1996, Why the Jahais view sharing as a 'passing-on of gifts'. Revision of a paper presented at the 7th International Conference on Hunting and Gathering Societies (CHAGS 7), Moscow, Russia, August 17–23, 1993.

—— 1999, The Jahai of northern peninsular Malaysia. In Richard B. Lee and Richard Daly, eds *The Cambridge encyclopaedia of hunter-gatherers*, 307–311. Cambridge: Cambridge University Press.

—— 2000, Gifts from the immortal ancestors: cosmology and ideology of Jahai sharing. In Peter P. Schweitzer, Megan Biesele and Robert K. Hitchcock, eds *Hunters and gatherers in the modern world: conflict, resistance and self-determination*, 427–454. New York: Berghahn.

Van Valin, Robert D. Jr and Randy J. LaPolla, 1997, *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.

Wazir Jahan Karim, 1996, Minorities of the minority: language death and patterns of cultural extinction. *Material culture and technological innovation*, 6–16. Toyota Foundation Report.

—— 2001, Minorities of the minority: language death and patterns of cultural extinction. In Razha Rashid and Wazir Jahan Karim, eds *Minority cultures of Peninsular Malaysia: survivals of indigenous heritage*, 69–74. Penang: Malaysian Academy of Social Sciences (AKASS).

Whitten, Jane, 1998, *Tropical wildlife of Malaysia and Southeast Asia.* Hong Kong: Periplus Editions.

Wilkinson, R.J., 1915, *A vocabulary of Central Sakai. Dialect of the aboriginal communities in the Gopeng Valley.* Papers on Malay Subjects, Second series No. 3. Kuala Lumpur: Government Press.

Zaharani b. Ahmad, 1991, *The phonology and morphology of the Perak dialect*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Zide, Norman and Vishvajit Pandya, 1989, A bibliographical introduction to Andamanese linguistics. *Journal of the American Oriental Society*, 109(4):639–651.