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.

FOUNDING EDITOR: Stephen A. Wurm
EDITORIAL BOARD: John Bowden, Malcolm Ross and Darrell Tryon (Managing Editors), I Wayan Arka, David Nash, Andrew Pawley, Paul Sidwell, Jane Simpson

EDITORIAL ADVISORY BOARD:
Karen Adams, Arizona State University
Peter Austin, School of Oriental and African Studies
Alexander Adelaar, University of Melbourne
Byron Bender, University of Hawai‘i
Walter Bisang, Johannes Gutenberg-Universität Mainz
Robert Blust, University of Hawai‘i
Lyle Campbell, Canterbury University
James Collins, Universiti Kebangsaan Malaysia
Bernard Comrie, Max Planck Institute for Evolutionary Anthropology
Soenjono Dardjowidjojo, Universitas Atma Jaya
Matthew Dryer, State University of New York at Buffalo
Jerold A. Edmondson, University of Texas at Arlington
Nicholas Evans, University of Melbourne
Margaret Florey, Monash University
William Foley, University of Sydney
Karl Franklin, Summer Institute of Linguistics
Charles Grimes, Universitas Kristen Artha Wacana Kupang
Nikolaus Himmelmann, Ruhr-Universität Bochum
Lillian Huang, National Taiwan Normal University
Bambang Kaswanti Purwo, Universitas Atma Jaya
Marian Klamer, Universiteit Leiden
Harold Koch, Australian National University
Frantisek Lichtenberk, University of Auckland
Patrick McConvell, Australian Institute of Aboriginal and Torres Strait Islander Studies
William McGregor, Aarhus Universitet
Ulrike Mosel, Christian-Albrechts-Universität zu Kiel
Claire Moyse-Faurie, Centre National de la Recherche Scientifique
Bernd Nothofer, Johann Wolfgang Goethe-Universität Frankfurt am Main
Ger Reesink, Universiteit Leiden
Lawrence Reid, University of Hawai‘i
Jean-Claude Rivierre, Centre National de la Recherche Scientifique
Melenaite Taumoefolau, University of Auckland
Tasaku Tsunoda, University of Tokyo
John Wolff, Cornell University
Elizabeth Zeitoun, Academia Sinica
A grammar of Jahai

Niclas Burenhult

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

## Abbreviations

## Preface and acknowledgements

### Chapter 1: Introducing Jahai

1. The Jahai
2. Linguistic classification and history
3. The sociolinguistic situation
   1.3.1 Idiolects, dialects and multilingualism
   1.3.2 Schooling and literacy
   1.3.3 Endangerment
4. Previous research
   1.4.1 Research on Jahai
      1.4.1.1 Schebesta's grammatical sketch
   1.4.2 Other research related to Aslian
      1.4.2.1 General
      1.4.2.2 Northern Aslian
      1.4.2.3 Central Aslian
      1.4.2.4 Southern Aslian
5. The present study
   1.5.1 Background and aims
   1.5.2 Fieldwork, methodology and data
   1.5.3 Outline
   1.5.4 Transcription, glossing and translation

### Chapter 2: Phonology

2. Preliminaries
2.2 Vowels
   2.2.1 Phonemic vowel nasality
   2.2.2 Phonetic description of vowel phonemes
   2.2.3 Environmentally conditioned nasalisation of vowels
      2.2.3.1 Directionality of nasalisation
      2.2.3.2 Variability in degree of nasalisation
2.3 Consonants
   2.3.1 Description of consonant phonemes and allophonic variation
      2.3.1.1 Stops
      2.3.1.2 Nasals
      2.3.1.3 Fricatives
      2.3.1.4 Liquids
      2.3.1.5 Approximants
2.4 Phonotactics
   2.4.1 Syllable structure and types
      2.4.1.1 Syllable structure
      2.4.1.2 Syllable types
   2.4.2 Word structure
      2.4.2.1 Monosyllabic words
      2.4.2.2 Sesquisyllabic words
      2.4.2.3 Disyllabic words
      2.4.2.4 Trisyllabic words
   2.4.3 Distribution of phonemes
      2.4.3.1 Final syllables
      2.4.3.2 Pre-final syllables
         2.4.3.2.1 The penultimate syllable in sesquisyllabic words
         2.4.3.2.2 The penultimate syllable in disyllabic words
         2.4.3.2.3 Pre-final syllables in trisyllabic words
   2.4.4 Reduplicative processes
      2.4.4.1 Phonetic realisation of coda copies
      2.4.4.2 Phonetic realisation of epenthetic vowels
   2.4.5 Syllabification
2.5 Prosodic features
   2.5.1 Stress
   2.5.2 Tone
2.6 The phonology of loanwords
   2.6.1 Phonetic adaptation
      2.6.1.1 Prestopping of word-final nasals
      2.6.1.2 Nasalisation of word-initial /b/
      2.6.1.3 Palatalisation of /s/
   2.6.2 Phonemic replacement
      2.6.2.1 Substitution of final syllable /a/
      2.6.2.2 Word-final glottal stop
      2.6.2.3 Voiceless labiodental fricative
   2.6.3 Reorganisation of syllabic structure
      2.6.3.1 Closure of syllables
      2.6.3.2 Reinterpretation of word-medial vowel sequences
      2.6.3.3 Deletion of voiced stops
2.6.3.4 Reinterpretation of disyllabic words as sesquisyllables 43
2.6.4 Relocation of stress 43
2.7 Summary 44

Chapter 3: Word formation 45
3.1 Morphological units 45
  3.1.1 Roots 45
  3.1.2 Lexemes 46
  3.1.3 Bases 46
  3.1.4 Affixes 46
  3.1.5 Clitics 46
3.2 Affixation 46
  3.2.1 Inner affixation 47
    3.2.1.1 Underspecified affixation 49
      3.2.1.1.1 Coda copy 50
      3.2.1.1.2 Onset copy + coda copy 52
    3.2.1.2 Partly prespecified affixation 53
      3.2.1.2.1 Prespecified consonants 53
      3.2.1.2.2 Prespecified vowels 55
    3.2.1.3 Prespecified affixation 58
      3.2.1.3.1 Prespecified infixation 58
      3.2.1.3.2 Prespecified prefixation 61
  3.2.2 Outer affixation 62
    3.2.2.1 Antepenultimate outer affixes 63
    3.2.2.2 Flexible outer affixes 63
  3.2.3 The origins of inner and outer affixation 64
  3.2.4 Total reduplication 64
3.3 Cliticisation 65
  3.3.1 Phonological and phonotactic characteristics 65
  3.3.2 Domain and location of attachment 67
    3.3.2.1 Clitics hosted by verbs 67
    3.3.2.2 Clitics hosted by the first constituent of the NP 68
    3.3.2.3 Clitics hosted by the first constituent of the core 68
    3.3.2.4 Clitics hosted by the first constituent of the clause 69
    3.3.2.5 Clitics hosted by any constituent 69
3.4 Summary 70

Chapter 4: Word classes 71
4.1 Nouns 71
  4.1.1 Collective plural 74
  4.1.2 Diverse plural 75
  4.1.3 Unitisation 75
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.4</td>
<td>Verb-to-noun derivation</td>
<td>77</td>
</tr>
<tr>
<td>4.1.4.1</td>
<td>Nominalising /ra/</td>
<td>77</td>
</tr>
<tr>
<td>4.1.4.2</td>
<td>Collective plural &lt;ra&gt; and &lt;a&gt;</td>
<td>79</td>
</tr>
<tr>
<td>4.1.4.3</td>
<td>Nominalising /ma/</td>
<td>79</td>
</tr>
<tr>
<td>4.1.5</td>
<td>Numeral-to-noun derivation</td>
<td>79</td>
</tr>
<tr>
<td>4.1.5.1</td>
<td>Nominalising /na/</td>
<td>79</td>
</tr>
<tr>
<td>4.1.5.2</td>
<td>Collectivising coda copy</td>
<td>80</td>
</tr>
<tr>
<td>4.2</td>
<td>Classifiers</td>
<td>80</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Unitisation</td>
<td>82</td>
</tr>
<tr>
<td>4.3</td>
<td>Personal pronouns</td>
<td>83</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Pronominal distinctions</td>
<td>83</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Allomorphs of personal pronouns</td>
<td>83</td>
</tr>
<tr>
<td>4.4</td>
<td>Demonstratives</td>
<td>84</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Basic demonstratives</td>
<td>84</td>
</tr>
<tr>
<td>4.4.1.1</td>
<td>Speaker-anchored accessible /?ah/</td>
<td>85</td>
</tr>
<tr>
<td>4.4.1.2</td>
<td>Addressee-anchored accessible /?on/</td>
<td>85</td>
</tr>
<tr>
<td>4.4.1.3</td>
<td>Speaker-anchored inaccessible /?ani?/</td>
<td>85</td>
</tr>
<tr>
<td>4.4.1.4</td>
<td>Addressee-anchored inaccessible /?ün/</td>
<td>85</td>
</tr>
<tr>
<td>4.4.1.5</td>
<td>Speaker-anchored exterior /?adeh/</td>
<td>86</td>
</tr>
<tr>
<td>4.4.1.6</td>
<td>Addressee-anchored exterior /?ni?/</td>
<td>86</td>
</tr>
<tr>
<td>4.4.1.7</td>
<td>Superjacent /?itih ~ ?otih/</td>
<td>86</td>
</tr>
<tr>
<td>4.4.1.8</td>
<td>Subjacent /?ujih/</td>
<td>86</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Nominal demonstratives</td>
<td>86</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Temporal demonstrative</td>
<td>97</td>
</tr>
<tr>
<td>4.5</td>
<td>Numerals and quantifiers</td>
<td>87</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Numerals</td>
<td>87</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Quantifiers</td>
<td>88</td>
</tr>
<tr>
<td>4.6</td>
<td>Interrogatives</td>
<td>89</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Person-questioning /maken/</td>
<td>89</td>
</tr>
<tr>
<td>4.6.2</td>
<td>Time-questioning /mapu?/</td>
<td>89</td>
</tr>
<tr>
<td>4.6.3</td>
<td>Item/situation/reason-questioning /mamej ~ mej/</td>
<td>89</td>
</tr>
<tr>
<td>4.6.4</td>
<td>Manner-questioning /ma?acin ~ ?acin/</td>
<td>90</td>
</tr>
<tr>
<td>4.6.5</td>
<td>Location-questioning /lbah/</td>
<td>90</td>
</tr>
<tr>
<td>4.7</td>
<td>Verbs</td>
<td>91</td>
</tr>
<tr>
<td>4.7.1</td>
<td>Aspect and Aktionsart</td>
<td>93</td>
</tr>
<tr>
<td>4.7.1.1</td>
<td>Imperfective</td>
<td>94</td>
</tr>
<tr>
<td>4.7.1.2</td>
<td>Progressive</td>
<td>96</td>
</tr>
<tr>
<td>4.7.1.3</td>
<td>Iterative</td>
<td>99</td>
</tr>
<tr>
<td>4.7.1.4</td>
<td>Continuative</td>
<td>100</td>
</tr>
<tr>
<td>4.7.1.5</td>
<td>Distributive</td>
<td>102</td>
</tr>
<tr>
<td>4.7.1.6</td>
<td>Reciprocal</td>
<td>104</td>
</tr>
</tbody>
</table>
4.7.1.7 Affix /m/  
4.7.2 Causatives  
4.7.2.1 Causative /p/-  
4.7.2.2 Causative /pi/-  
4.7.2.3 Causative /pr/-  
4.7.2.4 Causative /tr/-  
4.7.2.5 Causative /<ri>/  
4.7.3 Noun/numeral-to-verb derivation  
4.7.3.1 Noun-to-verb derivation  
4.7.3.2 Numeral-to-verb derivation  
4.7.4 Markers of modality  
4.7.4.1 Irrealis  
4.7.4.2 Desiderative  
4.7.4.3 Hortative  
4.8 Expressives  
4.8.1 Onomatopoeic forms  
4.9 Prepositions  
4.9.1 Location/instrument/subject /ka=/>  
4.9.2 Goal /ba=/>  
4.9.3 Source /can=/>  
4.9.4 Contrast /d=/>  
4.9.5 Equation /pon ~ pn=/>  
4.10 Auxiliaries and adverbs  
4.10.1 Temporal and modal auxiliaries  
4.10.1.1 Relational tense  
4.10.1.2 Root possibility  
4.10.2 Negation markers  
4.10.2.1 Negative  
4.10.2.2 Prohibitive  
4.10.3 Interrogative  
4.10.4 Emphasis  
4.10.5 Adverbs  
4.10.6 Other adverbial elements  
4.11 Conjunctions  
4.12 Co-ordinating morphemes in the NP  
4.12.1 Relative marker /k=/>  
4.12.2 Relative marker /t=/>  
4.12.3 Identification marker /l= ~ la=/>  
4.13 Summary
Chapter 5: Syntax

5.1 The structure of clauses and phrases

5.1.1 Statements

5.1.1.1 Common clauses

5.1.1.1.1 The verb and its arguments
5.1.1.1.2 Permutated arguments
5.1.1.1.3 Adjuncts
5.1.1.1.4 Detached phrases
5.1.1.1.5 Summary

5.1.1.2 Clauses without subject particle
5.1.1.3 Negated constructions
5.1.1.4 Subordination

5.1.2 Questions

5.1.2.1 Yes/no questions
5.1.2.2 WH questions

5.1.3 Relative clauses

5.1.3.1 Subordination by means of the identification marker

5.1.4 Noun phrases

5.1.4.1 The basic order of NP elements
5.1.4.2 WH modifiers of NPs
5.1.4.3 Nominalisations

5.1.5 Prepositional phrases

5.2 The arguments

5.2.1 The subject

5.2.1.1 The postverbal subject PP
5.2.1.1.1 Subject-marking /ka=/
5.2.1.1.2 Contrastive /d=/

5.2.2 The direct object

5.2.2.1 Oblique direct objects

5.2.3 Oblique arguments

5.2.4 The functions of prepositions: a summary

5.3 Valence and transitivity

5.3.1 The optional nature of arguments
5.3.2 Transitivity and imperfectivity

5.4 Valence-affecting operations

5.4.1 The causatives

5.4.2 The progressive as passive

5.5 Summary

Chapter 6: Expressive elaboration

6.1 Introduction

6.2 Phonological peculiarities
6.3 Morphological characteristics 162
6.4 Semantic characteristics 163
6.5 Syntactic characteristics 163
6.6 Summary 164

Chapter 7: Concluding remarks 165

Appendix I: Jahai rhyming glossary 167
Appendix II: English–Jahai finder list 209

References 239

Maps and figures
Map 1.1: Approximate distribution of Aslian languages 3
Figure 1.1: Aslian genetic relationships 5
Figure 2.1: Maximal syllable structure in Jahai 29
Figure 2.2: Syllabification of surface forms in Jahai 37

Tables
Table 2.1: Vowel phonemes in Jahai 20
Table 2.2: Consonant phonemes in Jahai 23
Table 2.3: Contrastive sets for consonant phonemes in Jahai 24
Table 2.4: Underlying syllable types in Jahai 29
Table 2.5: Syllable types in Jahai 29
Table 2.6: Word structure in Jahai 31
Table 2.7: Vowel distribution in Jahai 32
Table 2.8: Consonant distribution in Jahai 32
Table 4.1: Examples of meanings of nominalisations in Jahai 78
Table 4.2: Classifiers in Jahai 81
Table 4.3: Personal pronouns in Jahai 83
Table 4.4: In-law pronouns in Jahai 83
Table 4.5: Paradigm of verbal derivation in Jahai 94
## Abbreviations

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

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 Sabùm), 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).
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
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). 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.

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 [o]. 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 ke or de, or without any preposition at all. Locative and instrumental are marked with the preposition ke. The accusative, finally, is described as usually occurring after the verb without any preposition, but occasionally it is introduced by ke.

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 la (lā 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 Kruse (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
Introducing Jahai

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 ethnonyms 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.
Chapter 1

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. The latter includes e.g. hunting-stories and myths, particularly from genres referred to by the Jahai as /cnel/, 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 /yl/, whereas these are represented here as /j/ and /yl/ 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. 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

---

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 /u/ 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 /i/ is preferred to /u/.

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: [se’c'] /sec'/ ‘meat’, [lo’c’] /loc'/ ‘bow’, [rango’h] /rangop/ ‘mouth-harp’.

Table 2.1: Vowel phonemes in Jahai

<table>
<thead>
<tr>
<th></th>
<th>ORAL</th>
<th></th>
<th>NASAL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Front</td>
<td>Central</td>
<td>Back</td>
<td>Front</td>
</tr>
<tr>
<td>i</td>
<td>i</td>
<td>u</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
<td>o</td>
<td>e</td>
<td>a</td>
</tr>
<tr>
<td>e:</td>
<td>e</td>
<td>a</td>
<td>e</td>
<td>a</td>
</tr>
</tbody>
</table>

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 /o/, between /i/ and /a/, 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’
/gas/ ‘to apply make-up’
/gas/ ‘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 /a/ 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:

\[
\begin{array}{llll}
/kis/ & 'to dig' & /gis/ & 'to climb down' \\
/ges/ & 'smell' & /gos/ & 'to carve' \\
/ges/ & 'to descend' & /gas/ & '(skin disease)' \\
\end{array}
\]

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, e, a, u, o/ 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:

\[
\begin{array}{llll}
/kis/ & 'to dig' & /kis/ & 'ghost' \\
/pesk/ & 'to split' & /pesk/ & 'to sting' \\
/sit/ & 'honeycomb' & /sit/ & 'to rub oneself' \\
/titst/ & 'to stare' & /klist/ & 'to swallow' \\
/kawaw/ & 'bird' & /waw/ & '(a type of civet)' \\
/siruc/ & 'to slurp' & /gruc/ & 'slender-toed gecko' \\
/hakok/ & 'to throw' & /hokok/ & 'to burn off fur' \\
\end{array}
\]

2.2.2 Phonetic description of vowel phonemes

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

/i/ is the rather common nasal counterpart of /i/: [hiic] /hiic/ 'to rain', [kori] /kri/ 'wrist'.

/e/ is a close-mid front unrounded vowel [e], slightly more raised than [e]: [he?] /he?/ 'IP.INCL', [tek] /tek/ 'to sleep'. It has no nasal counterpart.

/e/ is an open-mid front unrounded vowel [e]: [tek] /tek/ 'to give', [ges] /ges/ 'to descend'.

/ê/ is the rather common nasal counterpart of /e/: [lapêk] /lapêk/ 'mud', [pece?] /pece?/ 'to be wet'.

/i/ is a close central unrounded vowel, which is slightly more backed than [i]: [hip] /hip/ 'forest', [sit] /sit/ 'honeycomb'. Before glottal consonants it becomes a slightly lowered [a]: [ba?] /ba?/ 'mother'. Occasionally it is in free variation with a slightly rounded [u]. Conditioned rounding is also evident following the bilabial approximant /w/: [wun] /win/ 'to crawl'.

/i/, the nasal counterpart of /i/, is infrequent: [sit] /sit/ 'to rub oneself', [tati] /tati/ '(a type of large bird)'.

/æ/ is a mid central unrounded vowel [æ]: [gos] /gos/ 'to carve', [ʔahoj] /ʔahoj/ 'to be small'.

/æ/ is the nasal counterpart of /æ/: [cænhã] /cænhã/ 'to be short', [sãh] /sãh/ 'to meet'.

/æ/ is an open central unrounded vowel [a]: [ʔap] /ʔap/ 'tiger', [cãh] /can/ 'foot'.
Chapter 2

/\l/, the nasal counterpart of /a/, has been identified in only a few lexical items: [jap\R] /jap\R/ ‘(name of Jahai group)’, [w\w\w]/w\w\w/ ‘(a type of civet)’.

/\u/ is a close back rounded vowel [u], frequently in free variation with a significantly advanced [\u]: [duk’]/duk’/ ‘to pounce upon’, [tu\R]/tu\R/ ‘to fell’.

/\u/ is the nasal counterpart of /\u/: [c\u\R]/c\u\R/ ‘blind’, [t\u]/t\u/ ‘to grub’.

/\o/ is a close-mid back rounded vowel [o]: [sop’]/sop’/ ‘lung’, [to\R]/to\R/ ‘tree-trunk’. It has no nasal counterpart.

/\a/ is an open-mid back half-rounded vowel [\a]: [g\a]/g\a/ ‘to live’, [\a\R]/\a\R/ ‘water’.

It is in free variation with an infrequent unrounded [\a].

/\i/ is the rather common nasal counterpart of /\o/: [\i\o’]/\i\o’/ ‘dog’, [\i\R]/\i\R/ ‘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: [c\u\R]/c\u\R/ ‘to flow’, [ha\R]/ha\R/ ‘stench’, [k\w\R]/k\w\R/ ‘bird’, [h\a]/h\a/ ‘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 §2.2.3.2). Progressive nasalisation from the onset, on the other hand, is allowed — [m\R]/m\R/ ‘eye’, [\R]/\R/ ‘throat’, [\R]/\R/ ‘nose’ — and may even be desirable, as suggested by Malay loanwords in which voiced stop onsets which follow homorganic nasals are assimilated and lost: [l\R]/l\R/ from Malay lembu ‘cattle’, [\R]/\R/ from Malay rendah ‘low’, [\R]/\R/ from Malay jangga ‘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: [m\R]/m\R/ ‘to wake up’, [\R]/\R/ ‘star’ (from Malay bintang). Anticipatory nasalisation of word-initial consonants becomes particularly apparent during infixation of /\o/ (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 /\o/ forms the onset of the following new syllable: [m\o\R]/m\o\R/ from /bawac/ ‘pig-tailed macaque’, [m\R]/m\R/ from /w\o\R/ ‘rat’, [\R]/\R/ from /\o\R/ ‘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 [ŋɔk’ ~ ɲɔk’] /ŋɔk/ ‘to sit’, in which the nasalisation of the vowel is not always apparent, and [bɔŋɔk’] /bŋɔk/ ‘(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: [mɔŋh] /mɔŋh/ ‘nose’.

2.3 Consonants

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

Table 2.2: Consonant phonemes in Jahai

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p, b</td>
<td>t, d</td>
<td>c, j</td>
<td>k, g</td>
<td>?</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>p</td>
<td>η</td>
<td>h</td>
</tr>
<tr>
<td>Fricative</td>
<td>φ</td>
<td>l</td>
<td>s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhotic</td>
<td>w</td>
<td></td>
<td>j</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 /φ/.

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

<table>
<thead>
<tr>
<th>INITIAL CONTRAST</th>
<th>FINAL CONTRAST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stop</strong></td>
<td></td>
</tr>
<tr>
<td>/p/</td>
<td>/pʰ/</td>
</tr>
<tr>
<td>/b/</td>
<td>/bʰ/</td>
</tr>
<tr>
<td>/t/</td>
<td>/tʰ/</td>
</tr>
<tr>
<td>/d/</td>
<td>/dʰ/</td>
</tr>
<tr>
<td>/c/</td>
<td>/cʰ/</td>
</tr>
<tr>
<td>/j/</td>
<td>/jʰ/</td>
</tr>
<tr>
<td>/l/</td>
<td>/lʰ/</td>
</tr>
<tr>
<td>/r/</td>
<td>/rʰ/</td>
</tr>
<tr>
<td>/n/</td>
<td>/nʰ/</td>
</tr>
<tr>
<td>/l/</td>
<td>/lʰ/</td>
</tr>
<tr>
<td>/m/</td>
<td>/mʰ/</td>
</tr>
<tr>
<td>/h/</td>
<td>/hʰ/</td>
</tr>
<tr>
<td>/r/</td>
<td>/rʰ/</td>
</tr>
<tr>
<td>/y/</td>
<td>/yʰ/</td>
</tr>
<tr>
<td><strong>Fricative</strong></td>
<td></td>
</tr>
<tr>
<td>/φ/</td>
<td></td>
</tr>
<tr>
<td>/s/</td>
<td>/sʰ/</td>
</tr>
<tr>
<td>/h/</td>
<td>/hʰ/</td>
</tr>
<tr>
<td>/n/</td>
<td>/nʰ/</td>
</tr>
<tr>
<td>/l/</td>
<td>/lʰ/</td>
</tr>
<tr>
<td>/m/</td>
<td>/mʰ/</td>
</tr>
<tr>
<td>/h/</td>
<td>/hʰ/</td>
</tr>
<tr>
<td><strong>Nasal</strong></td>
<td></td>
</tr>
<tr>
<td>/n/</td>
<td>/nʰ/</td>
</tr>
<tr>
<td>/l/</td>
<td>/lʰ/</td>
</tr>
<tr>
<td>/m/</td>
<td>/mʰ/</td>
</tr>
<tr>
<td>/h/</td>
<td>/hʰ/</td>
</tr>
<tr>
<td><strong>Lateral</strong></td>
<td></td>
</tr>
<tr>
<td>/l/</td>
<td>/lʰ/</td>
</tr>
<tr>
<td>/r/</td>
<td>/rʰ/</td>
</tr>
<tr>
<td><strong>Rhotic</strong></td>
<td></td>
</tr>
<tr>
<td>/r/</td>
<td>/rʰ/</td>
</tr>
<tr>
<td><strong>Approx.</strong></td>
<td></td>
</tr>
<tr>
<td>/w/</td>
<td>/wʰ/</td>
</tr>
<tr>
<td>/y/</td>
<td>/yʰ/</td>
</tr>
<tr>
<td><strong>Contrast</strong></td>
<td><strong>Contrast</strong></td>
</tr>
<tr>
<td>/pʰ/ 'to fan'</td>
<td>/kʰ/ 'to move hut'</td>
</tr>
<tr>
<td>/bʰ/ 'fruit'</td>
<td></td>
</tr>
<tr>
<td>/tʰ/ 'ancestor'</td>
<td>/kʰ/ 'to take'</td>
</tr>
<tr>
<td>/dʰ/ 'ipoh poison'</td>
<td></td>
</tr>
<tr>
<td>/cʰ/ 'to bite'</td>
<td>/kʰ/ 'to gnaw'</td>
</tr>
<tr>
<td>/jʰ/ 'to ascend'</td>
<td></td>
</tr>
<tr>
<td>/lʰ/ 'a type of tortoise'</td>
<td>/kʰ/ 'to call'</td>
</tr>
<tr>
<td>/gʰ/ 'manchild'</td>
<td></td>
</tr>
<tr>
<td>/lʰ/ 'to cough'</td>
<td>/m/kʰ/ 'egg'</td>
</tr>
<tr>
<td>/sʰ/ 'game animal'</td>
<td>/gʰ/ 'to live'</td>
</tr>
<tr>
<td>/hʰ/ 'jar'</td>
<td>/kʰ/ 'a type of tortoise'</td>
</tr>
<tr>
<td>/nʰ/ 'many'</td>
<td>/kʰ/ 'to copulate'</td>
</tr>
<tr>
<td>/lʰ/ 'outgrowth on tree'</td>
<td>/kʰ/ 'to sit'</td>
</tr>
<tr>
<td>/mʰ/ 'hand'</td>
<td>/gʰ/ 'to carry'</td>
</tr>
<tr>
<td>/hʰ/ 'allow'</td>
<td>/t/kʰ/ 'to cut'</td>
</tr>
<tr>
<td>/wʰ/ 'call'</td>
<td>/b/kʰ/ 'a type of tree'</td>
</tr>
<tr>
<td>/yʰ/ 'drop'</td>
<td>/t/kʰ/ 'sail-fin lizard'</td>
</tr>
</tbody>
</table>

### 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, j, g/ are usually realised as unaspirated bilabial, alveolar, palatal and velar stops. The palatals /c/ and /j/ are affricated [cʰ] and [jʰ], although affrication of /c/ does not usually occur before epenthetic [i] in pre-final syllables: [cijas] /cijas/ '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: [mʊŋker] /mʊŋker/ 'to wake up'. The voiced/voiceless contrast is illustrated by the following minimal pairs:

- [pʰ] /pʰ/ 'to fan'
- [tʰ] /tʰ/ 'soil'
- [cʰ] /cʰ/ 'to bite'
- [keʰ] /keʰ/ 'child'
- [bʰ] /bʰ/ 'fruit'
- [dʰ] /dʰ/ 'to make'
- [jʰ] /jʰ/ 'to ascend'
- [geʰ] /geʰ/ '(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 [p', t', c', k'] and symbolised phonemically by the voiceless /p, t, c, k/:

- [cʰp'] /cʰp/ 'to go'
- [dʰut'] /dʰut/ 'navel'
- [seʰc'] /seʰc/ 'meat'
- [sʰk'] /sʰk/ '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 /l/: [secʰ] /sec/ ‘meat’, [wekʰ] /wek/ ‘to return’. There is no regular pattern to suggest that this voiceless release has any contrastive significance.


2.3.1.2 Nasals

The nasal phonemes /m, n, ŋ, 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, ŋ, ŋ]:

- [məh] /məh/ ‘nose’
- [nɔj] /nɔj/ ‘to copulate’
- [ŋəp] /ŋəp/ ‘to chew’
- [ŋək] /ŋək/ ‘to sit’

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

- [ləkeʰm] /ləkeʔm/ ‘brain’
- [pətoʰm] /pətoʔm/ ‘day before yesterday’
- [wuʰn] /wuʔn/ ‘to crawl’
- [cədʰn] /cədʔn/ ‘foot’
- [teʰn] /teʔn/ ‘to plait’
- [piŋloʰŋ] /piŋloʔŋ/ ‘to sing’
- [tabəʰŋ] /tabəʔŋ/ ‘leaf monkey’
- [gahuʰŋ] /gahuʔŋ/ ‘cave’

Such nasal segments have been recorded in several Austronesian languages, as well as in many Austronesian languages, notably in Borneo (Blust 1997:154–169). They are varyingy 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 Kangiw 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, [ʰm, ʰn, ʰŋ] or [bm, dn, ŋn, ŋŋ], 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 [bʰ, tʰ, j, ɡ], 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. [tɕʰən] /can/ ‘foot’ with Vietnamese chan ‘foot’ and [tʰəŋ]/tʰəŋ/ ‘to plait’ with Khmuu tāap/ ‘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:

\[
\begin{align*}
[səmsɔ^b m] & /smsɔm/ & 'to buzz around a nest' \\
[hənho^d ən] & /hnɔɔn/ & 'to devour' \\
[jiŋje^o^d əŋ] & /jŋŋəŋ/ & 'to dream' \\
[j^2ənje^o^g əŋ] & /jŋŋəŋ/ & 'wide'
\end{align*}
\]

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

\[
\begin{align*}
[haja^b m] & \text{from ayam} & 'poultry' \\
[bula^d ən] & \text{from bulan} & 'moon' \\
[kuc^c əŋ] & \text{from kucing} & 'cat'
\end{align*}
\]

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:

\[
\begin{align*}
[kɔŋŋm] & /kɔŋŋm/ & 'to urinate' \\
[ɡonuŋ] & /ɡuŋuŋ/ & 'bamboo' \\
[ʃaməŋ] & /ʃaməŋ/ & 'to request' \\
[ʔaməŋ] & /ʔaməŋ/ & 'siamang'
\end{align*}
\]

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: [kəniŋ]/kəniŋ/ from Malay kambing
Phonology

'goat', [k.longa\] /k.m\]/ from Malay *kembang* ‘to swell’, [h.na\] /h.n\]/ from Malay *enggang* ‘hornbill’ and [pa.na\] /pa.n\]/ from Malay *panggang* ‘to roast’ (see also §2.6.3.3). Exceptions not yet accounted for include [tama\] /tam\]/ ‘(a type of small animal)’, [lu.a\] /lu.a]/ ‘binturong’, [ha.mi\] /h.mi]/ ‘(magic word)’, [bra.ma\] /bra.m\]/ ‘(a type of tree)’, [o.na\] /o.n\]/ ‘side’, [ge.no\] /ge.n\]/ ‘wooden material’ and [lam.o\] /lam.o]/ ‘(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.e.ki.m] ‘peacock-­pheasant’, [h.a.w.e.n] ‘wild boar’, [he.n] ‘mouth’ and [ka.ti.n] ‘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.e.ki.m] ‘lump’
[ji.n] ‘to hop on something’
[g.o.n] ‘outgrowth on tree’
[b.o.li.n] ‘upper arm’

[c.e.ki.m] ‘peacock-­pheasant’
[ji.n] ‘to stop’
[g.o.n] ‘to refuse to give’
[b.o.li.n] ‘long-­tailed parakeet’

This, 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:

[k.o.p\]/ /k.o.p/ ‘to move hut’
[k.o.b.m\] /k.o.m]/ ‘many’
[d.u.t\]/ /d.u.t/ ‘navel’
[d.u.n\]/ /d.u.n]/ ‘to cover’
[s.e.c\]/ /s.e.c/ ‘meat’
[s.e.j\]/ /s.e.j]/ ‘front’
[p.o.k\]/ /p.o.k/ ‘round object’
[p.o.g.n\] /p.o.n]/ ‘to tap poison’

2.3.1.3 Fricatives

The voiceless palatal fricative /s/ is commonly an alveolo-­palatal [c], 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. [s.o.m\] /s.o.m]/ ‘bird’s nest’, [ke.nsi.g.n\] /ke.nsi.n]/ ‘banded palm civet’ and [ko.s] /ko.s]/ ‘fire’.

The voiceless glottal fricative /h/ occurs frequently in syllable-­initial and, especially, word-­final position: [ho.k\]/ /h.o.k/ ‘to throw’, [si.h]/ /si.h]/ ‘to pound’, [pa.ah]/ /p.a.h]/ ‘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 /φ/ 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 [φ]. In pre-final syllables it is usually realised as an unreleased bilabial stop [p']. 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 [cεɛφ] /cɛɛφ/ ‘to fan fire’, [cɛɛνɛɛφ] /cνɛɛφ/ ‘tail feathers’, [niɛφ] /niɛφ/ ‘to blow smoke’, [kɛɔφ] /kɛɔφ/ ‘to splutter’, [bɛɔφ] /bɛɔφ/ ‘to be quick’, [lɔφ] /lɔφ/ ‘to leak’, and [pɔlɛφ] /plɛφ/ ‘(sound of blowpipe)’. Syllable-final /φ/ has been identified in other Northern Aslian languages as well, including Kensiw (Bishop 1996a:23) and Batek (Benjamin 1985b:10).

2.3.4 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: [lɔj] /lɔj/ ‘to run’, [halel] /halel/ ‘hungry’, [ɾɛl] /ɾɛl/ ‘to see’.


2.3.5 Approximants


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: [ʃ]. 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: [jɔk] /jɔk/ ‘to undress’, [ʃiʃ] /ʃiʃ/ ‘to carry in one’s hand’, [gɛʃ] /gɛʃ/ ‘to eat’, [kuj] /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]a. Onsets are obligatory but codas are not, and the minimal syllable is therefore [CV]a. Syllable structure is represented here as a flat consonant-vowel tier (see e.g. Clements and Keyser 1983):
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]₀, and heavy syllables are closed [CVC]₀. In the underlying form of both light and heavy syllables, however, the vowel nucleus may be either non-predictable and prespecified (/CV/, /ICVC/), or predictable and underspecified and thus omitted in phonemic transcriptions (/CI, /IC/). The following table separates the four underlying syllable types:

**Table 2.4: Underlying syllable types in Jahai**

<table>
<thead>
<tr>
<th></th>
<th>/C/</th>
<th>/CV/</th>
<th>/CC/</th>
<th>/CVC/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Prespecified nucleus</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

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**

<table>
<thead>
<tr>
<th></th>
<th>LIGHT</th>
<th>HEAVY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half</td>
<td>C</td>
<td>CC</td>
</tr>
<tr>
<td>Surface form</td>
<td>CV</td>
<td>CVC</td>
</tr>
</tbody>
</table>
Chapter 2

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. [cēp'] /cēp/ '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. [kone'c'] /knecl 'comb'. They are phonetically disyllabic, but the vowel of the penultimate syllable is epenthetic and predictable and therefore omitted in phonemic transcriptions. 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. [kawip'] /kawip/ 'sun bear'; or it may be closed and then have either a predictable or a non-predictable vowel nucleus: /CC.CVC/, e.g. [tɔmkal] /tmkal/ 'male'; or /CVC.CVC/, e.g. [kaltn]/kalton/ '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.
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. [kalənis] /kalənis/ ‘heart’; /C.CVC.CVC/, e.g. [cəmalpək] /cəmalpək/ ‘(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. [pəŋən] /pəŋən/ ‘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, [kuruhu] /kuruhu/ ‘(a type of owl)’, although this may be onomatopoeic and therefore uncertain.

As described for other Northern Aslian languages (see Asmah (1976:95) for Kintaq and Bishop (1996a:240) for Kensiw), trisyllabic roots are commonly Malay borrowings.

### Table 2.6: Word structure in Jahai

<table>
<thead>
<tr>
<th>Word type</th>
<th>Canonic structure</th>
<th>Example</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monosyllabic</td>
<td>/CVC/</td>
<td>/cepl</td>
<td>‘to catch’</td>
</tr>
<tr>
<td>Sesquisyllabic</td>
<td>/C.CV.CVC/</td>
<td>/knecl</td>
<td>‘comb’</td>
</tr>
<tr>
<td>Disyllabic</td>
<td>/C.CV.CVC/</td>
<td>/kawip/</td>
<td>‘sun bear’</td>
</tr>
<tr>
<td></td>
<td>/CC.CV/C/</td>
<td>/tmkal/</td>
<td>‘male’</td>
</tr>
<tr>
<td></td>
<td>/CVC.CVC/</td>
<td>/kalton</td>
<td>‘knee’</td>
</tr>
<tr>
<td>Trisyllabic</td>
<td>/C.CV.CVC/</td>
<td>/klañis</td>
<td>‘heart’</td>
</tr>
<tr>
<td></td>
<td>/C.CC.CV/C/</td>
<td>/prəŋən</td>
<td>‘pharynx’</td>
</tr>
<tr>
<td></td>
<td>/C.CV.CVC/C/</td>
<td>/cəmalpək/</td>
<td>‘(a type of millipede)’</td>
</tr>
</tbody>
</table>

### 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 /əl, əl/ 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

<table>
<thead>
<tr>
<th></th>
<th>Oral</th>
<th>Nasal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø</td>
<td>i e e i a u o o</td>
<td>i a a u u</td>
</tr>
<tr>
<td>Final</td>
<td>+ + + + + + + +</td>
<td>+ + + + + + + +</td>
</tr>
<tr>
<td>Penultimate</td>
<td>+ + + + + + + +</td>
<td>+ + + + + + + +</td>
</tr>
<tr>
<td>Antepenultimate</td>
<td>+ + + + + + + +</td>
<td>+ + + + + + + +</td>
</tr>
</tbody>
</table>

Table 2.8: Consonant distribution in Jahai

<table>
<thead>
<tr>
<th>Coda</th>
<th>p t c k ? b d j g φ s h m n n l r w j</th>
</tr>
</thead>
<tbody>
<tr>
<td>#CV#</td>
<td>+ + + + + + + + + + + + + + + + + + +</td>
</tr>
<tr>
<td>.CV#</td>
<td>- - - - - - - - + + + + + + + + + + +</td>
</tr>
<tr>
<td>#CV (V)C#</td>
<td>+ + + + + + + + + + + + + + + + + + +</td>
</tr>
<tr>
<td>#C (V)C#</td>
<td>+ + + + + + + + + + + + + + + + + + +</td>
</tr>
<tr>
<td>#C (V)C#</td>
<td>+ + + + + + + + + + + + + + + + + + +</td>
</tr>
<tr>
<td>Onset</td>
<td></td>
</tr>
<tr>
<td>#CV (V)C#</td>
<td>+ + + + + + + + + + + + + + + + + + +</td>
</tr>
<tr>
<td>#C (V)C#</td>
<td>+ + + + + + + + + + + + + + + + + + +</td>
</tr>
<tr>
<td>.CV (C)</td>
<td>- - - - - - - - - - - - - - - - - - -</td>
</tr>
</tbody>
</table>

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 /φ/. 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, j/, fricative /s/ and nasal /ɲ/ almost never occupy both onset and coda position, the one exception being /cOC/ '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, e, i, a/ 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 /u/ and /o/ 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-, j-, g-/. There are occasional exceptions to this pattern, e.g. /bɔk/ ‘to untie’, /dadɔl/ ‘(a type of reed snake)’, /piʃə/ ‘to fly up’, /gɔc/ ‘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 /Cl/, 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 /f/, 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/:

- [tij5k'] /tj5k/ ‘to point’
- [ci5s] /cjas/ ‘hand’
- [si5j9n] /sjin/ ‘to burn’
- [si5jul] /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: [keʔep'] /kʔep/ ‘centipede’, [taʔaʔ] /tʔaʔ/ ‘vegetable’. This pattern is much less apparent and [ə] is equally common in this environment: [cəʔaʔ] /cʔaʔ/ ‘to pour’. Furthermore, if a penultimate stop onset is combined with a liquid onset in the final syllable, especially /l/, [ə] is commonly dropped altogether in fast speech: [braʔ ~ braʔ] /braʔ/ ‘NEG’, [gəre9n ~ gre9n] /greq/ ‘(a type of monitor lizard), [pol5φ ~ pl5φ] /pl5φ/ ‘(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 /f/, which is never an onset. However, no unequivocal examples of the syllable-initially rare nasals /p/ and /l/ 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:

- /wawat/ ‘rat’
- /napuʔ/ ‘to die’
- /mamun/ ‘to beg’
- /dad5l/ ‘(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. [məneʔ] /bneʔ/ ‘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 /ɛ/ occur only sporadically. The remaining oral vowels, as well as the phonemically nasal vowels, are not found in this position.

/ralic/ 'to pass'
/timoʔ/ 'hard surface'
/cuhɛʔ/ 'to flow'
/pokah/ '(a type of gecko)'
/hokɔk/ '(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/:

/Rmpoŋ/ 'hole'
/Rnteŋ/ 'ear'
/Kaŋcɔʔ/ 'grandchild'
/Sanjkoh/ 'wreathed hornbill'
/Cnhɑt/ 'to be short'
/Knʔac/ 'father-in-law'

There are a few exceptions to this pattern, e.g. /tmket/ '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]: [hɔkɪt] /hɔkɪt/ 'night'. If the surrounding onset and coda are both alveolar or palatal, the epenthetic vowel frequently becomes a slightly lowered and fronted [ɛ]:

[tʰrɪhi] /trɪc/ '(a type of small bird)'
[ʃənˈlɛc] /snɪc/ 'blowpipe dart'
[ɛnˈdɪs] /cnrɛs/ 'nail'
[ʃəltuh] /sltuh/ 'to attack'
[ʃəˈməl] /ʃlmɔl/ 'mountain'
[ʃərwe] /ʃrwej/ '(a type of tree)'

There are also numerous examples of non-predictable /a/ in this position, e.g. /karwoʔ/ '(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: /lintes/ 'to go across', /cunfɪŋ/ 'temple'. One example of /ɔ/ has been identified, /hɔrjɨn/ '(a type of small animal)', but it is possible that this is in fact a compound of two free forms: /hɔr jɪn/. 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:
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 /kuruhji/ 'a type of owl'.

Occasional Malay loans fail to conform to the patterns outlined above. Thus, /dahagaʔ/ 'to be thirsty' (from Malay dahaga) and /punhulu̱/ '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. 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, l/, the fricatives /s, h/, liquids /l, t/ and the bilabial approximant /w/.

---

9 This is the terminology introduced by Kruspe (1999, 2004), see §3.2.
Chapter 2

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 \( /\phi / \) 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 \( [j] \):

\[
\begin{align*}
[\text{[o]p\text{-}l\text{-}\phi}] & \quad /\phi\text{-}\text{l\text{-}\phi}/ \quad \text{‘fontanelle’} \quad /\text{l\text{-}\phi}/ \quad \text{‘fontanelle’} \\
[\text{[h]or\text{-}n\text{-}p\text{e}j\text{\text{-}}}\text{\text{-}\text{\text{-}}}\text{\text{-}}} & \quad /\text{h\text{-}r\text{-}p\text{-}e}\text{j}/ \quad \text{‘goose-bumps’} \quad *\text{hpe}\text{j}/\text{hrpe}\text{j} \\
[\text{[p]n\text{-}s\text{e}j\text{\text{-}}}\text{\text{-}}} & \quad /\text{p\text{-}s\text{e}\text{-}j}/ \quad \text{‘to say’} \quad *\text{ps}\text{e}\text{j} \\
[\text{[h]l\text{-}ij\text{-}dij}] & \quad /\text{h\text{-}l\text{-}dij}/ \quad \text{‘flat ground’} \quad *\text{hdij} \\
\end{align*}
\]

\[\text{2.4.4.2 Phonetic realisation of epenthetic vowels}\]

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 \( /c, s, j, j/ \), the preceding vowel is realised as \( [i] \):

\[
\begin{align*}
[\text{[k]olic\text{\text{-}}ba\text{\text{-}}}c\text{\text{-}}c\text{\text{-}}] & \quad /\text{k}\text{-}lc\text{\text{-}bac}/ \quad \text{‘(a type of millipede)’} \quad *\text{kb}\text{ac} \\
\text{[ki\text{\text{-}sw\text{e}s}] & \quad /k<s>\text{w}\text{e}\text{s}/ \quad \text{‘to be sweeping’} \quad /\text{k\text{-}w\text{e}s}/ \quad \text{‘to sweep’} \\
\text{[pi\text{\text{-}n\text{\text{-}}l}\text{\text{-}}}s\text{\text{-}j\text{\text{-}}} & \quad /\text{p}\text{-}l\text{-}\text{-}s\text{\text{-}j}/ \quad \text{‘to sing’} \quad *\text{pl}\text{\text{-}j} \\
\text{[hi\text{\text{-}h}\text{\text{-}j}\text{\text{-}j}] & \quad /\text{h}\text{-}j\text{-}\text{h}\text{-}j}/ \quad \text{‘to be yawning’} \quad /\text{h}\text{-}h\text{\text{-}j}/ \quad \text{‘to yawn’} \\
\end{align*}
\]

In cases where the copy is a glottal \( /\text{\text{-}h} \), \( /\text{\text{-}h} \), the preceding vowel is realised as \( [a] \):

\[
\begin{align*}
\text{[sa\text{\text{-}}so\text{\text{-}}}j & \quad /s\text{-}\text{so}\text{-}j/ \quad \text{‘blood vessel’} \quad /\text{so}\text{-}j/ \quad \text{‘cubit’} \quad 10 \\
\text{[ba\text{\text{-}b}\text{\text{-}}}j & \quad /b\text{-}\text{b}\text{-}j/ \quad \text{‘to carry on one’s back’} \quad *\text{b}\text{o}\text{j} \\
\text{[ta\text{\text{-}h}\text{\text{-}t}\text{\text{-}}}j & \quad /\text{t}\text{-}\text{h}\text{-}t\text{-}j}/ \quad \text{‘oriental pied hornbill’} \quad *\text{t}\text{e}\text{h} \\
\text{[nah\text{\text{-}h}\text{\text{-}j} & \quad /\text{n}\text{-}\text{-h}\text{-}\text{-}j}/ \quad \text{‘height’} \quad /\text{\text{-}j}\text{\text{-}h}/ \quad \text{‘to ascend’} \\
\end{align*}
\]

With other copies, the preceding vowel is realised as \( [\text{\text{-}}] \):

\[
\begin{align*}
\text{[bo\text{\text{-}}}\text{t\text{-}et}] & \quad /\text{b}\text{-}\text{-}\text{t}\text{-}et}/ \quad \text{‘good’} \quad *\text{b}\text{-}\text{-}t\text{\text{-}et} \\
\text{[dok\text{\text{-}d}\text{\text{-}uk}] & \quad /\text{d}\text{-}\text{k}\text{-d}\text{-}uk}/ \quad \text{‘chest’} \quad /\text{d}\text{-}\text{uk}/ \quad \text{‘to pounce’} \\
\text{[som\text{\text{-}m}\text{\text{-}}}j & \quad /\text{s}\text{\text{-}m}\text{-}s\text{\text{-}m}/ \quad \text{‘to buzz around a nest’} \quad /\text{s}\text{-}\text{m}/ \quad \text{‘bird’s nest’} \\
\text{[sa\text{\text{-}n}\text{\text{-}o}\text{\text{-}}}j & \quad /\text{s}\text{-}\text{-n}\text{-}\text{-o}\text{\text{-}j}/ \quad \text{‘leafbird’} \quad *\text{sp}\text{o}\text{\text{-}j} \\
\text{[dol\text{\text{-}d}\text{\text{-}il}] & \quad /\text{d}\text{-}\text{-d}\text{-}il}/ \quad \text{‘heel’} \quad *\text{\text{-}d}\text{\text{-}il} \\
\text{[pr\text{\text{-}ber}] & \quad /\text{pr}\text{-}\text{-ber}/ \quad \text{‘lower arm’} \quad *\text{p}\text{\text{-}ber} \\
\text{[h\text{-}wh\text{\text{-}h}\text{\text{-}w}] & \quad /\text{h\text{-}wh\text{-\text{-}h}\text{\text{-}w}/ \quad \text{‘crested wood-partridge’} \quad *\text{h}\text{-}\text{h}\text{-\text{-}w} \\
\end{align*}
\]

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

\[
\begin{align*}
(1) \quad & \quad \emptyset > [i] / \quad _-C(palatal) + \\
(2) \quad & \quad \emptyset > [a] / \quad _-C(glottal) + \\
(3) \quad & \quad \emptyset > [\text{\text{-}}] / \quad _-C(\text{-}\text{-}palatal, \text{-}\text{-}glottal) + \\
\end{align*}
\]

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]₀, 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](image)

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]₀ syllables, which have precedence over minimal [CV]₀ 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 /CCCVCI/ 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.CVCI/ > [CVC.CVCI] and /CCCC.CVCI/ > [CVC.CVCI.CVCI]. It excludes the unattested forms */C.C.CVCI/, */CC.CVCI/ and */C.CC.CVCI/. 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 /CCCVCI/, /CCVCVCI/ 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.CVCI/ > [C.V.CVC], /C.CV.CVC/ > [C.CV.CV.CVC] and /C.CVC.CVCI/ > [C.V.CV.CV.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):

<table>
<thead>
<tr>
<th>HIGH LEVEL</th>
<th>LOW LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>'lo? 'part of a blowpipe'</td>
<td>_lo? 'bough'</td>
</tr>
<tr>
<td>'nus 'sleeping mat'</td>
<td>_nus 'lip'</td>
</tr>
<tr>
<td>'te? 'IS'</td>
<td>_te? '1D INCL'</td>
</tr>
<tr>
<td>'he? 'IP INCL'</td>
<td>_he:i '1D INCL'</td>
</tr>
<tr>
<td>'eg? 'to give'</td>
<td>_eg? 'belly'</td>
</tr>
<tr>
<td>'gus 'to rub fat into the face'</td>
<td>_gus 'to come down'</td>
</tr>
</tbody>
</table>

Contemporary Jahai:

| /jo?/ 'outer shaft of blowpipe' | /joh/ 'crown of tree' |
| /nis/ 'sleeping mat'           | /nus/ 'upper lip' |
| /je?/ 'IS'                     | /jeh/ '1D EXCL' |
| /he?/ 'IP INCL'                | /he:j/ '1D INCL' |
| /?ek/ 'to give'                | /?ec/ 'belly', 'excrement' |
| /gis–g?/ 'to apply make-up'     | /gis/ '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
Phonology

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 /hi/, the closure of open syllables by means of glottal /hi/ or /nt/, the realisation of the rhotic /rl/, 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 /krtih/ from kereta ‘car’, /kamirj/ from kambing ‘goat’, and /juwal/ from jual ‘to sell’. Others represent more basic vocabulary, e.g. /batu?/ from batu ‘stone’, /gahah/ from gajah ‘elephant’, and /kucen/ from kucing ‘cat’.

A number of words can be tentatively identified as loans from Temiar, a Central Aslian neighbour, e.g. /babj/ ‘woman’ and /1m0/ ‘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 /kbs/ ‘to die’ and /bgw/ ‘wind’.

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 /grinetJ/ ‘grenade’, /bubi? trep/ ‘booby-trap’, /pos/ ‘post’, and /trine?/ ‘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:

- [tanem] /tanem/ from tanam ‘to plant’
- [tomen] /tomen/ from toman ‘snakehead’
- [kunin] /kunin/ from kuning ‘yellow’
- [jam] /jam/ from jam ‘hour’
- [bulan] /bulan/ from bulan ‘moon’
- [gadi] /gadi/ from gading ‘elephant’s tusk’

2.6.1.2 Nasalisation of word-initial /bl/

According to a rather regular Jahai pattern, the voiced bilabial stop /b/ becomes nasalised [m] word-initially if the following consonant is nasal: [minte] /binten/ from Malay bintang ‘star’, [minat] /binat/ from Malay binatang ‘animal’. This process occasionally results in ambiguous forms, as in the case of [mne] /mne/ ‘to win’ (from Malay menang) or /bne/ ‘thread’ (from Malay benang).

2.6.1.3 Palatalisation of /sl/

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 /e/:

- /injet/ from ingat ‘to remember’
- /bun?i/ from bunga ‘flower’
- /tanem/ from tanam ‘to plant’
- /pge/ from pegang ‘to hold’
- /bile/ from bilang ‘to count’
- /tanel/ from tanggal ‘to fall’
- /lante/ from lantai ‘floor’
Sometimes /a/ is replaced by /ə/:

- /tmpat/ 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’.

- /kəpɔʔ/ from kapak ‘axe’
- /jəlɔʔ/ from jala ‘casting-net’
- /krəbɔʔ/ from kerbau ‘buffalo’
- /hiʃɔʔ/ from hijau ‘to be green’
- /pulɔʔ/ from pulau ‘island’
- /gurɔʔ/ from gurau ‘to jest’

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

- /tiŋkəp/ from tingkap ‘window’
- /kilət/ from kilat ‘lightning’
- /badaʔ/ from badak ‘Sumatran rhinoceros’
- /bras/ from beras ‘husked rice’
- /gajah/ 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: /tekl/ ‘to sleep’, /teʔl/ ‘soil’. The final /k/ in Malay is therefore consistently represented as phonemic /l/ in loanwords:

- /ʔiteʔl/ from itik ‘duck’
- /taseʔl/ from tasik ‘lake’
- /badaʔl/ from badak ‘Sumatran rhinoceros’
- /sjuʔl/ from sejuk ‘to be cold’
- /loboʔl/ from lobok ‘pool’
- /kəpɔʔl/ from kapak ‘axe’

### 2.6.2.3 Voiceless labiodental fricative

Standard Malay has a rather rare voiceless labiodental fricative phoneme /ʃ/ 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.
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'
/hjan/ from enggang 'rhinoceros hornbill'
/ʔanjet/ from angkit 'to take'
/ʔasin/ 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'
/limeʔ/ from lima 'five'
/lataʔ/ from lata 'waterfall'
/kritlhi from kereta 'car'
/punhuluʔ/ 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:

/halow/ from halau 'to scare'
/pulow/ from pulau 'island'
/lanteʔ/ 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), /p'rahu?/ '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'
/duwaw/ from dua 'two'
/bawun/ from baung 'a type of catfish'
/lajin/ from lain 'to be different'
/cjui/ 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'
/tmakaw/ from tembakau 'tobacco'
/mah/ from rendah 'to be low'
/banij/ from Banding '(toponym)'
/jarut/ from janggut 'beard'
/ñan/ 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'
/pni/ 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: /gajah/ from 'gajah 'elephant', /bilang/ 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  /tghe/    'oriental pied hornbill'
*twan /tgwn/    'blue coral snake'
*bo? /b?b?/  'to carry on one's back'
*ploŋ /ploŋ/    '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). 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
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 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. 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 /kjej/ 'to listen', the underlying structure of which is CCVC, and its disyllabic imperfective form /kjejjej/ 'to be listening'.

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

   \[
   \begin{array}{c|c|c|c|c|c}
   & & & & & \\
   C & C & V & C & \rightarrow & C & <C> & C & V & C \\
   k & j & e & j & & k & j & e & j \\
   \end{array}
   \]

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

   \[
   \begin{array}{c|c|c|c|c|c}
   & & & & & \\
   C & <C> & C & V & C & \rightarrow & k<\eta>jej \\
   k & j & e & j & & k & j & e & j \\
   \end{array}
   \]

3. **Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: */k<\eta>jej/ → */k<\eta>jej/ → /kok<\eta>gej/.[/]

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 /gulem/ ‘to carry’, the underlying structure of which is CVCVC, and its imperfective form /gu<ml>em/ ‘to be carrying’:

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

(2)  
\[
\begin{array}{|c|c|c|c|c|}
\hline
C & V & C & V & C \\
\hline
\end{array}
\rightarrow
\begin{array}{|c|c|c|c|c|}
\hline
C & V & <C> & C & V \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{gu} & \text{l} & \text{e} & \text{m} & \text{u} \\
\hline
\end{array}
\]

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:

\[
\begin{array}{|c|c|c|c|c|}
\hline
C & V & <C> & C & V \\
\hline
\end{array}
\rightarrow
\begin{array}{|c|c|c|c|c|}
\hline
\text{gu}<\text{m}>\text{lem} & \text{u} & \text{e} & \text{m} & \text{u} \\
\hline
\end{array}
\]

iii. Apply rules of syllabification in order to produce the correct surface form: /gu<ml>lem/ → /gu<ml>lem/ → [gumleЬm]

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. 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 /bakes/ ‘to grow up’ and its disyllabic imperfective form /b<s>kes/ ‘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)  
\[
\begin{array}{|c|c|c|c|c|}
\hline
C & V & C & V & C \\
\hline
\end{array}
\rightarrow
\begin{array}{|c|c|c|c|c|}
\hline
C & <C> & C & V & C \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{b} & \text{a} & \text{k} & \text{e} & \text{s} \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{b<s>} & \text{kes} & \text{b} & \text{a} & \text{k} & \text{e} & \text{s} \\
\hline
\end{array}
\]

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:

\[
\begin{array}{|c|c|c|c|c|}
\hline
C & <C> & C & V & C \\
\hline
\end{array}
\rightarrow
\begin{array}{|c|c|c|c|c|}
\hline
\text{b<s>kes} & \text{b} & \text{k} & \text{e} & \text{s} \\
\hline
\end{array}
\]

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

---

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:

\[
\begin{array}{c}
\text{C C C V C} \\
\text{k r l e p}
\end{array} \quad \rightarrow \quad \begin{array}{c}
\text{C C <C> C V C} \\
\text{k r l e p}
\end{array}
\]

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:

\[
\begin{array}{c}
\text{C C <C> C V C} \\
\text{k r l e p}
\end{array} \quad \rightarrow \quad \text{kr<p>lep}
\]

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /kr<p>lep/ \rightarrow /kr_r<p>.lep/ \rightarrow [kэрр'lep'].

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 /pis/ ‘to sweep’ and its disyllabic imperfective form /ps-pis/ ‘to be sweeping’.

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

\[
\begin{array}{c}
\text{C V C} \\
\text{p i s}
\end{array} \quad \rightarrow \quad \begin{array}{c}
\text{C C - C V C} \\
\text{p i s}
\end{array}
\]

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:

\[
\begin{array}{c}
\text{C C - C V C} \\
\text{p i s}
\end{array} \quad \rightarrow \quad \text{ps-pis}
\]
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.

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)  
\[
\begin{array}{cc}
C & V \\
| & |
\end{array}
\rightarrow
\begin{array}{cc}
C & C \\
| & |
\end{array}
C
\]

\[
\begin{array}{cc}
te?\
| & |
\end{array}
\rightarrow
\begin{array}{cc}
n & t & e\
| & | & |
\end{array}
\]

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

The same procedure may be applied to the monosyllabic verb root /pel/ 'to drip' and its disyllabic iterative form /liː-pel/ '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:

\[
\begin{array}{c}
\text{CVC} \\
| | | \\
\text{pel} \\
\end{array} \rightarrow \begin{array}{c}
\text{CVC} \\
| | | \\
\text{pel} \\
\end{array}
\]

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:

\[
\begin{array}{c}
\text{pel} \\
| | | \\
\end{array} \rightarrow \begin{array}{c}
\text{pel} \\
| | | \\
\end{array}
\]

iii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /liː-pel/ → /liː-.pel/ → [ləlpe]\[l\].

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 /tboːh/ 'to beat' and its trisyllabic nominalised form /t˂nh>boːh/ '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:

\[
\begin{array}{c}
\text{CVC} \\
| | | \\
\text{tboːh} \\
\end{array} \rightarrow \begin{array}{c}
\text{CVC} \\
| | | \\
\text{tboːh} \\
\end{array}
\]
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:

\[
\begin{array}{c}
\text{b c h} \\
\text{C} \quad \text{<C} \quad \text{C>} \quad \text{C} \quad \text{V} \quad \text{C} \\
\text{t n} \quad \text{b c h}
\end{array}
\rightarrow \text{t<nh>bch}
\]

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

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 /bdel/ ‘to shoot’ and its disyllabic affixed form /ml-del/:

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:

\[
\begin{array}{c}
\text{b d e l} \\
\text{C} \quad \text{C} \quad \text{V} \quad \text{C} \\
\text{m} \quad \text{d e l}
\end{array}
\rightarrow \text{ml-del}
\]

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:

\[
\begin{array}{c}
\text{d e l} \\
\text{C} \quad \text{C} \quad \text{V} \quad \text{C} \\
\text{m} \quad \text{d e l}
\end{array}
\rightarrow \text{ml-del}
\]

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

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 /sjør/ ‘to swim’ and its distributive form /s<ir>jør/ ‘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:

\[
\begin{align*}
(10) & \quad \text{C V C V} \quad \rightarrow \quad \text{C V C} \\
& \quad \text{s j a r} \quad \text{s i j a r}
\end{align*}
\]

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:

\[
\begin{align*}
(11) & \quad \text{C V C V} \quad \rightarrow \quad \text{C V C} \\
& \quad \text{t a n e m} \quad \text{t i n e m}
\end{align*}
\]

iii. Apply rules of syllabification in order to produce the correct surface form: /s<ir>jər/ \(\rightarrow\)/s<ir>.jər/ \(\rightarrow\) [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 /tanem/ ‘to plant’, a Malay loanword, and its distributive form /t<im>nem/ ‘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:

\[
\begin{align*}
(10) & \quad \text{C V C V} \quad \rightarrow \quad \text{C V C} \\
& \quad \text{t a n e m} \quad \text{t i n e m}
\end{align*}
\]

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:

\[
\begin{align*}
(11) & \quad \text{C V C V} \quad \rightarrow \quad \text{C V C} \\
& \quad \text{t i n e m}
\end{align*}
\]

iii. Apply rules of syllabification in order to produce the correct surface form: /t<im>nem/ \(\rightarrow\)/t<im>.nem/ \(\rightarrow\) [timnem].

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>lem/, from /gulem/ ‘to carry’, and /t<in>dun/, from /tudun/ ‘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 §3.2.1.2. It is illustrated here with the monosyllabic verb root /tel/ ‘to look’ and its disyllabic distributive form /til-?el/ ‘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

? e l i ? e 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:

? e l

| | |
C V C - C V C

i ? e l

?il-?el

iii. Apply rules of syllabification in order to produce the correct surface form: /til-?el/ → /til-.?el/ → [?il?el].

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:

\[
\begin{array}{c}
C \ V \ C \ \rightarrow \ \underline{C \ V - C \ V \ C} \\
\underline{g \ e \ j} \ \rightarrow \ \underline{a \ g \ e \ j}
\end{array}
\]

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:

\[
\begin{array}{c}
C \ V - C \ V \ C \ \rightarrow \ \underline{g a - g e j} \\
a \ g \ e \ j
\end{array}
\]

iii. Apply rules of syllabification in order to produce the correct surface form: \(/g a - g e j/ \rightarrow /g a - g e j/ \rightarrow [g a g e j].\)

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:

\[
\begin{array}{cccccccc}
C & C & C & V & C & \rightarrow & C & <C> & C & V & C \\
\mid & \mid & \mid & \mid & \mid & \rightarrow & \mid & \mid & \mid & \mid & \mid \\
t & m & k & a & l & & t & m & a & k & a & l
\end{array}
\]

\(\rightarrow tm<a>kal\)

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /tm<a>kal/ \(\rightarrow /t.m<a>.kal/ \rightarrow [t\text{maka}kal]\).

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 /smejn/ 'to ask', the underlying structure of which is CCVC, and its reciprocal form /s<a>mejn/ '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:

\[
\begin{array}{cccccccc}
C & C & V & C & \rightarrow & C & <V> & C & V & C \\
\mid & \mid & \mid & \mid & \mid & \rightarrow & \mid & \mid & \mid & \mid & \mid \\
s & m & e & j & n & & s & a & m & e & j & n
\end{array}
\]

\(\rightarrow s<a>mejn\)

ii. Apply rules of syllabification in order to produce the correct surface form: /s<a>mejn/ \(\rightarrow /s<a>.mejn/ \rightarrow [sam\text{ejn}]\).

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:

\[
\begin{array}{cccccccc}
C & V & C & V & C & \rightarrow & C & <C> & V & C & V & C \\
\mid & \mid & \mid & \mid & \mid & \rightarrow & \mid & \mid & \mid & \mid & \mid & \mid \\
b & a & b & o & ? & & b & n & a & b & o & ?
\end{array}
\]

\(\rightarrow b<n>abo?\)

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /b<n>abo?/ \(\rightarrow /b.<n>a.bo?/ \rightarrow [m\text{anabo}?]\).

Affixation of bases with a heavy penultimate syllable is illustrated here with the morphologically complex but synchronically unanalysable verb /ckwik/ 'to talk', where the
coda of the penultimate syllable consists of a fossilised instance of coda copy, and its nominalised form /c<n>kwik/ '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:

\[
\begin{array}{cccccc}
C & C & C & V & C & \rightarrow \\
| & | & | & | & | & \\
c & k & w & i & k & \\
\end{array}
\]

\[
\begin{array}{cccccc}
C & C & C & V & C & \rightarrow \\
| & | & | & | & | & \\
c & n & k & w & i & k & \\
\end{array}
\]

\[
\text{c<n>kwik}
\]

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /c<n>kwik/ → /c.<n>k.wik/ → [c<anak'wik'].

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 /ksep/ 'to adorn oneself with leaves for good luck' and its nominalised form /k<n>sep/ '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:

\[
\begin{array}{cccccc}
C & C & V & C & \rightarrow \\
| & | & | & | & | & \\
k & s & e & p & & \\
\end{array}
\]

\[
\begin{array}{cccccc}
C & C & V & C & \rightarrow \\
| & | & | & | & | & \\
k & n & s & e & p & \\
\end{array}
\]

\[
k<n>sep
\]

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

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 /<ni>/, 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 /tboh/ 'to hit' and its nominalised collective plural form /t<ra>boh/ 'fighters'.

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 /<ni>/, 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 /tboh/ 'to hit' and its nominalised collective plural form /t<ra>boh/ 'fighters'.
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 — /fpi-f, fpr-f and ftr-f (see §4.7.2) — as well as the rare allomorph /br-f 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 /fpr-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:

\[
\begin{array}{c}
\text{(21) } \\
\text{C V C } \rightarrow \text{ C C - C V C } \\
\hline
\text{h i t} \\
\text{p r h i t}
\end{array}
\]

ii. Apply rules of syllabification and vowel epenthesis in order to produce the correct surface form: /pr-hit/ → /pr-.hit/ → [pərhit'].
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-/ and causative /p-/ and may be attached to the left edge of sesquisyllabic 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'
/bde:lI/ 'to shoot' /l-bde:lI/ 'to shoot repeatedly'
/hgik/ 'to be afraid' /p-hgik/ 'to frighten'

(b) Disyllabic bases:

/kan?r/ 'to carry on shoulder' /b-kan?r/ 'to be carrying on shoulder'
/bj-baj/ 'to be digging' /b-bj-baj/ 'to be digging on'
/sa?ot/ 'to call someone' /l-sa?ot/ 'to call someone repeatedly'

3.2.2.2 Flexible outer affixes

Flexible outer affixes include relative-marking /t-/ and some instances of verbalising /b-/. It 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 sesquisyllabic 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'
/b?w/ 'to be big' /t-b?w/ 'big'
/t?m/ 'water' /b-t?m/ 'to have water'

(b) Trisyllabic base:

/b-tado?/ 'to be waiting' /t-b-tado?/ '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 /I/ 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 /I/, 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.
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 [-dn] 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 /ɾ/ 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 C(V). 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: [gịn]. 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-ŋpk
   IRR=PROG-IMPF-to.sit
   'I will be sitting.'

(b) wa= b-nasi?
   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.
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=hip
    SOURCE=forest
    'from the forest'

(b) can=jh0?
    SOURCE=tree
    'from the tree'

(c) cn=hawên
    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 (/\textipa{lja}=/ 'IRR', /\textipa{wa}=/ 'IRR.3S'), desiderative (/\textipa{lma}=/ 'DES') and hortative (/\textipa{lha}=/, /\textipa{lca}=/, /\textipa{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:

\[
\text{[(...) clitic=V (...)]} \text{CLAUSE}
\]

The following clauses exemplify the pattern:

(a) \textipa{wa=pi-?em} \textipa{wəŋ} \textipa{?o?}
    IRR.3S=CAUS-to.drink child 3S
    'She will suckle her baby.'

(b) \textipa{ja=çêφ} \textipa{ka=je?}
    IRR=to.fan.fire SUBJ=1S
    'I will fan the fire.'
Chapter 3

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=1/ 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=cip   ba=75h
     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=1 'GOAL'), source (/can=1 'SOURCE'), contrast (/d= 'CONTR') and equation (/pn=1 '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=wəŋ   ḳjih je? ṭ5h
     GOAL=child boy IS this
     'to this son of mine'

(b)  1=nej   ?5h
     ID=one here
     'the one here'

3.3.2.3 Clitics hosted by the first constituent of the core

The proclitic signalling relational tense (/ja=1 '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)  ja=je?  ?5et
     RT=IS to.know
     'I already know.'
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=moh ?t?et**  
Q=2S.FAM to.know  
‘Do you know?’

(b) **ha=wa=cip**  
Q=IRR.3S=to.go  
‘Is he leaving?’

(c) **ha=bt?et**  
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=hapa? kto? wej**  
REL=to.die day past  
‘who died yesterday’

(c) **k=pn=?³h**  
REL=EQU=here  
‘that [is] like this’
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).
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:

<table>
<thead>
<tr>
<th>Root</th>
<th>Nominal meaning</th>
<th>Verbal meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>/?ap/</td>
<td>'large feline'</td>
<td>'to encounter a large feline'</td>
</tr>
<tr>
<td>/?ec/</td>
<td>'belly; excrement'</td>
<td>'to defecate'</td>
</tr>
<tr>
<td>/ptis/</td>
<td>'pain; sickness'</td>
<td>'to be in pain; to be sick'</td>
</tr>
<tr>
<td>/was/</td>
<td>'fork; confluence'</td>
<td>'to split'</td>
</tr>
<tr>
<td>/?em/</td>
<td>'breast'</td>
<td>'to drink'</td>
</tr>
<tr>
<td>/knom/</td>
<td>'urine'</td>
<td>'to urinate'</td>
</tr>
<tr>
<td>/len/</td>
<td>'loincloth'</td>
<td>'to wear a loincloth'</td>
</tr>
<tr>
<td>/snp/</td>
<td>'silence'</td>
<td>'to be silent'</td>
</tr>
</tbody>
</table>

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:
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:

\[
/\text{hlak}k\text{et}/\quad \text{('a type of ant')}
\]
\[
/\text{klut}b\text{ot}/\quad \text{('a type of larva')}
\]
\[
/\text{kicbc}k/\quad \text{('a type of millipede')}
\]
\[
/\text{klict}k/\quad \text{('a type of terrestrial gastropod')}
\]

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

\[
/\text{krtl}s\text{t}/\quad \text{('kidney')}
\]
\[
/\text{smult}s\text{t}/\quad \text{('brain')}
\]
\[
/\text{dkd}k\text{uk}/\quad \text{('chest')}
\]
\[
/\text{s}k\text{so}\text{t}/\quad \text{('blood vessel')}
\]
\[
/\text{kmk}k\text{m}/\quad \text{('ankle')}
\]
\[
/\text{cqch}n/\quad \text{('area around the eyes')}
\]
\[
/\text{dld}l/\quad \text{('heel')}
\]
\[
/\text{prber}/\quad \text{('lower arm')}
\]
\[
/\text{jwjo}w/\quad \text{('Achilles tendon')}
\]

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
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. ?ap ?awej
   large.feline vine
   'leopard'

2. kit tom
   buttocks river
   'river mouth'

3. mit kto?
   eye sky
   'sun'

4. kdek ?abu?
   squirrel dust
   'Provost's squirrel'

5. ?em kaji?
   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. ?ikə? bawuŋ
   fish type.of.catfish
   '(a type of catfish)'

7. taŋu? sjul
   snake type.of.cobra
   '(a type of cobra)'

8. tom taŋu?
   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 /tkih/ 'backside', /kjсим/ 'underside', /seŋ/ 'front', /krpiŋ/ 'upper side', /kleŋ/ 'inside', /?naŋ/ 'outside' and /sirt/ 'side'. Examples are given in (9) and (10).

9. krpiŋ haje?
   upper.side house
   'upper side of house'/'above the house'
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<ar>kal/ ‘men’
/?ane?/ ‘girl’ /?ra>ne?/ ‘girls’
/kjih/ ‘boy’ /k<ra>jih/ ‘boys’
/bakes/ ‘adult’ /b<ra>kes/ ‘adults’
/bnswa?/ ‘daughter/son-in-law’ /bn<ar>swa?/ ‘daughters/sons-in-law’
/lami?/ ‘sister/brother-in-law’ /l<ra>mij/ ‘sisters/brothers-in-law’
/mnra?/ ‘person’ /mn<ar>ra?/ ‘persons’
/jahai/ ‘Jahai’ /j<ra>hai/ ‘Jahais’
/ples/ ‘Temiar’ /p<ra>les/ ‘Temis’
/knsiw/ ‘Kensiw’ /kn<ar>siw/ ‘Kensiws’
/knta?/ ‘Kintaq’ /kn<ar>ta?/ ‘Kintaqs’
/batek/ ‘Batek’ /b<ra>tek/ ‘Bateks’
/smej/ ‘Semai’ /s<ra>mej/ ‘Semais’
/tmwen/ ‘Temuan’ /tm<ar>wen/ ‘Temuans’

Collective-like forms of three non-human nouns have also been recorded:
Note that collective plural allomorphs associated with monosyllabic bases have not been documented.

From Schebesta (1928a:810-811) and onwards, the /<ra>/<a/> morpheme of Northern Asian 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. /he? /<ra>haji / 'we Jahai', /gin b<r<a>bo?1 / '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

/tmpot/ 'place' /tmpot-tmpot/ 'various (kinds of) places'
/pagi?1/ 'morning' /pagi?1-pagi?1/ 'various (kinds of) mornings'
/kritoh/ 'car' /kritoh-kritoh/ 'various (kinds of) cars'

Indigenous forms

/mnra?1/ 'people' /mnra?-mnra?1/ 'various (kinds of) people'
/ilmol15/ 'mountain' /ilmol-ilmol/ 'various (kinds of) mountains'
/ken/ 'child' /ken-ken/ 'various (kinds of) children'

4.1.3 Unitisation

Nouns modified by a numeral, the quantifier /km/ 'many' or the interrogative /mej si?/ '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.

This is possibly not an indigenous Jahai form but a borrowing of Temiar /ilmol/ 'mountain' (/jelmol/ 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.

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

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

The following examples illustrate the use of unitised nouns:

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(13)</td>
<td>After three days I will move to Kelap.</td>
</tr>
<tr>
<td>(14)</td>
<td>Don’t stay in one place.</td>
</tr>
<tr>
<td>(15)</td>
<td>How many mountains?</td>
</tr>
<tr>
<td>(16)</td>
<td>He gave away one piece of meat.</td>
</tr>
<tr>
<td>(17)</td>
<td>Two pieces of wood is nothing.</td>
</tr>
</tbody>
</table>

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 /nej nm-tom/ 'one [unit of] water' (from tom/ '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 unredrived 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).

| /bani/? | 'quiver' | /b<n>ani/? | '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:  
/cip/   'to go'      /np-cip/  'act of going'
/sam/   'to hunt'   /nm-sam/  'act of hunting'
/col/   'to tell'   /nl-col/  'act of telling'

Sesquisyllabic:  
/jhit/  'to smoke'  /j<n>hit/  'act of smoking'
/tbuh/  'to beat'   /t<nh>boh/ 'act of beating'
/rnj/   'to bathe'  /r<nj>naj/ 'act of bathing'

Disyllabic:  
/cwik/  'to talk'   /c<n>kwik/ 'act of talking'
/kaijl/ 'to fish'    /k<n>ajil/ 'act of fishing'
/pikirl/ 'to think'  /p<n>ikirl/ '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) \( c<n>\text{kwik} \ ?o? \ \text{ton lajin-lajin}, \ \text{bra? ja=?t?et} \)
\( \text{to.speak<NM> 3S that CONT-to.be.different NEG IRR=to.know} \)
‘That way of speaking of his is very different. I don’t understand [it].’

(19) \( \text{l}s\ \text{nk-jok japeh wek} \)
\( \text{after NM-to.move IP.EXCL to.go.back} \)
‘After moving we went back.’

(20) \( \text{tmpot nk-jok japeh pdæh} \)
\( \text{place NM-to.sit IP.EXCL to.be.near} \)
‘Our place of residence is nearby.’

(21) \( \text{slamat nk-wek bt?et} \)
\( \text{safety NM-to.go.back to.be.good} \)
‘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>
<thead>
<tr>
<th>ACTION</th>
<th>OBJECT</th>
<th>INSTRUMENT</th>
<th>LOCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>/np-cip/ from /cip/ ‘to go’</td>
<td>‘act/manner of going’</td>
<td>–</td>
<td>‘route’ ‘path’</td>
</tr>
<tr>
<td>/nc-kec/ from /kec/ ‘to cut’</td>
<td>‘act/manner of cutting’</td>
<td>‘thing cut’ ‘circumcised person’</td>
<td>‘cutter’ ‘knife’ ‘place of cutting’</td>
</tr>
<tr>
<td>/nm-com/ from /com/ ‘to burn’</td>
<td>‘act/manner of burning’</td>
<td>‘fuel’</td>
<td>‘fire’ ‘swidden’</td>
</tr>
<tr>
<td>/k&lt;n&gt;ajiU from /kajil/ ‘to fish’</td>
<td>‘act/manner of fishing’</td>
<td>‘fishing catch’ ‘fishing rod’</td>
<td>‘fishing place’</td>
</tr>
<tr>
<td>/h&lt;n&gt;aluh from /haluh/ ‘to shoot with blowpipe’</td>
<td>‘act/manner of shooting’</td>
<td>‘game’ ‘quarry’</td>
<td>‘blowpipe’ ‘hunting place’</td>
</tr>
</tbody>
</table>

Table 4.1: Examples of meanings of nominalisations in Jahai
The meanings of some concrete nominalisations may become very narrow and specific, such as the object interpretation of the nominalisation /nc-kec/ 'thing cut' (from /kecl/ '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.

/\hj\jut/ 'to be heavy' /\hj<a>jut/ 'heavy people'
/\hgi\k/ 'to be afraid' /h<ra>gik/ 'frightened people'
/\cr\j\l/ 'to be hungry' /c<ra>\ro\l/ 'hungry people'
/\k\b\s/ 'to be dead/to die' /k<ra>b\s/ 'dead people'
/\pti\s/ 'to be sick' /p<ra>tis/ 'sick people'
/\man\e\h/ 'to be old' /m<ra>\neh/ 'old people'
/\p\ca\h/ 'to break' /p<ra>\ca\h/ 'breakers'
/\t\b\h/ 'to hit' /t<ra>\bo\h/ '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:

/\wi\l/ 'left' /\m\?-\wi\l/ '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.
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.

<table>
<thead>
<tr>
<th>Numeral</th>
<th>Collectivised Numeral</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>/duwa?/</td>
<td>/d&lt;&gt;wa?/</td>
<td>'[group of] two'</td>
</tr>
<tr>
<td>/tiga?/</td>
<td>/ti&lt;&gt;ga?/</td>
<td>'[group of] three'</td>
</tr>
<tr>
<td>/spulohl</td>
<td>/spu&lt;&gt;loh/</td>
<td>'[group of] ten'</td>
</tr>
<tr>
<td>/sblas/</td>
<td>/sb&lt;s&gt;las/</td>
<td>'[group of] eleven'</td>
</tr>
</tbody>
</table>

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: /ken/ 'child' for human nouns, and /kmɔ? - kbi?/ 'fruit' for all non-human animate and many inanimate nouns. 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 /kmɔ? - kbi?/, 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 /kmɔ?/ is in free variation with /kbi?/, which is possibly a loan of Temiar /kabɔʔ/, which also means 'fruit' (Benjamin, pers. comm.).
### Table 4.2: Classifiers in Jahai

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Original meaning</th>
<th>Class characteristics</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indigenous forms</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ken</td>
<td>'child'</td>
<td>human</td>
<td>all human nouns</td>
</tr>
<tr>
<td>kmo? ~ kbi?</td>
<td>'fruit'</td>
<td>non-human</td>
<td>animals, fruit, various objects</td>
</tr>
<tr>
<td>tom</td>
<td>'tree'</td>
<td>vegetation</td>
<td>trees, plants, grass</td>
</tr>
<tr>
<td>hali?</td>
<td>'leaf'</td>
<td>small flat objects</td>
<td>leaves, paper sheets, notes</td>
</tr>
<tr>
<td>tunk3l</td>
<td>'stone'</td>
<td>spherical/cubical objects</td>
<td>stones</td>
</tr>
<tr>
<td>mit</td>
<td>'eye'</td>
<td>small objects</td>
<td>seeds, stones</td>
</tr>
<tr>
<td><strong>Forms borrowed from Malay</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bidan (from bidang)</td>
<td>'broad/flat piece'</td>
<td>large flat objects</td>
<td>mats, blankets, fields</td>
</tr>
<tr>
<td>bataq (from batang)</td>
<td>'tree-trunk', 'stick', 'shaft'</td>
<td>oblong objects</td>
<td>trees, pencils, cigarettes</td>
</tr>
<tr>
<td>kpiI (from keping)</td>
<td>'portion', 'piece'</td>
<td>flat objects</td>
<td>plates, helpings of food</td>
</tr>
<tr>
<td>biji? (from biji)</td>
<td>'seed'</td>
<td>small objects</td>
<td>seeds, stones</td>
</tr>
<tr>
<td>buwah (from buah)</td>
<td>'fruit'</td>
<td>spherical/cubical objects</td>
<td>houses</td>
</tr>
<tr>
<td>?ikor (from ekor)</td>
<td>'tail'</td>
<td>animal</td>
<td>all animal nouns</td>
</tr>
<tr>
<td>prdu? (from perdu)</td>
<td>'base of tree-trunk'</td>
<td>clusters of objects</td>
<td>clusters, bunches, banana stems</td>
</tr>
</tbody>
</table>

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) `je? b?bo? tomen, duwa? k<n?>mø? IS to.carry.on.back snakehead two CLF<UNIT> 'I carried snakeheads. Two of them.'

(23) `ja=je? bdil ?amøn k=tom maŋøh ?on leh, RT=IS to.shoot siamang LOC=river Mangga there EMP duwa? k<n?>mø? two CLF<UNIT> 'Then I shot siamangs by Mangga river! Two of them.'
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) je? bdil spuloh k<n?>m<> kasa?
   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.

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

/\?<n?>ikɔr/ from /\?ikɔr/ CLF: animal
/p<\?>rdu?/ from /prdu?/ 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>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Dual</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inclusive</td>
<td>Exclusive</td>
<td>Inclusive</td>
</tr>
<tr>
<td>1</td>
<td>je?</td>
<td>hej</td>
<td>jeh</td>
</tr>
<tr>
<td>2</td>
<td>Intimate</td>
<td>Familiar</td>
<td>Distant</td>
</tr>
<tr>
<td></td>
<td>mi?</td>
<td>moh</td>
<td>paj</td>
</tr>
<tr>
<td>3</td>
<td>?o?</td>
<td>wih</td>
<td></td>
</tr>
</tbody>
</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>
<thead>
<tr>
<th></th>
<th>Singular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>to father-in-law</td>
</tr>
<tr>
<td>2</td>
<td>gin</td>
</tr>
<tr>
<td>3</td>
<td>gin</td>
</tr>
</tbody>
</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 k'en/ ‘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 'je?'/ ‘my child’, /wɔŋ je?/ ‘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 unstered pronouns when these occur as preverbal subject agreement markers. This applies only to pronouns with a final glottal stop /ʔ/, which is usually dropped in this position: [je ~ feʔ] /jeʔ/ '1S', [ʔo ~ ʔoʔ] /ʔoʔ/ '3S', [he ~ hęʔ] /heʔ/ 'IP 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 unstered form of /gin/ '23P'. The unstered 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).

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.

/ʔəh/ ‘here’
/ʔon/ ‘there (you know)’
/ʔən/ ‘there (you don’t know)’
/ʔaniʔ/ ‘there (away)’
/ʔadeh/ ‘there (beyond me)’
/ʔniʔ/ ‘there (beyond you)’
/ʔitiʔ/ ‘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 (/ca=/) and similarity to (/pn=/) the location designated by the demonstrative (see also §4.9). The following examples illustrate such constructions:

(26) ja=wek ba=ʔəh leh
IRR=to.go.back GOAL=here EMP
‘I will return here!’

(27) ca=ʔon jeʔ jok ba=puʔəw tuʔoh
SOURCE=there IS 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 /japəh/ is in free variation with a reduced variant /pəh/ in the preverbal subject agreement position.
e.g. /k=t̪a/h/ ‘(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=t̪a/h/ ‘(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) kuciʔ ʔâɨn
cat there
‘the cat over there’

(30) slaj heʔ ʔâh
swidden 1S here
‘our swidden here’

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

4.4.1.1 Speaker-anchored accessible /ʔâh/

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 /ʔân/!

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 /ʔâniʔ/

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 /ʔân/!

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* /tadeh/

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* /tηī/ 

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* /tiih ~ tōih/

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 /tiih ~ tōih/ are in free variation and idiolectally determined.

4.4.1.8 *Subjacent* /tujih/

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 /l/ to form such nominal demonstratives.

- /tēlh/ 'this'
- /tōn/ 'that (you know)'
- /tūn/ 'that (you don’t know)'
- /taniʔ/ 'that (away)'
- /tadeh/ 'that (beyond me)'
- /tηīʔ/ 'that (beyond you)'
- /titih/ '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) hajēʔ tūn
    house that
    'that house'

(32) jhūʔ titih
    tree that
    'that [up] tree'
Occasionally nominal demonstratives occur pronominally as NP heads:

(33) ja=gej tɔh

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 /wej/ '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 wej

British past

'those British [of the past]'

(35) ?ej je? wej

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:

<table>
<thead>
<tr>
<th>Number</th>
<th>Jahai</th>
<th>Malay</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>nej</td>
<td>lapan</td>
</tr>
<tr>
<td>2</td>
<td>duwa?</td>
<td>smilan</td>
</tr>
<tr>
<td>3</td>
<td>tiga?</td>
<td>spuloh</td>
</tr>
<tr>
<td>4</td>
<td>?mpat</td>
<td>sblas</td>
</tr>
<tr>
<td>5</td>
<td>lime?</td>
<td>duwa? blas</td>
</tr>
<tr>
<td>6</td>
<td>nem</td>
<td>sratos</td>
</tr>
<tr>
<td>7</td>
<td>tu?oh</td>
<td>sribuh</td>
</tr>
</tbody>
</table>

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

(36) nej t<n>mkal

one man<UNIT>

'one man'

19 Diffloth (1976c:33) suggests that the only Mon-Khmer numeral in Jahai, /nej/ 'one', is a loan of Temiar /nej/ 'one' (or /ney/ in his transcription), the common Northern Aslian reflex being /nej/. However, Benjamin (1976a:113) gives the differing form /nej/ (or /ney/) for Temiar, and a sound shift of final syllable /a/ to /e/ appears to be a regular process that distinguishes Jahai from other Northern Aslian languages, so Diffloth’s suggestion must be regarded as uncertain.
Chapter 4

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

(38) spuloh nn-ken
ten UNIT-CLF
‘ten people’

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

(39) ja=gej nej
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 /nej/ ‘one’ may take the causative prefix /pi-/ (cf. §4.7.2.2) to signal ‘causation of oneness’, the resulting form /pi-nej/ meaning ‘one each’, ‘one by one’, ‘one at a time’, ‘one after another’ or the like. This is illustrated by the following example:

(42) o? oek pi-nej pi-nej k<n?><mo?
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: /bml/ ‘many’ and /pEW/ ‘other’. The former frequently co-occurs with a unitised noun or classifier (see §4.1.3 and §4.2.1).

(43) bml nh-boh
many UNIT-fruit
‘many fruits’

(44) pEW mnra?
other person
‘other person’

Other quantifiers are Malay loans and include /smwe?/ ‘all’ (from Malay semua ‘all’), /tiap-tiap/ ‘every’ (from Malay tiap-tiap ‘every’) and /masiŋ-masiŋ/ ‘each’ (from Malay masing-masing ‘each’). These are exemplified below.

(45) smwe? ?3t
all dog
‘all dogs’
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 /maken/

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

(48) ka=puni huluh ba=?ani2, maken
SUBJ=headdress GOAL=there who?
‘Who is the headman over there?’

(49) maken hajé?
whose? house
‘whose house?’

The reduplication /maken-maken/ represents an indefinite form meaning ‘whoever’, ‘anybody’.

4.6.2 Time-questioning /mapu2/?

The form /mapu2/ is a time-questioning interrogative and corresponds to English ‘when?’. There is no known root */pui/.

(50) cn=mapu?
SOURCE=when?
‘Since when?’

4.6.3 Item/situation/reason-questioning /mamej ~ mej/

The interrogative /mamej ~ mej/ 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) mamej k=hŋut
what? REL=to.be.heavy
‘What [is it that] is heavy?’

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

(53) mamej mōh jīm
why? 2S.FAM to.cry
‘Why do you cry?’

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

(54) mej si? nn-ken
what? number UNIT-CLF
‘How many people?’

The reduplication /mej-mej/ 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 mōh de?
how? 2S.FAM to.do
‘How do you do it?’

4.6.5 Location-questioning /lbahl/

The interrogative /lbahl/ 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-/.

(56) b=lbahl mi? ja=jok
GOAL=where? 2S.INT IRR=to.move
‘Where will you move?’

(57) lbahl paj ŋok
where? 2S.DIS to.sit
‘Where did you live?’

The reduplication /lbahl-lbahl/ 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:

/hchoc/ ‘to whistle’
/sksɛk/ ‘to devour’
/jkṣik/ ‘to breathe’
/ckwik/ ‘to talk’
/pkpśk/ ‘to smack’
/lkluk/ ‘to laugh’
/tkjok/ ‘to pluck’
/luʔleʔ/ ‘to roll’
/bʔbɔʔ/ ‘to carry on one’s back’
/tsdes/ ‘to stumble’
/lslɔs/ ‘to gnaw’
/kmjil/ ‘to taste’
/pnʔ calibrated ‘to hop’
/hpʃjan/ ‘to stand up’
/pfjamp/ ‘to sing’
/pŋʃen/ ‘to say’
/sʃʃen/ ‘to walk with a limp’
/knʃen/ ‘to flap wings’
/tanʃwoŋ/ ‘to carry on one’s shoulder’
/pnjum/ ‘to play a flute’
/krket/ ‘to yell’

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

/rtʔet/ ‘to know’
/hmʃɛm/ ‘to like’
/jnʃen/ ‘to dream’

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

/btʔet/ ‘to be good’
/bclac/ ‘to be smooth’
/jŋʃen/ ‘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'
- /?t?st/ 'to stroke an animal' /?st/ 'dog'
- /smɔm/ 'to buzz around a nest' /sɔm/ 'nest'
- /klniʔ/ 'to make sound' /klniʔ/ '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) japəh cip ktɔ? wej
1P.EXCL to.go day past
'We went yesterday.'

(59) jeʔ hŋik ba=taʔuʔ?
1S to.fear GOAL=snake
'I'm afraid of snakes.'

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

(61) hɔj
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=cip
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 kapij
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=cip
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. Some examples of property verbs follow:

```
/tmket/ 'to be cold'
/bkit/ 'to be hot'
/cnhâ/ 'to be short'
/btec/ 'to be long'
/pcê?/ 'to be wet'
/la?is/ 'to be bad'
/pdõh/ 'to be near'
/kriñ/ 'to be dry'
/ghel/ 'to be tired'
/bow/ 'to be big'
/ rêj/ '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.
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). 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

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Derivational morpheme} & \text{Monosyllabic} & \text{Sesquisyllabic} & \text{Disyllabic} \\
& /cip/ 'to go' & /kjen/ 'to listen' & /gulem/ 'to carry' \\
\hline
\text{Aspect/Aktionsart} (§4.7.1) & \text{Imperfective} & \text{cp-cip} & \text{k<j>jen} & \text{gu<m>lem} \\
& \text{Progressive} & \text{b-cpcip} & \text{b-kjen} & \text{b-gulem} \\
& \text{Iterative} & \text{lp-cip} & \text{l-kjen} & \text{l-gulem} \\
& \text{Continuative} & \text{cip-cip} & \text{kjen-kjen} & \text{gulem-gulem} \\
& \text{Distributive} & \text{cip-cip} & \text{k<j>jen} & \text{g<im>lem} \\
& \text{Reciprocal} & \text{ca-cip} & \text{k<a>jen} & - \\
& \text{Affix /m/} & \text{mp-cip} & \text{mjen (?)} & \text{mm-lem (?)} \\
& \text{Causative} (§4.7.2) & \text{pp-cip, pi-cip,} & \text{k<r>jen} & \text{g<r>lem} \\
& & \text{pr-cip, tr-cip} & & \\
& \text{Nominalisation} (§4.1.4.1) & \text{np-cip} & \text{k<njen} & \text{g<n>ulem} \\
& \text{Collective plural nomin.} (§4.1.4.2) & - & \text{k<r>a>jen} & \text{g<ra>lem (?)} \\
& \text{Relative} (§4.12.2) & \text{t-cip (?)} & \text{t-kjen (?)} & \text{t-gulem (?)} \\
\hline
\end{array}
\]

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.\(^{23}\)

\[
\begin{array}{ll}
/sut/ & 'to sob' \\
/k\text{ec}/ & 'to cut' \\
/we\text{c}/ & 'to exist' \\
/c\text{ip}/ & 'to hiss' \\
/p\text{is}/ & 'to sweep' \\
/s\text{am}/ & 'to hunt' \\
/r\text{in}/ & 'to sit down' \\
/t\text{r}/ & 'to plait' \\
/p\text{en}/ & 'to chop' \\
/d\text{or}/ & 'to balance' \\
/j\text{aw}/ & 'to wash' \\
/l\text{jo}/ & 'to run' \\
/p\text{kip}/ & 'to scorch' \\
/k\text{wac}/ & 'to swim' \\
/h\text{o}/ & 'to ask' \\
/b\text{rik}/ & 'to brake' \\
/b\text{li}/ & 'to buy' \\
/r\text{wis}/ & 'to cut grass' \\
/k\text{nom}/ & 'to urinate' \\
/s\text{mep}/ & 'to ask' \\
/k\text{jen}/ & 'to listen' \\
/b\text{del}/ & 'to blowpipe' \\
/c\text{j}/ & 'to sing' \\
/k\text{rej}/ & 'to forget' \\
/j\text{awap}/ & 'to answer' \\
/s\text{iruc}/ & 'to slurp' \\
/b\text{akes}/ & 'to grow up' \\
/t\text{urch}/ & 'to tap poison' \\
/g\text{ulem}/ & 'to carry' \\
/t\text{igil}/ & 'to go around'
\end{array}
\]

/st-sut/ 'to sobbing' /k\text{c}-k\text{ec}/ 'to cutting' /w?\text{-we}/ 'to existing' /c\text{ip}-c\text{ip}/ 'to hissing' /p\text{is}-p\text{is}/ 'to sweeping' /s\text{am}-s\text{am}/ 'to hunting' /r\text{in}-r\text{in}/ 'to sitting down' /t\text{r}-t\text{r}/ 'to plaiting' /p\text{en}-p\text{en}/ 'to chopping' /d\text{or}-d\text{or}/ 'to balancing' /j\text{aw-jaw}/ 'to washing' /l\text{jo-jo}/ 'to running' /p\text{kip-kip}/ 'to scorching' /k\text{wac-kwac}/ 'to swimming' /t\text{o}-t\text{o}/ 'to asking' /b\text{rik-brik}/ 'to braking' /b\text{li-bli}/ 'to buying' /r\text{wis-rwis}/ 'to cutting grass' /k\text{nom-knom}/ 'to urinating' /s\text{mep-smep}/ 'to asking' /k\text{jen-kjen}/ 'to listening' /b\text{del-bdel}/ 'to blowpiping' /c\text{j-cj}/ 'to singing' /k\text{rej-kej}/ 'to forgetting' /j\text{awap-jawap}/ 'to answering' /s\text{iruc-siruc}/ 'to slurping' /b\text{akes-bakes}/ 'to growing up' /t\text{urch-turch}/ 'to tapping poison' /g\text{ulem-gulem}/ 'to carrying' /t\text{igil-tigil}/ 'to 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-/ /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) ʔ oʔ  k<>< wac  hajiʔ  tɔm  taniʔ
     3S  to.swim<IMPF>  also  river  that
‘He was also swimming in that river.’

(66) ʔ oʔ  tk-tek  d=ʔej  ʔ oʔ  ton
     3S  IMPF-to.sleep  CONTR=father  3S  that
‘That father of his was sleeping.’

(67) gin  bj-baj  k=haʔeʔ  leh
     2/3P  IMPF-dig  LOC=house  EMP
‘They were digging by the house!’

(68) ʔa=jeʔ  cip  leh  d=jeʔ  ton,  jeʔ  dj-duj
     RT=1S  to.go  EMP  CONTR=IS  that  IS  IMPF-to.hunt
‘I [for my part] had already left. I was hunting.’

(69) ʔ oʔ  ks-kas  leh  ka=knʔac  ʔ oʔ  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.

/cip/  ‘to go’  /b-cpcip/  ‘to be going’
/kot/  ‘to take’  /b-ktot/  ‘to be taking’
/tek/  ‘to sleep’  /b-ttek/  ‘to be sleeping’
/weʔ/  ‘to exist’  /b-wʔweʔ/  ‘to be existing’
/caraʔ/  ‘to talk’  /b-caraʔ/  ‘to be talking’
/?imbusl/  ‘to ambush’  /b-ʔimbusl/  ‘to be ambushing’
Ikdihl 'to say' /b-kdihl/ 'to be saying'
Ib-lh 'to enter' /b-blh/ 'to be entering'
Ijm/ 'to cry' /b-jm/ 'to be crying'
I/ 'to cover' /b-dnd/ 'to be covering'
I-sam/ 'to put together' /b-sam/ 'to be putting together'
I-bdel/ 'to blowpipe' /b-bdel/ 'to be blowpiping'
I-kar/ 'to carry on shoulder' /b-kar/ 'to be carrying ...'
I-baj/ 'to dig' /b-baj/ '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'
/ek/ 'to give' /bk-ek/ 'to be giving'
/jok/ 'to move' /bk-jok/ 'to be moving'
/nok/ 'to sit' /bk-nok/ 'to be sitting'
/cah/ 'to cut' /bh-cah/ 'to be cutting'
/el/ 'to look' /bl-el/ 'to be looking'
/aj/ 'to run' /bj-aj/ '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) je? sšh kej kawip, ?o? b-pk-pok
1S to.encounter such bear 3S PROG-IMPF-to.stride
'I saw such a bear! It was striding along.'

(71) he? b-jk-jok ba=pšw hajē?
1P.INCL PROG-IMPF-to.move GOAL=other house
'We were moving on to another house.'
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) je? b-ch-cih slaj, japēh b-tanem padej
     1S PROG-IMPF-to.cut swidden IP.EXCL PROG-to.plant rice
     'I was clearing a swidden. We were planting rice.'

(77) je? ja=b-rihat leh d=je?, ja=b-tk-tek
     1S IRR=PROG-to.rest EMP CONTR=IS 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) japēh b-ŋk-ŋk k=slaj he? ʔəh wej
     IP.EXCL PROG-IMPF-to.sit LOC=swidden IP.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) ʔəs ton japēh b-jumpaʔ paj
     after that IP.EXCL PROG-to.meet 2S.DIST
     'After that we ran into you.'

(80) ʔo? b-kʔek ba=je?
     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.

\[
\begin{align*}
/cip/ & \quad \text{‘to go’} & /lp-cip/ & \quad \text{‘to go repeatedly’} \\
/hok/ & \quad \text{‘to throw’} & /lk-hok/ & \quad \text{‘to throw repeatedly’} \\
/pel/ & \quad \text{‘to drip’} & /ll-pel/ & \quad \text{‘to drip repeatedly’} \\
/kul/ & \quad \text{‘to call’} & /ll-kul/ & \quad \text{‘to call repeatedly’} \\
/gej/ & \quad \text{‘to eat’} & /lj-gej/ & \quad \text{‘to eat repeatedly’} \\
/sa?ot/ & \quad \text{‘to call someone’} & /l-sa?ot/ & \quad \text{‘to call someone repeatedly’} \\
/kdih/ & \quad \text{‘to say’} & /l-kdih/ & \quad \text{‘to say repeatedly’} \\
/bdel/ & \quad \text{‘to shoot’} & /l-bdel/ & \quad \text{‘to shoot repeatedly’}
\end{align*}
\]

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) \(\begin{array}{l}
\text{?o? It-sot} \quad \text{?o? It-but}, \quad \text{?o? sot,} \quad \text{?o? but,} \quad \text{?o? sot,} \quad \text{?o? but} \\
3S \quad \text{ITER-to.cut} \quad 3S \quad \text{ITER-to.eat} \quad 3S \quad \text{to.cut} \quad 3S \quad \text{to.eat} \quad 3S \quad \text{to.cut} \\
\text{‘He cut [repeatedly] and ate [repeatedly]. He cut, he ate. He cut, he ate.’} \\
\text{[Of a man who carved up meat and ate the pieces successively]}
\end{array}\)

(82) \(\begin{array}{l}
\text{?o? ll-?el} \quad k=jam \\
3S \quad \text{ITER-to.look} \quad \text{LOC=watch} \\
\text{‘He looked at his watch now and then.’}
\end{array}\)

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) \(\begin{array}{l}
pagi?-pagi? \quad japêh \quad lr-pir \\
\text{DP-morning} \quad \text{IP.EXCL ITER-to.court} \\
\text{‘We courted in the mornings.’}
\end{array}\)

(84) \(\begin{array}{l}
ha=môh \quad lj-gej \quad k=tâh \\
Q=2S.FAM \quad \text{ITER-to.eat} \quad \text{LOC=this} \\
\text{‘Do you usually eat here?’} \\
\text{[Uttered by a man who joined the author at a food stall in Grik]}
\end{array}\)
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:

```
/lc-?ec/ 'to defecate' /lc-?ec/ 'to defecate little by little'
/lk-wek/ 'to return' /lk-wek/ 'to return by stages'
/lm-?em/ 'to drink' /lm-?em/ 'to drink little by little'
/lj-baj/ 'to dig' /lj-baj/ 'to dig a little at a time'
/l-ckwik/ 'to talk' /l-ckwik/ 'to talk now and then'
/l-kje/ 'to listen' /l-kje/ 'to listen now and then'
```

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

(85) hej lc-?ec, bokan hej ja=?ec 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) ja=je? lp-cip leh ba=?sh
RT= lS 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 Aislian 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 Aislian 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:

```
/cep/ ‘to catch’ /cep-cep/ ‘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’
/ŋok/ ‘to sit’ /ŋok-ŋok/ ‘to keep on sitting’
/ŋampiʔ/ ‘to heal’ /ŋampiʔ-ŋampiʔ/ ‘to keep on healing’
/kwēs/ ‘to sweep’ /kwēs-kwēs/ ‘to keep on sweeping’
/blah/ ‘to enter’ /blah-blah/ ‘to keep on entering’
/jim/ ‘to cry’ /jim-jim/ ‘to keep on crying’
/tanem/ ‘to plant’ /tanem-tanem/ ‘to keep on planting’
```
It may also involve reduplication of derived forms such as imperfectives, iteratives and distributives. However, progressive forms do not reduplicate to form continuatives.

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/.

<table>
<thead>
<tr>
<th>Base</th>
<th>Distributive Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>/cip/</td>
<td>/cip-cip/ 'to go [here and there]'</td>
</tr>
<tr>
<td>/lwec/</td>
<td>/l&lt;ic&gt;we&lt;24/ 'to climb up [here and there]'</td>
</tr>
<tr>
<td>/vak/</td>
<td>/v&lt;ik&gt;-vak/ 'to give [here and there]'</td>
</tr>
<tr>
<td>/jok/</td>
<td>/j&lt;ik&gt;-jok/ 'to move [here and there]'</td>
</tr>
<tr>
<td>/njok/</td>
<td>/n&lt;ik&gt;-njok/ 'to sit [here and there]'</td>
</tr>
<tr>
<td>/swapuh/</td>
<td>/s&lt;ih&gt;puh/ 'to sweep [here and there]'</td>
</tr>
<tr>
<td>/boh/</td>
<td>/b&lt;ih&gt;-boh/ 'to put [here and there]'</td>
</tr>
<tr>
<td>/tboh/</td>
<td>/t&lt;ih&gt;-tboh/ 'to hit [here and there]'</td>
</tr>
<tr>
<td>/tanem/</td>
<td>/t&lt;im&gt;nem/ 'to plant [here and there]'</td>
</tr>
<tr>
<td>/kom/</td>
<td>/kim-kom/ 'to be many [here and there]'</td>
</tr>
<tr>
<td>/kriñ/</td>
<td>/k&lt;in&gt;-riñ/ 'to be dry [here and there]'</td>
</tr>
<tr>
<td>/kjen/</td>
<td>/k&lt;in&gt;-jen/ 'to listen [here and there]'</td>
</tr>
<tr>
<td>/bdil/</td>
<td>/b&lt;il&gt;dil/ 'to shoot [here and there]'</td>
</tr>
<tr>
<td>/tel/</td>
<td>/t&lt;il&gt;-tel/ 'to look [here and there]'</td>
</tr>
<tr>
<td>/jol/</td>
<td>/j&lt;il&gt;-jol/ 'to throw [here and there]'</td>
</tr>
<tr>
<td>/col/</td>
<td>/c&lt;il&gt;-col/ 'to tell [here and there]'</td>
</tr>
<tr>
<td>/sjari/</td>
<td>/s&lt;ir&gt;jar/ 'to swim [here and there]'</td>
</tr>
<tr>
<td>/bow/</td>
<td>/biw-bow/ 'to be big [here and there]'</td>
</tr>
</tbody>
</table>

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=jik-jok haden
      1.P.INCL IRR=DISTR-to.move tomorrow
      ‘We will move [in different directions] tomorrow.’

---

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<ic>we<24/ [lic"weç’].
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 /ril-?el/ (from /?el/ ‘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) mnrar ror hit-hût can=?lmol
person 3S DISTR-to.reveal.oneself SOURCE=mountain
‘The person revealed himself [here and there] on the mountain.’

(96) ?o? ñik-ŋak pn=?äh
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) ?o? lil-?el 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-hapl ‘to disperse’
/sih-sähpl ‘to meet [each other]’
/ñim-?aml ‘to embrace [each other]’
/∫<iw>niw/ ‘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’
/cəl/ ‘to tell’ /ca-cəl/ ‘to tell each other’
/gej/ ‘to eat’ /ga-gej/ ‘to eat together’

Sesquisyllabic:
/sme⁹/ ‘to ask’ /saɑ>meŋ/ ‘to ask each other’
/bdiU/ ‘to shoot’ /b<ɑ>dil/ ‘to shoot at each other’
/naj/ ‘to bathe’ /n<ɑ>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 Asian 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 [a] 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 ancien
tessness and importance to a story. Its deviant morphophonemic characteristics may support such an interpreta
tion. Still, it appears to be largely productive within its restricted domain. The following forms have been recorded:

\[
\begin{align*}
/\text{mp-cip}/ & \quad /\text{cip}/ \quad \text{‘to go’} \\
/\text{mt-get}/ & \quad /\text{get}/ \quad \text{‘to cut’} \\
/\text{mk-wek}/ & \quad /\text{wek}/ \quad \text{‘to go back’} \\
/\text{m?-ji}/ & \quad /\text{ji}/ \quad \text{‘to burn’} \\
/\text{m?-?i}/ & \quad /\text{?i}/ \quad \text{‘to pour’} \\
/\text{m?-bo}/ & \quad /\text{bo}/ \quad \text{‘to carry on one’s back’} \\
/\text{ms-des}/ & \quad /\text{des}/ \quad \text{‘to move’} \\
/\text{ms-les}/ & \quad /\text{les}/ \quad \text{‘to smear poison on blowpipe dart’} \\
/\text{ms-gos}/ & \quad /\text{gos}/ \quad \text{‘to carve’} \\
/\text{ms-was}/ & \quad /\text{was}/ \quad \text{‘to split’} \\
/\text{mh-?oh}/ & \quad /\text{?oh}/ \quad \text{‘to heat’} \\
/\text{mh-boh}/ & \quad /\text{boh}/ \quad \text{‘to put’} \\
/\text{mh-coh}/ & \quad /\text{coh}/ \quad \text{‘to bite’} \\
/\text{mm-?em}/ & \quad /\text{?em}/ \quad \text{‘to drink’} \\
/\text{mn-fin}/ & \quad /\text{fin}/ \quad \text{‘to take’} \\
/\text{ml-pel}/ & \quad /\text{pel}/ \quad \text{‘to cook’} \\
/\text{ml-del}/ & \quad /\text{del}/ \quad \text{‘to shoot’} \\
/\text{ml-?el}/ & \quad /\text{?el}/ \quad \text{‘to look’} \\
/\text{ml-pal}/ & \quad /\text{pal}/ \quad \text{[meaning unknown]} \\
/\text{ml-gol}/ & \quad /\text{gol}/ \quad \text{‘to carry’} \\
/\text{mr-?or}/ & \quad /\text{?or}/ \quad \text{‘to order’} \\
/\text{mj-gej}/ & \quad /\text{gej}/ \quad \text{‘to eat’} \\
/\text{mj-haj}/ & \quad /\text{haj}/ \quad \text{‘to follow’} \\
\end{align*}
\]

The following examples illustrate its use:

\[(98) \quad \text{je? mh-coh ?o?} \\
\quad \text{IS M-to.bite 3S} \\
\quad \text{‘I bit him.’} \]

\[(99) \quad \text{ja=je? ms-was} \\
\quad \text{RT=1S M-to.split} \\
\quad \text{‘Then I split [X].’} \]

/\text{m}/-affixed forms feed further derivational morphology in the form of progressive (§4.7.1.2), as shown by the following example: /b-mr-?or/ from /?or/ ‘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-cipl/, /pi-cipl/, /pr-cipl/ and /tr-cipl/ from /cip/ ‘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 /<cri>/ (see §4.7.2.5). With monosyllabic bases, the allomorph consists of a CC prefix composed of the prespecified Ipl 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:

<table>
<thead>
<tr>
<th>Base</th>
<th>Affix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>/cip/</td>
<td>/p-</td>
<td>‘to go’</td>
</tr>
<tr>
<td>/tek/</td>
<td>/pi-</td>
<td>‘to give’</td>
</tr>
<tr>
<td>/ji/</td>
<td>/jip-</td>
<td>‘to burn’</td>
</tr>
<tr>
<td>/c3l/</td>
<td>/pi-c3l/</td>
<td>‘to teach someone to speak’</td>
</tr>
<tr>
<td>/gej/</td>
<td>/pi-gej/</td>
<td>‘to feed’</td>
</tr>
</tbody>
</table>

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-?naj/ ‘to wash’ from /?naj/ ‘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:

<table>
<thead>
<tr>
<th>Base</th>
<th>Affix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>/cip/</td>
<td>/pi-</td>
<td>‘to go’</td>
</tr>
<tr>
<td>/kap/</td>
<td>/pi-</td>
<td>‘to bite’</td>
</tr>
<tr>
<td>/muc/</td>
<td>/pi-</td>
<td>‘to eat’</td>
</tr>
<tr>
<td>/tek/</td>
<td>/pi-</td>
<td>‘to sleep’</td>
</tr>
<tr>
<td>/wek/</td>
<td>/pi-</td>
<td>‘to go back’</td>
</tr>
<tr>
<td>/gis/</td>
<td>/pi-</td>
<td>‘to descend’</td>
</tr>
</tbody>
</table>
Word classes

The following examples illustrate the use of causative /pi-/:

(100) 3S CAUS-to.drink CONTR=child 3S
‘She suckled her baby.’

(101) RT=3S CAUS-to.descend water
‘Then he poured out the water.’

(102) 3S CAUS-to.go.back tail 3S this
‘He lifted its tail.’

(103) 2/3P CAUS-to.bathe other bamboo.tube other bamboo.tube other
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:

The following examples illustrate the use of causative /pr-/:

(104) 3S CAUS-to.eat GOAL=wife 3S that
‘He supported that wife of his.’

(105) 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:

\[
\begin{array}{lll}
/cip/ & 'to go' & /tr-cip/ & 'to train someone to walk' \\
/gim/ & 'to deliberate' & /tr-gim/ & 'to summon to a deliberation' \\
/loj/ & 'to run' & /tr-loj/ & 'to tell someone to run away'
\end{array}
\]

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:

\[
\begin{array}{lll}
/cip/ & 'to close one's eyes' & /c<ri>kip/ & 'to close someone's eyes' \\
/bkit/ & 'to be hot' & /b<ri>kit/ & 'to heat' \\
/hgik/ & 'to be afraid' & /h<ri>gik/ & 'to frighten' \\
/sju/ & 'to be cold' & /s<ri>ju/? & 'to cool something' \\
/kbis/ & 'to die/be dead' & /k<ri>bis/ & 'to kill' \\
/maneh/ & 'to be old' & /m<ri>neh/ & 'to make something old' \\
pkah/ & 'to snap/break' & /p<ri>kah/ & 'to break something' \\
/bloj/ & 'to be red' & /b<ri>loj/ & 'to make something red'
\end{array}
\]

The following examples illustrate the /<ri>/ allomorph:

(106) \[\text{ja}=\text{je? } k<ri>bis ?o? \]
RT=1S to.die<CAUS> 3S
'Then I killed it.'

(107) \[\text{he? } ja=p<ri>cah \text{ tawes} \]
IP.INCL IRR=to.break<CAUS> [type.of.tree]
'We will fell a tawes 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).

| /te?/ | 'earth/soil' | /b-te? ~ br-te?/ | 'to own/till land' |
| /ha?e?/ | 'house' | /b-ha?e?/ | 'to reside in a house' |
| /?ika?/ | 'fish' | /b-?ika?/ | 'to have fish' |
| /krit?/ | 'car' | /b-krit?/ | 'to be motorised' |
| /tom/ | 'water' | /b-tom/ | 'to have/use/contain water' |
| /hawen/ | 'pig' | /b-hawen/ | 'to have a pig/pork' |
| /twip/ | 'headache' | /b-twip/ | 'to have a headache' |
| /blawl/ | 'blowpipe' | /b-blawl/ | 'to have/use a blowpipe' |

The following examples illustrate its use:

(108) ?o? b-kne?h
3S PROP-wife
'He was married.'

(109) ?o? b-slaj ka?=u?n
3S PROP-swidden LOC=there
'He tilled a swidden there.'

(110) bra? wa=b-nasi?
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 /nej/ ‘one’ appears not to take this affix. Again, close parallels are found in Semelai (Kruspe 2004:152) and Perak Malay (Zaharani 1991:88–89).

| /duwa?/  | ‘two’     | /b-duwa?/  | ‘to be two’ |
| /tiga?/  | ‘three’   | /b-tiga?/  | ‘to be three’ |
| /?mpat/  | ‘four’    | /b-?mpat/ | ‘to be four’ |
| /lime?/  | ‘five’    | /b-lime?/ | ‘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:
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=cip ba=hip
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=knom
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 /bra/ (which is in free variation with the borrowed Malay negative marker /bukan/) and the irrealis-marked verb.

(120) c<n>kwik ?o? ton lajin-lajin, bra? ja=þt?et
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) bra? wa=mUrk lagi? ka=þap ton
NEG IRR.3S=to.wake.up again SUBJ=tiger that
'That tiger never woke up again!'

(122) bra? wa=lr-gir ka=þun
NEG IRR.3S=ITER.to.roll LOC=there
'There was no thunder.'

(123) bra? 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.

/cip/ 'to go' /ma=cip/ '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) japêh ma=pimic baraŋ japêh
IP.EXCL DES=to.take.back thing IP.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) kti? ma=hic
sky DES=to.rain
'It will rain.' (lit. 'The sky wants to rain. ')

A similar desiderative prefix /ma?-/ 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 procliticated 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.

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>/cip/</td>
<td>'to go'</td>
<td>/ha=cip/ 'please go'</td>
</tr>
<tr>
<td>/wek/</td>
<td>'to go back'</td>
<td>/ha=wek/ 'please go/come back'</td>
</tr>
<tr>
<td>/kjen/</td>
<td>'to listen'</td>
<td>/ka=kjen/ 'please listen'</td>
</tr>
<tr>
<td>/?el/</td>
<td>'to look'</td>
<td>/ca=?el/ 'please look'</td>
</tr>
</tbody>
</table>

Examples include the following:

(126) ha=cip ba=?5h
      HORT=to.go GOAL=here
      'Please come here!'

(127) wek ha=?el
to.go.back HORT=to.look
      'Come back and have a look!'

(128) ca=?el gamah ?op?
      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 Kruşpe (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:

\[
\begin{align*}
/\text{rop}/ & \quad \text{('sound of a muntjac deer')}
/\text{bat}/ & \quad \text{('sound of a bird landing on a branch')}
/\text{cik}/ & \quad \text{('sound of a blowpipe dart hitting muscle of prey')}
/\text{klık}/ & \quad \text{('sound of a person walking')}
/\text{klak}/ & \quad \text{('sound of something falling')}
/\text{pok}/ & \quad \text{('sound of an animal falling to the ground')}
/\text{chok}/ & \quad \text{('sound of running water or waterfall')}
/\text{kæk}/ & \quad \text{('sound of a hornbill')}
/\text{hokşk}/ & \quad \text{('sound of a leaf-monkey')}
/\text{ʔok}/ & \quad \text{('sound of boiling')}
/\text{riϕ}/ & \quad \text{('sound of flapping')}
/\text{japh}/ & \quad \text{('sound of flying or leaping')}
/\text{luϕ}/ & \quad \text{('sound of dashing')}
/\text{koph}/ & \quad \text{('sound of a blowpipe dart hitting stomach of prey')}
/\text{køϕ}/ & \quad \text{('sound of a small object falling to the ground')}
/\text{tøϕ}/ & \quad \text{('sound of a blowpipe dart hitting canopy')}
/\text{plɛʃ}/ & \quad \text{('sound of a blowpipe being fired')}
/\text{plɛs}/ & \quad \text{('sound of a blowpipe being fired')}
/\text{čeh}/ & \quad \text{('sound of a blowpipe dart hitting a vine')}
/\text{weh}/ & \quad \text{('sound of a blowpipe being fired')}
/\text{kran}/ & \quad \text{('sound of carving')}
/\text{pøw}/ & \quad \text{('sound of a shotgun')}
/\text{pɾəw}/ & \quad \text{('sound of an animal fleeing from one tree to another')}
\end{align*}
\]

In some cases there is a correlation between size/speed and vowel quality, as illustrated by the following set of forms:

\[
\begin{align*}
/\text{tik}/ & \quad \text{('sound of a small raindrop falling to the ground')}
/\text{tak} \sim /\text{tik}/ & \quad \text{('sound of a big raindrop falling to the ground')}
/\text{tak}/ & \quad \text{('sound of a big raindrop falling quickly to the ground')}
/\text{tuk} \sim /\text{tɔk}/ & \quad \text{('sound of a big raindrop falling slowly to the ground')}
\end{align*}
\]

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:
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.

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) luφ ka=bhəl
[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):
Chapter 4

(133) japêh ƞok k=klap
IP.EXCL to.sit LOC=Kelap
‘We stayed at Kelap.’

(134) je? p?êh k=bulô?
IS to.cook LOC=bamboo.tube
‘I cooked [it] in bamboo tubes.’

A more metaphorical reading is exemplified in the following sentence:

(135) ?o? ?el k=?awê? 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) je? b-pqsen ka=?ûn
IS PROG-to.talk LOC=there
‘I was talking over there.’

(137) cip ka=sep
to.go LOC=front
‘Walk in front!’

It is also used to mark the thematic role of instrument.

(138) je? cek ?aj ka=tâji?
IS 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) je? lkluk ka=je?
IS to.laugh SUBJ=IS
‘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) japêh ja=wek ba=hajê?
IP.EXCL IRR=to.go.back GOAL=house
‘We will go back to the house.’

It may also denote metaphorical motion to a recipient:

(141) ?o? ?ek dwi? ba=je?
3S to.give money GOAL=IS
‘He gave me money.’

The phonologically reduced form /b=/ is only found in combination with the demonstratives /?itih/ ‘there, above’ and /?ujih/ ‘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) je? jok can=moh cos tani?
1S to.move SOURCE=headwaters Cos that
'I moved from those headwaters of the river Cos.'

(143) ?o? ?el=?el ka=?ap cn=?ani?
3S to.look-CONT SUBJ=tiger SOURCE=there
'The tiger kept on watching from over there.'

(144) sec cn=kasa?
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) hej ja=kajil d=hej
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) ja=?o? bdil leh d=hawen
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) ?ek leh d=je?
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:
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 /ja=/ 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)  \( ja=je? \) ?t\( ?eh \)
RT=1S to.know
‘I already know.’

(151)  ha=ja=m\( \bar{oh} \) bdil ka=mutah
Q=RT=2S.FAM to.shoot INSTR=mortar
‘Have you ever used a mortar?’

(152)  ja=s\( \bar{oc} \) ka=sn\( \bar{oc} \)
RT=to.be.gone SUBJ=blowpipe.dart
‘The darts are already used up!’

(153)  hobi? ton ja=bp\( \bar{ji} \)
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)  ja=lo? bj ba=?ani?
RT=3S to.run GOAL=there
‘Then it fled in that direction.’
(155) je? wk-wk, ja=je? ges k=slaj
IS 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) ja=wa=soc tep tāh
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 /to/? to form a portmanteau morpheme /jo/? (RT.3S).

(157) jo? cip
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 je? cip
RP IS 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 /bra/? (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) bra? wa=bt?et
NEG IRR.3S=to.be.good
'It's not good.'

(161) bokan ja=b-ŋkŋk kampon
NEG IRR=PROG-IMPF-to.sit village
'I wasn’t living in a village.'
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 /bnit/, 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 /?aket/ (PROH) is placed before a verb to express negative imperative.

(163) ?aket lkluk
     PROH to.laugh
     ‘Don’t laugh!’

(164) ?aket rɔh bɔw
     PROH to.cut to.be.big
     ‘Don’t make a big clearing.’

(165) ?aket de? pn=ʔon
     PROH to.make EQU=that
     ‘Don’t do like that!’

(166) ?aket 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) ?aket
     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=1/ (Q) to the first constituent of the clause.

(168) ha=ja=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=?ûn
Q=IRR.3S=to.go.back GOAL=there
‘Will he return there?’

(171) ha=moh ja=haj tkih je?
Q=2S.FAM IRR=to.follow backside IS
‘Will you follow me?’

4.10.4 Emphasis

A particle /leh/ (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) japeh ŋok k=tûh leh
1P.EXCL to.sit LOC=this EMP
‘We stayed here!’

(173) pêh ŋok leh k=ton
1P.EXCL to.sit EMP LOC=that
‘We stayed there!’

(174) bra? wa=?ûic leh
NEG IRR.3S=to.burn EMP
‘He didn’t burn [X]!’

Affirmative use of /leh/ is exemplified in the following answer:

(175) ?o? leh
3S EMP
‘[It’s] him all right.’

If combined with imperative forms, /leh/ adds additional emphasis to the exhortation:

(176) cip leh
to.go EMP
‘Now, go!’

(177) ?ek leh d=je?
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 /leh/ described in §4.10.4.

(178) ?akет rûh bow
PROH to.cut to.be.big
‘Don’t make a big clearing.’
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 /lktwat/ 'to be quick', /lwes/ 'to be wide' and /la?is/ 'to be bad'.

### 4.10.6 Other adverbial elements

Additional adverbials include the following set of indigenous, postverbal forms:

- /haji?/ ‘also’, ‘too’
- /jaga?/ ‘only’
- /s?o?/ ‘just’, ‘only’
- /wel/ ‘again’
- /nuj/ ‘soon’

Also, the following forms borrowed from Malay occur frequently, usually clause-initially:

- /banu?/ ‘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 /?alo?/ ‘and’. These may be used to co-ordinate full clauses or phrases, typically noun phrases. The additive /?alo?/ 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)  
\[
\text{ja=cip ba=grik ?alo? ja=bli? bras}  
\text{IRR=to.go GOAL=Grik and IRR=to.buy rice}  
\text{‘I will go to Grik and I will buy rice.’ (elic.)}
\]

(182)  
\[
\text{ha=moh ha=je?}  
\text{or=2S.FAM or=1S}  
\text{‘you or me’}
\]
Subordinating conjunctions, of which three have been documented, are borrowings from Malay: a causal /sbap/ ‘because’ (from Malay sebab ‘cause’), a temporal /lpas/ ‘after’, ‘when’ (from Malay lepas ‘after’), and a conditional /kaluh ~ kalow/ ‘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 /lpas/, 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=1/ (§4.12.1) and a prefix /t-I/ (§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=1/

The relative marker /k=1/ (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) punhuhluh  k=hapaʔ  headman REL=to.die  
‘the headman who died’

(185) jaʔ  ʔoʔ  t̪ah  k=lawaʔ  ?ikan  grandmother 3S this REL=to.collect fish  
‘this grandmother of his, who caught fish’

(186) maken kolah  k=pnʔʔh  whose cup REL=like=here  
‘Whose cup [is it] that [looks] like this?’

(187) slaj  k=wih  rh-ɾoh  swidden REL=3D IMPF-to.clear  
‘the swidden that they were clearing’

Sometimes, the relative marker /k=1/ is replaced or supplemented by /jaʔ/, a borrowing of the Malay relative marker yang:
**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)  
\[
\begin{array}{l}
\text{mnra? t-kul} \\
\text{person REL-to.call} \\
\text{‘calling person’}
\end{array}
\]

(191)  
\[
\begin{array}{l}
\text{ken t-?a\(\)o?} \\
\text{child REL-to.be.small} \\
\text{‘small child’}
\end{array}
\]

(192)  
\[
\begin{array}{l}
\text{?o? t-b\(\)w} \\
3S REL-to.be.big \\
\text{‘the big one’} \\
[Avoidance term denoting elephant]
\end{array}
\]

(193)  
\[
\begin{array}{l}
\text{?o? t-des} \\
3S REL-to.move \\
\text{‘the moving one’}
\end{array}
\]

(194)  
\[
\begin{array}{l}
\text{?ap t-b-tad\(\)o?} \\
tiger REL-PROG-to.wait \\
\text{‘waiting tiger’}
\end{array}
\]

The following phrases exemplify the relative marker /t-/ with nouns and numerals:

(195)  
\[
\begin{array}{l}
\text{he? t-mnra?} \\
1P.INCL REL-person \\
\text{‘we humans’}
\end{array}
\]

(196)  
\[
\begin{array}{l}
\text{wih t-kjih ka=?\(\)un} \\
3D REL-boy LOC=there \\
\text{‘those [two] boys’}
\end{array}
\]
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) tæh, l=ksij ?o? ?ʌh
this ID=husband 3S here
‘This [is] her husband here.’

(199) tæh, l=nej ?ʌh
this ID=one here
‘this one here’

(200) ton, l=tahoh ?on
that ID=Tahoh there
‘That [is] Tahoh there.’

(201) tahanen, l=kneh ?o?
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 ?o?
ID=upper.side 3S
‘its upper side’

(203) l=seŋ ?o?
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=?ʌh
LOC=ID=here
‘here’

(205) k=la=?on
LOC=ID=there
‘there’
Given its affirmative and identifying function, it is likely that the identification marker bears some relationship to the phrase-final emphatic and affirmative particle /leh/, 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.
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) ?o? loj  
  3S to.run  
  'He ran.'

(2) ?o? hic  
  3S to.rain  
  'It rains.'

(3) je? hgik  
  1S to.be.afraid  
  'I'm afraid.'

(4) hej pek  
  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.
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.

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.

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).

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.

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) je? ?eck ba=moh soret
    IS 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 Permutated arguments

Two processes of argument permutation may alter the basic order of elements outlined in §5.1.1.1. 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.

    3S to.cut SUBJ=grandchild rattan
    'The grandchild cut the rattan.'

    3S to.go SUBJ=father IS past GOAL=here
    'My late father came here.'

(17) wih ja=c:l ka=wih ?os
    3D IRR=to.burn SUBJ=3D fire
    'They will make a fire.'

(18) je? ji? d=je?
    IS to.refuse CONTR=IS
    'I refuse.'

(19) je? lkluk ka=je?
    IS to.laugh SUBJ=IS
    'I laughed.'
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) bli? ?o? ?o? gantoŋ
    leg 3S 3S to.hang
    ‘He hung up its leg.’

(23) ?ikɔ? je? gj-gej
    fish 1S IMPF-to.eat
    ‘I was eating fish.’

5.1.1.3 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=tase?
    IRR=to.fish GOAL=lake
    ‘[I] will go fishing in the lake.’

(25) je? sāh ?ap kto? wej
    1S to.encounter tiger day past
    ‘I saw a tiger yesterday.’

(26) je? bdeŋ plēs
    1S to.shoot [sound of blowpipe]
    ‘I shot [sound]!’

(27) ?o? paŋkoh pn=?on ?aweŋ manɔw
    3S to.hold EQU=there vine manau
    ‘He held the manau vine like this.’

(28) lime? kto? japēh ŋok b=?itih
    five day IP.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).

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, je? bdil duwa? kbi?
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.

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=lawa? t?a?, t?a? paku?
IRR=to.collect edible.plant edible.plant paku
‘[I] will collect plants, paku plants.’

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) je? bdel ?ameŋ, duwa? k<n?>bi?
    IS to.shoot siamang two CLF<UNIT>
    'I shot siamangs, two of them.'

    IS 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).

5.1.1.5 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. /baw/ ‘to be big’, /hoj/ ‘to be small’ and /b?etl/ ‘to be good’), and (2) the existential verbs /soc/ ‘to be gone’ and /we?/ ‘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? bów
body 2S.INT to.be.big
‘Your body is big.’

(37) ghol ka=je?
to.be.tired SUBJ=1S
‘I’m tired.’

(38) pinej we?
betelnut to.exist
‘There are betelnuts.’

(39) we? ka=lwej
to.exist SUBJ=honey
‘There is honey.’

(40) soc ka=gin gop
to.be.gone SUBJ=2/3P stranger
‘Them strangers are gone.’

Examples (41) and (42) illustrate the same pattern, but now with an onomatopoeic form as predicating element.

(41) luφ ka=bhol
[sound] SUBJ=muntjac.deer
‘The muntjac deer [sound of dashing].’

(42) pow 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.’
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=qhel
    IRR=to.be.tired
    '[I] will be tired.'

(46) ?aj wa=s>3c
    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 /bra?/, which occurs in free variation with the borrowed negative marker /bukan/ (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) je? bra? ja=?t?et
    1S NEG IRR=to.know
    'I don’t know.'

(48) bra? he? ja=kje? kli3n krit?3
    NEG IP.INCL IRR=to.hear sound car
    'We didn’t hear the sound of cars.'

(49) kampo3 bra? wa=we?
    village NEG IRR.3S=to.exist
    'There was no village.'

(50) bra? ja=?em teh
    NEG IRR=to.drink tea
    '[I] didn’t drink tea.'

(51) bra? wa=bl?h ka=pe? 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=?o?
    NEG IRR.3S=to.eat CONTR=3S
    ‘He [as opposed to X] didn’t eat.’

(53) bra? japeh ja=pkt.;k d=gin kdek
    NEG IP.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 /ja=/ (§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 /lpos/ ‘after’/‘when’, and conditional /kaluh ~ kalow/ ‘if’. These may introduce either the initial or the final clause of such series.

(54) ja=gej sbap je? cr;?,
    IRR=to.eat because IS to.be.hungry
    ‘I will eat because I’m hungry.’ (elic.)

(55) lpos je? gej je? cip
    after IS to.eat IS to.go
    ‘I left when I had eaten.’

(56) kalow je? cr;?, ja=gej
    if IS 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=m=oh ?t;et
    Q=2S.FAM to.know
    ‘Do you know?’
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) maken but kmɔ?
who? to.eat fruit
‘Who ate the fruit?’ (elic.)

(62) mej paj d?-de?
what? 2S.DIS IMPF-to.do
‘What are you doing?’

(63) Ibah ?o? nɔk
where? 3S to.sit
‘Where does he live?’

(64) b=lbah mɔh cip
GOAL=where 2S.FAM to.go
‘Where did you go?’

(65) mapu? mɔh ja=wek
when? 2S.FAM IRR=to.go.back
‘When will you go back?’

Person-questioning /maken/ ‘who?’ and item-questioning /mej/ ‘what/which?’ typically replace subjects, direct objects or oblique arguments, time-questioning /mapu?/ ‘when?’ and location-questioning /lbah/ ‘where?’ typically replace temporal and locative adjuncts respectively, manner-questioning /ma?acin/ ‘how?’ replaces manner-related adverbial adjuncts, and reason-questioning /mej/ ‘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.
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) punhuluh ba=?ani? maken
    headman GOAL=there who?
    ‘Who is headman over there?’

(71) b=lbah ka=siput
    GOAL=where? SUBJ=snail
    ‘Where is the snail?’

(72) ma?acin ka=moh ton
    how? SUBJ=2S.FAM that
    ‘How about you there?’

(73) mej si? nn-ken 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) mnra? k=cip ba=hip
    people REL=to.go GOAL=forest
    ‘the people who went to the forest’ (elic.)

(75) slaj k=wih rh-roh
    swidden REL=3D IMPF-to.clear
    ‘the swidden that they were clearing’
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) hawen k=gej hobi?
    pig REL=to.eat root.crop
    ‘the pig that ate the root-crop’ (elic.)

(80) tmkal k=bdil kasa?
    man REL=shoot sambar.deer
    ‘the man who shot the sambar deer’ (elic.)

Relativised direct objects:

(81) hobi? k=je? tanem ktɔ? wej
    root.crop REL=1S to.plant day past
    ‘the root-crop that I planted yesterday’ (elic.)

(82) tmakɔw k=?o? ?ɔm ka=?o?
    tobacco REL=3S to.want SUBJ=3S
    ‘the tobacco that he wants’ (elic.)

(83) teh k=tmkal ?o? ?ek
    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=je? ?ek 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=jeʔ col ka=ʔun
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) t adolescents, 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 tiggyaʔ ʔoŋ jeʔ t-bow ton k=wek can=hip
2/3P three child IS 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) ʔoŋ kjih jeʔ t adolescents
child boy IS this
'this son of mine'

(89) gin bit
2/3P ant
'them ants'
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).

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.

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) ?ap ?o?
tiger 3S
'the tiger'

(101) mawë? ?o? ton
gibbon 3S that
'that gibbon'

(102) pëw ?o? 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) ka?ncö? ?o? ton
grandchild 3S that
'that grandchild of his'

bone thigh 3S this
'this thighbone of its'

5.1.4.2 WH modifiers of NPs

WH modifiers of NPs include possessor-questioning /maken/ 'whose?', item-questioning /mej/ 'what/which?' and number-questioning /mej si?/ '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 /maken/ 'whose?' replaces nominal or pronominal possessor modifiers; item-questioning /mej/ 'what/which?' replaces deictic modifiers, verbal modifiers, as well as relative clauses; and number-questioning /mej si?/ '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) maken ?5t
whose? dog
'whose dog?'

(106) mej hajë?
which? house
'which house?'
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 /h/ 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>kwik ?o? ton
to.speak<NM> 3S that
‘that way of speaking of his’

(110) nk-wek can=slaj crjkaj
NM-to.return SOURCE=swidden Cerikai
‘the return from Cerikai swidden’

(111) gin g<ra>hel tani?
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>anem hobi?
to.plant<NM> root.crop
‘tuber-planting’

(113) tmpot nk-nǒk
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=japèh
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 karei* ‘Karei’s lightning-flash’. In the present orthography this
would correspond to the following phrase:

(115) kilat d=karej
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). 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>kwik ?oto bt?et
to.speak<NM> 3S to.be.good
‘His way of speaking is good.’

(117) je? ?aŋket k<n>aajil
1S to.take to.fish<NM>
‘I took the fishing-rod.’

(118) je? sam ba=nm-sam ?aj
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) ?aket cip ba=hip, ja=np-cip ?ap 1eh
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) lpos nk-jok japēh wek
after NM-to.move IP.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 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 /da= ~ dso/ (Kruspe 2004:92), and Rischel (1995:140–141, 145)
describes possession-related usage of the words /di/ and /dso/ 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).
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=haʃe
    LOC=house
    'in the house'

(122) ba=woŋ jeŋ?
    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=woŋ ?oŋ?
    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. *je*  tboh  paj
   1S to.hit  2S.DIS
   'I hit you.'

b. *gin*  jij  bras
   2/3P to.carry rice
   'They carried rice.'

**EXPERIENCER:**

a. *?o*  ?el-?el  ka=??ap  cn=?ani?
   3S CONT-to.look  SUBJ=tiger  SOURCE=there.DIST
   'The tiger kept on watching from over there.'

b. *ha=mi*  ?t?et  ja?  ?o?
   Q=2S.INT to.know  grandmother  3S
   'Do you know his grandmother?'

**INSTRUMENT:**

a. *taji*  ?o?  get
   knife  3S to.cut
   'The knife made a cut.' (elic.)

b. *?o*  kac  ka=cnros  ?o?
   3S to.scratch  SUBJ=claw  3S
   'Its claw made a scratch.'

**FORCE:**

a. *ja=?o*  tpis  ka=tom
   RT=3S to.wash.away  SUBJ=river
   'Then the river washed [it] away.'

b. *?os*  ?o?  gjij
   fire  3S to.produce.smoke
   'The fire produced smoke.' (elic.)

**PATIENT:**

a. *?ap*  ?o?  kbis
   tiger  3S to.die
   'The tiger died.'

b. *ja= sóc*  ka=snlo
c
   RT=to.be.gone  SUBJ=blowpipe.dart
   'The darts are already used up!'

**THEME:**

a. *?o*  ??p  ka=pluruh  klen  hên  ?o?
   3S to.whiz  SUBJ=bullet inside mouth  3S
   'The bullet went into its mouth.'

b. *batu*  ?o?  res  ba=tom
   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) je? tek ka=je?
1S to.sleep SUBJ=1S
‘I slept.’

(125) ?o? sjiŋ leh ka=?ej je? wej slaj
3S to.burn EMP SUBJ=father IS 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/=1 and /Id/=1, the distribution of which appears to be determined by syntax and pragmatics respectively.

5.2.1.1.1 Subject-marking /ka=1/

Postverbal subject PPs are typically headed by the prepositional proclitic /ka=1, 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) ja=?o? loj wel ka=pew ba=?ani?
   RT=3S to.run again SUBJ=other GOAL=there
   ‘Then the other one fled over there again.’

(127) je? kul ka=je?
   IS to.call SUBJ=IS
   ‘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=1/. 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) ?o? get ka=taji?
   3S to.cut SUBJ=knife
   ‘The knife made a cut.’ (elic.)

(129) ?o? get ka=taji?
   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) ?o? get ka=tmkal ka=taji?
   3S to.cut SUBJ=man INSTR=knife
   ‘The man made a cut with the knife.’ (elic.)

In sum, argument-marking /ka=1/ 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) ?o? lobj d=?o?
    3S to.run CONTR=3S
    'He [as opposed to X] ran.'

(132) japêh wek leh d=japêh
    1P.EXCL to.go.back EMP CONTR=1P.EXCL
    'We [as opposed to X] went back.'

(133) ja=?o? cip leh d=?ap tâh
    RT=3S to.go EMP CONTR=tiger this
    'Then this tiger [as opposed to X] left.'

(134) je? gej d=je? duwa? kpiŋ
    IS to.eat CONTR=1S two CLF
    'I [as opposed to X] had two helpings.'

(135) tado? k=tâh d=mi?, ja=cip ja=?e1 ba=?ani?
    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 ?aj
    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 je?
    IRR=to.eat food 1S
    ‘[I] will eat my food.’

b. ?o? tamper je?
    3S to.claw 1S
    ‘It clawed me.’

THEME: a. wa=jij ?os k=tani?
    IRR.3S=to.carry.in.hand fire LOC=there
    ‘He will carry the fire over there.’

b. je? ?iŋ ka=je? leh wel bdil
    1S to.take SUBJ=1S EMP again gun
    ‘I took the gun again.’

PERCEPT: a. mi? ?el hajê? je?
    2S.INT to.see house 1S
    ‘You see my house.’

b. je? ji? ?o?
    3S to.dislike 3S
    ‘I dislike him.’

BENEFICIARY: a. japêh b-pimpin gin
    IP.EXCL PROG-to.guide 2/3P
    ‘We were guiding them.’

b. ja=pr-gej wɔŋ je?
    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)  ja=ʔoʔ bdil leh d=hawēn
        RT=3S to.shoot EMP CONTR=pig
    ‘Then he shot the pig [as opposed to shooting X].’

(138)  japēh kjeŋ d=mawēʔ ?oʔ ton
        IP.EXCL to.hear CONTR=gibbon 3S that
    ‘We heard that gibbon [as opposed to hearing X].’

(139)  braʔ japēh ja=ptsk d=gin kdek
        NEG IP.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 jeʔ
         IRR=to.eat food IS
     ‘[I] will eat my food.’

b.  pēh gej ba=tahɛʔ
       IP.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=baraŋ 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:**

?-aj !?isi? k=buloh game.animal to.insert LOC=bamboo.tube
‘Put the meat in the bamboo tube.’

**INSTRUMENT:**

mamej m::lh ja=d?-de? ka=?at ton what? 2S.FAM IRR=IMPF-to.do INSTR=stick that
‘What will you be doing with that stick?’

**GOAL:**

gin ja=hantar ba=kdah 2/3P IRR=to.send GOAL=Kedah
‘They were going to send [X] to Kedah.’

**RECIPIENT:**

ja=je? ?e=k ba=woŋ kŋih je? ton RT=1S to.give GOAL=child boy 1S that
‘Then I gave [X] to that son of mine.’

**SOURCE:**

?o? hantar krbow can=taj len 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. ?ek ba=w)kjih je?
to.give GOAL=child boy IS
‘Give [X] to my son!’
[with focus on change of location]

b. ja=je? ?ek d=w)kjih je? tāh
RT=IS to.give CONTR=child boy IS this
‘Then I gave [X] to this son of mine.’
[with focus on change of possessor]

c. ?ek k=w)kjih je?
to.give LOC=child boy IS
‘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=I or the goal-marking /ba=I.

(142) a. ?o? boh k=haje?
3S to.put LOC=house
‘He put [X] in the house.’ (elic.)
[with focus on location]

b. ?o? boh ba=haje?
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=I. 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=I ‘LOC’, /ba=I ‘GOAL’, /can=I ‘SOURCE’ and /ka=I ‘INSTR’ may be used to mark oblique arguments, whereas /k= ~ ka=I ‘LOC’, /ba=I ‘GOAL’ and /pn= ~ pon=I ‘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=I, 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 it is 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).
he:j pe:k can=?sh, boh, pe:k can=?sh, boh ba=?sh
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) loj ka=?aj pr:iw ba=?ani?
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 preverbal 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.

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=hip/ ‘to the forest’ is synonymous to ‘outside’]
(146) ja=gin b?bo?
RT=2/3P to.carry.on.back
'Then they carried [X].'
(147) ?ek ba=wo? kjih je?
to.give GOAL=child boy IS
'Give [X] to my son!'
(148) ?o? prise? ka=kn?ac ?on
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 /b?bo?rl '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 /rek/ '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) je? ?ek
  1S to.give
  'I gave [X to X].'

(150) gin ?ek gulah
  2/3P to.give sugar
  'They gave sugar [to X].'

(151) ?a=je? ?ek k=woŋ
  RT=1S to.give LOC=child
  'Then I gave [X] to the child.'

(152) ?ek d=je? ?oŋ
  to.give CONTR=1S fire
  'Give me the lighter!'

  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.
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) japêh ch-cih leh slaj
     IP.EXCL IMPF-to.cut EMP swidden
     ‘We were clearing a swidden.’

(158) ?ika? je? gj-gej
     fish IS 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. ?o? kbris
     3S  to.die
     ‘It died.’

b. je? k<ri>bis ?o?
    IS  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)  
\[ \text{je'? pr-hir leh} \]
\[ 1S \text{ CAUS-to.be.afraid EMP} \]
\[ 'I\text{ scared [X]}!' \]

(161)  
\[ \text{ja=je'? pi-dol k=la=hip} \]
\[ \text{RT}=1S \text{ CAUS-to.hide LOC}=1D=\text{forest} \]
\[ 'Then I hid [X] in the forest.' \]

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 /?ek/ ‘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)  
\[ \text{wa=bk-?ek ka=gtah, wa=bk-?ek ka=klapah sawit} \]
\[ \text{IRR.3S=PROG-to.give SUBJ=rubber IRR.3S=PROG-to.give SUBJ=oil.palm} \]
\[ '\text{Rubber will be given. Oil palms will be given.}' \]

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)  
\[ \text{?o? bk-?ek ba=je'?} \]
\[ 3S \text{ PROG-to.give GOAL}=1S \]
\[ 'He gave [X] to me [in passing].' \]

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.
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 /ũə/, which is not attested elsewhere, has been found in two instances of expressive elaboration:

/lîlũə/ 'the appearance of leeches moving'
/rînũə/ '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).

/glîwẽʔ/31 'the appearance of the fluttering wings of a butterfly'
/kpōtpẽ/ '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 /ϕ/, 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).

/klawẽhwũ/ 'the appearance of several crooked objects'
/prakɔpϕʁɛpũ/ 'the sound of a squirrel dropping empty nutshells'
/placiϕcɔϕũ/ 'the appearance of impressions in skin'

31 No root */gweʔ/ 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:** Ickipl 'to close one’s eyes'
- */<cr><kip><kip>/ ‘(appearance of many people closing their eyes)’

**Root:** /cnhēt/ ‘to be short’
- /<c><h><ēt>/ ‘(appearance of many short objects)’

**Root:** /hējjud/ ‘to be heavy’
- /<h><j><j><u><d>/ ‘(feeling of carrying many heavy objects)’

**Root:** /hāk/ ‘(sound of a big raindrop falling to the ground)’
- /<k><r><a><t><k><a>/ ‘(sound of many big raindrops falling to the ground)’

**Root:** /rīp/ ‘(sound of flapping)’
- /<r><i><p>/ ‘(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:** /kōb/ ‘(sound of a small object falling to the ground)’
- /<k><a><k><o><b>/ ‘(sound of empty nutshells being thrown to the ground by a squirrel, one after another)’

**Root:** /cōp/ ‘(sound of a blowpipe dart hitting canopy)’
- /<c><o><p>/ ‘(sound of blowpipe darts hitting canopy, one after another)’

**Root:** /gōb/ ‘to be dark’
- /<g><r><a><h><c><i><h>/ ‘(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 /i, a, o, ə/ 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 /e/ or its nasal counterpart. The following forms exemplify such correlation:

<table>
<thead>
<tr>
<th>Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>/kramitmɪt/</td>
<td>(appearance of large impressions in skin)</td>
</tr>
<tr>
<td>/kramitmɛt/</td>
<td>(appearance of small impressions in skin)</td>
</tr>
<tr>
<td>/kramitmɪt/</td>
<td>(appearance of tiny impressions in skin)</td>
</tr>
<tr>
<td>/cnahithɔt/</td>
<td>(appearance of many short objects)</td>
</tr>
<tr>
<td>/cnahithɛt/</td>
<td>(appearance of many very short objects)</td>
</tr>
<tr>
<td>/cnahithɪt/</td>
<td>(appearance of many very short, hardly visible objects)</td>
</tr>
<tr>
<td>/kla?icʔuc/</td>
<td>(appearance of large larvae crawling)</td>
</tr>
<tr>
<td>/kla?icʔɛc/</td>
<td>(appearance of small larvae crawling)</td>
</tr>
<tr>
<td>/kla?icʔɪc/</td>
<td>(appearance of tiny larvae crawling)</td>
</tr>
</tbody>
</table>

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) caʔel ton, klawɛhwɔh-klawɛhwɔh
    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.
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 $3 \times 3$ 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 (/lkaW/,
/l/kaW/,
/laJu/,
/lom/ and
/laWej/ respectively) followed by a proper, specifying name, e.g. /l/kaW bawunj, 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: [l]. In reality, however, the trill is in free variation with the approximant [l] in all positions and a flap [r] 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 /tP ~ 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, a, o, a, u, o, p, b, t, d, c, j, k, g, s, h, m, n, r, 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 (*).

<table>
<thead>
<tr>
<th>Phonemic form</th>
<th>Phonetic form</th>
<th>English translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>/-p/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tftp ~ tip</td>
<td>'tìp' ~ 'tip'</td>
<td>to be flat</td>
</tr>
<tr>
<td>ctip ~ ctip</td>
<td>c'ìtip' ~ c'ìtip'</td>
<td>to assemble; to grow together</td>
</tr>
<tr>
<td>sip</td>
<td>'sìp'</td>
<td>fruit without a seed</td>
</tr>
<tr>
<td>kawip</td>
<td>ka'wip'</td>
<td>sun bear <em>(Helarctos malayanus)</em></td>
</tr>
<tr>
<td>tep*</td>
<td>'tep'</td>
<td>cassette</td>
</tr>
<tr>
<td>pdep</td>
<td>pa'dep'</td>
<td>to peep</td>
</tr>
<tr>
<td>k?ep</td>
<td>ka?ep'</td>
<td>[a type of centipede]</td>
</tr>
<tr>
<td>klep</td>
<td>ka'lep'</td>
<td>[a type of tuber]</td>
</tr>
<tr>
<td>krlep</td>
<td>ka'lep'</td>
<td>to forget</td>
</tr>
<tr>
<td>titep</td>
<td>ti'tep'</td>
<td>opposite side</td>
</tr>
<tr>
<td>ktep</td>
<td>ka'tep'</td>
<td>to carry in one’s mouth</td>
</tr>
<tr>
<td>pdep</td>
<td>pa'dep'</td>
<td>to stare</td>
</tr>
<tr>
<td>cep</td>
<td>c'ep'</td>
<td>to catch</td>
</tr>
<tr>
<td>kakep</td>
<td>ka'kep'</td>
<td>to remember</td>
</tr>
<tr>
<td>ksep</td>
<td>ka'sep'</td>
<td>to adorn oneself with leaves for good luck</td>
</tr>
<tr>
<td>mep*</td>
<td>'mèp'</td>
<td>map</td>
</tr>
<tr>
<td>nemp</td>
<td>'nèp'</td>
<td>to be raw</td>
</tr>
<tr>
<td>lep</td>
<td>'lep'</td>
<td>to fold; to plait</td>
</tr>
<tr>
<td>halep</td>
<td>ha'lep'</td>
<td>to raft</td>
</tr>
<tr>
<td>pплеп</td>
<td>pip'lep'</td>
<td>to blaze; to twinkle</td>
</tr>
<tr>
<td>cпrep</td>
<td>c'er'ep'</td>
<td>babbler <em>(Malacopteron)</em></td>
</tr>
<tr>
<td>jeep</td>
<td>'jeep'</td>
<td>to count</td>
</tr>
<tr>
<td>pipjep</td>
<td>pip'jep'</td>
<td>to teach</td>
</tr>
<tr>
<td>snтip</td>
<td>san'tip'</td>
<td>to pound</td>
</tr>
<tr>
<td>cip</td>
<td>c'ip'</td>
<td>to go; to move</td>
</tr>
<tr>
<td>sjip</td>
<td>sa'jip'</td>
<td>to flow (of tears)</td>
</tr>
<tr>
<td>pkip</td>
<td>po'kip'</td>
<td>to scorch</td>
</tr>
<tr>
<td>crkip can</td>
<td>c'erkip' 'c'a'n</td>
<td>heel</td>
</tr>
<tr>
<td>kr?ip</td>
<td>ka'rip'</td>
<td>to crouch</td>
</tr>
<tr>
<td>hip</td>
<td>'hip'</td>
<td>wood/forest; outside</td>
</tr>
<tr>
<td>ranjip</td>
<td>ra'njip'</td>
<td>to crunch</td>
</tr>
</tbody>
</table>
Jahai rhyming glossary

grlip  
gərlip  to lay a floor

tap  
ˈtəp  to carry on one’s back

kr?əp  
karʔəp  to burp

harəp*  
haˈɾəp  to believe; to have faith

bap  
ˈbap  rice; food

sbap*  
səˈbap  because; but

ttap*  
təˈtap  to stay in one place; to settle down

stap  
səˈtap  to split

kap  
ˈkap  to bite

tiŋkap*  
tiŋˈkap  window

laŋkap*  
laŋˈkap  [a type of palm (Arenga obtusifolia)]

hagap  
haˈɡap  Sumatran rhinoceros (Dicerorhinus sumatrensis)

ʔap  
ʔap  [generic name for large felines (Panthera)]; to encounter a large feline

ʔap tmtum  
ʔapˈtəmˈtuʔm  black panther (Panthera pardus)

ʔap ?awej  
ʔapˈʔəˈwej  leopard (Panthera pardus)

maʔap*  
maʔap  pardon

hap ~ hiphap  
ˈhap ~ hipˈhap  to spread in different directions

blap  
bəˈlap  nothing

klap  
kəˈlap  [toponym: Kelap]

slap  
səˈlap  to die

jawai  
jaˈwäip  to answer

hawap  
haˈwap  [a type of large freshwater gastropod]

sajap*  
səˈjap  wing

tjap-tjap*  
tjapˈ-tiˈjap  every

saqup  
səˈqup  cloud; fog

jup  
ˈjup  to be flat

cukop*  
cuˈkop  to be full

gop  
ˈgop  stranger (derogatory)

ʔop  
ʔop  to stop

sop  
ˈsop  lung

rop  
ˈrop  to walk; [sound of walking; sound of muntjac deer]

hat5p  
haˈt5p  to be light (of weight)

dɔp  
dəˈɔp  to sneak; to whisper

kɔp  
kəˈɔp  to move hut

khɔp  
kəhˈɔp  to gobble down

nɔp  
ˈnɔp  to chew

karɔp  
kaˈɾɔp  dead bamboo

/ -t/  
ptpit  
pəˈpit  [a type of small animal]
cpit  
cəˈspit  to squeeze

bit  
ˈbit  [a type of ant]
Appendix I

ktit  kəfīt̚ to sting
kacīt  kəčit̚ to rattle
jit  'jit̚ to collect/gather
kit  'kit̚ bottom; buttocks
kit ktoʔ  'kit̚ kət̚ʔ mouth of river
kit tom  'kit̚'t̚m to fart
ktkit  kət̚'kit̚ birthmark
hit  'hit̚ to rattle; to tremble
it  'mit̚ eye; CLF: small objects
mit  'mit̚ sun
mit tajiʔ  'mit̚ tə'j̚ʔ blade of knife
mit ktoʔ  'mit̚ kət̚ʔ sun
mit knajil  'mit̚ kənajil fishing hook
mit ?haʔj  'mit̚ ?əhaʔj seed
hnit  hən̚it̚ [a type of fruit]
bnit  mən̚it̚ NEGATIVE MARKER
rnit  rən̚it̚ sandfly
plit  pəl̚iʔ to fade away; to extinguish by itself
ktl̚it  kəl̚'lit̚ to feel the sharpness of e.g. a knife
ktw̚it  kərt̚'wit̚ wrinkles
wtw̚it  wət̚'wit̚ [a type of red bird]
?mpet  ?əmp̚et̚ [a type of cobra]
kbet  kəb̚et̚ old woman
kbet clah  kəb̚et̚ cəlah old bachelor
puket*  pu̕ket̚ pocket
?aŋket*  ?aŋket̚ to get; to take
get  'get̚ to cut
lget  lə'get̚ mountain pass
?t̚iet  ?ət̚'iət̚ to know
pahet  pa'het̚ [a type of fruit]
plet  po'let̚ to be dense/heavy
wet  'wεt̚ to flow (of river); to run (of e.g. trail)
lumpet*  lum̚pet̚ to jump
bet  'bet̚ to lash; to slither
lbet  lə'bet̚ to be heavy (of rain)
rbet  ra'bet̚ to tie
kaltet  kal̚tet̚ to flip with one's finger
plancet  polaŋ'p̚eʔ̚et̚ mouse deer (Tragulus)
ket  'ket̚ to cut
?iket*  ?i̕'ket̚ to lash
?aket  ?a'ket̚ PROHIBITIVE
tmket  tam'ket̚ to be cold (of weather, objects etc.)
hlan̚ket  həlan̚'ket̚ [a type of ant]
haʔet  həʔeʔ̚ bad smell/stench
bt?et  boʔʔeʔ̚ to be beautiful/fine/good
Jahai rhyming glossary

set
'set'
to pour

set set
'set set'
sunbird (*Nectarinia; Anthrepetes*)

stset
'set set'
to pick one's teeth

lapset*
'lap set'
langsat (*Lansium domesticum*)

rnsët
'nsest'
to sob

bhet
'behet'
to be sweet

kmet
'kmet'
[a type of frog]; vulva

grinet*
'gri net'
grenade

?iinet*
'i net'
to believe; to think; to remember

bulet
'bullet'
to slither

bnolet
'manolet'
many

hrtlet
'hrretlet'
to be tired

paret
'paret'
[a type of large insect]

baret
'baret'
Malay tapir (*Tapirus indicus*)

jaret
'jaret'
[a type of small grasshopper]

soret*
'soret'
letter

stwet
'sawet'
[a type of fruit]

hwet
'hwet'
to stain

jet
'jet'
to ogle

sitjët
'sitセット'
to blow (of wind)

pit
'pit'
to blow (of person and wind);
to extinguish fire

kmit
'kmit'
egg

ddit'
'dedit'
[a type of bird]

jit
'jët'
to wipe

bkît
'be kit'
to be light; to be warm; to be hot

bnkit
'maingkit'
sweat

hrkit
'hrkit'
evening; night

b?it
'bebit'
to be sweet

sit
'sit'
honeycomb

sît ~ ?otsît
'sit ~ ?otsit'
to rub oneself

bhit
'behit'
bamboo

jhit
'jëhit'
to smoke; to suck

blhit
'be hit'
to be tasteless

kmit
'kmit'
gall-bladder; knee-cap

tnit
'tnit'
lower lip

ktlit
'këlit'
glow-worm

hmët
'hemët'
mythical cannibals

ptpot
'petpot'
to place one's hand on something

tmpot*
tomëpot'
place

sitkat
'sitkat'
to laugh

cokët
cëkët'
to kill/slash

ragot
'ragat'
to carry on one's back

hët
'hët'
to sting (of mosquito)
cnhät to be short
tšt to stare
klät to swallow
rat to tie
wawät rat
lktwät to be fast; quickly
ejt to watch/see
jät to hurt
?mpat ~ ?mpät* four
bat to sit (of bird); to roost; [sound of bird landing on a branch]
cat* to paint
sikat* comb
hakat to rub; to scratch
pukat* large net
gat to waylay
?at stick
sat to leave poisoned blowpipe dart to dry
rihat* rest; to rest
phat* to hammer
jahat* to be bad
hamat* to roll something
slamat* safety
kilat* lightning
calat [a type of lizard]
barat* west
karat* dirt
kawat* to march
kwat* to be strong
swat to build a hut
siput* snail
rumpüt* grass
but to eat
dut navel
hpjut to be heavy
?ût to argue
sut to sob
prasut to lose hair
hút to reveal oneself
janut* beard
?jüt throat
parüt* scar
surüt* to push
pot to suck
klutbot [a type of larvae]
<table>
<thead>
<tr>
<th>Sahai Rhyming Glossary</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>?ikot*</td>
<td>to follow</td>
</tr>
<tr>
<td>saʔot</td>
<td>to call someone</td>
</tr>
<tr>
<td>jəʔot</td>
<td>[toponym]</td>
</tr>
<tr>
<td>sot</td>
<td>to carve/cut; straight</td>
</tr>
<tr>
<td>kasot*</td>
<td>shoe</td>
</tr>
<tr>
<td>slot</td>
<td>to drown</td>
</tr>
<tr>
<td>lawot*</td>
<td>sea</td>
</tr>
<tr>
<td>kajot</td>
<td>to be pregnant</td>
</tr>
<tr>
<td>dbot</td>
<td>to climb a tree (of bear)</td>
</tr>
<tr>
<td>cbot</td>
<td>to devour</td>
</tr>
<tr>
<td>dst</td>
<td>vulva</td>
</tr>
<tr>
<td>kət</td>
<td>to take</td>
</tr>
<tr>
<td>hak5t</td>
<td>[a type of tuber]</td>
</tr>
<tr>
<td>?5t</td>
<td>dog</td>
</tr>
<tr>
<td>?t5t</td>
<td>to stroke an animal</td>
</tr>
<tr>
<td>h5t</td>
<td>to perceive taste</td>
</tr>
<tr>
<td>krt15t</td>
<td>kidney</td>
</tr>
<tr>
<td>smut15t</td>
<td>brain</td>
</tr>
<tr>
<td>w5t</td>
<td>to bend something down (e.g. a tree)</td>
</tr>
<tr>
<td>kaw5t</td>
<td>bird (generic)</td>
</tr>
<tr>
<td>kaw5t batuʔ*</td>
<td>broadbill; pitta (Psarismus; Pitta)</td>
</tr>
<tr>
<td>raj5t</td>
<td>[a type of small animal]</td>
</tr>
</tbody>
</table>

/-c/

<table>
<thead>
<tr>
<th>Sahai Rhyming Glossary</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>hic</td>
<td>to rain</td>
</tr>
<tr>
<td>btec</td>
<td>to be long</td>
</tr>
<tr>
<td>lwec</td>
<td>to ascend; to climb up</td>
</tr>
<tr>
<td>rmpec</td>
<td>hardwood tree (Turtur tigrinus)</td>
</tr>
<tr>
<td>kbéc</td>
<td>to spit slowly to the ground</td>
</tr>
<tr>
<td>kec</td>
<td>to cut</td>
</tr>
<tr>
<td>gěc</td>
<td>to scratch (of thorns)</td>
</tr>
<tr>
<td>?eć</td>
<td>belly; excrement; to defecate</td>
</tr>
<tr>
<td>sec</td>
<td>meat</td>
</tr>
<tr>
<td>knec</td>
<td>comb</td>
</tr>
<tr>
<td>nec</td>
<td>to tickle</td>
</tr>
<tr>
<td>lec</td>
<td>to miss target; to be wrong</td>
</tr>
<tr>
<td>taʔic</td>
<td>[a type of large bird]</td>
</tr>
<tr>
<td>krʔic</td>
<td>to heat</td>
</tr>
<tr>
<td>trhic</td>
<td>[a type of small bird]</td>
</tr>
<tr>
<td>sic</td>
<td>to sting</td>
</tr>
<tr>
<td>grsic</td>
<td>to feel</td>
</tr>
<tr>
<td>pimic</td>
<td>to take back</td>
</tr>
<tr>
<td>?ηic</td>
<td>to burn</td>
</tr>
<tr>
<td>?aʔic</td>
<td>to pass</td>
</tr>
<tr>
<td>haric</td>
<td>[magic word uttered to stop rain]</td>
</tr>
</tbody>
</table>
tök to follow
hchök to whistle
cmök boil; wound
plök to disappear
klcbök [a type of millipede]
kök to scratch
kn?ök father-in-law
bclök to be smooth; to be slippery
srök [magic word uttered to stop rain]
bawök pig-tailed macaque (*Macaca nemestrina*)
kwök to swim
hök to drink
muc to eat fish or meat
kmuc ghost
sirök to slurp
grök slender-toed gecko (*Cyrtodactylus*)
poc [a type of terrapin]
scbök to lick one's lips
t?ök to ask for something
soc to wash
hoc [a type of fruit]
poc to collect small objects
?mpoc salt
bök to tell a lie
tök to tickle by poking one's finger
coc ~ cac 'cöööc' ~ 'cöööc' scar
koc to gnaw
lkök fontanel
t?ök to climb
soc to be finished; to be gone; to stop
noc to move snout (of monkey)
loc bow
snlök blowpipe dart
canwök collared scops-owl (*Otus lempiji*)

/-k/
tik 'tik' [sound of raindrop]
mudök* 'mu'dök' to return
cik 'cöik' [sound of blowpipe dart hitting muscle of prey]
hgök to fear
lanik lie/haunt (of animal)
palök [a type of small animal]
clik 'cöölök' to keep a good lookout sideways while walking
Jahai rhyming glossary 175

ririk
ririk'  bee-eater (Merops)
brik*  bərik'  to brake
grik ~ grit  ga'rrik' ~ ga'rit'  [toponym: Gerik]
tek  'tek'  to lie down; to marry; to sleep
lnktek  lan'tek'  tongue
kdek  kə'dek'  squirrel
kdek ?abu?  ka'dek' ?a'bu?  Provost’s squirrel (Callosciurus prevosti)
kdek creh  ka'dek' cə'reh  plantain squirrel (Callosciurus notatus)
cek  'cə'ek'  to stab; to throw a spear
plek  pa'lə'ek'  to fall (of tree)
klek  kələ'ek'  quill of porcupine
wek  'wek'  to go back
piwek  pi'wek'  to lift
wikwek  wik'wek'  to turn around; to move around
pek  'pek'  to split; to chop
pęb  pęb'ek'  to prick (of thorns)
lapęb  la'pęb'ek'  mud
bek*  'bek'  bag
czbek  cə'sębek'  to spear
hrkbek  harak'bek' ~ harak'bąk'  to sink to the ground (of animal shot)
tek  'tek'  [ethnonym: other group of Jahai]; [sound of snails (?)]
batek  batek'  [ethnonym: Batek]
balttek  bal'teck'  friend
kktęk  kak'ęk'  to click (with the tip of the tongue towards the palate)
hnteck*  hą'ntęck'  to hit; to pound
kdek  ka'dek'  to be bitter
?nęck  ?ą'ńęc'ęk'  [a type of tree]
?ęk  ?ą'ęk'  to give
sksęk  sak'sęk'  to devour quickly
hęk  hę'k'  to snap and fall (of e.g. tree-branch)
hkhęk  hak'hęk'  to be ragged/torn (of cloth)
?aneck  ?ą'anjeck'  girl
laneck*  la'neck'  [a type of porcupine]
tulek*  tulek'  to push someone or something
blek  bə'leck'  to lick
snrek  san'dęck'  to go out
tawęk  tawęć'  butterfly (generic)
lajęk  la'jęk'  to tread
cik  'cık'  to wade across a river
jkjik  ją'kęj'ık'  to breathe
bkik  bə'kıć'  pigeon (generic)
sik  'sık'  to be caught
lık  'lɨk'  to breathe with difficulty
rhik  rhik'  red
lik  lik'  to swallow (of snake)
klik  k'ilik'  [sound of a person walking]
slik  s'lik'  to bind
rik  rik'  to stab
?awik  ?awuk'  [a type of small tree-dwelling mammal]
ckwik  c'ak'wuk'  to crackle; to bellow (of e.g. macaque, gaur, rhinoceros); to talk/chat

pkpok  p'ok'p'ok'  to smack
bok  b'ok'  to untie
kktok  k'ok't'ok'  to click (with the tip of the tongue towards the alveolar ridge)

carok*  c'a'rok'  stream-bed
pak  'pak'  to clap
napak  na'pak'  wild boar (Sus scrofa)
clapak  c'a'la'pak'  boar/male pig
ddkak  d'ak'dak'  fish-trap
p'ak  p'a'pak'  to cook
krhak  k'ahrhak'  to clear one's nasal cavity
lak  'lak'  to have sore eyes
calak  c'a'la'ak'  lizard (generic)
klak  k'alak'  [sound of something falling]
krak  k'arak'  [a type of plant/fruit]
kujak  k'ujak'  to butcher; to tear into pieces
kapuk  k'apuk'  flying squirrel (Petaurista)
tbuk*  t'buk'  to peck (of bird)
duk  'duk'  to pounce upon
taduk  t'aduk'  [a type of tree]
ddkduk  d'ak'duk'  chest
pucuk* loc  p'uc'uk' 'loc'  arrowhead
tuk  t'uk'  [a type of hornbill]
?uk  '?uk'  to blow
suk  'suk'  to light up
skujuk  s'ak'n'uk'  [a type of frog]
haluk  ha'lu'uk'  [a type of lizard]
juluk  j'u'lu'uk'  to keep a fire burning
lk luk  l'ak'lu'uk'  to laugh
rkruk  r'ok'r'uk'  to go along a watercourse
trjuk  t'ar'juk'  [ethnonym: different group of Jahai]
pok  'pok'  to break open; to stride; [sound of an animal falling to the ground]

cok  'c'ok'  to tear apart
ckcok  c'ak'c'ok'  Diard's trogon (Harpactes diardii)
jok  j'ok'  to move from one place to another
krkok  k'arkok'  casque of hornbill
chok  c'ahok'  [sound of running water or waterfall]
lahai rhyming glossary

Planok*  po'a'nok'  mouse deer (*Tragulus*)
Pok  'pok'  round object
Lumpok  lum'pok'  [a type of tree]
Cmalpok  c'e'amal'pok'  [a type of millipede]
Btjok  ba'tjok'  malkoha; cuckoo (*Phaenicophaeus; Cuculus*); papaya (*Carica papaya*)

Kluktok  kəluk'tok'  [a type of large terrestrial gastropod]
Ptjok  pətjok'  to cook meat in its own juice
Dok  'dok'  ipoh poison
Cok  'c'ok'  bag
Kucuk  ku'cuk'  [a type of bird]
Ckcoj  c'ek'c'e'ok'  banded palm civet (*Hemigalus derbyanus*)
Kok  'kok'  [sound of a hornbill]
Crikok  c'e'ri'kok'  to jabber/chatter (of monkey)
Hakoj  ha'koj'  to throw
Hokok  ho'koj'  to burn off fur/feathers from a killed animal
Ho'koj  ho'koj'  [sound of a leaf-monkey]
Kikkok  kik'koj'  to sit with one's knees pulled up
Tjok  təj'koj'  nape of the neck
Okoj  o'koj'  [sound of boiling]
Ju'koj  jə'u'koj'  tiger (*Panthera tigris*)
Sok  'soj'  hair
Sok Kawat  'soj' kə'wət'  feather; plumage; down
Sok Kuj  'soj' ku'j  hair of the head
Sksok  sə'koj'  to suck out (e.g. a snail from its shell)
Rkosok  re'koj'  keelback (generic) (*Xenochrophis, Amphiesma*)

Hok  'hok'  to throw
Bhok  bo'hok'  to vomit
Nok  'nok'  to sit; to stay; to reside
Snoj  se'noj'  [a type of bird]
Bmok  bər'mok'  [a type of toad]
Malok*  ma'loj'  to be embarrassed
Wok  'woj'  to dig tubers
Wawok  wa'woj'  to have a sad expression on one's face
Joj  'joj'  to undress; to stretch someone's arm
Tjok ~ Tkjok  təj'koj' ~ tək'joj'  to point
Tkjok  tək'joj'  to pull; to pluck

/-ʔ/  

Jampi?*  jə'mipi?  to heal
Labi?*  la'bi?  [a type of soft-shelled turtle]
Hobi?*  ho'bi?  root-crop; tuber (generic)
Hat?  ha'ti?  tail
Ptj?*  po'ti?  box
<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>bldi?*</td>
<td>bucket</td>
</tr>
<tr>
<td>ci?</td>
<td>to flare up; louse</td>
</tr>
<tr>
<td>?acî?</td>
<td>dog</td>
</tr>
<tr>
<td>cucî?*</td>
<td>to clean</td>
</tr>
<tr>
<td>küncî?*</td>
<td>key; to mend</td>
</tr>
<tr>
<td>bi:jî?*</td>
<td>seed; CLF: small objects</td>
</tr>
<tr>
<td>ta:jî?*</td>
<td>knife</td>
</tr>
<tr>
<td>bny:jî?</td>
<td>to be distant (horizontally)</td>
</tr>
<tr>
<td>pagî?*</td>
<td>morning</td>
</tr>
<tr>
<td>lagi?*</td>
<td>again; also; moreover; still; yet</td>
</tr>
<tr>
<td>si?</td>
<td>number</td>
</tr>
<tr>
<td>mej si?</td>
<td>how many?</td>
</tr>
<tr>
<td>?isi?*</td>
<td>to put inside; to insert</td>
</tr>
<tr>
<td>nasi?*</td>
<td>cooked rice</td>
</tr>
<tr>
<td>bsi?*</td>
<td>iron</td>
</tr>
<tr>
<td>bhi?</td>
<td>to be full/to have had enough to eat</td>
</tr>
<tr>
<td>mi?</td>
<td>PRONOUN 2S.INT</td>
</tr>
<tr>
<td>tani?</td>
<td>that (away)</td>
</tr>
<tr>
<td>?ani?</td>
<td>there (away)</td>
</tr>
<tr>
<td>skali?*</td>
<td>all; together</td>
</tr>
<tr>
<td>halî?</td>
<td>leaf; CLF: small flat objects</td>
</tr>
<tr>
<td>halî? gadiŋ</td>
<td>[a type of leaf]</td>
</tr>
<tr>
<td>bli?*</td>
<td>to buy</td>
</tr>
<tr>
<td>jli?</td>
<td>[toponym: Jeli]</td>
</tr>
<tr>
<td>jari?*</td>
<td>finger</td>
</tr>
<tr>
<td>mnri?</td>
<td>[ethnonym: Menriq]</td>
</tr>
<tr>
<td>wi?</td>
<td>left (side)</td>
</tr>
<tr>
<td>dwi?*</td>
<td>money</td>
</tr>
<tr>
<td>m?wi?</td>
<td>left-handed person</td>
</tr>
<tr>
<td>ji?</td>
<td>to cry (of infant); to dislike; to refuse; to reject</td>
</tr>
<tr>
<td>kajî?</td>
<td>bat</td>
</tr>
<tr>
<td>kajî?</td>
<td>little finger</td>
</tr>
<tr>
<td>ha jî?</td>
<td>also</td>
</tr>
<tr>
<td>te?</td>
<td>earth; ground; soil</td>
</tr>
<tr>
<td>?ite?*</td>
<td>duck; water-fowl</td>
</tr>
<tr>
<td>ste?*</td>
<td>to shoot with slingshot</td>
</tr>
<tr>
<td>de?</td>
<td>to make; to do</td>
</tr>
<tr>
<td>tase?*</td>
<td>lake</td>
</tr>
<tr>
<td>he?</td>
<td>PRONOUN 1P.INCL</td>
</tr>
<tr>
<td>male?</td>
<td>brown hornbill (<em>Ptilolaemus tickelli</em>)</td>
</tr>
<tr>
<td>cole?</td>
<td>cannibal</td>
</tr>
<tr>
<td>?le?</td>
<td>thorn</td>
</tr>
<tr>
<td>ti?le?</td>
<td>to point (with one's finger, hand or lips); to show</td>
</tr>
<tr>
<td>Term</td>
<td>Meaning</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>tare*</td>
<td>to pull</td>
</tr>
<tr>
<td>we?</td>
<td>to exist; to be found; to be present; to occur; to be available</td>
</tr>
<tr>
<td>pe?</td>
<td>older sibling</td>
</tr>
<tr>
<td>pce?</td>
<td>to be wet</td>
</tr>
<tr>
<td>joh wanke?</td>
<td>tree-branch</td>
</tr>
<tr>
<td>sise?</td>
<td>to dance</td>
</tr>
<tr>
<td>prise?</td>
<td>to keep</td>
</tr>
<tr>
<td>gase?</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>brase?</td>
<td>to remain; to stay behind</td>
</tr>
<tr>
<td>tahe?</td>
<td>[a type of salty tuber]</td>
</tr>
<tr>
<td>cuhe?</td>
<td>to flow; to spread</td>
</tr>
<tr>
<td>me?</td>
<td>female child</td>
</tr>
<tr>
<td>lime*</td>
<td>five</td>
</tr>
<tr>
<td>game?</td>
<td>to halt/stop</td>
</tr>
<tr>
<td>same*</td>
<td>with; likewise</td>
</tr>
<tr>
<td>cme?</td>
<td>to be sharp</td>
</tr>
<tr>
<td>cine*</td>
<td>[ethnonym: Chinese]</td>
</tr>
<tr>
<td>trine?</td>
<td>height</td>
</tr>
<tr>
<td>bane?</td>
<td>to give birth</td>
</tr>
<tr>
<td>bne?</td>
<td>size</td>
</tr>
<tr>
<td>kne*</td>
<td>to score a bull’s eye</td>
</tr>
<tr>
<td>tahe?</td>
<td>to ask</td>
</tr>
<tr>
<td>pune*</td>
<td>to have</td>
</tr>
<tr>
<td>nhe?</td>
<td>to kill an animal</td>
</tr>
<tr>
<td>sine?</td>
<td>[toponym: Singor]</td>
</tr>
<tr>
<td>bune*</td>
<td>flower</td>
</tr>
<tr>
<td>brile?</td>
<td>to pull off (e.g. a mat from the ground); to roll (of animal); to turn around</td>
</tr>
<tr>
<td>li?le?</td>
<td>to call for somebody; to rub against (of e.g. cat)</td>
</tr>
<tr>
<td>lu?le?</td>
<td>to roll (of animal)</td>
</tr>
<tr>
<td>g?le?</td>
<td>to say no; to decline</td>
</tr>
<tr>
<td>sle?</td>
<td>[a type of snake]</td>
</tr>
<tr>
<td>ciwe?</td>
<td>to open one’s eyes</td>
</tr>
<tr>
<td>?awe?</td>
<td>[a type of tortoise]</td>
</tr>
<tr>
<td>mawe?</td>
<td>gibbon (Hylobates)</td>
</tr>
<tr>
<td>smwe*</td>
<td>all</td>
</tr>
<tr>
<td>je?</td>
<td>PRONOUN IS</td>
</tr>
<tr>
<td>haje?</td>
<td>house; hut</td>
</tr>
<tr>
<td>hapi?</td>
<td>bag</td>
</tr>
<tr>
<td>bi?</td>
<td>mother</td>
</tr>
<tr>
<td>bi? masaj</td>
<td>sow</td>
</tr>
<tr>
<td>kbi?</td>
<td>fruit; CLF: animals, several objects</td>
</tr>
<tr>
<td>kbi? ?ntep</td>
<td>testicles</td>
</tr>
<tr>
<td>knbi?</td>
<td>[a type of rattan]</td>
</tr>
</tbody>
</table>
tic 'tə? to pound
tic siriŋ 'tə? siriŋ [a type of tree-dwelling mammal]
pti? po'tə? forehead
kti? kə'tə? skin
ti? ta?ta? old; grandparent
ji? 'jə? to burn; to make fire; to roast
ja'ji? ~ ja'ja? jə'ja'? ~ jə'ja'? old woman
bnji? mənjə'? to be distant (vertically); to be high; to be deep
ki? 'kə? to vomit
pi?ti? pə'tə? to put a child to sleep
ha?ti? ha'?ta? yes; right; correct
cə?i? cə'ə'? to pour fluid
bhi? bə'hə? to hug
mami? mə'mə'? [a type of colourful larvae]
bani? mənə'? quiver for blowpipe darts
ṭni? tə'ni'? that (beyond you)
ʔni? ʔə'ni'? there (beyond you)
bli? bə'lə? upper leg
pijə? pə'jə? to fly up
ʔikə? ʔi'kə? fish
jaŋka? jaŋkə? jaw/mandible
sipə?* si'pa? to kick
hapə? ha'pa? to die
ləpa? la'pa? [ethnonym: unidentified easterly group of Semang]
nampa?* nə'mpa? to see
jumpa?* jə'mpa? to meet
cuba?* cə'u'ba? to taste
ka? kə'ba? [a type of tuber]
rba? rə'ba? rotten wood
tə? tə? ancestor; ancestral spirit; father's father
mata? ma'tə? spear
lata?* la'tə? waterfall
bada?* ba'də? Sumatran rhinoceros (Dicerorhinus sumatrensis)
lada?* la'də? pestle
kuda?* ku'də? horse
lda? la'də? armpit
baca?* ba'cə'a? to talk
ci?ca? cə'i'cə'a? gecko
cə?ca? cə'cə'a? common kingfisher (Alcedo atthis)
səja?* sa'ja'? only
pja? po'ja'? to hide oneself
krJa?* kər'ja'? to work
lahai rhyming glossary

<table>
<thead>
<tr>
<th>Luka?</th>
<th>lu'ka?</th>
<th>to hit a target; to wound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pana?k</td>
<td>pa'nya?</td>
<td>to hit; to throw</td>
</tr>
<tr>
<td>Janya?</td>
<td>j'anya?</td>
<td>to go</td>
</tr>
<tr>
<td>Slanja?</td>
<td>sa'lan'ka?</td>
<td>collar-bone</td>
</tr>
<tr>
<td>Braka?</td>
<td>ma'nya?</td>
<td>molar tooth</td>
</tr>
<tr>
<td>Tiya?</td>
<td>ti'ya?</td>
<td>three</td>
</tr>
<tr>
<td>Jaga?</td>
<td>j'aga?</td>
<td>only</td>
</tr>
<tr>
<td>Dahaga?</td>
<td>daha'ga?</td>
<td>to be thirsty</td>
</tr>
<tr>
<td>Ragaa?</td>
<td>ra'ga?</td>
<td>basket</td>
</tr>
<tr>
<td>Tiga?</td>
<td>ta'ga?</td>
<td>to flex muscle</td>
</tr>
<tr>
<td>?a? Pok</td>
<td>ta'a?'pok'</td>
<td>[a type of mushroom]</td>
</tr>
<tr>
<td>Pasa?</td>
<td>pa'sa?</td>
<td>to live</td>
</tr>
<tr>
<td>Kasa?</td>
<td>ka'sa?</td>
<td>sambar deer (Cervus unicolor)</td>
</tr>
<tr>
<td>Brasaa</td>
<td>ba'ra'sa?</td>
<td>many</td>
</tr>
<tr>
<td>Piha?</td>
<td>pi'ha?</td>
<td>to be separate</td>
</tr>
<tr>
<td>Tuhaa</td>
<td>tu'ha?</td>
<td>to ripen</td>
</tr>
<tr>
<td>Cnha?</td>
<td>c3n'ha?</td>
<td>to jest</td>
</tr>
<tr>
<td>Ma?</td>
<td>'ma?</td>
<td>mother</td>
</tr>
<tr>
<td>Lma?*</td>
<td>la'ma?</td>
<td>intestines</td>
</tr>
<tr>
<td>La?</td>
<td>'la?</td>
<td>penis</td>
</tr>
<tr>
<td>Hala?</td>
<td>ha'la?</td>
<td>to fly</td>
</tr>
<tr>
<td>Tula?*</td>
<td>tu'la?</td>
<td>to push someone to go somewhere</td>
</tr>
<tr>
<td>Mula?*</td>
<td>mu'la?</td>
<td>to begin</td>
</tr>
<tr>
<td>Bla?</td>
<td>ba'la?</td>
<td>to be alone</td>
</tr>
<tr>
<td>Kira?</td>
<td>ki'ra?</td>
<td>to count</td>
</tr>
<tr>
<td>Pihira?</td>
<td>pihira?</td>
<td>to rear animals</td>
</tr>
<tr>
<td>Cara?</td>
<td>c'ara?</td>
<td>to talk</td>
</tr>
<tr>
<td>Pra?</td>
<td>pa'ra?</td>
<td>to live/reside</td>
</tr>
<tr>
<td>Bra?</td>
<td>ba'ra?</td>
<td>NEGATIVE MARKER</td>
</tr>
<tr>
<td>Pnra?</td>
<td>pa'n'ra?</td>
<td>[a type of sweet-smelling flower]</td>
</tr>
<tr>
<td>Mnra?</td>
<td>man'ja?</td>
<td>human being; people; person</td>
</tr>
<tr>
<td>?awa?</td>
<td>?a'wa?</td>
<td>elder</td>
</tr>
<tr>
<td>Lawa?</td>
<td>la'wa?</td>
<td>to collect/gather</td>
</tr>
<tr>
<td>Duwa?</td>
<td>du'wa?</td>
<td>two</td>
</tr>
<tr>
<td>Duwa? Puloh</td>
<td>du'wa? pu'loh</td>
<td>twenty</td>
</tr>
<tr>
<td>Brwa?</td>
<td>ba'rwaa?</td>
<td>wind</td>
</tr>
<tr>
<td>Ja?</td>
<td>'ja?</td>
<td>grandmother</td>
</tr>
<tr>
<td>Caja?</td>
<td>c'aja?</td>
<td>to believe</td>
</tr>
<tr>
<td>La'ja</td>
<td>la'ja?</td>
<td>leech</td>
</tr>
<tr>
<td>Sraja?</td>
<td>sa'ra'ja?</td>
<td>[a type of tree (Shorea)]</td>
</tr>
<tr>
<td>Buja?</td>
<td>bu'ja?</td>
<td>crocodile (Tomistoma, Crocodilus)</td>
</tr>
<tr>
<td>Hrja?</td>
<td>ho'ra'ja?</td>
<td>[a type of tuber]</td>
</tr>
<tr>
<td>Mapu?</td>
<td>ma'pu?</td>
<td>when?</td>
</tr>
<tr>
<td>?Ampu?</td>
<td>?a'm'pu?</td>
<td>to convey</td>
</tr>
<tr>
<td>Word</td>
<td>Meaning</td>
<td>Word</td>
</tr>
<tr>
<td>-------</td>
<td>---------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>kabu?</td>
<td>[a type of tuber]</td>
<td>ka'bu?</td>
</tr>
<tr>
<td>?abu?</td>
<td>dust</td>
<td>?a'bu?</td>
</tr>
<tr>
<td>?abu? ??os</td>
<td>ashes</td>
<td>?a'bu? ??os</td>
</tr>
<tr>
<td>labu?</td>
<td>bottle</td>
<td>la'bu?</td>
</tr>
<tr>
<td>batu?</td>
<td>door</td>
<td>ba'tu?</td>
</tr>
<tr>
<td>pintu?</td>
<td></td>
<td>pin'tu?</td>
</tr>
<tr>
<td>bantu?</td>
<td>to catch something that is falling</td>
<td>han'tu?</td>
</tr>
<tr>
<td>hantu?</td>
<td>to rest</td>
<td>b?ora'du?</td>
</tr>
<tr>
<td>prdu?</td>
<td>CLF: cluster</td>
<td>po'rdu?</td>
</tr>
<tr>
<td>ců?</td>
<td>to be blind</td>
<td>'ců'?</td>
</tr>
<tr>
<td>baju?</td>
<td>clothes</td>
<td>ba'jú'?</td>
</tr>
<tr>
<td>taju?</td>
<td>snake (generic)</td>
<td>ta'jú'?</td>
</tr>
<tr>
<td>sju?</td>
<td>to be cold (of weather, objects etc.)</td>
<td>sa'jú'?</td>
</tr>
<tr>
<td>paku?</td>
<td>Filex fern</td>
<td>pa'ku'?</td>
</tr>
<tr>
<td>taku?</td>
<td>to steal</td>
<td>ta'ku'?</td>
</tr>
<tr>
<td>blaků?</td>
<td></td>
<td>b?ela'ků'?</td>
</tr>
<tr>
<td>buku?</td>
<td>book</td>
<td>bu'ku'?</td>
</tr>
<tr>
<td>pusu?</td>
<td>termite mound</td>
<td>pu'su'?</td>
</tr>
<tr>
<td>hů?</td>
<td></td>
<td>'hu'?</td>
</tr>
<tr>
<td>prahu?</td>
<td>boat</td>
<td>po'ra'hu'?</td>
</tr>
<tr>
<td>jhů?</td>
<td></td>
<td>j'hů'?</td>
</tr>
<tr>
<td>bamu?</td>
<td>bamboo</td>
<td>m?amů'?</td>
</tr>
<tr>
<td>somu?</td>
<td>tapir’s snout; horn of rhinoceros (?)</td>
<td>so'mů'?</td>
</tr>
<tr>
<td>gm?u?</td>
<td>fat/grease</td>
<td>go'mů'?</td>
</tr>
<tr>
<td>lmů?</td>
<td>cattle</td>
<td>l?emů'?</td>
</tr>
<tr>
<td>sanu?</td>
<td>ghost</td>
<td>sa'nů'?</td>
</tr>
<tr>
<td>naaju?</td>
<td></td>
<td>na'nů'?</td>
</tr>
<tr>
<td>baru?</td>
<td>again; anew; until</td>
<td>ba'ru'?</td>
</tr>
<tr>
<td>babo?</td>
<td>female; woman</td>
<td>ba'bo'?</td>
</tr>
<tr>
<td>tabo?</td>
<td>big digit (thumb or big toe)</td>
<td>ta'bo'?</td>
</tr>
<tr>
<td>lobo?</td>
<td>plunge-pool beneath waterfall</td>
<td>lo'bo'?</td>
</tr>
<tr>
<td>to?</td>
<td>ancestor; grandparent; [ethnonym: To’]</td>
<td>'to'?</td>
</tr>
<tr>
<td>blato?</td>
<td>crimson-winged woodpecker (Picus puniceus)</td>
<td>b?elato'?</td>
</tr>
<tr>
<td>juo?</td>
<td>[toponym]</td>
<td>'jů'?</td>
</tr>
<tr>
<td>koko?</td>
<td>to crow (of cock)</td>
<td>ko'ko'?</td>
</tr>
<tr>
<td>pgo?</td>
<td>to cook</td>
<td>po'go'?</td>
</tr>
<tr>
<td>?o?</td>
<td></td>
<td>?o'?</td>
</tr>
<tr>
<td>s?so?</td>
<td></td>
<td>sa'so'?</td>
</tr>
<tr>
<td>so?</td>
<td></td>
<td>'so'?</td>
</tr>
<tr>
<td>s?so?</td>
<td></td>
<td>sa'so'?</td>
</tr>
<tr>
<td>tumo?</td>
<td></td>
<td>tu'mo'?</td>
</tr>
<tr>
<td>tano?</td>
<td></td>
<td>ta'nò'?</td>
</tr>
<tr>
<td>s?i?</td>
<td></td>
<td>s?i'no'?</td>
</tr>
<tr>
<td>s?i?</td>
<td></td>
<td>s?i?no'?</td>
</tr>
<tr>
<td>s?i?</td>
<td></td>
<td>s?i'nó'?</td>
</tr>
<tr>
<td>English</td>
<td>Jahai</td>
<td>Meaning</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>bamboo</td>
<td>bu'lo?</td>
<td>bamboo; bamboo tube</td>
</tr>
<tr>
<td>to decline</td>
<td>si'ro?</td>
<td></td>
</tr>
<tr>
<td>fish/meat</td>
<td>la'wo?</td>
<td></td>
</tr>
<tr>
<td>outer shaft of blowpipe; to carry on one's shoulders</td>
<td>jo?</td>
<td></td>
</tr>
<tr>
<td>mountain-top</td>
<td>'po?</td>
<td></td>
</tr>
<tr>
<td>axe; cheek</td>
<td>ka'po?</td>
<td></td>
</tr>
<tr>
<td>to carry on one's back</td>
<td>ba'bo?</td>
<td></td>
</tr>
<tr>
<td>day; sky</td>
<td>kə'to?</td>
<td></td>
</tr>
<tr>
<td>today</td>
<td>kə'to? təh</td>
<td></td>
</tr>
<tr>
<td>[toponym: Dok]</td>
<td>'də?</td>
<td></td>
</tr>
<tr>
<td>to wait</td>
<td>ta'do?</td>
<td></td>
</tr>
<tr>
<td>to rest one's chin on something</td>
<td>cə'ndo?</td>
<td></td>
</tr>
<tr>
<td>to sting</td>
<td>cə?</td>
<td></td>
</tr>
<tr>
<td>grandchild</td>
<td>ka'ni'cə?</td>
<td></td>
</tr>
<tr>
<td>a little; to be small</td>
<td>?ə'jə?</td>
<td></td>
</tr>
<tr>
<td>blanket; fabric</td>
<td>pə'ni'cə?</td>
<td></td>
</tr>
<tr>
<td>long-tailed shrike (Lanius schach)</td>
<td>cə'kəs?</td>
<td></td>
</tr>
<tr>
<td>trap/snare</td>
<td>ba'ko?</td>
<td></td>
</tr>
<tr>
<td>cup; glass</td>
<td>ta'ko?</td>
<td></td>
</tr>
<tr>
<td>egg</td>
<td>ma'ko?</td>
<td></td>
</tr>
<tr>
<td>to conceive; to be pregnant</td>
<td>mənkə?</td>
<td></td>
</tr>
<tr>
<td>[a type of bird]</td>
<td>ta'ko?</td>
<td></td>
</tr>
<tr>
<td>nail</td>
<td>kal'ko?</td>
<td></td>
</tr>
<tr>
<td>to burn</td>
<td>'sə?</td>
<td></td>
</tr>
<tr>
<td>rotten (of wood)</td>
<td>?ə'ni'sə?</td>
<td></td>
</tr>
<tr>
<td>jar</td>
<td>ho?</td>
<td></td>
</tr>
<tr>
<td>younger sister of parent</td>
<td>mə'ʃə?</td>
<td></td>
</tr>
<tr>
<td>hard surface (stone, paved road, brick wall etc.)</td>
<td>timə?</td>
<td></td>
</tr>
<tr>
<td>day after tomorrow</td>
<td>ca'mo?</td>
<td></td>
</tr>
<tr>
<td>fruit; pumpkin; CLF: animals, several objects</td>
<td>kə'mo?</td>
<td></td>
</tr>
<tr>
<td>[a type of short millipede]</td>
<td>ta'ho?</td>
<td></td>
</tr>
<tr>
<td>to look down</td>
<td>jə'ŋəs?</td>
<td></td>
</tr>
<tr>
<td>[toponym]</td>
<td>lə?</td>
<td></td>
</tr>
<tr>
<td>casting-net</td>
<td>jə'alə?</td>
<td></td>
</tr>
<tr>
<td>[a type of large freshwater gastropod]</td>
<td>kalə?</td>
<td></td>
</tr>
<tr>
<td>and; with</td>
<td>?ə'lə?</td>
<td></td>
</tr>
<tr>
<td>hole</td>
<td>jə'lo?</td>
<td></td>
</tr>
<tr>
<td>corpse</td>
<td>sa'ro?</td>
<td></td>
</tr>
<tr>
<td>to be hungry</td>
<td>cə'ro?</td>
<td></td>
</tr>
<tr>
<td>back (of person)</td>
<td>kə'ro?</td>
<td></td>
</tr>
<tr>
<td>[a type of terrapin]</td>
<td>karwo?</td>
<td></td>
</tr>
<tr>
<td>to roar (of tiger)</td>
<td>pi'jo?</td>
<td></td>
</tr>
</tbody>
</table>
### /-phi/

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ciϕ</td>
<td>to hiss (of e.g. bamboo in fire)</td>
</tr>
<tr>
<td>kicciϕ</td>
<td>to crackle/explode from heat</td>
</tr>
<tr>
<td>hίϕ ~ hέϕ</td>
<td>to whistle shrilly (with one’s fingers in one’s mouth); to blow a fire</td>
</tr>
<tr>
<td>lίϕ ~ lφίϕ</td>
<td>fontanelle</td>
</tr>
<tr>
<td>ciриϕ</td>
<td>to extinguish; to fade away; to put out fire with water</td>
</tr>
<tr>
<td>priϕ ~ piφriϕ</td>
<td>to flicker</td>
</tr>
<tr>
<td>cεϕ</td>
<td>to fan fire</td>
</tr>
<tr>
<td>cεεφ</td>
<td>tail feathers</td>
</tr>
<tr>
<td>nιϕ</td>
<td>to blow smoke</td>
</tr>
<tr>
<td>rιϕ</td>
<td>[sound of flapping]</td>
</tr>
<tr>
<td>jιϕ</td>
<td>to shake flapping (e.g. a piece of paper)</td>
</tr>
<tr>
<td>jaϕ</td>
<td>[sound of flying or leaping]</td>
</tr>
<tr>
<td>τυϕ</td>
<td>to spit</td>
</tr>
<tr>
<td>ιυϕ</td>
<td>to blow (e.g. a fire)</td>
</tr>
<tr>
<td>suφ</td>
<td>to attack</td>
</tr>
<tr>
<td>ιυφ</td>
<td>to haunt</td>
</tr>
<tr>
<td>luφ ~ ruφ</td>
<td>[sound of dashing]</td>
</tr>
<tr>
<td>cruφ</td>
<td>to descend</td>
</tr>
<tr>
<td>κοφ</td>
<td>[sound of blowpipe dart hitting stomach of prey]</td>
</tr>
<tr>
<td>κτοφ</td>
<td>to spit/splutter horizontally</td>
</tr>
<tr>
<td>cοφ</td>
<td>to cause a blister (of fire)</td>
</tr>
<tr>
<td>κοφ</td>
<td>[sound of small object falling to the ground]</td>
</tr>
<tr>
<td>bsοφ</td>
<td>to be quick</td>
</tr>
<tr>
<td>hοφ</td>
<td>to prick a hole in something (e.g. a leaf)</td>
</tr>
<tr>
<td>λοφ</td>
<td>to leak</td>
</tr>
<tr>
<td>τιλοφ</td>
<td>to dig</td>
</tr>
<tr>
<td>plοφ</td>
<td>[sound of a blowpipe being fired]</td>
</tr>
<tr>
<td>jοφ</td>
<td>to whiz</td>
</tr>
<tr>
<td>rοφ</td>
<td>[sound of a blowpipe dart hitting canopy]</td>
</tr>
</tbody>
</table>

### /-lis/ (cont.)

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>lipis*</td>
<td>to be thin</td>
</tr>
<tr>
<td>tpis</td>
<td>to blow away; to wash away</td>
</tr>
<tr>
<td>cbis*</td>
<td>to cleave</td>
</tr>
<tr>
<td>kbis</td>
<td>to die</td>
</tr>
<tr>
<td>tis batan*</td>
<td>[a type of insect]</td>
</tr>
<tr>
<td>britis*</td>
<td>[ethnonym: British]</td>
</tr>
<tr>
<td>ptis</td>
<td>to hurt; pain; sickness</td>
</tr>
<tr>
<td>ktis*</td>
<td>to break/snap</td>
</tr>
<tr>
<td>cscis</td>
<td>to scratch oneself</td>
</tr>
<tr>
<td>kis</td>
<td>to dig</td>
</tr>
</tbody>
</table>
lahai rhyming glossary

kīs  'kīs  ghost
ci 'īn's  ci 'īn's  to ask for something
hīs  'hīs  to blow one's nose; to snuffle
j 'īn's  j 'ān's  kind/sort
manjīs*  manjīs  mangosteen (*Garcinia mangostana*)
klanjīs  kela'njīs  heart
tīs  tēñ's  [a type of fish]
pīris  pīris  to go across a watercourse
hirīs*  hīrīs  to cut
wīs  'wīs  to split/to go different ways
blīwīs  bali 'wīs  to get out of bed
cawīs  cē 'wīs  [a type of small animal]
rwīs  rō 'wīs  to cut grass
tīs  tēñ's  to bounce back; to kick something to pieces
tsīses  tis'dēs  to stumble
ges  'ges  fragrance; odour; smell; stench
j 'ō'jes  j 'ē'jes  root
hampēs  ham'pēs  to live/reside
kampēs  kampēs  [a type of tree]
tēbēs*  tēbēs  to cut
līntēs*  lin'tēs  to go across
dēs  'dēs  to go and get
cscēses  cē 'scēses  to tear
kes  'kes  to put together
bakes  ba'kes  to grow up; to be adult
gēs  'gēs  to descend; to go downhill
?ēs  ?ēs  ice
mes*  'mēs  gold
plēses  pōles  to smear poison on blowpipe dart
plēses  pōlēses  [sound of blowpipe being fired]
res  'res  to fall
wēs  'wēs  frontal tuber (on forehead)
ciwares  cē 'wares  [toponym]
tawēses  tawēses  [a type of tree]
kwēses  kēwēses  to sweep with one's fingers
lwēses  la'wēses  to be wide
pjēses*  piji'jēses  R.P.S. (Rancangan Pengumpulan Semula; site of regroupment program)
pis  'pis  to sweep/wipe
tpis  tēpīs  valley
j 'ō'kis  j 'ā'kis  [a type of porcupine]
tanjīs  tanjīs  ?
gīs  'gīs  to descend/to climb down/to go down;
       to apply make-up
prgīs  pōrgīs  [a type of tree]
"is 'cis 'to die
la'is la'is 'to be bad
nis 'nis 'mat
la'jis la'jis [a type of palm leaf used for thatching]
lwos* lwos* after; to leave
k'os k'os 'to kill (?)
g'os g'os 'to carve
lo'os lo'os fatty tissue
kipas* k'ipas 'to wag tail
bas* 'bas 'bus
tas 'tas 'to cleave
kas 'kas 'to pinch somebody
gas ~ g'es 'gas ~ 'g'es [a type of skin disease]
hagas* ha'gas 'mosquito
?as ?as 'to believe; to have faith
has 'has 'to steam (intransitive)
lanas* la'nas pineapple (Ananassa)
tronas* t'oro'nas [toponym: Petronas petrol station at Banding]
balas* ba'las 'to answer
sblas* s'ab'las 'eleven
tlas 'tlas [a type of tree]
bras* b'oras 'husked rice
kras* k'oras 'to be tough (of meat)
wass 'was 'to split; fork/junction
wass tom 'was 'tobm 'river confluence
puwas* pu'was 'to be satisfied
?uswas ?us'was 'to meet (of rivers)
cjas ci'jas 'hand
?imbus* ?im'bus 'to ambush
t'us ~ t'mus t'us ~ t'am'us 'to grub; to grunt (of pig)
st'tus t'st'tus 'to lose hair
gus 'gus 'to be together
nus 'nus 'upper lip
tanus ta'nus 'boar's snout
barus* ba'rus 'camphor
kurus* ku'rus thin
pos* 'pos post
sratos* sara'tos hundred
dos 'dos 'to owe (?)
gos 'gos belch
phos po'hos 'to blow (of person)
hamos ha'mos 'to blow (of person)
ros 'ros 'liver
trros* t'ar'os at once; straightaway; straight
| tpos          | ṭepos          | [a type of fruit]                  |
| cós          | 'cêçó           | [toponym: Banun]                  |
| kãs          | 'kêças          | to scratch                        |
| gós          | 'gos            | to live/exist                     |
| ?ós          | 'ọs             | fire                              |
| lan?ós       | lan'ọs          | door                              |
| cnós ~ cnus  | c'Onos ~ c'Onus | ring                              |
| lslos        | lis'los         | to gnaw                           |
| cnrós         | c'an'róos      | nail; phalange                    |

| /-h/          |                  |                                   |
| pipih        | pi'pih           | [a type of rattan]                |
| ctih         | c'etihi          | to be easy                        |
| titih*       | t'itihi          | bridge                            |
| didih*       | d'idih           | to boil (intransitive)            |
| cih           | 'cihi            | to shoo away                      |
| gcîh         | ga'c'ihi         | to be black; to be dark           |
| ha?îh        | ha'îhi           | yes; right; correct               |
| sîh           | 'sihi            | to pound                          |
| gunîh*       | gun'îhi          | gunny bag; sack                   |
| malîh*       | mal'îhi          | [a type of palm-tree]             |
| tîh           | t'îhi            | to fall to the ground             |
| wih           | 'wih             | PRONOUN 3D                        |
| tujih         | tu'jîhi          | that (down)                       |
| ?ujih         | ?u'jîhi          | there (down)                      |
| lbêh          | la'bêh           | honey                             |
| puteh*       | pu'têh           | to be white                       |
| thîteh        | tah'têh          | oriental pied hornbill (*Anthracoceros albirostris*) |

| tadeh         | ta'dêh           | that (beyond me)                  |
| ?adêh         | ?a'dêh           | there (beyond me)                 |
| pihdêh        | pih'dêh          | to summon; to call for            |
| seh           | 'sêh             | to be good                        |
| kneh          | k'ênh            | wife                              |
| haleh         | ha'leh           | to be hungry                      |
| boleh*        | bo'lêh           | ROOT POSSIBILITY; to be able to   |
| koleh         | ko'lêh           | cup                               |
| ciwêh         | ci'wêh           | to climb up                       |
| jôweh         | j'o'weh          | [a type of bird]                  |
| rajêh         | ra'jêh           | to beat (e.g. a pillow)           |
| klîpeh        | kâli'peh         | to flay/skin                      |
| japêh ~ pêh   | ja'pêh ~ 'pêh    | PRONOUN IP.EXCL                    |
| beh           | 'beh             | younger brother of parent         |
| dibêh         | di'beh           | wallet                            |
| teh*          | 'teh             | tea                               |
cēh  'cēh  [sound of a blowpipe dart hitting a vine]
tərjēh  tərjēh  to fly up (of bird)
manēh  mānēh  to be long (of time); to be old
leh*  'leh  to go together; EMPHATIC PARTICLE
ʔa'leh  ʔa'leh  girl; virgin
bnaleh  māna'leh  to breastfeed
kraleh  kāra'leh  giant squirrel (Ratufa affinis) (?)
pleh  po'leh  [ethnonym: Temiar]
tureh  tu'reh  to tap poison from ipoh tree
weh  'weh  to cut off/to pick
wēh  'wēh  [sound of blowpipe]
jēh  jēh  PRONOUN 1D.EXCL
pihpīh  pih'pēh  to wake up somebody
titīh  ti'tēh  that (up)
ʔi'śćh  ʔi'tēh  there (up)
kdīh  kādēh  to say
śīh  'śēh  to fell/to cut down trees
jīh  'jēh  PRONOUN 2D
kjīh  kā'jēh  boy
tīh  ta'kēh  behind; backside
pō'śh  po'śēh  fluid
śi'h  'śēh  [a type of tuber (sweet potato?)]
rwīh  rā'wēh  [toponym]
klāph  kāla'pēh  shoulder
sruph*  sōrū'pēh  same
tāh  tō'śēh  this
plītāh*  pōlītāh  small fire; lamp
kritāh*  kārītāh  car
pō'āh*  po'āh  to be near
kāh  kēh  to lie down
pōkāh  po'kēh  [a type of gecko]
ʔāh  ʔō'śēh  here
pō'āh  po'āh  to cook/heat in a fire
sēh  sīsēh  to meet
hāh  ūh  to whistle with one's mouth open
maŋāh  māŋāh  [toponym: Mangga]
gulāh*  gu'lēh  sugar
blāh  bō'lıh  to enter
klāpah sawit*  kāla'pah saw'it'  oil palm
japāh  ja'pāh  [ethnonym: group of Jahai]
cpāh  cō'pāh  wood shavings from carving or scraping
kbah  kō'bah  to glare
lbah  la'bah  where?
lbah lwej  la'bah la'wej  honeybee
lahai rhyming glossary

mutah*  mu'tah  mortar

gtah*  gə'tah  rubber

kdah*  kə'dah  [toponym: Kedah]

cah  'cah  [a type of tree]; to cut

?acah*  ?a'cəah  to scold

pcah*  pə'cəah  to break; to move from one place to another

gajah*  ga'jəah  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 (?)  məŋ'kah  [toponym]

gagah  ga'gah  to walk

pʔah  paʔah  to kneel; to sit by fire

sisah*  sə'sah  remains

gamah*  ga'mah  photograph

lmah  lə'mah  to find; to meet

mah*  mə'nəah  to be low (of height, voice etc.)

kalah*  ka'lah  to lose (e.g. a fight)

hmalah  həməlah  young pig

blah*  bə'lah  to chop; to wane (of moon)

clah  cə'lah  childless person

llah  lə'lah  oriole (*Oriolus*)

rah  rə'ah  to vomit

darah*  də'rah  blood

carah  cə'rəah  barbet (*Megalaima*)

marah*  ma'ra'ah  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

situh  səltu'uh  to attack

duduh  dədu'uh  [a type of bird]

suduh*  su'du'uh  to scrape

juh  jʰu'uh  to blow away smoke

khkuh  kəh'kuh  [a type of hornbill]

susuh*  susu'uh  milk

bunuh*  bu'nuh  to kill

jnuh*  jʰənuh  too many/much; enough

jnuh ~ jnūh (?)  jʰənūh  [a type of small bird]

kaluh*  ka'luh  if

haluh  həlu'uh  to shoot

puhulu'uh*  puhul'u'uh  headman

ruh  rə'uh  [a type of wasp]

pluruh*  pləru'uh  bullet

biruh*  biru'uh  blue

krjuh  kər'juh  to slither
boh  'boh  to place/put
jatoh  j'atoh  to move; to wander
ctoh  c'atoh  phlegm; snot
coh  'c'oh  to refuse
tu'oh*  tu'j'oh  seven
pankoh*  panykoh  to hold
sanjoh  sanjoh  wreathed hornbill (Rhyticeros undulatus)
goh  'goh  manchild
pgoh  po'goh  to boil (transitive)
?oh  ?oh  to cough
soh  'soh  game animal
jnoh  j'noh  [a type of tree]
spuloh*  sap'loh  ten
ruoh  ru'roh  to shed leaves
woh ~ woh  'woh ~ 'woh  to call (of bird)
joh  'joh  crown of tree; tree branch
?oh  'po'oh  to fan; to hit with a flat hand
tap5h  ta'p5h  pelvis
boh  'boh  fruit (generic)
bruboh  baru'boh  yellow-vented bulbul (Pycnonotus goiavier)
tboh*  ta'boh  to beat/to hit
mantoh  män'loh  Malay person
lant5h  lan't5h  porcupine (Hystrix, Atherura)
klihd5h  blih'd5h  to meander (of river)
coh  'c'oh  to bite (of snake, bird etc.)
chcoh  cl:ah'c'oh  to mince
kalc5h  kal'cl:5h  to crack
joh  'j'oh  to ascend; to go uphill
nhjoh  nah'j'oh  height
koh  'koh  [a type of tortoise]
tang5h  tang5h  to ring
kuh?5h  küh'5h  [a type of turtle]
moh  'm5h  nose; PRONOUN 2S.FAM
moh tohm  'm5h tohm  headwater/source of river
knmoh  kən'm5h  name
kanoh  ka'noh  to grow (of plants)
lanoh  la'n5h  [ethnonym: Lanoh]
proh  po'r5h  [a type of tree]
sroh  sa'r5h  to give birth
roh  'roh  to cut; to clear
rawoh  rawoh  [a type of fruit tree]
joh  'joh  to drop
kajoh*  ka'joh  to paddle
<table>
<thead>
<tr>
<th>/-m/</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>cím</td>
<td>'cîm</td>
<td>to sizzle (of meat being roasted)</td>
</tr>
<tr>
<td>jím</td>
<td>'jîm</td>
<td>to cry/weep</td>
</tr>
<tr>
<td>bním</td>
<td>mânîm</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>smsîm</td>
<td>sâmîm</td>
<td>silver-eared mesia (<em>Leiothrix argentauris</em>)</td>
</tr>
<tr>
<td>harîm</td>
<td>hârîm</td>
<td>quill of porcupine’s tail</td>
</tr>
<tr>
<td>kmjîm</td>
<td>kmjîm</td>
<td>to taste</td>
</tr>
<tr>
<td>sbem</td>
<td>sōbê'm</td>
<td>to be full</td>
</tr>
<tr>
<td>tem</td>
<td>'tê'm</td>
<td>right (side)</td>
</tr>
<tr>
<td>mmtem</td>
<td>mâm'tê'm</td>
<td>right-handed person</td>
</tr>
<tr>
<td>lmtem</td>
<td>lâm'tê'm</td>
<td>to arrive; to become</td>
</tr>
<tr>
<td>tmtem</td>
<td>tam'dê'm</td>
<td>to go into the jungle and stay there overnight in search of food</td>
</tr>
<tr>
<td>museum*</td>
<td>mu'se'b'm</td>
<td>season</td>
</tr>
<tr>
<td>slem*</td>
<td>sâ'le'b'm</td>
<td>coin</td>
</tr>
<tr>
<td>jrem*</td>
<td>jô're'b'm</td>
<td>rapid</td>
</tr>
<tr>
<td>krem</td>
<td>kô're'm</td>
<td>to play an instrument</td>
</tr>
<tr>
<td>dêm*</td>
<td>'dêm</td>
<td>lake</td>
</tr>
<tr>
<td>l kem</td>
<td>lô'ke'b'm</td>
<td>brain</td>
</tr>
<tr>
<td>?êm</td>
<td>?êm</td>
<td>breast; to drink</td>
</tr>
<tr>
<td>?êm kajì?</td>
<td>?êm kajì?</td>
<td>[a type of thorn]</td>
</tr>
<tr>
<td>pi?êm</td>
<td>pi?êm</td>
<td>to breastfeed</td>
</tr>
<tr>
<td>masêm*</td>
<td>mâ'sêm</td>
<td>to be sour</td>
</tr>
<tr>
<td>krhém</td>
<td>kôrhîm</td>
<td>to clear one’s throat</td>
</tr>
<tr>
<td>nêm*</td>
<td>'nêm</td>
<td>six</td>
</tr>
<tr>
<td>tanem*</td>
<td>ta'nêm</td>
<td>to plant</td>
</tr>
<tr>
<td>ñem</td>
<td>'nêm</td>
<td>to eat</td>
</tr>
<tr>
<td>gulem</td>
<td>gu'le'b'm</td>
<td>to carry on a stick</td>
</tr>
<tr>
<td>karem*</td>
<td>kà're'b'm</td>
<td>to sink</td>
</tr>
<tr>
<td>sabim</td>
<td>sa'bi'b'm</td>
<td>[ethnonym: Sabûm]</td>
</tr>
<tr>
<td>patim</td>
<td>pati'b'm</td>
<td>to pound</td>
</tr>
<tr>
<td>ktim</td>
<td>kôti'b'm</td>
<td>to hit with side of fist</td>
</tr>
<tr>
<td>ckim</td>
<td>cô'ki'b'm</td>
<td>lump; chip</td>
</tr>
<tr>
<td>cîm</td>
<td>cô'kîm</td>
<td>peacock-pheasant (<em>Polyplectron</em>)</td>
</tr>
<tr>
<td>gim</td>
<td>'gî'b'm</td>
<td>to deliberate</td>
</tr>
<tr>
<td>lgîm</td>
<td>lô'gî'b'm</td>
<td>valley floor</td>
</tr>
<tr>
<td>bhîm</td>
<td>bô'hîm</td>
<td>blood</td>
</tr>
<tr>
<td>rhîm</td>
<td>tôrhî'b'm</td>
<td>[a type of large beetle]</td>
</tr>
<tr>
<td>kmîm</td>
<td>kômîm</td>
<td>to smoke with the cigarette inside one’s mouth</td>
</tr>
<tr>
<td>tmîm</td>
<td>tômî'b'm</td>
<td>to shoot</td>
</tr>
<tr>
<td>lîm</td>
<td>'lî'b'm</td>
<td>to be big; [toponym: Long]</td>
</tr>
<tr>
<td>pêm</td>
<td>'pô'b'm</td>
<td>to sit (of bird)</td>
</tr>
<tr>
<td>katom*</td>
<td>kà'tô'b'm</td>
<td>to cut rattan</td>
</tr>
<tr>
<td>tmtoôm</td>
<td>tam'tôôm</td>
<td>to buzz (of insect)</td>
</tr>
</tbody>
</table>
kəm 'kə bm to rub one's hands
kmkəm kəm'kəm ankle
ʔəm ~ ʔimʔəm ʔə bm ~ ʔimʔə bm to hug
məm 'məm mother's milk
gərm* gə rə m salt
pam* 'pə bm to pump
lampam* lam'pa bm [a type of carp (Puntius)]
kbam kə 'ba bm to smoke with the cigarette the wrong way round

padam* pa'da bm to close one's eyes
jam* 'ja bm hour; wristwatch
tkam ta'ka bm [toponym: Tekam]
sam 'sa bm to go hunting
paham* pa'ha bm to understand
mam 'məm to drink milk from mother's breast
tamam ta'ma bm [a type of small animal]
lam 'la bm to squeeze
wam 'wa bm [a type of rattan]
sjam si'ja bm Thailand
hajam* ha'ja bm poultry
piɲjam* piɲ'ja bm to borrow
prjam pe'ja bm time; when
pum 'pu bm to lie face down
kutum ku'tu bm flower bud
cdum o'du bm to carry in one's arms
rdùm ~ rdup ra'du bm ~ ra'dup bm to be cloudy
kmum kə'nu bm [a type of tree]
blum bo'lu bm [toponym: Belum]
jarum* ja'ru bm needle
tom 'to bm tree-base; CLF: vegetation
tom kəʔ to'bm kəʔ bm east
ptom po'to bm day before yesterday
lantom la'nto bm to sting
mpjəlom* məŋjəlo bm [toponym: Mendelum]
hajom ha'jo bm young green tree python (Chondropython viridis)
tom 'to bm fluid; river; water
tom mit 'to bm 'mət bm tear
cəm 'eə bm to burn something; to make fire
kəm 'kə bm many
ʔəm 'ʔə bm to want
jʔəm j'ərʔə bm to be dirty
krʔom kərʔə bm to cook in ashes or fire
səm 'sə bm bird's nest
smsəm səm'sə bm to buzz around a nest (of birds, bees)
<table>
<thead>
<tr>
<th>Term</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>hnhn</td>
<td>həmnʰo̞m</td>
<td>to like</td>
</tr>
<tr>
<td>knom</td>
<td>kənʰɔm</td>
<td>urinary bladder; urine; to urinate</td>
</tr>
<tr>
<td>rlim</td>
<td>rəlɔ̞m</td>
<td>[toponym]</td>
</tr>
<tr>
<td>kjom</td>
<td>kəˈjo̞b̊m</td>
<td>below; lower side; underside</td>
</tr>
<tr>
<td>pimpin*</td>
<td>pimˈpi̞d̪n</td>
<td>to guide</td>
</tr>
<tr>
<td>bin</td>
<td>ˈbɪ̞dn</td>
<td>to sleep on one's stomach</td>
</tr>
<tr>
<td>tin</td>
<td>ˈti̞d̪n</td>
<td>to kick open</td>
</tr>
<tr>
<td>?acin</td>
<td>ʔaˈc̞i̞d̪n</td>
<td>how?</td>
</tr>
<tr>
<td>gin</td>
<td>ˈgi̞d̪n</td>
<td>PRONOUN 2/3P</td>
</tr>
<tr>
<td>risin*</td>
<td>riˈsi̞d̪n</td>
<td>ration</td>
</tr>
<tr>
<td>masin*</td>
<td>maˈsi̞d̪n</td>
<td>to be salty</td>
</tr>
<tr>
<td>brsin*</td>
<td>bəˈsi̞d̪n</td>
<td>to sneeze</td>
</tr>
<tr>
<td>jamin*</td>
<td>jəˈɑ̈m̡i̞n̡</td>
<td>to guarantee</td>
</tr>
<tr>
<td>mnm</td>
<td>mənˈm̡i̞n̡</td>
<td>to play (games)</td>
</tr>
<tr>
<td>nnpin</td>
<td>nənˈɲi̞n̡</td>
<td>to hop on something</td>
</tr>
<tr>
<td>jilin</td>
<td>ɬi̞l̡i̞n̡</td>
<td>to jump over a high obstacle</td>
</tr>
<tr>
<td>tanlin</td>
<td>tanˈlin̡</td>
<td>[a type of civet]</td>
</tr>
<tr>
<td>lajin*</td>
<td>laˈjɪ̞d̪n</td>
<td>to be different</td>
</tr>
<tr>
<td>ben</td>
<td>ˈbe̞d̪n</td>
<td>to tie around</td>
</tr>
<tr>
<td>hadn</td>
<td>haˈde̞d̪n</td>
<td>tomorrow; near future; soon</td>
</tr>
<tr>
<td>?ncen</td>
<td>ʔənˈc̞e̞d̪n</td>
<td>to cook</td>
</tr>
<tr>
<td>ken</td>
<td>ˈkɛ̞d̪n</td>
<td>to hear</td>
</tr>
<tr>
<td>bnken ~ mnken (?)</td>
<td>mənˈkɛ̞d̪n</td>
<td>to astonish</td>
</tr>
<tr>
<td>munken*</td>
<td>mʊnˈkɛ̞d̪n</td>
<td>maybe</td>
</tr>
<tr>
<td>kawen</td>
<td>kəˈwe̞d̪n</td>
<td>to plant</td>
</tr>
<tr>
<td>blawen</td>
<td>bələˈwe̞d̪n</td>
<td>mango</td>
</tr>
<tr>
<td>pɛn</td>
<td>ˈpɛn̡</td>
<td>to hit with an object</td>
</tr>
<tr>
<td>ben</td>
<td>ˈbe̞d̪n</td>
<td>to wrap a blanket around one's body</td>
</tr>
<tr>
<td>taben</td>
<td>taˈbe̞d̪n</td>
<td>[a type of fruit]</td>
</tr>
<tr>
<td>kmten</td>
<td>kəˈrənte̞d̪n</td>
<td>wrinkles</td>
</tr>
<tr>
<td>ken</td>
<td>ˈkɛ̞d̪n</td>
<td>child; CLF: human</td>
</tr>
<tr>
<td>sriken</td>
<td>sərɨˈke̞d̪n</td>
<td>[a type of leaf]</td>
</tr>
<tr>
<td>maken</td>
<td>maˈkɛ̞d̪n</td>
<td>who?; whose?</td>
</tr>
<tr>
<td>gen</td>
<td>ˈge̞d̪n</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>bagen</td>
<td>baˈge̞d̪n</td>
<td>[a type of monitor lizard (Varanus)]</td>
</tr>
<tr>
<td>tomien*</td>
<td>toˈmɛ̞n̡</td>
<td>snakehead (Channa micropeltes)</td>
</tr>
<tr>
<td>tahanen</td>
<td>tahaˈnɛ̞n̡</td>
<td>[a type of rattan]</td>
</tr>
<tr>
<td>prmen</td>
<td>pərˈnɛ̞n̡</td>
<td>to refuse to give</td>
</tr>
<tr>
<td>len</td>
<td>ˈle̞d̪n</td>
<td>loincloth; to wear a loincloth</td>
</tr>
<tr>
<td>dore*</td>
<td>doˈre̞d̪n̡</td>
<td>durian (Durio zibethinus)</td>
</tr>
<tr>
<td>hawen</td>
<td>həˈwɛ̞n̡</td>
<td>wild boar (Sus scrofa); pig</td>
</tr>
<tr>
<td>tmwen*</td>
<td>toˈm̡wɛ̞n̡</td>
<td>[ethnonym: Temuan]</td>
</tr>
<tr>
<td>kin</td>
<td>ˈki̞n̡</td>
<td>hole</td>
</tr>
<tr>
<td>Word</td>
<td>Transliteration</td>
<td>Meaning</td>
</tr>
<tr>
<td>------</td>
<td>----------------</td>
<td>---------</td>
</tr>
<tr>
<td>gìn</td>
<td>'gi?n</td>
<td>to pour</td>
</tr>
<tr>
<td>rìn</td>
<td>'ri?n</td>
<td>to sit down</td>
</tr>
<tr>
<td>hmirìn</td>
<td>h?om?iri?n</td>
<td>to extinguish (by itself)</td>
</tr>
<tr>
<td>win</td>
<td>'wu?n</td>
<td>to crawl</td>
</tr>
<tr>
<td>jìn</td>
<td>'jìn</td>
<td>to hop on something to test its supporting capacity</td>
</tr>
<tr>
<td>jìn</td>
<td>'jìn</td>
<td>to stop</td>
</tr>
<tr>
<td>don</td>
<td>'ða?n</td>
<td>to run short/to run out</td>
</tr>
<tr>
<td>hnhòn</td>
<td>h?on'hò?n</td>
<td>to devour</td>
</tr>
<tr>
<td>papan*</td>
<td>pa'pa?n</td>
<td>plank; wall; [a type of snake]</td>
</tr>
<tr>
<td>lapan*</td>
<td>la'pa?n</td>
<td>eight</td>
</tr>
<tr>
<td>sampan*</td>
<td>sam'pa?n</td>
<td>boat</td>
</tr>
<tr>
<td>dan</td>
<td>'da?n</td>
<td>to take a hit</td>
</tr>
<tr>
<td>can</td>
<td>'c?a?n</td>
<td>foot</td>
</tr>
<tr>
<td>jan</td>
<td>'j?a?n</td>
<td>to throw something to the ground</td>
</tr>
<tr>
<td>ïikan*</td>
<td>ïi'ka?n</td>
<td>fish</td>
</tr>
<tr>
<td>bokan*</td>
<td>bo'ka?n</td>
<td>NEGATIVE MARKER</td>
</tr>
<tr>
<td>bagan*</td>
<td>ba'ga?n</td>
<td>[a type of fruit]</td>
</tr>
<tr>
<td>tagan</td>
<td>ta'ga?n</td>
<td>adult green tree python (<em>Chondropython viridis</em>)</td>
</tr>
<tr>
<td>tasan</td>
<td>ta'sa?n</td>
<td>flying squirrel (<em>Petaurista</em>)</td>
</tr>
<tr>
<td>tahan*</td>
<td>ta'ha?n</td>
<td>to endure</td>
</tr>
<tr>
<td>pìjan*</td>
<td>pi'ja?n</td>
<td>helping; serving (of food)</td>
</tr>
<tr>
<td>lunan*</td>
<td>lu'ja?n</td>
<td>binturong (<em>Arctitis binturong</em>)</td>
</tr>
<tr>
<td>smilan*</td>
<td>s?omi'ja?n</td>
<td>nine</td>
</tr>
<tr>
<td>bolan*</td>
<td>bo'la?n</td>
<td>moon</td>
</tr>
<tr>
<td>blan</td>
<td>bo'la?n</td>
<td>[toponym]</td>
</tr>
<tr>
<td>hiran</td>
<td>hi'ra?n</td>
<td>to be surprised</td>
</tr>
<tr>
<td>daran*</td>
<td>da'ra?n</td>
<td>[a type of tuber]</td>
</tr>
<tr>
<td>lawan*</td>
<td>la'wa?n</td>
<td>to kill</td>
</tr>
<tr>
<td>tûn</td>
<td>'tûn</td>
<td>that (you don't know)</td>
</tr>
<tr>
<td>dun</td>
<td>'du?n</td>
<td>to cover; to fill</td>
</tr>
<tr>
<td>ïûn</td>
<td>ïûn</td>
<td>there (you don't know)</td>
</tr>
<tr>
<td>banun*</td>
<td>mânûn</td>
<td>[toponym: Banun]</td>
</tr>
<tr>
<td>gnun</td>
<td>g?ônûn</td>
<td>bamboo</td>
</tr>
<tr>
<td>krlun</td>
<td>k?arlù?n</td>
<td>to retreat</td>
</tr>
<tr>
<td>run</td>
<td>'ru?n</td>
<td>to fall to the ground (of e.g. fruit, leaf)</td>
</tr>
<tr>
<td>pon*</td>
<td>'po?n</td>
<td>like; also</td>
</tr>
<tr>
<td>jpon*</td>
<td>j'spo?n</td>
<td>[ethnonym: Japanese]</td>
</tr>
<tr>
<td>sabon*</td>
<td>sa'bo?n</td>
<td>soap</td>
</tr>
<tr>
<td>ton</td>
<td>'to?n</td>
<td>that (you know)</td>
</tr>
<tr>
<td>ton ~ tnton</td>
<td>'to?n ~ t?an'to?n</td>
<td>to climb up (e.g. a rope); to balance</td>
</tr>
<tr>
<td>ïon</td>
<td>ïo?n</td>
<td>there (you know)</td>
</tr>
<tr>
<td>pan?on</td>
<td>pan?o?n</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>lon</td>
<td>'lo?n</td>
<td>to push</td>
</tr>
<tr>
<td>English</td>
<td>Jahai</td>
<td>Notes</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>talon</td>
<td>ta'lo'n</td>
<td>reticulated python (<em>Python reticulatus</em>)</td>
</tr>
<tr>
<td>tawon*</td>
<td>ta'wo'n</td>
<td>year</td>
</tr>
<tr>
<td>bon</td>
<td>'bo'n</td>
<td>[a type of fruit]</td>
</tr>
<tr>
<td>don</td>
<td>'do'n</td>
<td>great grandparent</td>
</tr>
<tr>
<td>/-t/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tawin</td>
<td>ta'wi'n</td>
<td>[a type of small spider]</td>
</tr>
<tr>
<td>twiñ</td>
<td>to'wi'n</td>
<td>headache</td>
</tr>
<tr>
<td>peñ</td>
<td>'pe'n</td>
<td>daylight; sunshine</td>
</tr>
<tr>
<td>hrñpêñ</td>
<td>harñj'peñ</td>
<td>goose bumps</td>
</tr>
<tr>
<td>teñ</td>
<td>'te'n</td>
<td>to plait; to make a net</td>
</tr>
<tr>
<td>peñ</td>
<td>'pe'n</td>
<td>to shoot (with bow or slingshot)</td>
</tr>
<tr>
<td>bñññ</td>
<td>bol?e'n</td>
<td>green</td>
</tr>
<tr>
<td>señ</td>
<td>'se'n</td>
<td>before; front</td>
</tr>
<tr>
<td>sññ kññ?</td>
<td>'se'n kññ?</td>
<td>west</td>
</tr>
<tr>
<td>hêñ</td>
<td>hêñ</td>
<td>tooth</td>
</tr>
<tr>
<td>smen</td>
<td>sa'mêñ</td>
<td>to ask for; to request</td>
</tr>
<tr>
<td>pñññ</td>
<td>to're'n</td>
<td>to climb</td>
</tr>
<tr>
<td>rñññ</td>
<td>rin'reñn</td>
<td>to cause pain</td>
</tr>
<tr>
<td>jñññ</td>
<td>jip'jëññ</td>
<td>to dream</td>
</tr>
<tr>
<td>tñññ</td>
<td>to'pi'n</td>
<td>part of honeycomb (?)</td>
</tr>
<tr>
<td>tagññ</td>
<td>ta'gi'n</td>
<td>wood (material); firewood</td>
</tr>
<tr>
<td>tñññ</td>
<td>t'sgi'n</td>
<td>to tear something apart with one's teeth</td>
</tr>
<tr>
<td>hñññ</td>
<td>h'amîññ</td>
<td>[magic word uttered to redress broken taboo]</td>
</tr>
<tr>
<td>darññ</td>
<td>da'ti'n</td>
<td>termite; small worm causing tooth decay</td>
</tr>
<tr>
<td>hñññ</td>
<td>hîñj'aññ</td>
<td>to rise; to stand</td>
</tr>
<tr>
<td>mamññ</td>
<td>mâ'mùñ</td>
<td>to beg</td>
</tr>
<tr>
<td>bññ</td>
<td>'bo'n</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>gññ</td>
<td>'go'n</td>
<td>[a type of outgrowth on trees]</td>
</tr>
<tr>
<td>ranññ</td>
<td>ran'go'n</td>
<td>mouth harp</td>
</tr>
<tr>
<td>trññ</td>
<td>tar'po'n</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>gñññ</td>
<td>'gëññ</td>
<td>to refuse to give something</td>
</tr>
<tr>
<td>ñññ</td>
<td>ñññ</td>
<td>to smell something</td>
</tr>
<tr>
<td>pñññ</td>
<td>pñññco'n</td>
<td>to sing</td>
</tr>
<tr>
<td>/-ñ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td>kñññ</td>
<td>kñpi'n</td>
<td>portion; piece; CLF: flat objects</td>
</tr>
<tr>
<td>pñññ</td>
<td>pøñpany</td>
<td>Philippine glossy starling (<em>Aplonis panayensis</em>)</td>
</tr>
<tr>
<td>kñññ</td>
<td>kør'pi'n</td>
<td>above; top; upper side</td>
</tr>
<tr>
<td>cñññ</td>
<td>cüntíñ</td>
<td>temple</td>
</tr>
<tr>
<td>gñññ</td>
<td>ga'di'n</td>
<td>boar's tusk; elephant's tusk</td>
</tr>
<tr>
<td>tñññ</td>
<td>ta'c'i'n</td>
<td>worm</td>
</tr>
<tr>
<td>cñññ</td>
<td>c'æñ'c'ññ</td>
<td>biorbital area; spectacle</td>
</tr>
<tr>
<td>Word</td>
<td>Meaning</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>jīn</td>
<td>to take</td>
<td></td>
</tr>
<tr>
<td>?asin*</td>
<td>to be different; other</td>
<td></td>
</tr>
<tr>
<td>masi-masin*</td>
<td>each</td>
<td></td>
</tr>
<tr>
<td>pusin*</td>
<td>to move in circles</td>
<td></td>
</tr>
<tr>
<td>kpsin</td>
<td>banded palm civet (<em>Hemigalus derbyanus</em>)</td>
<td></td>
</tr>
<tr>
<td>cinihin</td>
<td>to turn head (of bird)</td>
<td></td>
</tr>
<tr>
<td>kamin*</td>
<td>goat</td>
<td></td>
</tr>
<tr>
<td>bani*</td>
<td>[toponym: Banding]</td>
<td></td>
</tr>
<tr>
<td>kunin*</td>
<td>yellow</td>
<td></td>
</tr>
<tr>
<td>pinin*</td>
<td>headache; confusion</td>
<td></td>
</tr>
<tr>
<td>bali</td>
<td>tiger (<em>Panthera tigris</em>)</td>
<td></td>
</tr>
<tr>
<td>malin*</td>
<td>to steal</td>
<td></td>
</tr>
<tr>
<td>bli</td>
<td>upper arm</td>
<td></td>
</tr>
<tr>
<td>klin</td>
<td>language; sound</td>
<td></td>
</tr>
<tr>
<td>cinhin</td>
<td>to look over one's shoulder</td>
<td></td>
</tr>
<tr>
<td>grunglin</td>
<td>to roll down</td>
<td></td>
</tr>
<tr>
<td>knin</td>
<td>to make sound: e.g. howl/croak/bark</td>
<td></td>
</tr>
<tr>
<td>brin</td>
<td>long-tailed parakeet (<em>Psittacula longicauda</em>)</td>
<td></td>
</tr>
<tr>
<td>barin</td>
<td>[toponym: Baring]</td>
<td></td>
</tr>
<tr>
<td>krin*</td>
<td>to be dry</td>
<td></td>
</tr>
<tr>
<td>cnrin</td>
<td>to hold one's wrist</td>
<td></td>
</tr>
<tr>
<td>cawin</td>
<td>eyebrow</td>
<td></td>
</tr>
<tr>
<td>rinwin</td>
<td>to dangle one's arms when walking</td>
<td></td>
</tr>
<tr>
<td>kalwin</td>
<td>[a type of fruit]</td>
<td></td>
</tr>
<tr>
<td>sajin</td>
<td>friend</td>
<td></td>
</tr>
<tr>
<td>horjin</td>
<td>[a type of small animal]</td>
<td></td>
</tr>
<tr>
<td>ten</td>
<td>bamboo slat</td>
<td></td>
</tr>
<tr>
<td>?ntej</td>
<td>ear</td>
<td></td>
</tr>
<tr>
<td>kucen*</td>
<td>cat</td>
<td></td>
</tr>
<tr>
<td>jen</td>
<td>to hear</td>
<td></td>
</tr>
<tr>
<td>ken</td>
<td>to pull</td>
<td></td>
</tr>
<tr>
<td>panken*</td>
<td>chair</td>
<td></td>
</tr>
<tr>
<td>jjen*</td>
<td>bone; skeleton</td>
<td></td>
</tr>
<tr>
<td>pipen</td>
<td>fluid</td>
<td></td>
</tr>
<tr>
<td>barjen*</td>
<td>[a type of tortoise]</td>
<td></td>
</tr>
<tr>
<td>kensen</td>
<td>to dance</td>
<td></td>
</tr>
<tr>
<td>ljen</td>
<td>phlegm; saliva</td>
<td></td>
</tr>
<tr>
<td>jlen</td>
<td>to put inside</td>
<td></td>
</tr>
<tr>
<td>jren</td>
<td>soul</td>
<td></td>
</tr>
<tr>
<td>pen*</td>
<td>to chop vegetables</td>
<td></td>
</tr>
<tr>
<td>tampen</td>
<td>to ascend</td>
<td></td>
</tr>
<tr>
<td>baten*</td>
<td>[a type of rattan]</td>
<td></td>
</tr>
<tr>
<td>binten*</td>
<td>star</td>
<td></td>
</tr>
<tr>
<td>tunten</td>
<td>to slither</td>
<td></td>
</tr>
<tr>
<td>jen</td>
<td>whiskers</td>
<td></td>
</tr>
</tbody>
</table>
lahai rhyming glossary

197

to be wide
to hold
to be in pain; to scratch (of monkey hit by blowpipe dart)
to say; to speak; to talk; to tell
to walk or run with a limp
siamang (Hylobates syndactylus)
betelnut
training
[a type of tree]
thread
to be easy
to win
to count
[a type of tuber]
inside
bird of prey
crested serpent-eagle (Spilornis cheela)
brahminy kite (Haliastur indus)
to see at a distance
egret (Egretta); [a type of cobra (Naja)]
[a type of monitor lizard (Varanus)]
[a type of fruit tree]
to lean one's head
[a type of python (Python)]
scorpion (generic)
great hornbill (Buceros bicornis)
to be awake
to flap wings
big permanent house
to stretch something
small hut
[a type of wasp]
[a type of tuber]
to think of someone
to be blunt (of knife)
[toponym: Chiong]
to burn
[a type of insect]
leaf monkey (Semnopithecus)
skull
pharynx
[a type of small animal]
fish spear
to turn head (of bird)
<table>
<thead>
<tr>
<th>Noun</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>caban</td>
<td>branch</td>
</tr>
<tr>
<td>caban t'om</td>
<td>tributary</td>
</tr>
<tr>
<td>kaban</td>
<td>[a type of fruit]</td>
</tr>
<tr>
<td>kaban</td>
<td>mud</td>
</tr>
<tr>
<td>bata'</td>
<td>CLF: oblong objects</td>
</tr>
<tr>
<td>ganta'</td>
<td>potato</td>
</tr>
<tr>
<td>t'anta'</td>
<td>greater racket-tailed drongo (<em>Dicturus paradiseus</em>)</td>
</tr>
<tr>
<td>bidan</td>
<td>CLF: large flat objects</td>
</tr>
<tr>
<td>padan</td>
<td>field</td>
</tr>
<tr>
<td>cadan</td>
<td>to be ready to do something</td>
</tr>
<tr>
<td>sladan</td>
<td>wild ox/gaur (<em>Bos gaurus</em>)</td>
</tr>
<tr>
<td>pance' ~ pance'</td>
<td>to light up; to shine</td>
</tr>
<tr>
<td>cinca'</td>
<td>to mince</td>
</tr>
<tr>
<td>kanka'</td>
<td>Asian horned toad (<em>Megophrys monticola</em>)</td>
</tr>
<tr>
<td>pisa'</td>
<td>banana</td>
</tr>
<tr>
<td>lasa'</td>
<td>taste</td>
</tr>
<tr>
<td>man</td>
<td>to prevent something from falling</td>
</tr>
<tr>
<td>sima*</td>
<td>to make magic (?)</td>
</tr>
<tr>
<td>braman</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>kma*</td>
<td>to swell</td>
</tr>
<tr>
<td>?na'</td>
<td>side</td>
</tr>
<tr>
<td>pada*</td>
<td>to roast</td>
</tr>
<tr>
<td>han*</td>
<td>rhinoceros hornbill (<em>Buceros rhinoceros</em>)</td>
</tr>
<tr>
<td>hila*</td>
<td>to spin; to whirl</td>
</tr>
<tr>
<td>bala*</td>
<td>bottle</td>
</tr>
<tr>
<td>kala'</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>balan</td>
<td>[a type of large grasshopper]</td>
</tr>
<tr>
<td>baran</td>
<td>thing</td>
</tr>
<tr>
<td>karan</td>
<td>[a type of tree-growing fern]</td>
</tr>
<tr>
<td>sara'</td>
<td>roof</td>
</tr>
<tr>
<td>pra'</td>
<td>to shoot</td>
</tr>
<tr>
<td>kran</td>
<td>[sound of carving]</td>
</tr>
<tr>
<td>sinra'</td>
<td>to chatter</td>
</tr>
<tr>
<td>tiran</td>
<td>[a type of fish]</td>
</tr>
<tr>
<td>swan</td>
<td>to go around</td>
</tr>
<tr>
<td>twan</td>
<td>blue coral snake (<em>Maticora bivirgata</em>)</td>
</tr>
<tr>
<td>kwan</td>
<td>flying fox</td>
</tr>
<tr>
<td>jan</td>
<td>RELATIVE MARKER</td>
</tr>
<tr>
<td>tjan ~ tjin</td>
<td>to shake something</td>
</tr>
<tr>
<td>ljun'</td>
<td>to beat (of heart)</td>
</tr>
<tr>
<td>tun</td>
<td>to fell; to bring down</td>
</tr>
<tr>
<td>katu'</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>cantun</td>
<td>to stamp one’s feet; to play bamboo tubes</td>
</tr>
</tbody>
</table>
gantu [a type of large poisonous spider]
tudu [to cover one's eyes from the sun]
ganju [a type of civet]
ki [a type of tree]
lu [mortar; [toponym]]
lu [cave]
ji [to move ears before attacking (of predator)]
hajam [a type of fruit]
tul [a type of millipede]
ju [flat-headed cat (Felis planiceps)]
wu [a type of catfish (Mystus planiceps)]
jwu [to move in circles]
ju [umbrella]
ju [myna (Acridotheres)]
ju [fang]
ju [to play flute]
po [village]
po [to float]
poj [to roll oneself]
tjo [collared kingfisher (Halcyon chloris)]
kalo [knee]
to [a type of snake]
ko [honey (?)]
rin [a type of frog]
ko [to trumpet (of elephant)]
ho [to bark]
hanho [to blow (of person); to whistle]
samo [to put something together]
no [wooden material]
lo [a type of snake]
ro [a type of lizard]
wo [to build a house]
kwon [great argus (Argusianus argus)]
tanw [to carry on one's shoulder]
lao [a type of tree]
po [to tap poison from ipoh tree]
mpo [hole]
mpo kit [anus]
sp [leafbird (Chloropsis)]
kat [claw; nail]
kat cjas [finger nail]
kat can [toe nail]
c [to hit from above (e.g. of falling fruit)]
\begin{verbatim}
200 Appendix I

kaŋcon [kaŋ'cəŋŋ] [a type of insect]
jilong [j'ɔŋŋ] to smell (of rotten wood)
koŋ [kəŋ] to plait
ʔiŋ [ʔiŋ] water
pʰiŋ [pʰiŋ] to heat poisoned blowpipe dart in fire
pahšiŋ [pahšiŋ] thrush \textit{(Zoothera)}
rəŋhlyŋ ŋos [rəŋhlyŋ ŋos] charcoal
lamon [la'mɔŋŋ] [a type of fruit]
tanong [ta'nɔŋŋ] neuropteran (generic)
rəŋ [rəŋ] [a type of snake]
cnong [cə'nɔŋŋ] beak/bill/nib
ʔiŋ [ʔiŋ] masked civet \textit{(Paguma larvata)}
balong [ba'lɔŋŋ] [a type of tree]
dalong [da'lɔŋŋ] [a type of lizard]
clon [cə'lɔŋŋ] wild dog \textit{(Cuon alpinus)}
glon [gə'lɔŋŋ] wide path; way
snlon [sə'nɔŋŋ] to remain in one place to make a fire for a returning hunting party

barong [ba'rɔŋŋ] Malay tapir \textit{(Tapirus indicus)}
manrong [ma'nɔŋŋ] skink \textit{(Emoia)}
wong [wɔŋŋ] child; offspring
kwon [kə'wɔŋŋ] dove
brwɔŋ [brwɔŋŋ] [a type of frog]
jɔŋ [jɔŋ] [a type of salty tuber]
pjɔŋ [pjɔŋŋ] [a type of turtle]
kaŋjɔŋ [kə'ŋjɔŋŋ] elbow
brjon [br'jɔŋŋ] to burn

/-l/

kpil [kə'pil] [a type of turtle]
tampil [tam'pil] slow lori \textit{(Nycticebus coucang)}
knabil [kə'na'bil] [a type of large spider]
til [tɪl] to be hot
cantil [cə'ntil] to lose one’s footing
bdil* [bə'dil] gun; to shoot with gun
klidil [kə'lidil] [a type of snake]
cil [c'ɪl] to blink
wakil [wa'kil] younger sibling
saŋkil [saŋ'kil] to peck (of bird)
cilcil [c'ɪl'ɪl] to look around
kulkil [kʊl'kil] to fall
tigil [tɪ'gil] to go along a mountain-side
sil [sɪl] [a type of tortoise]
rañsil [rañ'sil] [a type of tree]
khil [kə'hl] to pound; to hit somebody with one’s elbow
\end{verbatim}
<table>
<thead>
<tr>
<th>Jahai rhyme glossary 201</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>karil</strong></td>
</tr>
<tr>
<td><strong>kril</strong></td>
</tr>
<tr>
<td><strong>klwil</strong></td>
</tr>
<tr>
<td><strong>kajil</strong></td>
</tr>
<tr>
<td><strong>pel</strong></td>
</tr>
<tr>
<td><strong>tel</strong></td>
</tr>
<tr>
<td><strong>mhel</strong></td>
</tr>
<tr>
<td><strong>wel</strong></td>
</tr>
<tr>
<td><strong>pel</strong></td>
</tr>
<tr>
<td><strong>tabel</strong></td>
</tr>
<tr>
<td><strong>cntel</strong></td>
</tr>
<tr>
<td><strong>bdel</strong></td>
</tr>
<tr>
<td><strong>gel</strong></td>
</tr>
<tr>
<td><strong>Fel</strong></td>
</tr>
<tr>
<td><strong>rasel</strong></td>
</tr>
<tr>
<td><strong>psel</strong></td>
</tr>
<tr>
<td><strong>kisel</strong></td>
</tr>
<tr>
<td><strong>hel</strong></td>
</tr>
<tr>
<td><strong>ghel</strong></td>
</tr>
<tr>
<td><strong>mhel</strong></td>
</tr>
<tr>
<td><strong>tanel</strong></td>
</tr>
<tr>
<td><strong>klel</strong></td>
</tr>
<tr>
<td><strong>hnel</strong></td>
</tr>
<tr>
<td><strong>wel</strong></td>
</tr>
<tr>
<td><strong>criwel</strong></td>
</tr>
<tr>
<td><strong>kjel</strong></td>
</tr>
<tr>
<td><strong>karjel</strong></td>
</tr>
<tr>
<td><strong>tampil</strong></td>
</tr>
<tr>
<td><strong>bil</strong></td>
</tr>
<tr>
<td><strong>til</strong></td>
</tr>
<tr>
<td><strong>pital</strong></td>
</tr>
<tr>
<td><strong>titil</strong></td>
</tr>
<tr>
<td><strong>dldil</strong></td>
</tr>
<tr>
<td><strong>cil</strong></td>
</tr>
<tr>
<td><strong>kil</strong></td>
</tr>
<tr>
<td><strong>gil</strong></td>
</tr>
<tr>
<td><strong>jil?il</strong></td>
</tr>
<tr>
<td><strong>jil</strong></td>
</tr>
<tr>
<td><strong>?ipal</strong></td>
</tr>
<tr>
<td><strong>kmbarbol</strong></td>
</tr>
<tr>
<td><strong>k bol</strong></td>
</tr>
<tr>
<td><strong>tankal</strong></td>
</tr>
<tr>
<td><strong>mal</strong></td>
</tr>
<tr>
<td><strong>gimol</strong></td>
</tr>
<tr>
<td>Word</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>mlml</td>
</tr>
<tr>
<td>balol</td>
</tr>
<tr>
<td>kapal*</td>
</tr>
<tr>
<td>smpal</td>
</tr>
<tr>
<td>gibal</td>
</tr>
<tr>
<td>bdal*</td>
</tr>
<tr>
<td>kal</td>
</tr>
<tr>
<td>tmkal</td>
</tr>
<tr>
<td>hankal</td>
</tr>
<tr>
<td>mkal</td>
</tr>
<tr>
<td>kpral*</td>
</tr>
<tr>
<td>buwal*</td>
</tr>
<tr>
<td>juwal*</td>
</tr>
<tr>
<td>kwal</td>
</tr>
<tr>
<td>kilwal</td>
</tr>
<tr>
<td>trbul*</td>
</tr>
<tr>
<td>tul</td>
</tr>
<tr>
<td>kul</td>
</tr>
<tr>
<td>cankul*</td>
</tr>
<tr>
<td>gul</td>
</tr>
<tr>
<td>timul*</td>
</tr>
<tr>
<td>tunul</td>
</tr>
<tr>
<td>sjul</td>
</tr>
<tr>
<td>brul</td>
</tr>
<tr>
<td>critol</td>
</tr>
<tr>
<td>btol*</td>
</tr>
<tr>
<td>dol ~ pidol</td>
</tr>
<tr>
<td>kol</td>
</tr>
<tr>
<td>kalkol</td>
</tr>
<tr>
<td>gol</td>
</tr>
<tr>
<td>pr?ol ~ pr?il</td>
</tr>
<tr>
<td>bisol*</td>
</tr>
<tr>
<td>?nsol</td>
</tr>
<tr>
<td>brol</td>
</tr>
<tr>
<td>jol</td>
</tr>
<tr>
<td>butol*</td>
</tr>
<tr>
<td>dad5l</td>
</tr>
<tr>
<td>col</td>
</tr>
<tr>
<td>hkol</td>
</tr>
<tr>
<td>snk5l</td>
</tr>
<tr>
<td>tunk5l</td>
</tr>
<tr>
<td>gol</td>
</tr>
<tr>
<td>p?ol</td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>bhol</td>
</tr>
<tr>
<td>jilmol</td>
</tr>
<tr>
<td>brawol</td>
</tr>
<tr>
<td>hiljol</td>
</tr>
<tr>
<td>/prpr/</td>
</tr>
<tr>
<td>pikir*</td>
</tr>
<tr>
<td>sir</td>
</tr>
<tr>
<td>srsir</td>
</tr>
<tr>
<td>hir</td>
</tr>
<tr>
<td>prhir</td>
</tr>
<tr>
<td>samir</td>
</tr>
<tr>
<td>crmir</td>
</tr>
<tr>
<td>ger taji?</td>
</tr>
<tr>
<td>her</td>
</tr>
<tr>
<td>trwer</td>
</tr>
<tr>
<td>tamper*</td>
</tr>
<tr>
<td>?amper</td>
</tr>
<tr>
<td>ber</td>
</tr>
<tr>
<td>prber</td>
</tr>
<tr>
<td>jader</td>
</tr>
<tr>
<td>pancer*</td>
</tr>
<tr>
<td>munker</td>
</tr>
<tr>
<td>krker</td>
</tr>
<tr>
<td>pher</td>
</tr>
<tr>
<td>tmer</td>
</tr>
<tr>
<td>jawer</td>
</tr>
<tr>
<td>kwer*</td>
</tr>
<tr>
<td>pir</td>
</tr>
<tr>
<td>crbir</td>
</tr>
<tr>
<td>tir</td>
</tr>
<tr>
<td>gir ~ gar</td>
</tr>
<tr>
<td>higir</td>
</tr>
<tr>
<td>sagir</td>
</tr>
<tr>
<td>grgir</td>
</tr>
<tr>
<td>bla'gir</td>
</tr>
<tr>
<td>sir</td>
</tr>
<tr>
<td>bhir</td>
</tr>
<tr>
<td>hawir ~ hawor</td>
</tr>
<tr>
<td>tapor</td>
</tr>
<tr>
<td>tar</td>
</tr>
<tr>
<td>sitor ~ hitor</td>
</tr>
</tbody>
</table>
hantər*  
?akər  
kanər*  
sajar  
sjar  
prjər  
tapar* cjas  
star  
gntar*  
cicar  
?a.jar*  
kjər  
tukar*  
har  
kmər*  
lənjar*  
wər  
gugur*  
pir?ur  
brhur  
simur*  
smur*  
'dor  
dapor* ?ɛm  
gor  
hor  
smlər  
pər  
kubər*  
sənantər  
dor  
tac5r  
?ikər*  
tgər*  
?or  
sər  
bsər  
hər  
jər  
snpər  
lər  
wər  

hantər: to send  
?akər: [a type of snake]  
kanər: to carry on one's shoulder, e.g. an animal  
sajar: flock/herd  
sjar: to swim  
pərjər: to accompany  
tapər: cjas palm of hand  
star: [a type of fruit]  
gəntər: to shake; to vibrate  
cicər: [a type of tree]  
?ajar: to learn  
kəjər: [toponym: Kejar]  
tukər: to exchange  
'har: small path/trail  
kmər: twins  
lənjar: to crash  
wər: plaintive cuckoo (*Cacomantis merulinus)  
gugur: to lose hair  
pir?ur: to growl (of stomach)  
brhur: [a type of snake]  
simər: east  
smur: to spray water  
'dor: to balance  
dapor: ?ɛm nipple  
gor: lower leg  
hor: striped tit-babbler (*Macronous gularis); to have a running nose  
smlər: [toponym: Semelor]  
pər: tiger shrike (*Lanius tigrinus)  
kubər: grave  
sənantər: [a type of snake]  
dor: to spread something  
tac5r: treeshrew (*Tupia)  
?ikər: CLF: animals  
tgər: to allow (?)  
?or: to order; to request  
sər: [a type of cockroach]  
bsər: [a type of tree]  
hər: to drill or carve a hole  
jər: coconut palm  
snpər: silence; to be silent  
lər: to hiss  
wər: to go around
<table>
<thead>
<tr>
<th>Root</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>bgiw</td>
<td>wind</td>
</tr>
<tr>
<td>knsiw</td>
<td>[ethnonym: Kensiw]</td>
</tr>
<tr>
<td>hwhiw</td>
<td>crested wood-partridge (<em>Rollulus rouloul</em>)</td>
</tr>
<tr>
<td>?iwunjw ~ ?uwunjw</td>
<td>to gaze; to look</td>
</tr>
<tr>
<td>putew</td>
<td>tiger (<em>Panthera tigris</em>)</td>
</tr>
<tr>
<td>pew</td>
<td>different; other</td>
</tr>
<tr>
<td>jritew</td>
<td>to squat</td>
</tr>
<tr>
<td>sew</td>
<td>[a type of small mammal]</td>
</tr>
<tr>
<td>hew</td>
<td>to eat (vegetables, fruit)</td>
</tr>
<tr>
<td>mew</td>
<td>[a type of civet]</td>
</tr>
<tr>
<td>?amew</td>
<td>cat</td>
</tr>
<tr>
<td>jew</td>
<td>to have a runny nose</td>
</tr>
<tr>
<td>pow</td>
<td>[sound of a shotgun]</td>
</tr>
<tr>
<td>bow</td>
<td>to be big</td>
</tr>
<tr>
<td>jwunjw</td>
<td>Achilles tendon</td>
</tr>
<tr>
<td>gow</td>
<td>to go along the crest of a ridge</td>
</tr>
<tr>
<td>b?ow</td>
<td>whip snake (<em>Ahaetulla</em>)</td>
</tr>
<tr>
<td>haw</td>
<td>[a type of tuber (<em>Discorea pentaphylla?</em>)]</td>
</tr>
<tr>
<td>maw</td>
<td>to make sound (of monkey)</td>
</tr>
<tr>
<td>praw</td>
<td>[sound of an animal fleeing from one tree to another]</td>
</tr>
<tr>
<td>paw</td>
<td>side of body</td>
</tr>
<tr>
<td>rabaw</td>
<td>[toponym: Raba]</td>
</tr>
<tr>
<td>hntaw</td>
<td><em>petai</em> fruit</td>
</tr>
<tr>
<td>daw</td>
<td>to reach (of e.g. tree branch)</td>
</tr>
<tr>
<td>cdaw</td>
<td>rainbow</td>
</tr>
<tr>
<td>kakaw</td>
<td>[a type of monitor lizard (<em>Varanus</em>)]</td>
</tr>
<tr>
<td>bukaw</td>
<td>flower (generic)</td>
</tr>
<tr>
<td>gaw</td>
<td>wild boar (<em>Sus scrofa</em>)</td>
</tr>
<tr>
<td>mnsaw</td>
<td>daughter/son-in-law</td>
</tr>
<tr>
<td>blaw</td>
<td>blowpipe</td>
</tr>
<tr>
<td>waw</td>
<td>[a type of civet]</td>
</tr>
<tr>
<td>kawaw</td>
<td>bird (generic)</td>
</tr>
<tr>
<td>rjaw</td>
<td>[a type of wasp]</td>
</tr>
<tr>
<td>snow</td>
<td>[a type of cobra]</td>
</tr>
<tr>
<td>rampow</td>
<td>long-tailed macaque (<em>Macaca fascicularis</em>)</td>
</tr>
<tr>
<td>krbow*</td>
<td>buffalo</td>
</tr>
<tr>
<td>litow</td>
<td>boy; young bachelor</td>
</tr>
<tr>
<td>kacow*</td>
<td>to work</td>
</tr>
<tr>
<td>hijow*</td>
<td>to be green</td>
</tr>
<tr>
<td>kjow</td>
<td>to take off; to fly up</td>
</tr>
<tr>
<td>sikow</td>
<td>to roar; to meow (of leopard, cat)</td>
</tr>
<tr>
<td>bakow</td>
<td>[a type of tree]</td>
</tr>
<tr>
<td>tmakow*</td>
<td>tobacco</td>
</tr>
</tbody>
</table>
App. I

manòw*  mə'nòw  manau rattan
kalòw*  kə'lòw  if
salòw  sə'lòw  to split
halòw*  hə'lòw  to scare
pulòw*  pu'lòw  island
pulòw tujoh*  pu'lòw tu'joh  [toponym: Pulau Tujuh]
baròw  bə'ròw  oriental white-eye (Zosterops palpebrosus)
gurow*  gu'ròw  to joke/to jest
smrow*  səm'ròw  to heal
cowrow  cə'cw'ròw  [a type of bird]
jòw  'jòw  to take

/-j/

lamij  la'mij  sister/brother-in-law
gej  'gẹj  to eat
pregej  pə'r'gej  to feed
?j?ej tom  ?i'j?eʃ tə'm  steam
sej  'sej  to raise something
lej  'lej  body
pantej*  pan'teʃ  sand
lantej*  lan'teʃ  [toponym]
krtej  kər'teʃ  to ascend to a mountain pass
dej  'deʃ  rice in the husk; unharvested rice;
     [a type of small green snake]
padej*  pa'deʃ  shop
kdej*  kə'deʃ  such; [a type of tuber (yam?)]
cej  'ceʃ  to be big
cankej  cə'arı'keʃ  [a type of frog]
srankej  sə'raʃ'keʃ  [a type of large spider]
?ej  ?eʃ  father
sej  'seʃ  to shovel
hej  'heʃ  PRONOUN 1D.INCL
mej  'meʃ  what?
mej ~ mèj (?)  'mèʃ  to comb
kmej  kə'meʃ  [a type of larvae]
smej  sə'meʃ  [ethnonym: Semai]
nej  'nɛʃ  one
manej  mə'nɛʃ  otter (Lutra sumatrana)
bralej ~ blalej*  bə'ralə'ʃ ~ bə'la'leʃ  elephant’s trunk
gulej  gu'lə'ʃ  to whirl
karej  ka'reʃ  thunder; an evil human-like mythical being
bcrej*  bəcə'ɾeʃ  to repudiate one’s wife
sorej  so'ɾeʃ  [a type of tortoise]
<table>
<thead>
<tr>
<th>Morej* Batu?</th>
<th>Mo'rej Ba'tu?</th>
<th>Magpie Robin (<em>Copsychus saularis</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wej</td>
<td>'Wej</td>
<td>Past</td>
</tr>
<tr>
<td>Gawej*</td>
<td>Ga'wej</td>
<td>To play; to work</td>
</tr>
<tr>
<td>?Awej</td>
<td>?A'wej</td>
<td>Rattan; root; rope; string; vine</td>
</tr>
<tr>
<td>Lawej</td>
<td>La'wej</td>
<td>Leech</td>
</tr>
<tr>
<td>Lwej</td>
<td>Lo'wej</td>
<td>Honey</td>
</tr>
<tr>
<td>Jwej</td>
<td>Ja'wej</td>
<td>[A type of tree]</td>
</tr>
<tr>
<td>Jej</td>
<td>Jej</td>
<td>Fly</td>
</tr>
<tr>
<td>Kapij</td>
<td>Ka'pij</td>
<td>To fly</td>
</tr>
<tr>
<td>Tij</td>
<td>Ti'j</td>
<td>Thunder; river source</td>
</tr>
<tr>
<td>Sdiij</td>
<td>Sa'diij</td>
<td>Vomit</td>
</tr>
<tr>
<td>Hldij</td>
<td>Ha'li'dij</td>
<td>Flat ground</td>
</tr>
<tr>
<td>Hakij</td>
<td>Ha'kij</td>
<td>To be slow</td>
</tr>
<tr>
<td>Gij</td>
<td>Gi'j</td>
<td>To produce/blow smoke</td>
</tr>
<tr>
<td>Sij</td>
<td>Si'j</td>
<td>To abound (of game)</td>
</tr>
<tr>
<td>Ksjj</td>
<td>Ka'sij</td>
<td>Husband</td>
</tr>
<tr>
<td>Chij</td>
<td>Ca'hij</td>
<td>Moon</td>
</tr>
<tr>
<td>Brij</td>
<td>Ba'rij</td>
<td>Late afternoon</td>
</tr>
<tr>
<td>Jij</td>
<td>Ji'j</td>
<td>To carry in one's hand; to drive (a car); to transport</td>
</tr>
<tr>
<td>Poj</td>
<td>Po'j</td>
<td>To produce/blow smoke</td>
</tr>
<tr>
<td>Mantaj</td>
<td>Man'ta'j</td>
<td>Pangolin (<em>Manis</em>)</td>
</tr>
<tr>
<td>Poj</td>
<td>Po'j</td>
<td>Cave</td>
</tr>
<tr>
<td>?Haj</td>
<td>?Ha'ja'j</td>
<td>To be small</td>
</tr>
<tr>
<td>HjHaj</td>
<td>Hi'ja'j</td>
<td>To yawn</td>
</tr>
<tr>
<td>Paj</td>
<td>Pa'j</td>
<td>New; PRONOUN 2S.DIS</td>
</tr>
<tr>
<td>Baj</td>
<td>Ba'j</td>
<td>To dig</td>
</tr>
<tr>
<td>Lambaj*</td>
<td>Lam'ba'j</td>
<td>To wave</td>
</tr>
<tr>
<td>Ja'j</td>
<td>Ja'j</td>
<td>Banana</td>
</tr>
<tr>
<td>Kaj</td>
<td>Ka'j</td>
<td>[A type of tree]</td>
</tr>
<tr>
<td>Ka'j</td>
<td>Ka'j</td>
<td>To heal</td>
</tr>
<tr>
<td>Pakaj*</td>
<td>Pa'ka'j</td>
<td>To eat a lot</td>
</tr>
<tr>
<td>Crkaj</td>
<td>Ca'rej'ka'j</td>
<td>[Toponym]</td>
</tr>
<tr>
<td>Ga'j</td>
<td>Ga'j</td>
<td>To turn/roll over in one's sleep</td>
</tr>
<tr>
<td>Tungaj</td>
<td>Tur'ga'j</td>
<td>Headman</td>
</tr>
<tr>
<td>?Aaj</td>
<td>?A'aj</td>
<td>Game animal</td>
</tr>
<tr>
<td>Ca'aj</td>
<td>Ca'a'aj</td>
<td>To sing</td>
</tr>
<tr>
<td>Pli'saj</td>
<td>Poli'sa'j</td>
<td>Antenna (of fish)</td>
</tr>
<tr>
<td>Ha'j</td>
<td>Ha'j</td>
<td>To follow</td>
</tr>
<tr>
<td>Jaha'j ~ Jhaj (?)</td>
<td>Ja'a'ha'j</td>
<td>[Ethnonym: Jahai]</td>
</tr>
<tr>
<td>Rmaj</td>
<td>Ro'ma'j</td>
<td>[A type of fruit]</td>
</tr>
<tr>
<td>?Naaj</td>
<td>?Na'a'j</td>
<td>To bathe</td>
</tr>
<tr>
<td>Blaj</td>
<td>Bo'la'j</td>
<td>To look upwards</td>
</tr>
<tr>
<td>Slaj</td>
<td>Sa'la'j</td>
<td>Clearing; swidden</td>
</tr>
<tr>
<td>Caraj</td>
<td>Ca'ra'j</td>
<td>Gliding frog (<em>Rhacophorus</em>)</td>
</tr>
</tbody>
</table>
manraj man\textsuperscript{d}ra\textsuperscript{j} [a type of tuber]
waj 'wa\textsuperscript{j} to cut
rwaj ra\textsuperscript{w}a\textsuperscript{j} soul
duj 'duj to go hunting/gathering late afternoon
kuj 'ku\textsuperscript{j} head
kuruhuj kuru'hu\textsuperscript{j} [a type of owl]
\textnuj ?\textnu\textsuperscript{j} soon
tan\textnuj ta\textnu\textsuperscript{j} rambutan (*Nephelium lappaceum*)
pruj pe\textnu\textsuperscript{j} to sprinkle food to animals
toj 'toj older brother of parent
toj bi? toj 'bi? older sister of parent
?oj ?o\textsuperscript{j} to open; to pull off (e.g. clothes)
huroj hu'ro\textsuperscript{j} to become full
poj 'po\textsuperscript{j} to dry in the sun
ha'poj ha'po\textsuperscript{j} hut
smpoj mit s\textsuperscript{m}poj 'mit' eyelid
bjboj bi'boj [a type of insect]
btoj bo\textsuperscript{t}oj to be red
kad\textsuperscript{d}oj ka'd\textsuperscript{d}oj [a type of long-snouted insect]
takoj ta'ko\textsuperscript{j} sail-fin lizard (*Hydrosaurus*)
gr\textsuperscript{d}oj ra'go\textsuperscript{j} to be red
?oj ?o\textsuperscript{j} INTERJECTION: hey!
th\textsuperscript{d}oj t\textsuperscript{d}ho\textsuperscript{j} to be red
cnh\textsuperscript{d}oj c\textsuperscript{\textnu}nh\textsuperscript{d}oj noon
noj 'no\textsuperscript{j} to copulate
la\textsuperscript{\nu}oj la\textsuperscript{\nu}oj' shadow
loj 'lo\textsuperscript{j} to run; to flee; to hurry
prloj po'lo\textsuperscript{j} star
roj 'ro\textsuperscript{j} to leave
woj 'wo\textsuperscript{j} banded kingfisher (*Lacedo pulchella*)
tw\textsuperscript{o}j t\textsuperscript{\nu}w\textsuperscript{o}j evening/night
kw\textsuperscript{o}j k\textsuperscript{\nu}w\textsuperscript{o}j [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)  sij
accompany  prjor
Achilles tendon  jwjaw
adorn  ksep
after  lpos*
afternoon  brij
again  wel
aim (with blowpipe or gun) (v.)  btol*
aircraft  kapal*
all  smwe?*
allow (?)  tgor*
alone  bla?
also  haji?
ambush (v.)  ?imbus*
ancestor  to?
and  ?alɔ?
angry  marah*
animal (game)  soh, ?aj
animal, type of  ptpit
animal, type of  rajɔt
animal, type of  paļk
animal, type of  ?awik
animal, type of  ti? siriŋ
animal, type of  cawis
animal, type of
animal, type of
animal, type of
animal, type of
animal, type of
animal, type of
ankle
ankle-joint
answer (v.)
ant, type of
ant, type of
ant, type of
antenna (of fish)
anus
apple
argue
arm, lower
arm, upper
armpit
arrive
arrowhead
ascend
ashes
Asian horned toad (*Megophrys monticola*)
ask
assemble
astonish
attack (v.)
awake
axe
babbler (*Malacopteron*)
bachelor, old
back (body part)
backside
bad
bag
balance (v.)
bamboo
bamboo slat
banana
banded kingfisher (*Lacedo pulchella*)
banded palm civet (*Hemigalus derbianus*)
barbet (*Megalaima*)
bark (v.)
basket
tamam
họrjĩn
raŋ
raŋel
hŋrel
sèw
kmkɔm
kril
jawap*, balas*
bit
hlaŋket
tampil
pljsaj
?mŋŋ kit
?iŋ*l
?uŋ
prber
blĩŋ
lda?
Imtem
pucuk* loc
joh, tampeŋ, dej
?abu?* ?os
kŋkaj
t?oc, tane?*, ciʔis
ctíp ~ ctíp
bnken ~ mnken (?), mhel
suʔ, sltuh
cŋwẹŋ
kapɔ?*
cprép
kbet clah
kroʔ
tkih
jahat*, laʔis
bek*, cɔk, hapi?
ton ~ tnton, dor
karɔp, bhit, bulo?*, bamu?*, gnun
tẹn
pisaŋ*, jaj
woj
ckçk, kpsinŋ
carah
jhoŋ
raga?*
bat
Batek (ethnonym)
bathe
beak
bear (Helarctos malayanus) (n.)
beard
beat (v. tr.)
beat (v., intr., of heart)
because
bee-eater (Merops)
beetle, type of
beg
begin
belch (n.)
believe
belly
bend (v.)
betelnut
big
big toe
bind (v.)
binturong (bearcat; Arctitis binturong)
bird
bird, type of
bird, type of
bird, type of
bird, type of
bird, type of
bird, type of
bird, type of
bird, type of
bird, type of
bird, type of
bird's nest
bird of prey
birthmark
bite (v.)
bitter
black panther (Panthera pardus)
blade (of knife)
blanket
blind
blink (v.)

kaji?, tapər
batek
?naj
cnoŋ
kawip
janut*
rajeh
ŋpuŋ
sbap*
ririk
*him
mamun
mula?*
gos
harap*, caja?, ?as
?ec
wɔt
pineŋ*
lim, bəw, ckej
tabo?
sklik
luŋan
kawɔt, kawaw
wtwít
ddit'?
taʔic
trhic
ku'cɔk
sŋok
tŋka?
jɔweh
duduh
jnuh - jnûh (?)
brul
cəwɾəw
som
klęŋ
ksit
kap, coh
kdeŋ
mit taji?
pjəŋ?
cuʔ?
cil
blood
blood vessel
blow (of person)

blow (of smoke)
blow (of wind)
blowpipe
blowpipe dart
blowpipe, shaft of
blue
blue coral snake (*Maticora bivirgata*)
blue-eared kingfisher (*Alcedo meninting*)
blunt
boat
body
boil (n.)
boil (v., intr.)
boil (v., tr.)
bone
book
borrow
bottle (n.)
bottom (body part)
bow (n.)
box (n.)
boy
brahminy kite (*Haliastur indus*)
brain
brake (v.)
branch (n.)
break (v., intr)
break (v., tr.)
breast
breastfeed
breathe
bridge (n.)
British (ethnonym)
broadbill (*Psarisomus*)
brother of parent, older
brother of parent, younger
brother-in-law
brown
brown hornbill (*Ptilolaemus tickelli*)
bucket
buffalo
darah*, bhîm
s?so?
pit, ?uk, ?uf, phos, hamos, þuh, hanþhoñ
niñ, gi, poj
sitjët
blaw
snloc
jo?
biruh*
tqwañ
smpal
majin
prahu?*, sampan*
lej
cmoc, bisol*
didih*
pgoh
j?eñ
buku?*
piñjam*
labu?*, balanñ, butol*
kit
loc
pti?*
kjih, litow
klëñ siawah
smutlñt, lkem
brik*
cabanñ*
ktis*, pcah*
rigih, pok
?ëm
bnaleh, pi?èm
jkjik, hik
titih*
britis*
kawñt batu?*
toj
beh
lamij
blakû?
mañ?
bldi?*
krñw*
build (a house)
build (a hut)
bullet
burn (v.)
brerp (v.)
bus
butcher (v.)
butterfly
buttock
buy
buzz (v.)
call (v.)
call (v., of bird)
camphor
cannibal
car
carp, type of
carp, type of (Puntius)
carry (in one's arms)
carry (in one's hand)
carry (in one's mouth)
carry (on one's back)
carry (on one's neck)
carry (on one's shoulder)
carry (on stick)
carve
casque (of hornbill)
cassette
casting-net
cat
catch (v.)
catfish, type of (Mystus planiceps)
cattle
caught
cause a blister (of fire)
cause pain
cave
centipede, type of
chair
charcoal
chat (v.)
chatter (v.)
cheek
chest
chew
woŋ
swat
pluruh*
?ŋic, hok5k, ʒiʔ, sɔʔ, sjiŋ, brjŋ, com
krTŋp
bas*
kujak
tawék
hawir ~ hawór
bliʔ*
tmtőm, smsm
sáʔot, liʔleʔ, kul, pir
woh ~ wōh
barus*
coleʔ, hmɓit
kritŋh*
irbul*
lampam*
cdum
jij
ktep
tap, rągət, bʔbəʔ, tąŋkəl, gəl
critol
joʔ, tąŋwoŋ, kanər*
gulem
gəs, sot
krkək
tep*
jaʔtəʔ*
kucəŋ*, ?əməw
cəp, bantuʔ*, maŋ
bawuŋ*
lmuʔ*
sik
cəf
ŋruŋ
ghaŋuŋ*, jəŋ
kʔeŋ
pəŋkəŋ*
ŋəŋŋəŋəʔ?əs
buwəl*
siŋraŋ, prihkah, crik5k
kəps?
dkduk
nɔŋ
child
childless person
Chinese (ethnonym)
chop (v.)
circumcision
civet, type of
civet, type of
civet, type of
civet, type of
clap (v.)
claw (n.)
claw (v.)
clean (v.)
clear (one’s nasal cavity)
clear (one’s throat)
clear (v.)
cleave
CLF: animals
CLF: cluster
CLF: large flat objects
CLF: oblong objects
CLF: spherical/cubical objects
click (v.)
climb (down)
climb (v.)
climb up
close one’s eyes
clothes
cloud
cloudy
cobra, type of
cobra, type of
cobra, type of
cobra, type of
cockroach, type of
coconut palm
coin
cold
collarbone
collared kingfisher (Halcyon chloris)
collared scops-owl (Otus lempiji)
collect (v.)
comb (n.)
comb (v.)
come

ken, wοŋ
clah
cine?*
peŋ, blah*
snkɔl
tanlin
gnɡuŋ
mew
wǎw
pak
klko?, cnrɔs, katŋŋ
tamper*
cuci?*
krhak
krhɛm ~ krhɪm
roh
cbis*, tas
ʔikɔr*
prdu?*
bidaŋ*
bataŋ*
buwah*, tuŋkɔl
kktɔk, kktɛk
gis
dبوت, tʔɔc, treŋ
lwec, ciweh
padam*
baŋu?*
sagup
rdʊm ~ rdup
sjul
ʔmpet
snow
preŋ
sor
ŋɔr
sleŋ*
tmket, sju?*
slanjaʔ*
tʃtoŋ
canwɔc
poc, jit, lawa?
sikat*, knɛc
meŋ ~ mɛŋ (?)
klwɪl
come out (of leaves)
common kingfisher (*Alcedo atthis*)
convey
cook (v.)
copulate
corporal
corpse
cough (v.)
count (v.)
cover (v.)
cover one's eyes
crack (v.)
crush (v.)
crawl (v.)
crested serpent-eagle (*Spilornis cheela*)
crested wood-partridge (*Rollulus rouloul*)
crimson-winged woodpecker (*Picus puniceus*)
crocodile (*Tomistoma, Crocodylus*)
crouch (v.)
crow (v.)
crunch (v.)
cubit
cuckoo (*Cuculus*)
cup
curl (v.)
cut (v.)
dance (v.)
dangle (one's arms)
dark
daughter-in-law
day
day after tomorrow
day before yesterday
daylight
decline (v.)
deep
defecate
deliberate (v.)
dense
descend
devour
Diard's trogon (*Harpactes diardii*)
die
crbir
c?ca?
?ampu?*
p?ak, pktok, pgo?, p?eh, kr?om,
?ncen, pel
noj
kpral*
sarö?
?oh
jecp, kira?*, bilen* 
dun, bil, harkal
tudul*
kalc5h
laŋar*
win
klęŋ hip
hwhiw
blato?
buja?*
kr?ip
koko?
ranjip
so?
btok
koleh, takö?
crilwel
gel, ket, kec, hiris*, tbes*, cah, waj,
  rwis, katæm*
sise?, ksonen
rinwïn
gcïh
mïnsaw
ktï?
camo?
ptom
pëñ
gïle?, si?ro?
bïgi?
?ec
gim
plet
cruφ, ï?il, ges
cbot, sksëk, hññan
ckcook
slap, hapa?, naŋu?, kbis, ?is
different

dig (v.)
dirt
dirty
disappear
do
dog
door
dove
dragonfly
dream (v.)
drill (v.)
drink (milk from mother's breast)
drink (v.)
drip (v.)
drop (v.)
drown (v. intr.)
dry
dry (v. tr.)
duck
durian (Durio zibethinus)
dust
each
ear
earth
east
easy
eat
egg
egret (Egretta)
eight
elbow
elder
elephant (Elephas maximus)
eleven
embarrassed
EMPHATIC PARTICLE
endure
enough
enter
every
exchange (v.)
excrement
exist
expensive
explode
extinguish (v. intr.)
extinguish (v. tr.)
eye
eyebrow
eyelid
fall (v.)
fan (v.)
fang
far
fart (v.)
fast
fat (n.)
father
father's father
father-in-law
fear (v.)
feather
feed
feel
fell (v.)
fern (*Filix*)
fern, type of
field
fight (v.)
fill (v., intr.)
finger
fingernail
finish (v.)
finished
fire
fireplace
fish (n.)
fish (v.)
fish, type of
fish, type of
fishing hook
fish spear
fish trap
five
flap (wings)
flat
flat ground
flat-headed cat (*Felis planiceps*)
flee (v.)
kiciφ
hmirin
plit, cirφ
mit
cawīŋ
smpoj mit
plek, rēs, tǐlī, run, kulkīl, tāŋel*, gibal
cēφ, pōh
sjuŋ*
bŋjīi?
ktkit
lktwɔt, bɔcφ
gmu?*, Ί?̃s
?eį
ta?
kn?ac
hgiķ
sōk kawɔt
prgej
grsic, ktlīt
cīh, tuŋ
paku?*
karan
padaŋ*
tumo?*
kseļ, huroj
jariį*
katɔŋ cjas
jil
sɔc
?ɔs
tuŋul
?ikɔ?, ?ikan*
kajil*
tŋis
tŋran
mit knajil
sramparį*
dkdak
lime?*
kŋjeŋ
tŋ - tŋ, jup
hljdiŋ
juluŋ
loj
flex (v.)
flicker (v.)
flip (v.)
float
floor
flow (v.)
flower
flower bud
flower, type of
fluid
fly (n.)
fly (v.)
fly up
flying fox
flying squirrel (*Petaurista*)
fold (v.)
follow (v.)
fontanelle
food
foot
forehead
forest
forget
fork (of river) (v.)
four
friend
frighten
frightened
frog, type of
frog, type of
frog, type of
frog, type of
frog, type of
front
frontal tuber
fruit
fruit, type of
fruit, type of
fruit, type of
fruit, type of
fruit, type of
fruit, type of
fruit, type of
fruit, type of
fruit, type of
fruit, type of

fruit, type of
fruit, type of
fruit, type of
fruit, type of
fruit, type of
fruit tree, type of
fruit tree, type of
full
gall-bladder
gastropod, type of
gastropod, type of
gastropod, type of
gaur (Bos gaurus)
gaze (v.)
gecko, type of
gecko, type of
get
go (of bed)
ghost
giant squirrel (Ratufa affinis) (?)
gibbon (Hylobates)
girl
give (v.)
give birth
gleare (v.)
gliding frog (Rhacophorus)
glow-worm
gnaw
go (across)
go (along crest of a ridge)
go (along mountain-side)
go (along watercourse)
go (alongside)
go (around)
go (back)
go (out)
go (together)
go (v.)
goat
gobble (v.)
gold
good
goose bumps
grandchild

kaban
balũŋ hajam*
lamøŋ
rŋkal
star
rmaj
rawoh
?qrenŋ
cukop*, bhi?, sbem
kmit
hawap
kluktok
kal5?
sladanŋ*
?qwnjw - ?iwŋaw
ci?ca?
pokoh
des
biwiśs
kmuc, hantu?*, sanu?, kīs
kraleh
mawë?
?anek, me?, ?aleh
?ek
bane?, sr̂ho
kbah
caraj
ktlit
kɔc, lslɔs
piris, lintes*
gɔw
tigil
rkruk
sr̂sr
swan, wor
wek
snrek
leh
cip, ĵanja?*
kaminŋ*
khɔp
mes*
bə?t, seh
hr̂peŋ
kapco?
grandmother
grass
grass, type of
grasshopper, type of
grasshopper, type of
great argus (Argusianus argus)
great grandparent
great hornbill (Buceros bicornis)
greater racket-tailed drongo (Dicrurus paradiseus)
green
green tree python (Chondropython viridis)
grenade
grow (of animates)
grow (of plants)
growl (of stomach)
grun (v.)
guarantee (v.)
guide (v.)
gun
hair
halt (v.)
hammer (v.)
hand
hang
hard surface
hardwood tree (Turtur tigrinus)
haunt (v.)
have
have a running nose
have a sad expression on one’s face
have sore eyes
head (n.)
headache
head-hair
headman
headwater
heal (v., tr.)
hear
heart
heat (v.)
heat poisoned blowpipe dart in fire
heavy
heavy (of rain)
heel
crkip can, didil
g
trine?
th
	tuluŋ*
		piŋan*
		saŋar
		ʔaŋ
		ʔoŋ

h
pya?
hide (v. intr.)
dol ~ pidol
hide (v. tr.)
byjį?
hight

ciŋ, l5r
hit (v.)
luka?*, tboh*, ktim, pēn, c5ŋ, khil
ho (v.)
caŋkul*
hone
pankoh*, pgen*
honey
caŋŋ *
honeycomb
toŋin
hop (v.)
jilo?, kin, tμρaŋ
horn
lbeh, kong, lwej
hornbill, type of
lbah lwej
hornbill, type of
sit, tpiŋ
horse
țiŋ
horse
mej si?
hundred
	tano?*
how?
	tkuk
human
khkuh
hungry
	kuda?*
hunt (v.)
bkit, til
hunt (v., intr.)
jam*
husband
	kong
husband
	jista?
husband
	ptis
husband
	ksiŋ
ice
	haŋʔi*, ŋŋkiŋ, haŋŋoŋ
if
	ʔes
insect, type of
	kaluh*, kalow*
insect, type of
	paret
insect, type of
	tis bataŋ*
insect, type of
	krpajŋ
insect, type of
	kapcοŋ
insect, type of
	kmarbɔl
insect, type of
insect, type of
inside
intestines
ipoh poison
iron
island
Jahai (ethnonym)
Japanese (ethnonym)
jar
jaw
jest (v.)
jump (v.)
just
keelback (Xenochrophis, Amphiesma)
keep (fire burning)
keep (v.)
Kensiw (ethnonym)
key
kick (v.)
kidney
kill (v.)
kind (n.)
Kintaq (ethnonym)
knee
knee-cap
kneel
knife
knife handle
know
krait, type of (Bungarus)
lake
lamp
langsat (Lansium domesticum)
language
Lanoh (ethnonym)
larvae, type of
larvae, type of
larvae, type of
lash (v.)
laugh
lay (a floor)
leaf
leaf monkey (Semnopithecus)
leaf, type of
leaf, type of
leafbird (*Chloropsis*)
leak (v.)
lean (one's head)
learn
leave (v.)
leech
left (side)
left-handed person
leg, lower
leg, upper
leopard (*Panthera pardus*)
letter
lick (v.)
lie (of animal) (n.)
lie down
lift (v.)
light (of weight)
light (v.)
lightning
like (prep.)
like (v.)
limp (v.)
lip, lower
lip, upper
little
little finger
live
liver
lizard
lizard, type of
lizard, type of
lizard, type of
lizard, type of
loincloth
long
long-tailed macaque (*Macaca fascicularis*)
long-tailed parakeet (*Psittacula longicauda*)
long-tailed shrike (*Lanius schach*)
look (v.)
lose
lose (hair)
lose (one's footing)
louse

sriken
snipoŋ
lɔŋ
ʔiʔŋoʔ, wəŋ ~ wįŋwɛŋ
ʔaʔəɾ*
ɾəj
lajaj, lawɛj
wɨ?
mɨwi?
gor
bliʔ?
ʔap təwej
soret*
scbc, blek, kal
lanik
tek, kəh, pum, bin, kol
piwek
hatʃp
suk, paŋcaŋ ~ paŋçon
kilat*
pon*
hmhɔm
sŋsɛŋ
tnit
nus
ʔaŋʔ, ʔəŋj
kaʃʔ?
pasaʔ, ɡəs, praʔ, hampes
ros
calak
calat
haluk
taroŋ
daŋŋ
len
btec
rampow
brlîŋ
ći’k5ʔ
cliŋ, ɬəŋʔ*, ciŋliŋ, cilkil, ʔel, trwer,
sitor ~ hitor, blaj
kalah*
prasut, tstus, gugur*
cantil
ciʔ
low  
lump  
lung  
magpie robin \((Copsychus saularis)\)  
make (v.)  
make sound  
Malay  
malkoha \((Phaenicophaeus)\)  
man  
manchild  
mango  
mangosteen \((Garcinia mangostana)\)  
many  
map  
march (v.)  
masked civet \((Paguma larvata)\)  
mat  
maybe  
meander (v.)  
meat  
meet  
mend (v.)  
Menriq (ethnonym)  
milk  
millipede, type of  
millipede, type of  
millipede, type of  
millipede, type of  
mince (v.)  
miss (v.)  
molar tooth  
money  
monitor lizard, type of \((Varanus)\)  
monitor lizard, type of \((Varanus)\)  
monitor lizard, type of \((Varanus)\)  
moon  
morning  
mortar (household utensil)  
mortar (weapon)  
mosquito  
mother  
mountain  
mountain pass  
mountain-top  
mouse deer \((Tragulus)\)  

\textit{man}  
ckim  
sop  
morej* batu?*  
de?  
hū?, knjiŋ, məw  
mantoŋ  
bток  
tmkal  
goh  
blawen  
manjis*  
bnolet, brasa?, kom  
meŋ*  
kawat*  
?iIšŋ  
nis  
munken*  
klihdšh  
sec, lawo?*  
jumpa?*, ?uswas, səh ~ sihsəh, lmah  
kunci?*  
mnri?  
susuh*, məm  
klcbac  
cmalpɔk  
taŋp?  
taluŋ  
choəh, ciŋcaŋ*  
lec  
bŋka?  
dwi?*  
bagen  
greŋ  
kakaw  
bolan*, chij  
pagi?*  
lsuŋ  
mutah*  
hagas*  
bi?, ma?  
ylmoł  
lget  
pə?, tul  
plancet, planok*
mouth harp
move (ears)
move (in circles)
move (snout)
move (v.)
mud
muntjac (barking) deer (*Muntiacus muntjac*)
mushroom, type of
myna (*Acridotheres*)
nail
name (n.)
navel
near
neck, nape of
needle
NEGATIVE MARKER
net (n.)
new
night
nine
nipple
noon
nose
nothing
number (n.)
ogle (v.)
oil palm
old
one
only
open (one’s eyes)
open (v.)
opposite side
order (v.)
oriental pied hornbill (*Anthracoceros albirostris*)
oriental white-eye (*Zosterops palpebrosus*)
oriole (*Oriolus*)
other
otter (*Lutra sumatrana*)
outgrowth (on tree)
outside
overgrown
overnight (in forest) (v.)
owe (?)
owl, type of

*range*  
*sinu*  
*pusin*, *linwij*
*noc*  
*kop*, *jok*, *jatoh*
*lapék*, *kobaŋ*  
*bhol*  
*t?aj? pok*  
*ciunj*  
*klko*, *cnros*, *katŋ*  
*knmo*  
*dut*  
*pdah*  
*tŋko*  
*jarum*  
*bŋit*, *bra?*, *bokan*  
*pukat*  
*paj*  
*tw5j*, *hrkit*  
*smilan*  
*dapor*, *ʔem*  
*cnhŋ*  
*moh*  
*blap*  
*ʔi?*  
*jet*  
*klapah sawit*  
*maneh*  
*nej*  
*saja?*, *jaga?*  
*ciwę*  
*ʔoj*  
*titep*  
*ʔor*  
*thteh*  
*barow*  
*lla*  
*pęw*  
*mane*  
*goŋ*  
*hip*  
*bhir*  
*tmdem*  
*dos*  
*kuruhuj*
Pacific swallow (*Hirundo tahitica*)
paddle (v.)
pain (n.)
paint (v.)

dead (of hand)
palm leaf, type of
palm, type of
palm, type of (*Arenga obtusifolia*)
pangolin (*Manis*)
pant (v.)
papaya (*Carica papaya*)
paradise tree snake (*Chrysopelea paradisi*)
pardon
pass (v.)
past
path
peacock-pheasant (*Polyplectron*)

dead (of bird) (v.)
peep (v.)
pelvis
penis
people
pestle
petai fruit
phalange
pharynx

Philippine glossy starling (*Aplonis panayensis*)
photograph
pick (one’s teeth)
pick (v.)
pig
pigeon
pig-tailed macaque (*Macaca nemestrina*)

pinch (v.)
pineapple (*Ananassa*)
pitta (*Pitta*)


place (n.)
place (one’s hand)
plaintive cuckoo (*Cacomantis merulinus*)

plait (v.)
plank
plant (v.)
plant, type of
plantain squirrel (*Callosciurus notatus*)

play (games)
play (instrument)
pluck (v.)
pocket
point (v.)
pool (n.)
porcupine (*Hystrix, Atherura*)
porcupine, type of
porcupine, type of
portion
post (military) (n.)
potato
poultry
pounce upon
pound (v.)
pour
pregnant
prick (v.)
PROHIBITIVE
PRONOUN 1D.EXCL
PRONOUN 1D.INCL
PRONOUN 1P.EXCL
PRONOUN 1P.INCL
PRONOUN 1S
PRONOUN 2/3P
PRONOUN 2D
PRONOUN 2S.DIS
PRONOUN 2S.FAM
PRONOUN 2S.INT
PRONOUN 3D
PRONOUN 3S
Provost’s squirrel (*Callosciurus prevosti*)
pull (blowpipe apart)
pull (v.)
pump (n., v.)
push
put
put (inside)
pull (to sleep)
pull (together)
python, type of (*Python*)
quill (n.)
quiver
raft (v.)
ragged
rain (v.)
rainbow
raise (v.)
rambutan (Nephelium lappaceum)
rapid (n.)
rat
ration
rattan
rattan, type of
rattan, type of
rattan, type of
rattan, type of
rattan, type of
rattan, type of
rattle (v.)
raw
reach (v.)
ready
rear (v.)
red
reed snake, type of (Macrocalamus)
refuse (v.)
RELATIVE MARKER
remain (v.)
remains
remember
repudiate
request (v.)
rest (v.)
reticulated python (Python reticulatus)
retreat (v.)
return (v.)
reveal
rhinoceros (Dicerorhinus sumatrensis)
rhinoceros hornbill (Buceros rhinoceros)
rice (cooked)
rice (husked)
rice (in the husk)
right (side)
right-handed person
ring (n.)
ring (v.)
ripen
river confluence
river-mouth
roar (v.)
cdaw
hjo, sej
tanuj
jrem*
wawat
risin*
?awej
manow*
knbi?
pipih
wam
tahanen
baten*
kačit
jep
daw
cadan*
pihira?
rhik, btoj, rgoj, th5j
dad5l
coh, ji?, prmen, g5n
jaŋ
brase?, snlön
sisah*
kakep, ?iŋet*
bcreej*
smeŋ
rihat*, bradu?, cundo?, kil, gol
talon
krlun
mudik*
hüt
hagap, bada?*
hŋuŋ*
nasi?*
bras*
padej*
tem
mmttem
cnos ~ cnus
tangd5h
tuha?*
was tōm
kit tōm
piʔjoʔ, sikow
<table>
<thead>
<tr>
<th>English</th>
<th>Jahai</th>
</tr>
</thead>
<tbody>
<tr>
<td>roast (v.)</td>
<td>panjan*</td>
</tr>
<tr>
<td>roll (of thunder) (v. intr.)</td>
<td>gir ~ gar</td>
</tr>
<tr>
<td>roll (v. intr.)</td>
<td>lu?le?, grunlin, cri?ro?</td>
</tr>
<tr>
<td>roll (v. tr.)</td>
<td>hamat*, mœl</td>
</tr>
<tr>
<td>roof</td>
<td>sarañ*</td>
</tr>
<tr>
<td>root</td>
<td>?es</td>
</tr>
<tr>
<td>root-crop</td>
<td>hobí?*</td>
</tr>
<tr>
<td>rope</td>
<td>?ns?</td>
</tr>
<tr>
<td>rotten (of wood)</td>
<td>pok</td>
</tr>
<tr>
<td>round object</td>
<td>sit ~ ?otsit, kœm, kles</td>
</tr>
<tr>
<td>rub (v.)</td>
<td>gtañ*</td>
</tr>
<tr>
<td>rubber</td>
<td>bœj</td>
</tr>
<tr>
<td>run (v.)</td>
<td>don</td>
</tr>
<tr>
<td>run short</td>
<td>sabim</td>
</tr>
<tr>
<td>Sabüm (ethnonym)</td>
<td>gunih*</td>
</tr>
<tr>
<td>sack (n.)</td>
<td>slamat*</td>
</tr>
<tr>
<td>safety</td>
<td>takœj</td>
</tr>
<tr>
<td>sail-fin lizard (<em>Hydrosaurus</em>)</td>
<td>lheñ</td>
</tr>
<tr>
<td>saliva</td>
<td>?mpoc, garœm*</td>
</tr>
<tr>
<td>salt</td>
<td>masin*</td>
</tr>
<tr>
<td>salty</td>
<td>kasa?</td>
</tr>
<tr>
<td>sambar deer (<em>Cervus unicolor</em>)</td>
<td>srupœh*</td>
</tr>
<tr>
<td>same</td>
<td>pantej*</td>
</tr>
<tr>
<td>sand</td>
<td>mit</td>
</tr>
<tr>
<td>sandfly</td>
<td>puwas*</td>
</tr>
<tr>
<td>satisfied</td>
<td>kdiñ</td>
</tr>
<tr>
<td>say</td>
<td>parut*, cœc ~ cœc</td>
</tr>
<tr>
<td>scar</td>
<td>halœw*</td>
</tr>
<tr>
<td>scare (v.)</td>
<td>crœjir ~ crœjor</td>
</tr>
<tr>
<td>scent (v.)</td>
<td>?acah*</td>
</tr>
<tr>
<td>scold (v.)</td>
<td>pkip</td>
</tr>
<tr>
<td>scorch (v.)</td>
<td>kñœñ*</td>
</tr>
<tr>
<td>score (v.)</td>
<td>jœweñ</td>
</tr>
<tr>
<td>scorpion</td>
<td>suduh*</td>
</tr>
<tr>
<td>scrape (v.)</td>
<td>hakat, gœc, kac, cscis, kœs, gluñœñ</td>
</tr>
<tr>
<td>scratch</td>
<td>lawot*</td>
</tr>
<tr>
<td>sea</td>
<td>musem*</td>
</tr>
<tr>
<td>season</td>
<td>cntel</td>
</tr>
<tr>
<td>seat bone</td>
<td>jœt, nampa?*, tñleñ, ?œl</td>
</tr>
<tr>
<td>see</td>
<td>biçi?*, mit ?hœj</td>
</tr>
<tr>
<td>seed</td>
<td>juwale*</td>
</tr>
<tr>
<td>sell</td>
<td>smej</td>
</tr>
<tr>
<td>Semai (ethnonym)</td>
<td>hantor*</td>
</tr>
<tr>
<td>send</td>
<td>piha?</td>
</tr>
</tbody>
</table>
settle  ttap*
seven   tujoh*
shadow lanjoj
shake (v., tr.) jiϕ, ?unjaϕ ~ ?unjin, gntar*
sharp     cmɛ?
shaving (from carving) cpah
shed (v.) ruroh
shiver (v.) gqrir
shoe     kasɔt*
shoo (v.) cih
shoot (v.) ste?*, haluh, tunim, peɲ, praɲ*, bdil*,
          bdel*
shop (n.) kdej*
short     cnhɔt
shoulder klapah
shovel (v.) sej
siamang (Hylobates syndactylus) ?ameɲ
sibling, older pɛ?
(sibling, younger) wakil, ber
sibling, younger ?naɲ, sir
side      paw
side (body part) sɲr
silence   sɲr
silent    smslim
silver-eared mesia (Leiothrix argentauris) pɲloɲ, cʔaj
sing      hrkbek ~hrkbək, karem*, higir
sink (v.) toj bi?
sister of parent, older mɔ?
sister of parent, younger lamij
sister of parent, younger kikkɔk, ɲok, rin, kilwal
sister-in-law pɛm, bat
sit        nem*
sit (of bird) bne?
six        cɪm
size       kti?
sizzle    klipɛh
skin (n.)  gas ~ gɛs
skin (v.)  manrɔn
skin disease, type of labɔn
skink (Emoia) ktsɔ?
skull      tek
sky        gruc
sleep (v.)  sir
slender-toed gecko (Cyrtodactylus) bulet, krjuh, tunɛn, gimɛl
slide
slither
slow      hakij
slow lori (Nycticebus coucang)  tampîl
slurp (v.)  siruc
smack (v.)  pkpûk
smear (v.)  pleś
smell (n.)  ges
smell (v.)  ?ûn, ?ûnôn
smoke (n.)  ?j?e? ?ôs
smoke (v.)  jhîit, kmim, kbam
smooth  bclac
snail  siput*
snake  tajû?
snake, type of  papan
snake, type of  padej*
snake, type of  sle?
snake, type of  tdoû
snake, type of  rloû
snake, type of  ranoû
snake, type of  kldûl
snake, type of  akôr
snake, type of  brhur
snake, type of  slantor
snakehead (Channa micropeltes)  tomen*
snap (v.)  hêk
sneak (v.)  dop
sneeze (v.)  brsin*
snot  ctoh
snout (of boar)  tanus
snout (of tapir)  somu?
snuffle (v.)  hîs
soap (n.)  sabon*
sob (v.)  rpsêt, sut
son-in-law  mpsaw
soon  ?nuj
soul  ?rens, rwaj
sound  klin
sound (animal fleeing from one tree to another)  prow
sound (blowpipe dart hitting a vine)  cêh
sound (blowpipe dart hitting canopy)  rôf
sound (blowpipe dart hitting muscle of prey)  cik
sound (blowpipe dart hitting stomach of prey)  kôf
sound (blowpipe)  plês
sound (blowpipe)  plês
sound (boiling)  ?ok
Appendix II

sound (carving)  krāŋ
sound (dashing)  luf – ruľ
sound (falling)  klak
sound (flapping) riľ
sound (flying or leaping) jař
sound (hornbill) kōk
sound (leaf-monkey) hōkįk
sound (muntjac deer) rop
sound (raindrop) tik
sound (running water or waterfall) chok
sound (shotgun) pōw
sound (small object falling to the ground) rop, klik
sound (walking) kōľ
sour masēm*
sow (n.) bi? masaj
speak pįseŋ
spear (n.) mataŋ
spear (v.) cįćįŋ
spectacle tawįŋ
dspider, type of tįtűŋ
dspider, type of knambil
dspider, type of sranįkej
dspit (v.) kbec, tuf, ktoľ, pąncęr*
dsplit (v.) stap, pek, wįs, was, salow
sspray (v.) smur*
dspread (v.) hap ~ hiphap, dor
dspring (v.) tikah
dsprinkle (v.) psel, pruj
dquat (v.) jrițew
dsqueeze (v.) cįćįt, lam
dsquirt kdek
dstab (v.) cek, rįk
dstain (v.) hwęt
dstand (v.) hŋjaŋ
dstar bintęŋ*, prloj
dstare (v.) pdep, tltat
dsteal takų? ~ ?aku?, malįŋ*
dsteam (n.) ?įʔeŋ tôm
dsteam (v., intr.) has
dstench haʔęt
dstep over laŋkah*
dstick (n.) ?at
dsting (v.) kft, hāt, sic, c5?, lantom
dstomach ?ec
stone
stop (v.)
straight
stranger
stream-bed
stretch (v. intr.)
stretch (v. tr.)
striped tit-babbler (Macronous gularis)
stroke (v.)
strong
stumble
such
suck (v.)
sugar
summon
sun
sunbird (Nectarinia; Anthreptes)
surprised
swallow (v.)
sweat (n.)
sweep (v.)
sweet
swell (v.)
swidden
swim (v.)
tail
tail feathers
take (a hit)
take (back)
take (v.)
talk (v.)
tap (poison) (v.)
tapir (Tapirus indicus)
taste (n.)
taste (v.)
tasteless
tattoo
tea
teach
tear (n.)
tear (v.)
tell
tell a lie
Temiar (ethnonym)
temple (body part)

batu?*, balal
?op, jin
tros*
gop
carək*
?t?et
cičiŋ
hor
?t?5t
kwat*
tsdes
kej
jhit, pot, sks5k
gulah*
pihdeh
kit kto?, mit kto?
setset
hiran
klāt, lik
bŋkit
kwēs, pis, sapuh*, kwer*
bhet, b?ıt
kmanŋ*
slaj
kwac, sjør
haft
ceño
dan
pimic
kot, fiŋ, ?aŋket*, jow
ckwik, baca?*, cara?
tureh, pŋ
baret, banson
lasanŋ
hōt, cuba?*, kmjim
blhit
cil
teh*
pipjep
tom mit
cok, csses, tging
cōl
boc, bër
pleh, tmer
cunŋ
Temuan (ethnonym)

Temuan (ethnonym)

ten

termite

termite mound

terrapin, type of

terrapin, type of

testicles

Thailand

that (away)

that (beyond me)

that (beyond you)

that (down)

that (up)

that (you don’t know)

that (you know)

there (away)

there (beyond me)

there (beyond you)

there (down)

there (up)

there (you don’t know)

there (you know)

thick

thick-billed pigeon (Treron curvirostra)

thin

thing

think

thirsty

this

thorn

thorn, type of

thousand

thread (n.)

three

throat

throw (v.)

thrush (Zoothera)

thumb

thunder

thunder-spirit

tickle (v.)
tie (v.)
tiger (Panthera tigris)
tiger shrike (Lanius tigrinus)
time
tired
toad, type of
tobacco
today
toe nail
together	
tomorrow
tongue
tooth
tortoise, type of
tortoise, type of
tortoise, type of
tortoise, type of
tortoise, type of
touch (v.)
tough (of meat)
track (n.)
training
trap (n.)
tread (v.)
tree
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
tree, type of
ghel, hrtlet
brnjæk
tmakow*
ktɔ? tɔh
katɔŋ can
skaliʔ*, gus
haden
lnæk
hɛn
ʔawɛʔ
koh
barʔɛŋ
sil
sorej
pitil
kraŋ*
tel
trineŋ*
bakoʔ
lajɛk
jhuʔ
ʔpcɛk
taduk
lumpok
gaseʔ
kmpɛs
tawes
prɡis
tlas
ŋnoh
prh
bnim
knum
gen
panʔon
bɔŋ
trpɔŋ
ʔaneŋ
bramaŋ
knaŋ
katŋ
kʔuŋ
lajŋ
balŋ
raŋsil
tree, type of

broil

?amper

jader

saqir

cicar

bsor

bakow

jirwej

kjaj

cah

sraja?

tom

joh, joh wankē?
tac3r

hit

cabanj* tom

btol*

kr?onj

bralej ~ blalej*

gil

hobi?*

kej

klep

hak5t

tahe?

kba?

hrja?

kabu?

daran

sih

tolej

ci?in

j5n

haw

manraj

kw5j

wikwek, brile?, gaj

clinhîn, clinpanj

labi?*

kuh5sh

pj5n

kpil

gadinj*

duwa? puloh*

piplēp

turn (v. tr. and intr.)
turn head (of bird)
turtle, type of
turtle, type of
turtle, type of
turtle, type of
tusk
twenty
twinkle (v.)
twins
two
umbrella
underside
understand
undress
untie
until
upper side
urinary bladder
urinate
urine
valley
vegetable
village
vine
viper, type of *(Trimeresurus)*
vomit (n.)
vomit (v.)
vulva
wade (across)
wag (v.)
waist (body part)
wait (v.)
wake (up) (v. intr.)
wake (v. tr.)
walk (v.)
wallet
wane (v.)
want (v.)
wash (v.)
wash away
wasp, type of
wasp, type of
wasp, type of
water
water strider *(Gerridae)*
waterfall
wave (v.)
wax (v.)
waylay
weep (v.)
west
wet
what?

kmär*
duwa?*
pajuŋ*
kJom
paham*
jok
bāk
baru?*
kripŋ
knom
knöm
tpis, lgim
l?ä?
kampŋ*
?awɛj
bla?ir
sdij
bhok, ki?, rah
dot
cik
kipas*
gel
tado?
muyker
pihpīh
rop, gagah, pher
dibeh
blah*
?om
soc
tpis
ruh
?īŋ
rjaw
tom, ?ōŋ
samir
lata?*
lamba?*
wēl
gat
jim
barat*, seŋ kto?
pce?*
mej
when?
where?
whip snake (*Ahaetulla*)
whirl (v.)
whiskers
whistle (v.)
white
white-rumped shama (*Copsychus malabaricus*)
whiz (v.)
who?
whose?
wide
wife
wild boar (*Sus scrofa*)
wild dog (*Cuon alpinus*)
win (v.)
wind
wind (v.)
window
wing
wipe (v.)
with
woman
woman, old
wood (material)
wood (rotten)
work (v.)
worm
wrap (v.)
wreathed hornbill (*Rhyticeros undulatus*)
wrinkle (n.)
wrist
wrong
yawn (v.)
year
yell (v.)
yellow
yellow-vented bulbul (*Pycnonotus goiavier*)
yes
yet

mapu?
ibah
b?ow
hilan, gulej
jη
hčac, hīf ~ hēf, həh
puteh
kwal
jlof
maken
maken
lwes, ḫnej
kneh
napak, clapak, hmalah, hawēn, gaw
cloη
mnej
brwa?, bgiw
mlmèl
tŋkap
sajap
jit
same?
babo?
kbet, faji? ~ faja?
tagin, gnoη
rba?
krja?, kacow?, ga'wej
taciη
ben
sangoh
krtvit, krinten
krīl
lec
hjhaŋ
tawon
krker
kunη
bruboh
ha'i?, ha'rīh
lagi?
References


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, so-called 'Sakai' or 'Semang' at Trang Province, 310–335. Bangkok: Silpakorn University, Faculty of Archaeology.


References


References


References


van der Sluys, Comelia 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.


