USE OF THESES

This copy is supplied for purposes of private study and research only. Passages from the thesis may not be copied or closely paraphrased without the written consent of the author.
Topics in the Phonology and Morphosyntax of Balinese

based on the dialect of Singaraja, North Bali

Adrian Clynes

A thesis submitted for the degree of Doctor of Philosophy of the Australian National University

December 1995
Except where otherwise indicated, this thesis is my own work.

[Signature]
In memory of my parents,

Jean and Peter Clynes
Acknowledgements

This thesis has been a long time in the making, and owes much to many people. I therefore approach the writing of acknowledgements both with a feeling of pleasure, and one of concern, knowing that there will be names I should remember to thank and will forget.

I thank first, and most of all, my wife, Ketut Maseni Widiastuti Clynes, of Banjar Peguyangan, Singaraja, Buleleng, Bali, and, latterly, also of Australia and Brunei. Without your support and faith in me Tut, I could never have started this, let alone brought it to a close. You supported me not only in my studies towards this PhD, but also (I blush to say) the MA which I had to get before I could go on to this. When times were hard, in a new, strange country, you went out and got jobs, to help us through. And you had, and have, the energy to mother three beautiful children along the way. You have made a crucially important contribution to the content of this thesis, through your countless hours teaching me about your language. Your presence is felt on every page. Very truly, this thesis would not be here without you. Matur suksma pisan, Tut.

I owe a debt of gratitude to 'Beli' Ketut Mandi Pinatih, of Banjar Peguyangan. Beli Mandi is a retired army officer, and former leading administrator in the local government (DPR). When he learnt why I had come back to Peguyangan with my family, Beli Mandi quietly volunteered his services. For nearly eleven months, almost every day, I met with Beli Mandi, for at least an hour, usually two or more. He taught me his language patiently, and with humour and insight, for no other reason than a love of learning, and a desire to help me. His teaching informs every page of this work.

Another who generously gave me his time was Jero Dalang Penarukan, shadow puppet master, of Penarukan village, 3 km to the east of Peguyangan. Jero Dalang Penarukan permitted me to record several of his performances, and spent several sessions answering questions I put to him about the texts, and about Balinese in general.

My mother-in-law, Ibu Nyoman Resi, and my father-in-law, the late Bapak Ketut Mantra, gave constant moral and linguistic support, as did many members of the extended family, and the wider Arya Wang Bang Pinatih clan. Sadly, Bapak is not here today to see this work finally produced.

It is a true pleasure also to thank Tut's brothers and sisters and their spouses: Gde Dermawan (Beli Awan) & Mbok Komang, Mbok Koming & Ida Bagus Putra; Ketut Sulendra (Robet) & Kadek; Dr Luh Putu Supadmi (Mbok Mamik) & Dr Made Siria, Made Widiartini (Dek Nik) & Ketut Budiada, Nyoman Suciarini/Jero Puspa (Ucik) & GUSTI Made Darwin. In many different ways you all made life happier and easier for us in Bali, often putting us up, transporting us: Mamik & Siria, who were serving as government doctors in Irian, lent us their car for the duration of our stay in Bali. For many kindnesses and good times spent with you all, thank you.

Other patient teachers included the late Guru Natih, and the late Guru Ketut Geledag, both also now sadly deceased. Ibu ('Mbok') Ketut Supani spent several hours recording
folk tales for me, and sent me further written tales when I was back in Australia. Maman Madé Bija, Beli Putu Legawa and Mbok Luh Semita have always given generous help, in many ways, as have Mbok Putu Sukmayani and Putu Wijaya. Bapak Gde Dharma, playwrite, cultural animator, of Sukasada (5 km to the north of Peguyangan), managed to fit me into a busy schedule, to answer questions about his work.

Many other people helped me to learn about Bali, in the years before I began this study. I will mention only those with whom I was privileged to live for long periods: Dr Ida Bagus Oka & family, of Banjars Bindu & Ceramcam, Kesiman, Bapak Ketut Ebu, of Banjar Abian Kapas Tengah, and Bapak Nyoman Purna & Ibu Culik of Banjar Abian Kapas Tengah.

I am indebted for my linguistic training first and foremost to the staff of the Department of Linguistics, Faculty of Arts, Australian National University. I was fortunate to study for my MA in that department, and to be taught by several demanding, inspiring teachers. My thanks to Avery Andrews, Bob Dixon, Bill Foley, Harold Koch, Karl Rensch, Phil Rose, Tim Shopen, and Anna Wierzbicka. It was a special privilege to have studied with you all. The influence of your teaching will be apparent, I hope, throughout this thesis.

The Department of Linguistics, RSPAS, gave me shelter and support for the writing of this thesis. Thanks especially to my fellow students, Chuck Grimes, Rob Early Bob Bugenhagen, and Jeff Marck for moral and intellectual support. Chuck and Robert helped 1000 times with the computer and, with their families, shared good times. Helpful comments on various chapters in Part I were given by Sander Adelaar, Avery Andrews (of the Faculties' department), and Andy Pawley. I thank Avery especially, for help beyond the call, and for encouragement. Sander's input greatly improved the presentation and content of several chapters. Andy's comments were always valuable. Valuable comments, which improved the content of chapter 3, were given by Carl James at Universiti Brunei Darussalam, and by Juliette Blevins and other participants at a seminar given at UWA. Yvonne Pittelkow, of the Computer Centre, RSPAS, gave important help with the statistical procedures. Macolm Ross read a late draft of the whole thesis; his important suggestions have led to many improvements in both form and content.

The fieldwork on which this thesis is based was carried out under permission of the Lembaga Ilmu Pengetahuan Indonesia (Surat Izin Penelitian 2795/sk/1990). I am grateful to staff at the Pusat Penelitian of Udayana University, especially Drs Ida Wayan Oka Granoka, who helped me in my initial attempts to find a fieldword site in Karangasem. I benefitted from an Australian Government Postgraduate Research Award during the time I was studying full time at the ANU.

Lastly, thanks to all the members of my family who, together with Tut, lived with this thesis for so long. My ever-patient father Peter Clynes gave both moral and financial support along the way, but sadly did not live to see the finished work. Special thanks to Rob and Kate, for all your help to us as a family, equally to Rosie and Bill, and to Anna and Don. Finally, thank you to Putu Jeannie Arimbi and to Madé James Arya, for reminding me to go and finish my work, and to Komang Rosie Amalina, for keeping me from doing too much of it.
Abstract

This study describes features of the phonology, verbal morphology, and morphosyntax of the Balinese of Singaraja, in Buléléng district, north Bali.

In chapter two I describe the phoneme repertoire, and their realisations. I give evidence for meaningful postlexical rules, of the kind described by Woodbury (1987). In chapter 3 I show that there is a regular interaction between semantics and phonology at a pre-lexical level, contra the double articulation hypothesis, and in support of the findings of Fudge (1970). I propose an (ultimately pre-linguistic) explanation for this interaction. The findings support and enrich optimality theory, which seeks to formalise the gradient nature of wellformedness (Prince & Smolensky 1993). The semantic evidence helps explain why non-optimality is tolerated in the phonology.

While the phoneme is an important unit of organisation at an intermediate level, it is neither a phonological primitive, nor the largest recurrent phonological unit. I give evidence in chapter 4, and in chapter 7, for complex phonological formants more complex than the phoneme. These can be 'segmental', or involve processes such as reduplication (pace the treatment of reduplication in Prince 1987).

In chapter 5 various levels of prosodic organisation are described. The syllable, the foot and the phonological word are each relevant to word-formation processes (though no evidence was found for the mora). I give evidence in this chapter, and in chapters 6 and 8, that all morphemes must be well-formed at foot level during the lexical phonology. These data support, and can be accounted for in terms of, the Prosodic Morphology theories developed by McCarthy and Prince (eg 1990).

In chapter 6 constraints on the structure of morphs and morphemes are described. Balinese shows similar coocurrence restrictions to Arabic (Greenberg 1960, McCarthy 1986) and Javanese (Uhlenbeck 1949, Yip 1989), dispreffering the occurrence of more than one consonant of a given place of articulation, underlyingly. Morphemes showing exceptions to this 'preferred rule' predictably belong to the expressive semantic classes defined in chapter 3.

Non-prosodic phonological processes involved in morpheme- and word-formation (including concatenation and morphophonemic alternations) are described in chapter 7. Chapter 8 catalogues the various formal types of reduplication. Two previously undescribed surface reduplication types in Balinese are identified: Foot-reduplication and internal reduplication. As with the restrictions on consonant coocurrence morpheme-internally (chapter 6), reduplication processes involving vowel overwrite give evidence for a non-linear organisation of phonological structure. A variety of evidence is presented to show that phonological or 'inherent' reduplication is a
synchronic process in Balinese, (despite previous claims that there is no evidence for this). Finally, I describe theoretically problematical reduplication processes where non-distinctive material is transferred (cf Steriade 1988), and offer an explanation for them in terms of the Prosodic Morphology theory of McCarthy and Prince.

The second part of the thesis describes aspects of the morphosyntax, in particular word formation processes applying to derive verbs in Balinese, and related aspects of clause syntax. After an overview of basic features of morphosyntax and terminology in chapters 9 and 10, I describe in detail the functioning of verbal affixes, in chapters 11 to 13. In chapter 14 I give evidence that Balinese shows 'undergoer primacy' in a variety of ways, at the same time arguing against an ergative analysis of the syntax (pace Blake 1993).

My major conclusions relate to the function of verbal morphology, and the way meaning is assigned to lexemes and clauses. The meaning of a given root-affix combination is broadly predictable, once the semantics of the root morpheme is known. To that extent, morphology signals primarily semantic information in Balinese. The Actor/Undergoer distinction which informs the morphological and syntactic patterning of other Austronesian languages (Foley & Van Valin 1984, Durie 1984, 1987) is an important one in Balinese. Prefixes (or their absence) signal whether the subject of both transitive and intransitive verbs is an Actor or an Undergoer (chapter 11). The semantic type of the Undergoer is further specified by verbal suffixing (chapter 12).

Although formally Balinese is close to the ideal of an agglutinating language, the meaning of a lexeme is not simply the sum of the meanings of its component morphemes (cf Koch 1990). Meaning assignment is best treated using a model such as those in the word & paradigm tradition (Beard 1988, Aronoff 1994). Affixes are not in themselves meaningful: they serve primarily to signal the presence of meaning elements inherent in the lexemes of which they form part. As such some lexical bases always occur with affixes, even though they have not undergone derivational processes. (I argue against the 'precategorial' analysis of such lexical bases in chapter 10). Where affixes participate in derivational processes, the meaning of the resulting lexeme is assigned by the derivational process as a whole. Indeed, I show that in some cases an adequate account of the meaning conveyed by a given lexeme can only be achieved if discourse and other contextual factors are also taken into account (Chapter 13).

The evidence from Balinese supports the view of (for example) Wierzbicka and Givon that transitivity is fundamentally a semantic notion. The primary function of affixing on verbs in Balinese is to signal the presence of entities in the semantic structure of the lexeme; these may or may not be represented in the syntax, according to contextual and other factors. Semantically and formally transitive verbs thus often behave as though they are syntactically intransitive, and affixes often described as 'transitivising' are so as long as one accepts this semantic interpretation of the term. This approach enables a wider explanatory coverage than previous descriptions of Balinese verb morphology, or of similar facts in related languages (Chung 1976, Kana 1986, Alsagoff 1992).
# Table of Contents

Acknowledgements\textsuperscript{v}  
Abstract\textsuperscript{vii}  
Table of Contents\textsuperscript{ix}  
List of Tables\textsuperscript{xiv}  
Abbreviations\textsuperscript{xv}  
Note on orthography\textsuperscript{xvi}  

1 Introduction\textsuperscript{1}  
  1.0 Aims\textsuperscript{1}  
  1.2 Balinese in its wider context\textsuperscript{1}  
  1.3 The dialect described\textsuperscript{2}  
    1.3.1 Dialectal variation and speech styles\textsuperscript{5}  
  1.4 Previous studies of Balinese\textsuperscript{6}  
  1.5 Organisation of the thesis\textsuperscript{7}  

I - Phonology\textsuperscript{11}  

2 Phonemes and other segmental phonological formants\textsuperscript{11}  
  2.0 Introduction\textsuperscript{11}  
  2.1 Consonant phonemes and allophony\textsuperscript{12}  
    2.1.1 Consonant allophony\textsuperscript{13}  
    2.1.2 Meaningful phonological rules 1: consonants\textsuperscript{16}  
    2.1.3 The status of glides\textsuperscript{17}  
      2.1.3.1 Epenthesis of non-phonemic glides\textsuperscript{19}  
    2.1.4 Prenasalised consonants?\textsuperscript{20}  
    2.1.5 Phonetic reduction of nasal-stop clusters\textsuperscript{22}  
  2.2 Vowel phonemes\textsuperscript{23}  
    2.2.1 Vowel allophony\textsuperscript{25}  
    2.2.2 A rule targeting /u/, /ʊ/, /e/ and /o/\textsuperscript{27}  
    2.2.3 Allophony of the mid, non-central vowels /e, o/\textsuperscript{28}  
    2.2.4 Allophony of the central vowels, /ə/ and /ɑ/\textsuperscript{30}  
    2.2.5 Deletions of underlying schwa\textsuperscript{31}  
    2.2.6 Non-phonemic schwa: optional epenthesis\textsuperscript{33}  
    2.2.7 Spreading nasalisation of vowels\textsuperscript{34}  
    2.2.8 Meaningful phonological processes 2: vowels\textsuperscript{34}  

3 Phonology and Semantics\textsuperscript{37}  
  3.0 Phonology and semantics: introduction\textsuperscript{37}  
  3.1 Phonology and semantics: examples\textsuperscript{39}  
  3.2 Analysis\textsuperscript{41}  
    3.2.1 Fudge's approach, and the present one\textsuperscript{41}  
    3.2.2 The 'expressive' semantic categories defined\textsuperscript{45}  
    3.2.3 Statistical tests: central versus peripheral morphemes\textsuperscript{48}  
    3.2.4 Synchronic evidence for the productive interaction of phonology and semantics - I: Nicknames\textsuperscript{51}  
    3.2.5 Synchronic evidence for the productive interaction of phonology and semantics - II: Dorsal-dorsal interactions\textsuperscript{52}  
    3.2.6 Diachronic evidence for the interaction of phonology and semantics: lenition of /r/\textsuperscript{54}  

ix
3.3 Identifying Preferred Rules  
3.4 Explaining the correlations  
  3.4.1 Explaining the correlation 1: Phonological sub-classes  
  3.4.2 Explaining the correlation 2: Marginal phonological structures generate meaning  
  3.4.3 On revising 'revised double articulation'  
3.5 Conclusions  

4 Complex, non-prosodic phonological formants  
4.0 Introduction  
4.1 Morphs: Segmental complex formants  
4.2 Inherent reduplication as a purely phonological process  

5 Prosodic structure  
5.0 The prosodic hierarchy  
5.1 Syllable structure  
  5.1.1 Contrastive syllable weight?  
5.2 Onset-internal constraints  
  5.2.1 The onset slot  
  5.2.2 The L slot  
  5.2.3 Ci-L interactions  
5.3 Rhyme-internal structure  
  5.3.1 The nucleus  
  5.3.2 The syllable coda  
  5.3.3 Nucleus-coda interactions  
5.4 Onset-rhyme interactions  
  5.4.1 Ci-Coda restrictions  
  5.4.2 Onset-nucleus interactions  
5.5 Syllabification rules  
5.6 The Foot  
5.7 The Phonological word  
5.8 Stress patterns and interaction with prosodic structure  

6 Constraints on morph and morpheme structure  
6.0 Introduction: terminology  
6.1 Preferred morph and morpheme structure: number of syllables  
  6.1.1 Maximum size of morphs, and of morphemes  
6.2 Morph-level constraints on consonants  
  6.2.1 The 'one per place per morph' constraint  
  6.2.2 Morph-level constraints on the L slot  
  6.2.3 Morph-level constraints on the N slot  
  6.2.4 Constraints on morph-final consonants  
  6.2.5 Cross-syllable consonant clusters  
6.3 Morph-level constraints on vowels  
  6.3.1 Constraints on vowels in adjacent syllables  
  6.3.2 Morpheme-final vowels: non-occurrence of /a/ finally  
  6.3.3 A constraint on identical vowel sequences  
6.4 Complex morphemes
7 Morphophonological processes

7.0 Morphophonological processes 137
7.1 Concatenation 137
7.2 Nasal assimilation, and other processes associated with {N} 138
7.3 Alternation of /a/ and /o/ 141
7.4 Lowering of high vowels to mid, in the 'genitive' construction 142
7.5 Alternation of /a/ and /a/ in lexical morphemes 144
7.6 Alternations in the shape of prefixes and prepositions: Ca ~ Cə ~ C 144

8 Reduplication

8.0 Reduplication, introductory overview 149
8.1 The direction of reduplication 152
8.2 Partial reduplication 153
  8.2.1 CV-reduplication 153
  8.2.2 C-reduplication 154
8.3 Full-base reduplication 156
8.4 'Final syllable reduplication' 156
8.5 Foot reduplication 159
8.6 Reduplication across morpheme-boundaries 161
8.7 Reduplication plus /a/ 'overwrite' 162
  8.7.1 Lexically specified vowel overwrite 165
8.8 Inherently reduplicated morphemes as synchronic reduplications 166
8.9 Ordering: Lexical versus Postlexical Reduplication 170
  8.9.1 Derivational vs Inherent reduplication 176

Part 1 - Concluding remarks 177

II - Morphosyntax

9 Overview of Morphosyntax

9.0 Introduction 179
9.1 Typological sketch 179
9.2 Simple clause structure 180
  9.2.1 Ellipsis of Noun Phrases 182
9.3 Basic verbal morphology 183
  9.3.1 Terminology 1: Actor and Undergoer 184
  9.3.2 Verbal prefixing 185
  9.3.3 Verbal suffixing 188
  9.3.4 Terminology 2: on the use of the term 'transitive' 189
9.4 Actor Pivot/Undergoer Pivot diatheses 190
  9.4.1 Actor Pivot 190
  9.4.2 Bare stem Undergoer Pivot 190
  9.4.3 ke-Undergoer Pivot 192
9.5 On Subjects, Objects, and Core NPs 193
  9.5.1 The Subject NP 193
  9.5.2 The Object NP 194
  9.5.3 Core and Non-core NPs 197
  9.5.4 The External Topic NP 198
10 Verbal morphology - approach

10.0 Introduction
10.1 Collection of data
10.2 Zero derivation, bound morphemes, and free morphemes
10.3 Derivation and Inflection
10.4 Are bound morphemes 'precategorial'?
  10.4.1 Inherent and derived transitivity values

11 Verbal Suffixes

11.0 The Transitive stem, 3 forms: '-∅, -in, -ang
11.1 Monomorphemic ('-∅suffixed') transitive stems
11.2 Verbs with suffix -in
  11.2.1 Verbs with suffix -in, root is a noun
  11.2.2 Verbs with suffix -in, root is an intransitive verb
    11.2.2.1 Verbs of form intrans-in: U is locative
    11.2.2.2 Verbs of form intrans-in: U is a 'causee/experiencer'
  11.2.3 Verbs with suffix -in, root is a transitive verb
11.3 Verbs with suffix -ang
  11.3.1 Realisations of -ang
  11.3.2 Verbs with suffix -ang, root is an intransitive verb
    11.3.2.1 Verbs of form intrans-ang, undergoer is a causee
    11.3.2.2 Verbs of form intrans-ang, undergoer is a 'body product'
    11.3.2.3 Verbs of form intrans-ang, undergoer is 'goal of feeling'
  11.3.3 Verbs with suffix -ang, root is a transitive verb
11.4 The derivation of 'stative intransitive stems' of shape -root(-an)

12 Verbal Prefixes

12.0 Verbal prefixes: 'zero prefix', /NI-/, and me-
12.1 'Zero prefix': bare stem intransitive verbs
12.2 The agentive prefix {N}
  12.2.1 {N} on monomorphemic intransitive verbs
12.3 The semantics of 'N-' and '/' verbs: Balinese as a 'split-S' language
12.4 The prefix me-: introduction
  12.4.1 me- on detransitive stems
  12.4.2 me- on detransitive stems, with an incorporated undergoer NP
  12.4.3 me- on intransitive stems
  12.4.4 me- on nominal stems
  12.4.5 me- as a realisation of {N}
12.5 The confix me-[noun]-an

13 Pragmatics and morphosyntax: the meanings of actorless Undergoer Pivot constructions

13.0 Introduction
13.1 The two Undergoer Pivot types in Balinese
  13.1.1 Bare stem-Undergoer Pivot verbs
  13.1.2 'Actorless' bare stem-Undergoer Pivots
  13.1.3 ke-Undergoer Pivot
13.2 Are the two Actorless Undergoer Pivots freely interchangeable?
  13.2.1 Previous descriptions of Actorless Undergoer Pivots
  13.2.2 Can the two UP types be freely interchanged?
13.2.3 Where the two Undergoer Pivots are interchangeable, with no meaning difference 279
13.2.4 Interchangeable: but meaning difference 280
13.2.5 Not interchangeable: only one form grammatical 281

13.3 Other Complicating Factors
13.3.1 Style 282
13.3.2 Exceptional Verb classes 282
  13.3.2.1 Verbs of perception 283
  13.3.2.2 Verbs of commanding 284
13.3.3 Interplay of Aspect & Modality 285
13.3.4 Clause type: main versus subordinate clauses 286
  13.3.4.1 Relative clauses 286
  13.3.4.2 Infinitival use of bare Undergoer Pivots? 286
13.3.5 Other discourse factors affecting the interpretation of ke-UPs 288
  13.3.5.1 Interaction with basic principles permitting ellipsis 288
  13.3.5.2 Other contextual factors 289

13.4 Conclusions 290

14 Undergoer primacy and the morphosyntax of Balinese
14.0 Introduction 293
14.1 Undergoer Pivot as the basic voice in Balinese 293
14.2 Syntactic orientation of Balinese: ergative, nominative or neither? 297

15 Verbal Morphology in Balinese - some conclusions
15.0 What does affixing do, and not do, in Balinese? 301
15.1 Meaning as a function of lexemes and derivations 301
15.2 The primary function of affixes in Balinese: signalling, not deriving 302
15.3 Transitivity and Affixation 305
  15.3.1 Semantic transitivity versus syntactic transitivity 306
  15.3.2 Evidence for transitivity as a semantic notion - 1 308
  15.3.3 Evidence for transitivity as a semantic notion - 2 311
  15.3.4 Conclusions, and comparison with other analyses 313

Epilogue 317

Appendices

Appendix A Semantics of phonologically 'optimal' morphemes: control sample 319
Appendix B Morph-internal interactions between consonants 322
Appendix C Morphemes with final /r/ 344
Appendix D Headwords in Warna et al (1990) of shape (σ)σ[red-CVC] 345
Appendix E Semantic classes of inherently reduplicated morphemes 346
Appendix F Potentially monomorphic lexemes of four or five syllables 348
Appendix G Analyses of mono- and poly-syllabic morphemes in Warna et al 1990 351
Appendix H CL Clusters 356
Appendix I Exceptions to [+ATR] height harmony 358
Appendix J Internal reduplications: data set for morphemes containing non-initial /...IV1V2.../ and /...rV1rV2.../ 362
Appendix K Diachronic Evidence for Reduplication 363
Appendix L Other pronoun-avoidance strategies 364

Bibliography 365
Tables

Table 2.1 Consonant phonemes of Balinese
Table 2.2 Vowel phonemes of Balinese
Table 2.3 Main allophones of Balinese vowel phonemes
Table 3.1 Usage and Restriction of syllables with short vowels, of the form C₁xVyC₂ (from Fudge 1970:169)
Table 3.2 Some recurrent semantic and phonological types associated with peripheral morphemes in Balinese
Table 3.3 Semantic classes of morphemes in control sample
Table 3.4 Semantic classes of some morpheme-types breaking the Preferred Rule 'one Cₚ place per morpheme'
Table 3.5 Numbers of morphemes of structure /#C_initial... C_final#/2
Table 3.6 Semantic types of /#C_dorsal... C_dorsal#/ morphemes
Table 3.7 Analysis of a random sample of morphemes of shape /k... k/
Table 3.8 Semantic (etc) classes of a sample of otherwise phonologically 'neutral' /r/- final morphemes
Table 4.1 Semantic classes of inherently reduplicated disyllabic bases
Table 4.2 Semantic classes of inherently reduplicated monosyllabic bases
Table 5.1 Frequencies of occurrence of syllable-types in morpheme-initial position
Table 5.2 Frequencies of occurrence of syllable-types in morpheme-final position
Table 5.3 Frequency of occurrence of consonants morpheme-initially
Table 5.4 Initial consonants in final closed syllables
Table 5.5 Frequencies of CL sequences, morpheme-initially
Table 5.6 CL sequences, intervocically
Table 5.7 Morpheme-final VC sequences
Table 5.8 CGV sequences
Table 5.9 C₁-V interactions, morpheme initially
Table 5.10 Chi-square values for the preceding table
Table 6.1 ...C(V)C_coda interactions
Table 6.2 /C_initial... C_final/ combinations
Table 6.3 Occurrence of complex onsets in morphs of comparable structure
Table 6.4: Occurrence of disyllabic morph types - with and without final consonants
Table 6.5 Shapes of monosyllabic morphemes
Table 6.6 Occurrence of morpheme-final consonants
Table 6.7 ...V₁V₂... sequences
Table 6.8 ...VV... sequences
Table 6.9 ...V₁CV₂... combinations
Table 6.10 Sample of V₁CV₂ combinations
Table 8.1: Numbers of morphemes containing medial /(...)IV₁V₂(....)/ and /(...r)ΙV₁rV₂(...)/ sequences
Table 8.2: Inherently reduplicated morphemes where vowels in copy and base differ
Table 8.3: Numbers of morphemes containing Cₐ - Cₐ sequences, by shape
Table 9.1 Examples of different root-affix combinations in Balinese
Table 9.2 Paradigmatic contrasts in the prefix slot
Table 9.3 Paradigmatic contrasts in the suffix slot, transitive verbs
Table 13.1: Person of implied Actor of actorless Undergoer Pivots, and verb affix shape, in simple clauses unmarked for aspect; a first approximation
Table 15.1 Interactions of semantic- and syntactic transitivity in Balinese
# Abbreviations & Conventions

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1ACT</td>
<td>1st-person Actor</td>
</tr>
<tr>
<td>2ACT</td>
<td>2nd-person Actor</td>
</tr>
<tr>
<td>A</td>
<td>Actor</td>
</tr>
<tr>
<td>AP</td>
<td>Actor Pivot, Agentive Prefix</td>
</tr>
<tr>
<td>APP</td>
<td>'Applied' suffix</td>
</tr>
<tr>
<td>BI</td>
<td>Bahasa Indonesia (Indonesian)</td>
</tr>
<tr>
<td>CAUS</td>
<td>causative</td>
</tr>
<tr>
<td>C</td>
<td>consonant</td>
</tr>
<tr>
<td>CRD</td>
<td>crude</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
</tr>
<tr>
<td>DFR</td>
<td>deferential</td>
</tr>
<tr>
<td>Du</td>
<td>Dutch</td>
</tr>
<tr>
<td>DUR</td>
<td>durative aspect</td>
</tr>
<tr>
<td>G</td>
<td>glide</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>HI</td>
<td>high/formal vocabulary</td>
</tr>
<tr>
<td>HON</td>
<td>honorific vocabulary</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
</tr>
<tr>
<td>Jav</td>
<td>Javanese</td>
</tr>
<tr>
<td>L</td>
<td>/l, r, w/ or /y/</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative</td>
</tr>
<tr>
<td>LOW</td>
<td>low/informal vocabulary</td>
</tr>
<tr>
<td>N</td>
<td>nasal consonant</td>
</tr>
<tr>
<td>NOM</td>
<td>nominal affix</td>
</tr>
<tr>
<td>OCP</td>
<td>Obligatory Contour Principle</td>
</tr>
<tr>
<td>OJ</td>
<td>Old Javanese</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>P-word</td>
<td>phonological/prosodic word</td>
</tr>
<tr>
<td>RED</td>
<td>Full base reduplication</td>
</tr>
<tr>
<td>RED&lt;sub&gt;cv&lt;/sub&gt;</td>
<td>Partial reduplication</td>
</tr>
<tr>
<td>RED&lt;sub&gt;a&lt;/sub&gt;</td>
<td>Reduplication, plus overwrite</td>
</tr>
<tr>
<td>S</td>
<td>Subject/pivot NP of a syntactically intransitive verb</td>
</tr>
<tr>
<td>S&lt;sub&gt;A&lt;/sub&gt;</td>
<td>S with Actor semantic role</td>
</tr>
<tr>
<td>S&lt;sub&gt;U&lt;/sub&gt;</td>
<td>S with Undergoer semantic role</td>
</tr>
<tr>
<td>Skt</td>
<td>Sanskrit</td>
</tr>
<tr>
<td>U</td>
<td>Undergoer</td>
</tr>
<tr>
<td>UP</td>
<td>Undergoer Pivot</td>
</tr>
<tr>
<td>VBL</td>
<td>verbal affix</td>
</tr>
</tbody>
</table>

- **VV** Verb of Violence
- **ZV** 'zero verb, trans. verb not bearing a suffix (U = true patient)'
- **< >** orthographic form
- **< >** morpheme boundary
- **< >** phonological word boundary
- **< >** syllable boundary
- **ka'dek** stress the syllable following the apostrophe
- **kadék** stress the underlined syllable
- **(*)** 'unacceptable if the element in brackets is not present'
- **??()** 'of doubtful acceptability if the element in brackets is not present.'
- **(*)** 'unacceptable if the element in brackets is present'
- **!** strikingly unacceptable, gibberish

**root:** lexical morpheme to which derivational affixes concatenate
**stem:** root morpheme, plus suffixes

**morph, morpheme, simple morpheme** and **complex morpheme** are defined in section 6.0. **complex formant** is defined in 4.2
Note on orthography and glosses:

The orthography used to represent Balinese largely follows that of Warna et al 1990, though with some divergences, eg in the representation of glides (2.1.3). The orthography is always in bold typeface, with phonetic/phonemic representations in plain (IPA) typeface: *lumbalumba* [lumb3lumb3] 'porpoise'.

Angled brackets are sometimes used to stress that orthographical symbols are being used. The laminal nasal */n/* is represented by *<ny>* before a vowel, and *<fi>* before a consonant: *nyakan* 'cook rice', *ěičéh* 'liquid (v)'. The standard orthography uses *<n>* in the latter case; the *aksara* orthography writes (its equivalent of) *<fi>*.

Orthographic *<ə>* represents schwa, IPA [ə], and *<e>* represents IPA [e], thus *me-ketël* /maketel/ 'drip'.

Within slash brackets and square brackets, symbols have their normal IPA values: [e] = mid-high front vowel, and so on. I have had to represent the mid-low back vowel as [ø]; [o] has its usual value, though is sometimes transcribed as [ö] for clarity.

I write all inherently reduplicated morphemes without a hyphen, thus *cakcak* 'chop up', *kupukupu* 'butterfly, kebaratkebirit' 'run about wildly'. The standard orthography inserts a hyphen where the base is di- or trisyllabic: *kupu-kupu, kebarat-kebirit* but not when it is a monosyllable: *cakcak*

I write prefixes *pe-* and *me-* where the standard orthography uses *pa-* and *ma*; cf 7.6.

Pronouns do not distinguish number, or in most cases, gender. For simplicity when glossing I usually translate with the singular, and generally use he/him to gloss 3rd person pronominals.
Introduction

1.1 Aims

This thesis aims to describe features of the phonology and morphosyntax of the Balinese spoken in the banjar ('hamlet') of Peguyangan, Singaraja, Bali. As such, it is intended as a contribution towards Balinese and Austronesian descriptive linguistics. At the same time, I hope that it will be easily used by those interested in the description of languages and language generally.

The analysis is not heavily oriented in favour of one linguistic school over another, except insofar as I have looked for interactions between semantics and other aspects of linguistic structure. I have aimed to produce analyses that are clearly argued for, and so easily falsifiable by new data. Explicitness has taken priority over breadth of analysis: what was originally intended as a 'reasonably complete' reference grammar of a dialect of Balinese has thus become just a part of one.

I have based the analysis on a solid body of language data, taken from a wide variety of discourse genres. The data corpus includes informal and formal conversational material, oratory, oral literature, including folk tales, shadow puppet performances and folk drama, and a variety of written texts, including modern literary works, school texts, and transcribed folk tales.

My original, and continuing motivation, has been a long-term fascination with Balinese, as a complex artefact shaped by a multiplex culture. I hope that the analysis, even in the phonology, gives some idea of the way that language has interacted with other aspects of culture in Bali. I agree with those who believe that language cannot be adequately described if it is divorced from its socio-cultural context.

1.2 Balinese in its wider context

Balinese is the principal language spoken on the island of Bali, Indonesia. In 1980, the last census for which language spoken was surveyed, just under two and a half million people indicated they spoke Balinese in the home. At current rates of population increase, there are now more than 3 million speakers of Balinese. Of these, around 80,000 live on the neighbouring island of Lombok while, largely due to transmigration, 40,000 speakers are found in southern Sumatra, and 60,000 in Sulawesi.
Balinese is an Austronesian language, and shares many features with Western Indonesian languages like Malay and Javanese - no doubt due in part to long term areal contacts (see below). Balinese was sub-grouped with the neighbouring languages Sasak and Sumbawanese by Esser (1938), based on shared lexis and grammatical features, 'especially the [verbal] suffixes -ang and -in' (Noorduyn 1963:336), and also by Dyen (1982), on lexicostatistical grounds. Blust (1985) refers to a 'Java-Bali-Sasak' group, though without giving the basis for this analysis. I have not had access to Mbete (1990), a work said to follow Esser's hypothesis (Beratha 1992).

Balinese has been written down for more than a thousand years. The earliest known texts are the royal decrees inscribed on copper/bronze plaques from the 9th to the 11th centuries. The language of these texts (about ten thousand words in all) is referred to as Old Balinese (OB). Beratha (1992) is a useful source of data on Old Balinese. Goris (1954) contains almost all known OB texts, with his translations into Dutch, as well as a glossary. The texts show considerable influence from Sanskrit and Old Javanese.

As a result of intimate contact in the Old Balinese period, Balinese shows marked influence from Javanese - though borrowings largely ceased after the 16th century (Clynes 1989, 1994). There is also a significant Sanskrit element in modern varieties of Balinese, as well as borrowings from other languages such as Malay, Dutch, Portuguese and Chinese languages. Most of the latter borrowings can be presumed to have come via Malay, recently.

Most people under the age of, say, 60, also speak the national language, Indonesian, which has had an increasing influence, encroaching on many domains where previously only Balinese was spoken. In the capital, Denpasar, I have met teenagers who speak mainly Indonesian with their parents, and whose command of Balinese seems far diminished compared with that of their parents. There has been a growing concern in the last ten years about developments such as these (see papers presented at the Third Congress on Balinese (Bagus 1992)).

1.3 The dialect described

The variety of Balinese (chiefly) represented in this study is that spoken in the banjar ['hamlet', 'village'] of Peguyangan, which forms part of Singaraja, a port town and regional capital of the Buléléng district in north Bali. Fieldwork was carried out there from June 1990 to April 1991, for approximately eleven months. This was not my only period of data collection. I had previously lived in Bali for 5 years, beginning in 1979, and during that time collected a good deal of written and oral data (though I had no training in linguistics at the time). As well, my wife is Balinese, and from Singaraja: a good deal of Balinese is used in the household by family members, so data collection has continued on a daily basis.
There were several other good reasons for choosing the Singaraja/Buleleng dialect. Singaraja was the administrative capital in Bali in the Dutch colonial era, and as such has long been a centre for scholarly work on the language and culture. These have been documented by a series of outstanding Dutch scholars (Vickers 1989), the first of whom, the linguist Van der Tuuk, and the scholar-administrator Liefrinck, arrived in the 1870's.

The important Kirtya Liefrink-Van der Tuuk (later renamed Gedong Kirtya) library of traditional literary works was established in Singaraja in the 1920's. Since its establishment, officials of the Gedong Kirtya have collected and transcribed oral and written texts, many of which originate locally.

Singaraja has been an educational centre for most of this century. The first secondary school on Bali was established here, and the IKIP (teachers' college) has produced generations of teachers for schools in Bali. A series of authors from Buleleng have produced textbooks for the teaching of Balinese in primary and high schools. In the post-war era these have included dedicated scholars like I Gusti Bagus Sugriwa, I Gusti Ketut Ranuh, I Ketut Ginarsa, I Wayan Simpen, I Nengah Tinggen, I Made Denes, and I Gusti Putu Antara, to mention just some. Influential writers, such as Gde Srawana and A.A. Panji Tisna (who wrote some of his work in Balinese), and a later generation represented in works like Bagus and Ginarsa (1978), also originate from the Singaraja area. Until the late 1980's, the government Balai Penelitian Bahasa [Language Research Bureau] was also based in Singaraja, publishing a variety of literary works and linguistic studies.

There exists therefore a wide variety of published texts written in the Singaraja dialect, which has thus played an important part in the shaping of 'standard Balinese', an adequate grammar of which has yet to be written.

I stress however that I have not tried to describe, or prescribe, Standard Balinese (bahasa Bali baku [BI]) in the present work. Participants at the 1974 Congress on the Standardisation of Balinese were in agreement that a standard variety of Balinese should be promoted, and that the varieties spoken in the Buleleng and Klungkung districts should be taken as representative of that standard (Bagus 1975:44, cited in Sulaga et al 1985). Given the traditional regional rivalries in Bali, this was a bold recommendation. Not surprisingly, as Ginarsa et al. (1975) point out, no single regional variant is acceptable to all Balinese as a model for standard Balinese: none is entirely free of unsatisfactory (kurang memuaskan [BI]) features, at least to Balinese from other areas.¹

Still, as a result of the 1974 congress, Buleleng and Klungkung dialects have become the basis for descriptive work by government linguists such as Udayana University's Fakultas Sastra (1977), Warna et al (1983:2), and Sulaga et al (1985:8), all of which aim at defining the standard. Bawa and Jendra (1981) even include a folk tale by the late

¹The recommendation also passes in silence over the striking dialectal variation to be found locally in both Klungkung and Buleleng. From the examples given in the papers presented at the Congress, it is clear that the model was expected to be from the Lowland group of dialects, rather than a Bali Aga (1.3.1) dialect.
Chapter 1

Madé Narsin, from Peguyangan, Singaraja in an appendix, as a sample of standard Balinese.²

There are no doubt, therefore, many features of the dialect described in this thesis which most Balinese would find acceptable in a 'standard' variety of Balinese. However there are also other features which would definitely be rejected by most as non-standard. These include for example much of the vowel allophony, as well as the alternation of /a/ and /o/ in certain contexts (7.3), the lowering of high vowels in others (7.4), and aspects of the syntax of relative clauses (13.3.4.1). I stress that I am trying to give a coherent synchronic account of a single variety of Balinese, rather than proposing a model for emulation.

Linguistically, the Singaraja dialect is an ideal one for studying verbal affixation processes, since it makes full use of affixes. In some dialects, such as those in the Denpasar area, suffixes are often omitted, blurring distinctions such as those described in chapter 9.

Another good reason to study the Singaraja dialect is because some of the most sensitive, and the most influential, of previous work on Balinese grammar, that of the Dutch missionary Kersten (1948, n.d., 1970, 1984), was based on this dialect. Kersten 1984 also includes a medium-sized dictionary of Singaraja Balinese. The Singaraja dialect is thus probably the best described of the modern dialects (while still being in many ways a linguistic unknown, given the many very basic gaps in all descriptions that I am aware of). I was thus fortunate to have Kersten's work as a starting point and a stepping stone to further investigations.

Although the core of the description is of the dialect spoken by people living in Banjar Peguyangan, Singaraja. I have also used data from speakers living in surrounding areas, to a radius of approximately 5km, including the villages of Penarukan in the east, Sukasada in the south and Pemaron in the west. All of these villages are part of 'greater Singaraja'. I have also occasionally used texts from the towns of Banjar and Seririt (respectively 20 and 30km to the east), where the grammar accords largely with Peguyangan usage. These varieties share a great many features common to lowland Buléleng dialects.

There is of course variation, sociolinguistic and geographical, even across this quite limited area, but for the purposes of data collection I have operated as though people in these various areas spoke a single dialect. I have not tried to study variation in detail, though both cross-generational and geographical variation is occasionally mentioned. I found that Peguyangan speakers aged below around fifty showed a variety of (usually minor) linguistic differences from their elders in a range of respects: phonology, morphology, syntax, conventions of usage (eg politeness conventions) and so on. The focus of much of the study of verbal morphology is on the variety of older Peguyangan

²Unfortunately this text was poorly proofread; it contains many typographical errors, and occasional missing words/morphemes. It cannot be used as a representative sample of Singaraja Balinese.
Introduction

speakers (over 50), since I spent many hours working with people of that generation, Bapak Ketut Mandi Pinatih in particular. There was nonetheless basic agreement in usage between the older and younger speakers.

The most significant, and often commented on, change in Peguyangan is in the loss of the more 'refined' speech styles among young people. While many members of the older generation still have an easy command of the high styles, the younger people's grasp is more limited, both in the control of high vocabulary, and the ability (and, no doubt, the desire) to use a sustained elaborate high style in conversation.

My data base includes many folk tales collected in the Singaraja area by workers at the Gedong Kirtya, often in the 1930's and 1940's. In most respects these texts accord closely with the variety spoken in Singaraja today, though occasionally they contain elements which seem old-fashioned or archaic to present-day speakers. I have not included such elements in this study.

All in all, my data base is much more limited than that used by many writers on Balinese, some of whom apparently consider any published text as being acceptable data. I use only texts from the area defined above. As a result my detailed claims about usage apply only to Singaraja Balinese; on the other hand, because I use an homogenous data base, clearer generalisations emerge from the data than would otherwise be the case.

1.3.1 dialectal variation and speech styles

There is phonological and lexical variation across the island, though all varieties are largely mutually intelligible. The principal dialectal study is Bawa (1983), also Jendra (1976). A distinction is made between the so-called 'Bali Aga' varieties, found largely but not entirely in the central mountainous area, and the 'mainstream' lowland varieties. Bali Aga varieties lack certain phonological, morphological and lexical innovations found in the lowland varieties, which were influenced by contact with Old Javanese. Lowland varieties are often further distinguished according to the present-day regional districts, which were formerly independent political units: Klungkung, Karangasem, Buléleng [with capital Singaraja], Gianyar, Badung [with capital Denpasar], Tabanan and Jembrana. The situation is in fact more complex than this, with noticeable lexical and phonological variation often occurring over distances of less than a few kilometres (cf Bawa, Jendra).

The greatest difference within the lowland varieties is between those spoken in the north and east (Buléleng, Klungkung, Karangasem), on the one hand, and the varieties of the Denpasar area in the south, on the other. Denpasar Balinese is characterised by a loss or neutralisation of verbal and nominal suffixes found elsewhere.

Lowland varieties are characterised by 'speech levels' or 'speech styles' - distinct formal and informal registers (Bagus 1979, Clynes 1989). In formal registers social distance
and differences in the relative status of the speaker and the person referred to are expressed by lexical choices. Almost any utterance is marked in terms of the perceived social relationship obtaining between the speaker and the addressee/person referred to. Bagus (1979) discusses the effect of social change on use of styles. Speech styles were borrowed from Javanese, and the membership of all but the low lexical class consists virtually entirely of Javanese loanwords (Clynes 1989, 1994). Styles usage is not a primary focus of the present work, though related avoidance phenomena, and their effect on morphosyntax even in the low style, are (Chapter 13).

1.4 Previous studies of Balinese

Previous work on Balinese includes short sketch grammars by Kersten (1948, and its successive revisions, no date (c.1960), 1970 and 1984), Barber (1977, largely derivative of Kersten), Warna et al (1983), and Oka Granoka et al (1985). All of these are practical pedagogical grammars. The pedagogical grammar of Tinggen (1984) is the only one I am aware of which is written in Balinese. Artawa (1994) studies aspects of the syntax. Historical change is studied by Hunter (1988) and Beratha (1992), who provides a useful critical survey of previous work, with accurate evaluations. Ward (1973) analyses aspects of the phonology and morphology, but does not advance significantly on the analyses of Kersten.

Dictionaries include Warna et al (1990) and Kersten (1984), both Balinese-Indonesian, as well as Barber (1979) and Shadeg (1977), both Balinese-English, and Balai Penelitian [...] (1975), Indonesian-Balinese. Simpen (1985) is the only monolingual Balinese dictionary. None of these dictionaries is comprehensive enough for a study like the present one: in particular, the absence of a root-affix combination from the dictionaries does not necessarily mean that that combination is not a regularly formed word, in everyday use.

A fault of many previous studies is that they do not specify which dialect they are describing (though Kersten's work appears to be based largely on material from Buléléng speakers). Work done by some writers mixes data from a variety of sources, on the assumption that it all represents 'standard Balinese'. This is a problematic assumption, given the existence of significant grammatical variation, even among the lowland dialects (see for example comments by Anom 1975:77).

All previous grammars are limited in scope and all leave many questions of interest to linguists (and non-linguists) unanswered, including even very basic ones such as what word classes are needed to describe Balinese, what are the patterns of vowel allophony, what is the structure of the noun phrase, or the pronominal system. Beratha (1992) is more ambitious than most of the others, and is written in the vocabulary of modern linguistics. I have not been able to consult Artawa (1994).
1.5 Organisation of the thesis

This thesis is organised in two parts: Part I (chapters 2 to 8) describes principal features of the phonology, while Part II (chapters 9 to 15) deals with aspects of verb morphology and morphosyntax. In chapter 2 segment-level phonology is described. The description of the phonology departs from usual practice in often making use of semantic evidence; this approach is justified in chapter 3. In chapter 4 I argue for the existence of morphs, phonological formants more complex than the segment. Chapter 5 describes the prosodic units which play a significant role in the phonology, including the syllable, the foot and the phonological phrase. In chapter 6, phonotactic constraints on morph and morpheme structure are described, while chapter 7 discusses non-prosodic morphophonemic alternations. A separate chapter is then devoted to reduplication, because of the complexity of the issues involved, and because of the key role that evidence from reduplication processes has played in the development of phonological and morphological theory in the eighties and nineties. The description, while not heavily committed to any particular theory, assumes a constraint-based model. In the latter chapters of Part I a variety of evidence is presented in support of the prosodic morphology hypotheses of McCarthy & Prince.

Part 2 of the thesis deals with topics in the morphosyntax of Balinese. Chapter 9 prevents an overview of basic morphosyntax, while chapter 10 describes some analytical decisions. Chapter 11 describes the principal verbal prefixes, while chapter 12 presents the suffixes. In chapter 13 I show that some very basic aspects of the morphosyntax cannot be described adequately without taking into account the discourse context. In chapter 14 evidence for Undergoer-primacy in Balinese is presented; I argue against the view that the syntax of Balinese shows an ergative orientation. In chapter 15 I present conclusions about the functioning of affixation in Balinese.

The description where it touches upon issues of semantics (including linguistic pragmatics) is heavily influenced by the ideas of Wierzbicka. I find that this approach, which seeks always to explain the complex in terms of simpler concepts, is well suited to the analytical task undertaken.
PART I

PHONOLOGY
2 Segmental phonology

2.0 Introduction

Balinese is most easily described assuming a hierarchy of phonological complexity, going from the primitive units, phonological features, to sets of associated features (segments), which in turn combine to form maximally complex non-prosodic phonological units, morphs.¹

(1) morph
    | segment
    | feature

This chapter is primarily devoted to a description of the segmental repertory. Consonant and vowel phonemes, and their phonetic realisations, are described successively, in sections 2.1 and 2.2. I argue that certain realisations are the product of meaningful phonological rules (2.1.2) and (2.2.8).

Evidence for the morph as a phonological unit is presented in chapter 4.

At several stages in the description I refer to two components in the phonology, the lexical phonology and the postlexical phonology. Balinese presents a variety of evidence in favour of this distinction; see especially (8.9). Phonological rules typically apply differently in the two components. Kaisse & Hargus (1993:16) list (amongst others) the following typical, though not diagnostic, differences between the two types:

<table>
<thead>
<tr>
<th>Lexical</th>
<th>Postlexical</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. may not apply across words</td>
<td>may apply across words</td>
</tr>
<tr>
<td>b. may refer to word-internal structure</td>
<td>do not refer to word-internal structure</td>
</tr>
<tr>
<td>c. structure-preserving [AC: don't generate non-distinctive elements]</td>
<td>need not be structure-preserving [AC: do generate, eg, allophones]</td>
</tr>
<tr>
<td>d. may have exceptions</td>
<td>do not normally have exceptions</td>
</tr>
<tr>
<td>e. precede all postlexical rules</td>
<td>follow all postlexical rules</td>
</tr>
</tbody>
</table>

The description is based mainly on the speech of people aged in their forties and over, though the usage of younger speakers where it departs from that of their elders is also

---

¹I take it that the phoneme, like the 'segment', is not a phonological primitive. Archangeli & Pulleyblank (1994:45-47) marshal a good deal of evidence to this effect. At the same time, they note the central importance that this unit has been accorded in virtually all recent phonological work, and that there may be a place for the segment as a 'derivative' organisational unit ((1994:46)).
occasionally recorded (e.g., 2.2.3). The description of phonetics is necessarily limited in that it is based on impressionistic rather than instrumental data.

### 2.1 Consonant phonemes and allophony

The 18 consonant phonemes of the Peguyangan dialect are set out in Table 2.1, in the symbols of the orthography used in this thesis. Equivalent IPA symbols are given in brackets in some cases. In addition, evidence is given in 2.1.4 for the existence at an underlying level of '/N/': a nasal consonant unspecified for place, found in homorganic nasal-stop clusters. The glides /w/ and /y/ do not contrast with the high vowels /u/ and /i/, underlyingly (2.1.3).

<table>
<thead>
<tr>
<th></th>
<th>labial</th>
<th>apical</th>
<th>laminal</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>nasal stops</td>
<td>m</td>
<td>n</td>
<td>ny, ň² (/ŋ/)</td>
<td>ng (/ŋ/)</td>
</tr>
<tr>
<td>oral stops</td>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
</tr>
<tr>
<td>vclss</td>
<td>b</td>
<td>d</td>
<td>j</td>
<td>g</td>
</tr>
<tr>
<td>vcd</td>
<td></td>
<td></td>
<td>s</td>
<td>h</td>
</tr>
<tr>
<td>fricatives</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>liquids</td>
<td></td>
<td></td>
<td>l, r</td>
<td></td>
</tr>
<tr>
<td>glides</td>
<td>w</td>
<td></td>
<td>y (/j/)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1 Consonant phonemes of Balinese

Minimal pairs/sets contrasting the consonant phonemes of Balinese are given in the following sets:

(2) Labials:

| *pépék* 'vagina' | *apa* 'what' | *ngaap* 'smart, sting (vi)' |
| *bébék* 'duck'   | *aba* 'carry' | *raab* 'roof'              |
| *mémé* 'mother'  | *tama* 'inherit' | *maam* 'eat [childish]' |
| *wéwéh* 'give.HI' | *sawa* 'corpse.HI' | (-)                        |

(3) Apicals:

| *tatah* 'file (vt)' | *bati* 'profit' | *ngaat* 'smart, sting (vi)' |
| *dadah* 'heat up (vt)' | *kadi* 'like.HI' | *laad* 'remains' |
| *nanah* 'pus' | *bani* '(be) brave' | *baan* 'get' |
| *lalah* 'spicey, hot' | *bali* 'Bali' | *maal* 'expensive' |
| *raras* 'style, bearing' | *kari* 'still.HI' | *daar* 'eat' |

---

2See note on orthography, page xiv.
3I take forms like *pépék* and *blabar* to be underlyingly reduplicated monosyllables (8.2.1).
### Segments

<table>
<thead>
<tr>
<th>Initial</th>
<th>Medial</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Laminals:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cucu 'grandchild'</td>
<td>pacah 'kind of flower'</td>
<td>(-)</td>
</tr>
<tr>
<td>jujuh 'reach (v)'</td>
<td>-ajah 'teach'</td>
<td>(-)</td>
</tr>
<tr>
<td>susu 'breast (HI)'</td>
<td>asah 'level (v)'</td>
<td>kadas 'clean'</td>
</tr>
<tr>
<td>yuyu 'crab'</td>
<td>ayah 'help (v)'</td>
<td>(-)</td>
</tr>
<tr>
<td>nyuh 'coconut'</td>
<td>anyar 'fresh, new'</td>
<td>(-)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initial</th>
<th>Medial</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Back:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kahkah 'rough (skin)'</td>
<td>akah 'root'</td>
<td>baak 'robber'</td>
</tr>
<tr>
<td>gagah 'open (vt)'</td>
<td>lagas 'freely'</td>
<td>baag 'red (of face)'</td>
</tr>
<tr>
<td>ngahngah 'sore'</td>
<td>langah 'rare'</td>
<td>baang 'give (vt)'</td>
</tr>
<tr>
<td>[hyang 'divine title']</td>
<td>saahasa 'suddenly']</td>
<td>baah 'flare up (of infection)'</td>
</tr>
</tbody>
</table>

All consonants may occur syllable-initially in very formal styles, though in everyday speech /h/ does not occur in onsets: forms like hyang and saahasa 'suddenly' above are yang /yan/ and saasa /saasa/ in the everyday language. All may occur syllable-finally, except the laminal stops /J1, c, j/ and the glides /w, vi; the liquids and glides may occur as the second element in a syllable-initial cluster. More detailed descriptions of phonotactic constraints are given in chapters five and six.

### 2.1.1 Consonant allophony

Unless commented upon specifically, consonant phonemes can be assumed to be realised with the usual IPA values. The labial /p/, usually [pl, is occasionally realised by some speakers as [j], a bilabial fricative, both initially and intervocally.

Both apical and laminal series are articulated in Catford’s (1977) alveolar region. The apicals /t/, /d/ and /n/ all involve contact of the tip of the tongue with the alveolar ridge, close behind but not contacting the teeth.  

The laminals /c/, /j/ and /n/ involve laminal contact on the convex curve of the alveolar ridge. /n/ has two main realisations: before vowels, as [n] (a laminal stop, with a

---

4 This variety of Balinese, like most, has no dental phones, /t/ being alveolar like all the other apicals. (In much of southern Bali the apical series is quite strongly retroflexed (Bawa 1983), but this is not the case in Buleleng). The alveolar realisation of [t] is contrary to the pattern found in many Western Indonesian languages, which have a dental /t/ patterning with an otherwise apico-alveolar phoneme series (Henderson 1965). However Bawa (1983) reports that some isolated Balinese dialects do have the dental /t/ as part of an otherwise alveolar-apical phoneme class. A speaker from Bukti (about 20km to the east of Singaraja) but living in Peguyangan had occasional apico-dental realisations of /t/ eg tatakan 'base, support' [tatakan]; I also noticed her occasionally realise the /pl/ cluster as laminodental eg in Pojok Batu (placename). This non-phonemic variation between dental and alveolar realisations of both apicals and laminals reflects the basically active-articulator orientation of phoneme classes: Sagey (1990 [1986]), Ladefoged and Halle (1988).

5 For /c/, one speaker described the closure as sometimes apical-laminal, closing the entire alveolar ridge, sometimes as purely laminal, and making contact only with the convex part of the alveolar ridge. Others described it only as the latter.
laminal off-glide); and before consonants, as [n], (a laminal stop, realised in the same area of the alveolar ridge as apical [n], the main allophone of /n/). 6

/s/ has a laminal stricture, also on the alveolar ridge region, with the sides of the tongue touching the teeth behind the canines. 7

The distribution of high vowels /i/ and /u/ and their glide counterparts /y/ and /w/ is predictable and never contrastive (2.1.3). The surface distinction between these phonemes derives from syllabification processes (5.5, Hayes 1989).

A single natural class 'back' is assumed, grouping both phonetically dorso-velar and glottal phonemes. The two groups pattern together in allophony (see below); evidence from morpheme structure constraints also supports the grouping (6.2). This patterning together of laryngeal and dorsal consonant phonemes as a single natural class is not well captured in current models of feature geometry such as Sagey (1990), Clements & Hume (1995). McCarthy (1994) gives evidence for a [pharyngeal] zone-like place of articulation feature in Semitic languages, shared by, amongst others laryngeal and uvular, but not velar, consonants.

The dorsal stop phonemes have fronted (post-palatal) allophones before /i/ and /e/, and are lowered to post-velar or uvular position before /o/ and /a/.

Allomorphy rules for /k/ are: 8

(6) /k/ |?| p-word
    > [k] elsewhere

/k/ is realised as the oral stop [k] in all environments except phonological-word-finally, where it is realised as [ʔ]. This includes before pause and before consonant-initial clitics.

An apparent exception to the above rule is the realisation of /k/ as [k] word-finally in 'expressive' uses; I take this to be the product of an optional meaningful rule (see 2.1.2 and 3.2.1).

---

6Auditorily too, laminal [n] and apical [n] are very similar, particularly since the opposition only occurs before a following stop consonant - [nc, nj] versus [nt, nd]. However speakers stress that the area of realisation on the tongue is quite distinct in each case: [n] cannot occur before [t] or [d] and [n] cannot occur before [c] and [j].

7Some speakers indicated an apical-laminal stricture, with the tip and blade of the tongue near the curve of the alveolar ridge, others a purely laminal one, with the tip of the tongue near the lower teeth.

8These rules are specific to the Peguyangan dialect. Quite different rules of allophony (eg realisation of word-final /k/ as [k]) are said to be found in varieties only a few kilometres to the east, cf Bawa (1983).
(7) kuluk   'dog'      [kuluʔ]
kuluk-é 'the dog'   [kuluke]
kuluk-né 'his dog'  [kuluʔné]
jemak  'take'       [jamaʔ]
jemak = a 'taken by him' [jemak3]
kauk-kauk 'call out (vi)' [kauʔkauʔ]

/k/ sometimes occurs syllable-finally as the first element in a consonant cluster, though only in loanwords. In this case, the 'elsewhere' realisation, [k] occurs:

(8) me-bakti 'pray' (Skt bhakti) [məbakti]
pemaksan 'congregation' (Skt paksā) [pemaksan]

As well, for some younger speakers at least, /k/ is optionally realised as [ʔ] when morpheme-final but occurring intervocally due to suffixing or cliticisation:

(9) kauk-in 'call to' [kaukin] ~ [kauʔin]
kuluk-é 'dog-DEF' [kuluke] ~ [kuluʔe]
jemak = a 'taken by him' [jemak3] ~ [jamaʔ]

The occurrence of [ʔ] in this position may be relatively recent: KMP (aged in his sixties) rejected it as 'affected', however both realisations were common in the speech of people aged in their 20's. Those who optionally realise /k/ in this way can be said to alternate between the rule in (6) above and the following variant:

(6)b /k/ > [ʔ] / _ _ _ 

(where the conditioning environment ']' represents any morpheme boundary).

Underlyingly, the back fricative /h/ occurs only morpheme-finally in inherited lexemes. It surfaces medially in morphemes consisting of a reduplicated /C(L)Vh/ sequence, eg kihkih 'grate (vt)', the first syllable being here analysed as a phonetic copy from an underlying monosyllable, see 8.8. Usually /h/ has the glottal realisation [hl.

(10) pasih 'sea' [pasih]
uluh 'swallow whole' [uluh]
yeh 'water' [yeh]

The oral allophones - velar fricative [x] varying with uvular [χ] after the low vowels /o/ and /a/, and with palatal [ʃ] after /l/ - occur obligatorily before other consonants in loanwords:

(11) cihna 'sign, omen' (Skt cihna-) [ciːn3]
brahma 'Brahma' (Skt brahma-) [braːx3]

There are other occurrences of the oral allophones of /h/, which I take to be the products of a meaningful postlexical rule; this is discussed in the following sub-section, 2.1.2, and also in 3.2.1). Given the very marked nature of /hC/ consonant clusters, the oral allophones above are perhaps best described as products of that rule too.
2.1.2 Meaningful phonological rules 1: consonants

Certain optional realisations of the phonemes /k/ and /h/ are the products of meaningful phonological rules, of the type identified by Woodbury (1987). Postlexical or allophony rules are manipulated to convey 'expressive' meaning.\(^9\) The nature of the expressive meaning component added by rules such as (15) and (17) is proposed in section 3.4.2.

The allophony of /k/ described in rule (6), applies only optionally; where words are uttered with great emphasis or 'feeling', word-final /k/ will be realised as [k], rather than the [?] predicted by that rule. In examples (12) and (13), the first sentence shows the usual realisation, while the second shows the marked usage:

(12) a. bes cenik baju-né
    too small shirt-DEF
    The shirt is too small
b. awak nu cenik!
    body still small
    You're still so small! [too young]

(13) a. kayu-né berek
    wood-DEF rotten
    The wood is rotten
b. berek gati bikas cai-né
    rotten very behaviour you-DEF Your behaviour is rotten!

It should be emphasized that the [k] realisations in (12b) and (13b) are optional, though the use of the normal [?] allophone would not carry the same expressive force.

Where the /k/ phoneme occurs finally in an inherently reduplicated morpheme, the usual realisation is again [k]:\(^{10}\)

(14) cakcak 'chop up' [cakcak]
     cekcek 'house lizard' [cekcek]
     ciplak- ciplak 'smack one's lips' [ciplakciplak]
     noktok %N-toktok% 'AP-knock.on' [nôktôk]

I take it that this realisation of word-final /k/ is again semantically motivated, given the typically 'expressive' meanings of this morpheme class, which include various iconic meanings (onomataopoeia, repeated movements), and names of plants and animals (see 3.2.1).

\(^9\) Detailed arguments in favour of direct interactions of phonology and semantics are presented in chapter 3. On the distinction between lexical and postlexical phonology, see 2.0. The meaningful postlexical rules described here and in 2.2.8 are atypical of this rule type in being optional.

\(^{10}\) The 'correct' realisation of cekcek according to KMP is [cekcek], [ceʔceʔ] being 'usually said to children'.
Rather than introduce a semantic component into the environment of rule (6), I follow Woodbury (1987) in concluding that the [kl] realisations are the product of a distinct, meaningful, but purely phonological, process:

(15) \[ k \rightarrow [k] \quad /\_\text{Foot} \]

Meaning: 'adds expressive meaning' (see 3.4.2)

Postlexical realisation of /h/ is equally best described by two rules, firstly the automatic, meaningless rule already described in 2.1:

(16) \[ h \rightarrow [h] \]

and secondly an optional meaningful, but purely phonological, postlexical rule:

(17) \[ h \rightarrow [\chi, \chi] \]

Meaning: 'adds expressive meaning' (see 3.4.2)

This rule describes optional realisations of /h/ in the same expressive semantic contexts determining the realisation of /k/:

(i) when words are uttered emphatically, 'with feeling':

(18) \[ \text{luh!} \quad 'female; term of address' \quad [\text{lu}\chi] \quad (\sim [\text{luh}]) \]
    \[ \text{matah!} \quad 'pathetic! half-baked!' \quad [\text{mata}\chi (:)] \]

When used non-emphatically, with its primary meaning, 'raw', matah is realised as [matah].

(ii) In morphemes with 'iconic' semantics - see 3.2.1 - and consisting of a reduplicated /C(L)Vh/ sequence:

(19) \[ \text{kihkih} \quad 'grate (v)' \quad [\text{kihkih}] \quad \sim [\text{ki}\chi\text{kichi}] \quad \sim [\text{ki}\chi\chi\chi] \]
    \[ \text{kohkoh} \quad 'sound of coughing' \quad [k\chi\text{ochi}] \quad \sim [k\chi\text{ochi}] \quad \sim [k\chi\text{ohk oh}] \]

2.1.3 The status of glides

The high vowels /i/ and /u/ only occur in syllable nuclei, and their glide counterparts /y/ and /w/ never do. There is no evidence for an underlyingly distinction between the two. Differences in phonological behaviour, such as these different phonotactic distributions, derive from processes assigning prosodic structure to morphs. Syllabification (5.5)) and Foot Wellformedness (5.6) rules impose disyllabic structure where possible, and in the process, assign surface consonant or vowel status to underlying %i% and %u%. 

17
This analysis differs from most others in positing that orthographic \(<i>\) and \(<u>\) should generally be analysed as phonemic glides, not vowels, where they occur before vowels in morphemes containing at least one other potential syllable nucleus, thus:

\[(20)\] satua /satwa/ 'folktale'
-kебiar kебyar / 'flash (vi)'
buaya /bowaya/ 'crocodile'
biaung /byaug/ 'k.o. plant'

Such morphemes should thus be analysed as disyllabic. Evidence for this analysis comes from vowel allophony, reduplication processes, syllabification tests, and from phonotactics.

1. Vowel allophony:

The vowel /e/ is realised as [el] before [i] in a following syllable, independently of the number of intervening consonants: see the allophony rule (51), eg setra 'cemetery' /se'tra/ [setra]. In forms such as délia and wésia, the segment represented by orthographic \(<i>\) behaves like a consonant:

\[(21)\] <délia> /de'tyə/ [de'ty3] 'k. o. deity (Skt)'
* [de'ti3]

<wésia> /wesya/ [wesy3] 'waisya caste (Skt)'
* [wesi3]

If <délia> had the structure /de'tia/, we would expect it to surface as * [de'ti3], since /e/ is realised as [e] before [i]: Kedi [kedil] (personal name), and examples (54).

2. Reduplication: the morpheme \{pe\} /pa/ '(do) repeatedly' occurs with (sometimes optional) reduplication of the final syllable of the stem:

\[(22)\] ketel pe-ketēl-tel 'drip'
ceptul pe-ceptul-bul 'repeatedly drip'

cetebug pe-cetebug-bug 'sound of object falling to the ground'
'id. (of many things)'

With morphemes of shape \(<CVCVVC>\) in the standard orthography, the whole \(<CVVC>\) sequence is reduplicated, indicating it constitutes a single syllable of structure CGVC:

\[(23)\] -kebiar kебyar / 'flash (v)'
pekebiarbiar [pekебyarbiyar] 'id. (of many things)'

-cebuah /cебwah/ [cебwah] 'shimmer'
pecebuahbuah [пеcебwahbueb] 'id. (of many things)'

11 Kersten (1984:15) seems also to prefer the glide analysis proposed here, though the wording is unclear. This analysis also seems to apply in other Western Indonesian languages, such as Indonesian.
3. In syllabification tests, such bases occurring with the morpheme \{pe\} are invariably syllabified as though the first \(<V>\) is a syllable-internal glide: \texttt{ke.byar, *ke.bi.ar, pe.ke.byar.byar, ce.bwah pe.ce.bwah.bwah * pe.ce.ah.ah.bu.ah}.

Inconsistent results are obtained with similar forms not occurring with \{pe\}: \texttt{sisia 'student, follower (Skt)' si.sia ~ si.sya, satua 'story' sa.twaa ~ sa.tu.a}. Where the \(<VV>\) sequence occurs before \(<CV>\), the results are again mixed: \texttt{me-biayag-an 'scattered about' me.bya.ya.gan ~ me.bi.ya.ya.gan} These mixed results are probably due to influence of the orthographies: until recently the aksara orthography wrote all high vocoids occurring before vowels as glides, and now follows the the roman orthography in writing most of them as vowels.

4. Phonotactics: in disyllabic morphemes of shape /CVV(C)/ a wide variety of \(V_1V_2/\) sequences are allowed (6.3.1). We would expect the same possibilities of \(V_1V_2/\) combination to occur in 'orthographically tri-syllabic' morphemes of shape \(<CVVCV(C)>\) and \(<CVCVV(C)>\). However, the data in Warna et al (1990) shows that \(<iV>\) and \(<uV>\) orthographic \(VV\) sequences predominate by far, all other 'vowels' being illegal in the first slot, with a small number of exceptions.\(^{12}\) This indicates that \(<i>\) and \(<u>\) are here functioning not as vowels but as glides: they can occur in this position precisely because they do not cause a departure from the preferred disyllabic morpheme structure (5.6, 6.1).

On the other hand, CGV sequences do not occur in mono-syllabic morphemes (6.2.2); the imposition of disyllabic status again 'drives' surface realisation of vocoids, though this time with the reverse effect.\(^{13}\)

2.1.3.1 Epenthesis of non-phonemic glides

Non-distinctive homorganic off-glides optionally occur between \(VV\) sequences where the first vowel is one of the '[+ATR] vowels', /i/, /e/, /u/ or /o/:

\[
\begin{align*}
\text{iis} & \rightarrow [i(yl)s] \quad \text{ 'slice'} \\
\text{léak} & \rightarrow [le(yl)a?] \quad \text{ 'witch'} \\
\text{ued} & \rightarrow [u(w)ed] \quad \text{ 'origin'} \\
\text{tuak} & \rightarrow [tu(w)a?] \quad \text{ 'palm wine'} \\
\text{suud} & \rightarrow [su(w)ud] \quad \text{ 'finished'}
\end{align*}
\]

and, less often, where the [+ATR] vowel occurs second in the \(VV\) sequence. This occurs in peripheral morphemes only (see 3.1)):

\[
\begin{align*}
\text{baos} & \rightarrow [ba(w)os] \quad \text{ 'speech. Hi'} \\
\text{séos} & \rightarrow [se(w)os] \quad \text{ 'different. Hi'} \\
\text{rauh} & \rightarrow [ra(w)uh] \quad \text{ 'arrive. Hi'}
\end{align*}
\]

\(^{12}\) Warna et al (1990) list 141 'orthographical tri-syllabics' containing <VV> sequences. Of these 103 or 73% have /i/ or /u/ as the first vowel (59 /iV/, 44 /uV/). The remaining 38 <VV> (where \(V_1\) is not /i/ or /u/) I take to be 'true' trisyllabics - most are expressives, loans or dialectal variants (eg all forms containing /oV/ have <uV> forms in most dialects, including Peguyangan).

\(^{13}\) The one exception noted is the onomatopoeic \texttt{byur} 'splash!'.

19
Evidence that these are not part of the underlying representation of the morpheme includes:

1. In reduplicated forms with vowel alternation (8.7), the off-glide is not copied to the reduplicated element: 14

\[
\begin{align*}
\text{(26)} & \quad \text{diis-diis} & [dii\text{l}sdi\{y\}ts] & \text{'pant'} \\
& \quad \text{daas-diis} & [daas\text{sdi\{y\}ts}] & \text{'pant heavily'} \\
& \quad \text{uug-ang} & \{aagu\{w\}ug\text{u}g\} & \text{'in.ruins-APP, destroy} \\
& \quad \text{aag-uug-ang} & \{aawagu\{w\}ug\} & \text{'destroy randomly'}
\end{align*}
\]

2. The absence or presence of these glides is non-significant, speakers often using and accepting both forms. (This however is weaker evidence: /a/ occurring underlyingly can be freely omitted in rapid speech, see 2.2.5.)

It can be concluded that these optional glides are either non-phonemic, or if phonemic for some speakers, are added by optional late rule. (Syllabification tests give inconsistent results, e.g. JDP syllabified puus 'speech (of a god)' as pu.wus but suud 'finished' as su.ud, and kuud 'young coconut' as either ku.ud or ku.wud. Different speakers give different analyses of the same forms, within the different possibilities.)

2.1.4 Prenasalised consonants?

In the native vocabulary, the only consonant clusters allowed across syllables morph-internally (cf 6.0) are homorganic nasal-stop sequences (see also 6.2.5):

\[
\begin{align*}
\text{(27)} & \quad \text{amba}h & \text{'make one's way'} \\
& \quad \text{muntis} & \text{'pomelo'} \\
& \quad \text{kaficing} & \text{'button'} \\
& \quad \text{angeh} & \text{'pay by credit'}
\end{align*}
\]

or, where the second syllable begins with /s/, /s/:

\[
\begin{align*}
\text{(28)} & \quad \text{dangsek} & /\text{dan}sak/ & \text{'shove'} \\
& \quad \text{bungsil} & /\text{bu}ngsil/ & \text{'coconut bud'}
\end{align*}
\]

I take these sequences to be two distinct phonemes, rather than a single pre-nasalised stop (which they apparently are in the related language Javanese, cf Clynes 1995).

The pre-nasalised analysis would remove two apparent anomalies of the 'bi-phonemic' one:

---

14 The lexical dyad mu{w)ah 'and (LOW style)', and its high-style alternant miwah suggests that such off-glides do have phonemic status, since the HI form, historically derived from the LOW, retains the /w/ segment. Both members of the dyad were in fact borrowed from Old Javanese (Clynes 1989), hence do not constitute counter-evidence to the analysis here.
1. /p/, like all laminal stops, may not close a syllable morpheme-finally, but apparently does so morpheme-internally, thus:

(29)  
banjar  'village organisation'
ěńceh  'urine'

2. The rule dispreferring the occurrence of two consonants of the same place of articulation within the same morpheme 6.2 must under the bi-phonemic analysis, (apparently) carry the qualification 'unless they occur in a cluster'.

On the other hand, a pre-nasalised analysis would have to account in a principled way for the apparently non-homorganic /ŋs/ sequence (see also below).15

The evidence of vowel allophony is inconclusive, since in this dialect the relevant allophony rules (2.2.2, 2.2.3) apply irrespective of whether the syllables preceding such <NC> sequences are to be considered closed or open. Compare:

(30)  
buku  [buku]  'book'
sungkan  [sunŋkan]  'sick.HON'
Bukti  [bukti]  [place name, Skt]
songko  [səŋkō]  'hat'
toko  [tōkō]  'shop'
tongos  [tōŋōs]  'place'
jongkok  [jōŋkōʔ]  'squat'

There is moreover no unambiguous evidence that these sequences occur syllable-initially, since they are never found morpheme- or word-initially, only morpheme-internally after a vowel. Morphemes of orthographic shape <eNCV>, like embok 'elder sister', ento 'that' are often realised without a perceptible initial schwa vowel: [mbōʔ], [ntō], however the initial nasal segment can be considered to be syllabic here. Moreover, given the overwhelmingly disyllabic nature of lexical morphemes (6.1), if these were underlyingly /NCVC../ we would expect most to have the shape /NCVCV(C)/, graphemically <eNCVCV(C)>. Warna et al (1990) lists no morphemes of this shape, but many of type <eNC(L)V(C)>, suggesting that the initial schwa is phonemic here and is optionally deleted (see 2.2.5).

Finally, in syllabification tests, speakers always treat these sequences as though they were composed of two units: sampun 'already' [sam.pun], inggih 'yes (HI)' [ĩŋ.gih], and so on, though this may be due to the influence of both the aksara and modern orthographies.

In short, the available evidence from surface allophony and phonotactic patterns is inconclusive, while speaker intuitions go against a pre-nasalised analysis. In other dialects, vowel allophony patterns suggest that the syllables before such clusters are treated as closed, favouring the bi-segmental analysis (see eg Jendra 1976). Analysing

---

15Such an account is in any case required for /ŋs/ sequences in Javanese.
these sequences pre-nasalised phonemes would give a larger phoneme inventory (requiring nine extra phonemes, including /ŋs/), but a simpler phonotactic structure.

The disparities in distribution of the palatal stops - /c, y, j/ never occur morpheme-finally, though /p/ does occur syllable-finally, only in such clusters - are just what we would expect if the nasal consonant is a distinct segment, but unspecified for place of articulation, gaining its laminal place of articulation by a spreading rule (cf Goldsmith 1990:129-131 on this as a common constraint cross-linguistically). In the absence of strong evidence for unitary pre-nasalised segments, I adopt the bi-segmental analysis, and assume the following assimilation rule (discussed more fully in 7.2):

(31) \[
\begin{array}{c c c c c}
\text{C}_1 & \cdot & \cdot & \cdot & \cdot \\
\text{[+nasal]} & \text{[place]} & | & \text{[place]} & | \\
\text{[+son]} & | & | & | & \\
\end{array}
\]

To account for the aberrant /ŋs/ rather than non-occurring */ŋs/ clusters (e.g. gangsa 'k.o. metallophone', mangsit 'smell of urine'), I assume the following negative constraint on well-formedness:16

(32) *\[
\begin{array}{c c c c c}
\text{[laminal]} & \text{[+cont]} \\
\text{C}_1 & \text{C}_2 & | \\
\end{array}
\]

Where \(C_2\) is /s/, this filter rejects the output of (31). Instead, /N/ surfaces as /ŋ/ by the following default rule:

(33) \[C_1 [\text{[+nasal]}] > /ŋ/\]

Both (31) and (33) are independently required to account for the allomorphy of the agentive prefix, '/N/-', see 7.2.

2.1.5 Phonetic reduction of nasal-stop clusters

As Court (1970) reports for other Western Indonesian languages, in rapid speech the oral stop following a nasal tends to be lost:

(34) bañjar \(\approx\) [bɐnˈjø] 'village organisation'
anggon \(\approx\) [æŋˈɡø] 'use'

\[16\text{Selkirk (1990, reported in Inkelas & Cho 1993) claims that nasals never assimilate in place to fricatives. Inkelas & Cho cite languages which show exceptions to this claim, while agreeing that it is an infrequent occurrence. This proposed filter may therefore have some motivation, at least as a preference, in universal grammar.}\]
This lenition seems to be restricted to voiced oral stops. The lenited form in such pairs is not used in careful speech, though occasionally the process has apparently given rise to true doublets:

(35) \textit{kamben, kamen} \quad 'loincloth'

The form \textit{lumbur} alongside \textit{lumur} 'glass' (from Dutch \textit{roemer} 'rummer') is apparently a backformation deriving by analogy from such doublets.

### 2.2 Vowel phonemes

There are six underlying vowel phonemes, shown in Table 2.2:

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>Mid-high</td>
<td>e</td>
<td>ø</td>
<td>o</td>
</tr>
<tr>
<td>Low</td>
<td>ø</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2 Vowel phonemes of Balinese

As well, evidence for a further, underspecified, vowel, contrasting with underlying /æ/, is given in 2.2.5. The case for assigning phoneme status to /æ/ when it occurs in \textit{næ}, [næ], an allomorph of the third person pronominal is discussed in 2.2.3.

A minimal set which illustrates these phonemes in contrast is:

(36) /bik/ 'aunt' /bubuk/ 'powder'
     /bebek/ 'duck' /bobok/ 'palm-leaf torch'
     /babek/ 'k.o. worm' /babak/ 'grazed (of skin)'

The individual vowel phonemes are assumed to bear the following marking values of phonological features, prior to the operation of rules of allophony: \(^{17}\)

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>e</th>
<th>ø</th>
<th>a</th>
<th>ð</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Low</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Back</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>[ATR]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

\(^{17}\) The question of to what degree the underlying representations of these phonemes are specified for these features (cf Archangeli (1988), Mohanan (1991)) is not addressed in this thesis.
All vowels occur in all positions, morpheme-internally, except /a/ which cannot occur in a final open syllable, underlyingly. On the other hand, /a/ may occur phonological-word finally due to morphophonemic processes. The distribution of /a/ is the inverse of that of /a/ : /a/ does occur morpheme-finally underlyingly, but cannot occur phonological-word-finally: /a/ always replaces %a% when it occurs phonological-word-finally (see 7.5):

(37)  
\[ \text{lima} \quad %\text{lima}% \quad /\text{lime}/ \quad '\text{hand}' \]
\[ \text{lima-} \quad /\text{limane}/ \quad '\text{hand-DEF}' \]

Each syllable contains only one vowel. Phonetically long vowel sequences occur eg /luu /luu/ [luu] 'rubbish', /oot /oot/ [dőt] 'rice husks'. There are apparent 'minimal pairs' where a difference in vowel length corresponds to a meaning difference:

(38)  
\[ \text{adéng} \quad [\text{adeŋ}] \quad '\text{slow-HI}' \]
\[ \text{dééng} \quad [\text{deŋ}] \quad '\text{member of a procession}' \]
\[ \text{cong} \quad [\text{cɔŋ}] \quad '\text{card game}' \]
\[ \text{coong} \quad [\text{cɔ̃ŋ}] \quad '\text{wan, pale}' \]
\[ \text{bang} \quad [\text{baŋ}] \quad '\text{bank}' \]
\[ \text{baang} \quad [\text{baŋ}] \quad '\text{give}' \]

However, speakers consistently syllabify such sequences as disyllables. The following confirms that they represent two distinct phoneme segments, rather than phonemic long vowels:

1. In the process of final-vowel lowering when morphemes cooccur with the suffix -n 'GEN' or clitic -né '3rd person' (7.4), only the second vowel in such sequences is lowered: 19

(39)  
\[ \text{jempiit} \quad [\text{jampiit}] \quad 'k. o. bat' \]
\[ \text{jempiit-} \quad /\text{jampiit-ne}/ \quad [\text{jampiitne}] \quad 'his/her/their bat' \]
\[ \text{kepii} \quad [\text{kəpii}] \quad 'k. o. bug' \]
\[ \text{kepii-} \quad /\text{kepiit-ne}/ \quad [\text{kepiitne}] \quad 'bug-GEN-3, his bug' \]
\[ \text{luu} \quad [\text{luu}] \quad 'rubbish' \]
\[ \text{luu-} \quad /\text{luo-n-ne}/ \quad [\text{luo̯ne}] \quad 'his/her rubbish' \]

2. \{N\} allomorphy: the morpheme \{N\} 'actor pivot' has two distinct allomorph sets (7.2), varying with the number of syllables in the verb root to which it attaches. Roots of structure /CVV(C)/ never take the allomorph which occurs with unambiguously mono-syllabic roots, /ŋa/. Compare:

18 'Central' corresponds to Catford's (1977) 'antero-dorsal' and 'back' to his 'postero-dorsal'.
19 This apparent violation of OCP constraints is further discussed in 7.4.
3. Where stress (which falls on the final syllable, see 5.8) falls on such a sequence, it is always on the second vowel:

(41) jempiit 'k.o bat'  
     [jempi’it]  
     *[jem’pi1t]  

suud 'finished'  
     [su’ud]  
     *[suud]  

On the other hand, both vowels are 'replaced' in reduplication plus vowel overwrite (8.7), suggesting they constitute a single unit in some way:

(42) seeb-seeb %RED_v-seeb %  'peer (vi)'  
    saab-seeb 'peer this way and that'  
    uug-ang 'be.in.ruins-APP'  
    aag-uug-ang 'destroy randomly'  

It turns out however that this is the case, even where there is an intervening consonant (kebarat-kebirit 'run about wildly), so that this overwrite process does not constitute evidence for phonemic vowel length. Similarly, alternations such as the following are not evidence of phonemic long vowels (see also 7.5):

(43) paa %paa% /paæ/  'thigh'  
    paa-né /paane/  'thigh-DEF'  

2.2.1 Vowel allophony

Vowel phonemes can be divided into various sub-classes according to their allophonic behaviour. Historically, there are two groupings: the four 'oldest' phonemes, /i, u, a/ and /a/ (those reconstructed for proto-Austronesian) show less complex variation than the more recently acquired /e/ and /o/. The high vowels (/i, u/) in turn pattern together as a sub-class, as do the central vowels, /a/ and /æ/. The possibility of giving separate morpheme status to the phone [e], when it occurs in one morpho-syntactic context only, is discussed in 2.2.3.

All vowels can have phonetically longer realisations when occurring in a stressed open syllable:

---

20 An autosegmental representation where sequences of identical vowels are linked underlyingly at the melody level provides a simple explanation for the facts of (42) and (43).
The length difference is particularly noticeable when both vowels of a morpheme are /a/, the first unstressed /a/ often being elided, while the final one remains long:

(45) **teka** [takː:] ~ [tkː:] 'arrive'

The high vowels /i, u/ have 'lower' ([i, u1] and 'higher' allophones ([i, u]). The lower allophones sound closer to their higher alternants than do the English counterparts. The mid, non-central vowels also have lower ([e, ð]) and higher ([e, ð]) allophones.

The central vowel /a/ varies from [æ] to a mid-high central vowel [3] ~ [i]. /a/ has mid [a] to back [α] allophones, the back allophone being rounded.

This allophone distribution can be represented as in Table 2.3, with dotted lines grouping members of a single phoneme:

<table>
<thead>
<tr>
<th></th>
<th>high</th>
<th>mid</th>
<th>low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>i</td>
<td>e</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>u</td>
<td>3</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>i</td>
<td>ø</td>
<td>a</td>
</tr>
</tbody>
</table>

Table 2.3 Main allophones of Balinese vowel phonemes

Rules of vowel allophony are discussed and exemplified in the following sub-sections. The main rules are summarised here, with the numbering they have in their respective sub-sections:

1. A rule affecting 'non-central' vowels /i/, /u/, /e/ and /o/:

   (47) \( V_{[+ATR]} > [-ATR] / \_C\_p-word \)

2. Rules affecting only /e/ and /o/:

   (51) \( V_{[-HI, +ATR]} > [-ATR] / \_C_0 V_{[+HI, +ATR]} \)

   (52) \( V_{[-HI, +ATR]} > [-ATR] / \_C_0 V_{[-HI, -ATR]} \)

3. A rule affecting /a/:

   (65) \( V_{[-HI, -ATR, -LOW]} > [+ATR] / \_p-word, \)

   \( / \_A \)
As well, semantically motivated allophony of /a/ is discussed in 2.2.8.

Feature values of surface allophones after the operation of postlexical rules are:

(46) 

<table>
<thead>
<tr>
<th>i</th>
<th>e</th>
<th>ə</th>
<th>a</th>
<th>o</th>
<th>a</th>
<th>o</th>
</tr>
</thead>
<tbody>
<tr>
<td>[high]</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[low]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>[back]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>[ATR]</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

2.2.2 A rule targeting /i/, /u/, /e/ and /o/ 

The following rule applies both to /i/ and /u/ and to their non-high counterparts, /e/ and /o/: 

(47) \[ V[+ATR] > [-ATR] \] 

When occurring in a phonological-word-final closed syllable, high vowels are realised as [i] and [u], mid-vowels as [e] and [o].

(48) cenik [kanik] ‘small’
    pucil [puclip] ‘bud’
    kuluk [kuluk] ‘dog’
    intuk [intuk] ‘pound’
    tompelor [tompelor] ‘birthmark’
    bedeg [bedeg] ‘bamboo wall’
    tegor [tegor] ‘bite’
    tongos [tongos] ‘place’

Apparent exceptions to rule (47), involving ‘expressive’ meanings, are produced by a meaningful postlexical rule described in 2.2.8. In all other contexts, /i/ and /u/ are realised as [i] and [u] respectively, independently of factors such as stress:

(49) jagur [ja’gur] ‘punch’
jagur = a [ja’gur3] ‘punch-3ACT’
jagur-in [jagur’in] ‘punch-LOC’
jagur-in = a [jagur’in3] ‘punch-LOC-3’

or the nature (closed or open) of the syllable:

---

21 The two realisations of /a/ differ in roundedness, [a] being [+round]; [ə] is distinguished from [ø] in the same way.
In this last respect, /i/ and /u/ differ from /e/ and /o/, as sub-section 2.2.3 shows.

### 2.2.3 Allophony of the mid, non-central vowels /e, o/

The mid vowels /e/ and /o/ have more complex allophony than the four others. The occurrence of the marked (lower/-ATR) allophones is as follows. The first rule applies also to /i/ and /u/ and is repeated from (47) above:

\[(47) \ V_{[+ATR]} > [-ATR] \quad / \quad _C_l\text{-}p\text{-}w o r d\]

The second and third rules, (51) and (52), have the same effect of causing the 'marked' [-ATR] allophones of /e/ and /o/ to surface:

\[(51) \ V_{[-HI, +ATR]} > [-ATR] \quad / \quad _C_o \quad V_{[+HI, +ATR]}\]

\[(52) \ V_{[-HI, +ATR]} > [-ATR] \quad / \quad _C_o \quad V_{[-HI, -ATR]}\]

The parallel environments of rules (51) and (52) invites a restatement as a single rule:

\[(53) \ V_{[-HI, +ATR]} > [-ATR] \quad / \quad _C_o \quad V_{[a HI, a ATR]}\]

'/e/ and /o/ are realised as [e] and [ø] when occurring before either [+ ATR] realisations of high vowels or [- ATR] realisations of non-high vowels, in the immediately following syllable.'

I have chosen not to do this, since, though the same net result is achieved in each case, the motivations appear to be different (though complimentary): dissimilation away from a following high vowel in (51), and assimilation to a following non-high vowel in (52).

Rule (51) then is one of 'ATR-dissimilation', /e/ and /o/ being realised as [e] and [ø] before [+ ATR] realisations of /i/ and /u/ (on the loanwords kopi and roti see also chapter 3, footnote 5):

\[(54) \ \text{lécî(né)} \ [lëci\{ne\}] \ 'lychee(-DEF)' \]

\[\text{kéni} \ [këni] \ '({personal name})' \]

\[\text{kopi(né)} \ [kôpi\{ne\}] \ 'coffee(-DEF)' \]

\[\text{roti} \ [ròti] \ 'bread'\]
When occurring before [-ATR] allophones of /i/ and /u/ in the same environment, the unmarked values surface:

(55) séh-in [sein] 'change-LOC'
croring [cro'rın] 'kind of fruit'
éling [eilı̃] 'remember.HON'
oih [olı̃] 'by.HI'

Rule (52) produces a kind of ATR-harmony, when /e/ and /o/ occur before [-ATR] realisations of non-high vowels (ie, before [e, ø, ø] or [a]) in the following syllable:

(56) tongos [tό̊gό̊s] 'place'
léak(-é) [lěak(ø)] 'witch(-DEF)'
délem(-é) [dělem(ø)] 'personal name(-DEF)'
sénter-é [sentere] 'torch-DEF'
méong [me̩ŋ] 'cat'
déwa-né [dewane] 'god-DEF'

When occurring before [+ATR] non-high phones, the unmarked values again surface:

(57) tongos-é [tό̊gό̊s-e] 'place-DEF'
méong-é [me̩ŋ-e] 'cat-DEF'
déwa [dewa] '(caste title)'
gésgés=a [gesges̩] 'scratched-3'

The two rules lead to 'conflicting' effects, according to whether the triggering vowels are [+ATR/+HI] or [-ATR/-HI], producing alternations when suffixes and clitics are added to bases:

22 That /e/ is realised as [e] rather than [ø], before /a/ in contexts such as léak-é, shows that /a/ does assimilate to the [+ATR] of a following vowel. The same observation applies to /a/, cf délem-é. The two central vowels thus behave as a natural class in this respect.

23 Variants with [-ATR] allophones such as the following are accepted in elicitation, at least among speakers aged under around 50 years old:

gésgés=a [gesges̩] ~ [gesges̩] 'scratched-3ACT'
corcor =a [corcor̩] ~ [corcor̩] 'poured.water.on-3ACT'
adol =a [adol̩] ~ [adol̩] 'sold-3ACT'

gésgés =a [gesges̩] ~ [gesges̩] 'scratched-3ACT'
To my knowledge, rules (47), (51) and (52) are sufficient to describe the realisation of /o/ and /e/ in all environments except one. The morpheme, né, realised as /n/, is a third person pronoun occurring as an enclitic to noun heads (and preceded by the {GEN} 'genitive' morpheme, /n/, where the noun ends in a vowel):

(59) peka-k-né [pékak-ne] 'his/her grandfather'
désa-n-né [desanne] 'village-GEN-3'

This is the only case where a [-A TRJ vowel occurs word-finally in Balinese. One possible analysis then is that [e] here represents a distinct phoneme, of extremely limited distribution. I have assumed instead that there is morphological conditioning (7.4). Nevertheless, in distinguishing pairs such as:

(60) mémé-n-né [memennel] 'mother-GEN-3, his mother'
mémé-n-né [memennel] 'mother-GEN-3, his mother'

vowel height provides a major clue to the meaning difference. This could suggest that [e] is on the verge of acquiring phoneme status in this variety of Balinese. While people in their thirties and above generally commented on the presence of two /n/’s in this environment, some younger speakers seemed occasionally to reduce the surface geminate in forms such as [memennel] to a single [n], and in elicitation some even said that there was ‘only one n’. On the other hand, other variant realisations among young people allow free interchange of [e] and [e], and of [ó] and [ô], see footnote (23).

2.2.4 Allophony of the central vowels, /a/ and /a/

The mid vowel, schwa, /a/ has two principal allophones, [a] the unmarked allophone, occurring in most contexts, and [e], a central mid-high vowel, occurring in two main environments:

1. phonological word-finally in open syllables.

24 The 'definite' morpheme has realisations /e/ after a consonant, and /ne/ after a vowel. Third-person -né is invariant, but is always preceded by a consonant. The stress patterns of both words are identical (5.8).

25 That [a] is not the expected allophone word-finally was demonstrated clearly when two speakers independently responded in the same way to my attempts to pronounce meka /meke/ 'mirror' as [meke], saying that it sounded as though there was an /h/ at the end of the word.
The low vowel /a/ generally has a central-to-back realisation, [ə]:

(66) *katak* [katə?] ‘frog’
*baat* [baat] ‘heavy’
*gati* [gati] ‘very’

An optional rule, adding expressive meaning, produces the backed, rounded allophone, [al] (2.2.8).

In the analysis adopted above underlying /a/ goes to /ə/ ([ə]) P-word-finally (7.5). A strong case can nonetheless be made for an alternative analysis, where there is no alternation, and [ə] realises final /a/, not /ə/, (6.3.2). I have chosen the phonetically more natural solution, rather than the phonologically more parsimonious, and even more explanatory, one.

2.2.5 Deletions of underlying schwa

Recent analyses of related languages have proposed that the vowel ə does not occur underlyingly (Schlindwein 1988, on (East) Javanese, cited in Archangeli & Pulleyblank...
1994; Cohn 1989, on Indonesian). That \( \text{a} \) does occur underlingly in Balinese is indicated by the following:

1. As discussed in 8.7, when reduplication plus vowel overwrite applies to a base with identical vowels, both are replaced with /a/:

\[(67) \quad \text{srandang-srendeng} \quad \text{‘reel from side to side’} \\
\text{daas-diis} \quad \text{‘sigh repeatedly’}\]

Forms where only the final vowel is replaced are ungrammatical:

\[(68) \quad *\text{srendeng-srendeng} \\
*\text{grudag-grudug}\]

Only where the base contains two /a/ vowels, at least on the surface, is the situation more complicated. With some bases, only the final vowel of the base can be overwritten:

\[(69) \\
\text{-celep} \quad \text{‘enter’} \\
\text{-dlekep} \quad \text{‘press one’s face against sthg’} \\
\text{-gremek} \quad \text{‘mumble’} \\
\text{-keblet} \quad \text{‘flicker’} \\
\text{-gremeng} \quad \text{‘obscured’} \\
\text{-kremek} \quad \text{‘feel, squeeze’} \\
\text{-krebek} \quad \text{‘speak unclearly’} \]

However, with others, both vowels must be overwritten:

\[(70) \\
\text{-gremeng} \quad \text{‘obscured’} \\
\text{-kremek} \quad \text{‘speak haltingly’} \\
\text{-gremet} \quad \text{‘feel, squeeze’} \\
\text{-krebek} \quad \text{‘speak unclearly’} \\
\text{gramang-gremeng} \quad \text{‘speak haltingly’} \\
\text{kramak-kremek} \quad \text{‘speak haltingly’} \\
\text{gramat-gremet} \quad \text{‘speak unclearly’} \\
\text{krebak-krek} \quad \text{‘speak unclearly’} \]

These contrasting patterns can be accounted for if (i) the penultimate vowel in the bases in (69) is underspecified underlingly, bearing just the feature [+syllabic], while the final vowel is specified as /a/, and (ii) bases in (70) are specified for /a/ underlingly. This account then requires a distinction between underlying /a/ and an underspecified vowel in the ‘phoneme’ inventory.

2. Where morphemes containing /a/ occur with {N} ‘actor pivot’ the allomorph which occurs is that associated with disyllabic (or longer) roots, not that occurring with monosyllabics, \( \text{nge-} \), cf 7.2:

\[(71) \\
\text{belus} \quad \text{‘wet’} \\
\text{melusang} \quad \%\text{N-belus-APP}\% \quad \text{‘wet (vt)’} \\
\text{*ngebelusang} \\
\text{beler} \quad \text{‘misbehave’} \\
\text{memeler} \quad \%\text{N-belar}\% \quad \text{‘misbehave’} \\
\text{*ngebler} \\
\text{teka} \quad \text{‘come, arrive’} \\
\text{nekain} \quad \%\text{N-teka in}\% \quad \text{‘visit (vt, AP)’} \\
\text{*ngetkain} \]

The crucial element in \{N\}-allomorphy is the number of syllables in the root, rather than in the stem (7.2): compare monosyllabic cét ‘paint’ (noun) and the derived verb cét-in
'paint-LOC; paint (vt), with Actor Pivot form nge-cét-in (not *nyétin). /a/ is thus present at a very early stage in the derivation of lexemes such as melusang and memeler, otherwise *ngeblusang and *ngebler would be generated.  

3. To assume [a] is not present underlyingly in forms such as belus, and teka would greatly complicate, and render implausible, the description of syllable phonotactics, allowing virtually any consonant pairing to occur in initial clusters, underlyingly, but a far more restricted set on the surface. Absence of schwa underlyingly would go against the conclusion reached elsewhere that underlying forms are subject to the same constraints of prosodic well-formedness as surface forms (5.1, 6.1.1).

4. Syllabification tests invariably give a two syllable analysis. In the following subsection it is argued that true epenthetic vowels are inconsistently syllabified, even by the same speaker.

Underlying /a/ may be optionally deleted in the first (unstressed) syllable of disyllabic (or longer) roots:

(72) beler [b(a)lar] 'badly behaved' 
berek [b(a)ra?] 'rotten' 
peték [p(a)te?] 'count (vt)' 
enduk [t(a)ndu?] 'feeble, limp' 
emba [(a)mb3] 'fried onion'

The following subsection deals with the opposite phenomenon, optional epenthesis of non-underlying schwa.

2.2.6 Non-phonemic schwa: optional epenthesis

Non-distinctive [a] optionally breaks up consonant-liquid clusters in the first syllable of a disyllabic (or longer) morpheme:

(73) klangsah [klaosah] ~ [kelaosah] 'woven coconut leaf' 
clagi [calagi] ~ [ca.lagi] 'tamarind seed' 
crukcuk [carukcuk] ~ [carukcuk] 'k.o. bird'

In syllabification tests, the same person will give or accept two different analyses: thus KMP accepted the following as equally good: croring 'k.o. fruit' [ca.ro.rrol] ~ [cro.rrol]; plawa 'shrub leaves used to make offerings' [pa.la.w3] ~ [pla.w3].

I take [a] here to be non-distinctive, based on the evidence of syllabification, and on the evidence from well-formedness constraints. For example, where trisyllabic morphemes contain an NC cluster (jempiring 'gardenia', sengguu 'k.o. priest'), it invariably occurs at the juncture of the ante-penultimate and penultimate syllables (6.2.3). To give [a] 

26 In the Lexical Phonology framework, if the underlying forms are %blus% and %blr% (ie if all schwas are absent in URs), this would require the introduction of a non-distinctive segment during the lexical phonology, violating the the principle of Structure Preservation (Kiparsky 1985).
phoneme status would be to create a whole class of exceptions to this statement. It
would also add significantly to the numbers of (dispreferred) trisyllabic morphemes in
the lexicon.

If \( \text{l} \) does occur underlyingly here the reduplication pattern associated with morphemes
like ‘*cerucuk’ would be bizarre: reduplication would copy only the initial consonant,
and the final syllable rhyme of underlying ‘*%ceruk\%', and suffix it to the base. This
pattern would implausibly apply only to morphemes containing \( \text{l} \text{r} \) and \( \text{l} \text{ll} \) in the initial
syllable. Strong evidence that morphemes like cruucuk derive from underlying
%crucuk\% is given in 8.8.

2.2.7 Spreading nasalisation of vowels

All vowels are non-distinctively nasalised after a nasal consonant:

(74) \text{nyuh} [n\text{u}h] 'coconut'
\text{nyama-n-né} [n\text{äm}m\text{än}n\text{é}] 'sibling-GEN-3, his family'

Nasality spreads across a sequence of vowels following a nasal:

(75) \text{nyai} [n\text{a}i] '2 PRO FEM'
\text{ngaë} [n\text{a}ë] 'make (AP)'

This progressive spread of nasality appears to be similar to that reported by Robins
(1957) and Cohn (1990) for Sundanese, and described as an areal feature by Court
(1970). Nasality also appears to spread regressively,

(76) \text{baang} [b\text{a}a\text{n}] 'give'
\text{kacang} [k\text{a}c\text{a}ŋ] 'bean, legume'

Instrumental data were not collected, and conditions on the spreading and blocking of
nasal spreading were not investigated in detail.

2.2.8 Meaningful phonological processes 2: vowels

As with consonants (2.1.2), vowel allophony in some cases is manipulated to convey
pragmatic meaning. Rule (47), repeated as (77), is regularly 'broken' where vowels of
type \( V_{ [+\text{HI}, +\text{ATR}]} \) occur in onomatopoeia, as in (77a):

(77) \( V_{ [+\text{HI}, +\text{ATR}]} > [\text{-ATR}] \_\text{Clp-word} \)

a. \text{byur} [byur] 'splash!'
\text{tin} [ti(:)n] 'sound of car horn'
That these 'exceptions' occur in a purely semantic environment indicates that they are better treated as products of a meaningful, but purely phonological, rule.27

\[(78) \quad V_{[+HI,-ATR]} \Rightarrow [+ATR] \quad /_CIP\text{-word}\]

Meaning: 'adds expressive meaning'

A second meaningful phonological rule realises /a/ with a rounded allophone, [a], again in 'expressive' uses only. Contrast matah [matah] 'raw', and its figurative use [matoh] 'pathetic!'. Other examples:

\[(79) \quad \begin{array}{ll}
gedé\ gati! & \text{[gedega(:)ti]} \quad 'big\ very; (it's)\ very\ big' 
baat! & \text{[baat]} \quad '(it's\ really)\ heavy!' 
(s)aja-an! & \text{[(s)ajaan]} \quad 'really!^{28}\end{array}\]

In 3.4.2 I attempt to formulate the kind of meaning added by meaningful phonological rules.

---

27| I have formulated this assuming that rule (47) applies normally, but is then overridden. Uhlenbeck (1978) describes a parallel rule in Javanese.

28| The last two examples show how two /a/'s (like any two identical non-high phonemes) have the same allophony when strictly contiguous. I have transcribed unstressed initial /a/ in examples like [matoh] 'pathetic!' with the usual phonetic value [a]; it is possible that here too the backed/rounded allophony occurs in expressive contexts; this needs further confirmation.
3.0 Phonology and semantics: introduction

In this chapter my principal aim is to show that in Balinese semantics interacts with phonology, in regular, productive ways. This interaction influences not only allophony patterns (2.1.2, 2.2.8), but also motivates 'exceptions' to a variety of phonotactic constraints, at the level of the syllable, and above. Since I make this claim repeatedly in subsequent chapters, a demonstration of the correctness of my position is therefore required.

This claim runs counter to the generally accepted view of the relationship between semantics and phonology - that there is, and can be, no interaction between the two (Archangeli & Pulleyblank 1994:5, 433). The standard view is often referred to as the Double Articulation principle: language has a double articulation into form (sound) and content (meaning or function), with a strict separation between the two.

Over the years a significant minority of linguists, including Trubetzkoy (1969 [1939]), Hymes (1974), Bolinger (1977), Jakobson and Waugh (1979), Woodbury (1987), and Hinton et al (1994), have argued in various ways against this position. Woodbury, for example, presents strong evidence that postlexical phonological processes are used to signal 'linguistically significant expressive or other pragmatic meaning' (1987:685). He therefore proposes a theory of 'Revised Double Articulation' (RDA) (1987:690), whose central tenet is that:

(1) Optional rules of postlexical phonology can bear meaning

In Balinese a variety of semantically motivated postlexical rules are found (see 2.1.2, 2.2.8). In general, they do encode the types of expressive/pragmatic meanings Woodbury predicts will be associated with them. Woodbury's RDA proposal is however limited to postlexical phonological processes. In this chapter I show that, in Balinese, semantics also interacts with phonology at the 'pre-lexical' level, that of constraints on underlying phonological representations, such as syllable- and morpheme-structure constraints. (Semantic effects at this level are typically referred to

1 'Post lexical processes...other pragmatic, rather than strictly referential meaning. Among them we should find expressers of emphasis or degree [...]; information-organizing notions in sentences and discourse (e.g. topicality, focus, presupposition and assertion); affective meaning; indexation and manipulation of social categories pertaining both to the speaker and to others in the speech situation [...]; and disguised speech or play-language forms whose use is optional and socially meaningful.' (Woodbury, 1987:692). On the distinction between lexical and postlexical phonology, see 2.0.

2 Woodbury is not categorical on this issue. Balinese derivational reduplication is one (often) postlexical process which signals non-expressive meaning (see though 8.9).
as 'sound symbolism'; I do not use this term as it suggests a narrower range of semantic effects than is actually found (see 3.2.2.)

If such an interaction can be demonstrated to exist, this represents a further challenge to the traditional double articulation view. At the same time, I believe that the Balinese case is important in that it demonstrates or supports the following views:

1) In word formation processes, phonology interacts with semantics in regular, predictable and productive ways: the relationship between signifiant and signifié is often not an arbitrary one, pace Saussure. (Data cited in Fudge 1970 and in Hinton et al. 1994 indicate that this may well be true for all languages.)

2) Evidence from semantics supports the view that languages permit a continuum of well-formedness in phonological representations, from prototypically well-formed, through to marginal. (Alternatively, the evidence shows that many phonological rules are 'violable', Fudge 1970.)

This 'continuum' view of well-formedness has been formalised and developed in recent unpublished papers by Smolensky, and Prince and Smolensky as 'Optimality Theory' (see for example Goldsmith 1990, Archangeli & Pulleyblank 1994). As I have not had access to the work of Prince and Smolensky, in this thesis I express this notion in terms of Preferred Rules: these specify certain structures as prototypically or optimally well-formed. Preferred Rules are violable, and hence exceptions to them occur. Rules holding without exception I refer to as Absolute Rules. My concern is to demonstrate that the phenomenon exists, and interacts productively with semantics, rather than to develop the formal notation.

3) Consequently, semantic evidence is as important in phonological analysis as it is in any other branch of linguistics. Many, if not all, Preferred Rules can probably only be identified using semantic and phonological evidence in combination (3.3).

A strength of the theory proposed in 3.4.2 and 3.4.3 is that it has explanatory value diachronically as well as synchronically: the mechanism which accounts for the link between phonological peripherality and semantics synchronically is probably also a prime motivating factor in sound change over time.

The phenomenon discussed in this chapter cannot be dismissed as 'peripheral' or marginal: like English (Fudge 1970), Balinese makes extensive, productive, use of peripheral phonological structures, to create expressive vocabulary, which moreover constitutes a major element in the lexicon. This latter fact is true not only for Balinese but, apparently, for many western Indonesian languages, having been often commented upon for Austronesian languages like Javanese (Gonda 1947) and Malay (Gonda 1947, Carr 1966). The far-fetched theories of Maxwell (cf Gonda 1940) start from the correct

---

3 I received Hinton et al (1994) too late to incorporate its findings into this study. The conclusions reached by several of its authors appear to agree with those reached here. The present study, I believe, goes further than any of the papers there, in its use of statistical tests, and in attempting to explain the correlations observed 3.4.2.
Phonology and Semantics

observation that Malay has a great many expressive words (formed from 'violations' of rules very similar to those shaping morphemes in Balinese and Javanese); see also Carr (1966) and Collins (1979) on the high percentage of expressive lexis in Malay. Eisman (1990) and Scherzer (1993) both write on the large number of Balinese expressives.

The chapter has the following structure. The claims made are outlined and exemplified in 3.1. I then discuss the (to my knowledge) only attempt at a rigorous statistical study of the phonotactics-semantics correlation: Fudge 1970, showing how my approach differs from his. Statistical and other evidence for the phonotactics-semantics correlation is presented in 3.2.3, 3.2.4, 3.2.5, and 3.2.6. Criteria for identifying preferred rules are discussed in 3.3. An explanation for this interaction is proposed in 3.4, and especially 3.4.2. The question of whether 'double articulation' as an axiom of linguistic theory is tenable is discussed in 3.4.3. Finally, in 3.5, I sum up, and draw some more general conclusions.

3.1 Phonology and semantics: examples

In this section the principal claims made about the relationship of phonology and semantics in Balinese are introduced and exemplified, before evidence in favour of them is presented in subsequent sections.

In Balinese, phonological rules are of two types: those which apply without exception, and those which hold in general, but which have exceptions, or apparent exceptions which are in fact semantically motivated (see below). Exceptionless rules include many morphophonemic processes, and some morpheme structure conditions 4.

Rules with exceptions include many morpheme structure conditions (illustrated below), and also some allophonic/postlexical rules. Since 'exceptions' of the latter type occur in purely semantic environments, I have treated them as products of distinct, meaningful rules: see 2.1.2 and 2.2.8. For example, the usual realisation of /k/, is as [k], except phonological word-finally, where it surfaces as [ʔ] (2.1). An optional meaningful phonological rule 'overrides' the usual allophony, thus:

\[(2) \quad k > [k] / \_P\text{-word}\]

Meaning: 'adds expressive meaning'

Rule (2) is discussed and illustrated in 2.1.2, together with other such rules. In addition to interactions of meaning with the postlexical phonology, Balinese shows a correlation which recurs again and again in the world's languages: where morphemes have phonologically exceptional phonotactic structures, they regularly belong to one of three broad classes. Either they are members of one of a small set of 'expressive' semantic

---

4 An example of the former type is the realisation rule for the {N} morpheme (7.2); an exceptionless morpheme structure condition is the rule barring all non-continuant laminal obstruents morpheme-finally, (6.2.4).
classes, or they are loanwords, or they are grammatical morphemes (Trubetzkoy 1969 [1939], Fudge 1970).

Balinese in fact uses peripheral phonological structures productively to coin expressive vocabulary. To give a first illustration, the following morpheme structure constraint (6.3.1) is often 'broken' for this purpose in Balinese:

(3) Prefer morphemes where cooccurring [+ATR] vowels share the same value for the feature [HI]

Rule (3) 'bars' either of the mid-vowels /o/ and /e/ from cooccurring in a morpheme with either of the high vowels /i/ and /u/ (6.3.1). Strong phonological evidence that this rule holds as a synchronic constraint in the language comes from loanwords, which are regularized. Except for éling (Javanese), the examples in (4) come, probably via Malay, from a variety of languages:

(4) Source > Balinese
    roti  ruti  'bread'
    polisi  pulisi  'police'
    kopi  kupi  'coffee'
    éling  éling ~ iling  'remember'

Other evidence that the 'regular' structure is preferred comes from productivity. A count of morphemes containing V[+ATR]CV[+ATR] sequences in Warna et al (1990) showed that regular sequences outnumber 'irregular' by 10 to 1:

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Tokens</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>hi-hi</td>
<td>1305</td>
<td>2222</td>
</tr>
<tr>
<td>mid-mid</td>
<td>917</td>
<td>regular</td>
</tr>
<tr>
<td>mid-hi</td>
<td>102</td>
<td>221</td>
</tr>
<tr>
<td>hi-mid</td>
<td>119</td>
<td>irregular</td>
</tr>
</tbody>
</table>

Of a sample of 97 morphemes containing the 'disfavoured' sequences iCé, iCo, uCo, and éCi, 84, or 87% fell into the recurrent expressive semantic classes (defined in 3.2.2) BAD, or PLANT/ANIMAL, or were NAMES, or loan morphemes. Examples of morphemes containing these disfavoured sequences, with their semantic/functional type, include:

(5) a. BAD:
    bebéki 'mischievous'
    cété  'pimp'
    ingé  'careless'
    bisén  'attack from behind'

b. PLANT AND ANIMAL names:
    bekicot 'snail'
    bligo  'kind of gourd'
    coring 'k.o. fruit'
    wéni  'k.o. fruit'

c. Personal NAMES:
    Séli, Kédi, Kéni, Maséni, Réli, Wéli,

8People below the age of about 50 no longer regularize loan words like this, no doubt due to the pervasive influence of Indonesian, which virtually all but the oldest Balinese speak as a second language.
Phonology and Semantics

d. High style/honorific terms and other loanwords
   ipén 'sibling-in-law' HON Jav.  icén 'grant' HON Jav
   biséka 'royal title' HON Skt  wirén 'because' HI Jav.

e. Grammatical words:
   tidong 'not'

This then is a first example of the correlation between phonology and semantics in pre-lexical word-formation processes: there is independent phonological evidence for the marginality of the phonological structure in question, as well as a high proportion of 'expressives' among the morphemes having that structure.

The statistical procedure used to support the latter claim is described in 3.2.3. A detailed analysis of $V_1[+ATR]C[V_1[+ATR]]$ height harmony is given in Appendix I. Other examples of the correlation can be found in 6.2.1, and 6.2.2.

Similar recurrent patterns in the related language Javanese led Uhlenbeck (1949, 1950) to distinguish between 'central' and 'peripheral' root morpheme classes. Central morphemes are phonologically totally regular, while peripheral root morphemes are both phonologically unusual in some way, and belong to recurrent semantic or functional types similar to those above: loans, dialectal words, affective-expressive, abbreviations, onomatopoeia, adhortatives, names of plants and animals, sometimes deictic and krama [High/Honorific] morphemes (1950:32).

I have followed Uhlenbeck's general descriptive approach. I use the term central morphemes to refer to the phonologically regular subset of lexical morphemes. This definition is a purely phonological one, and makes no predictions about the semantics of such morphemes (though see 3.2.3). Peripheral morphemes are those which 'break' Preferred Rules. They are either grammatical morphemes, loanwords, or lexical morphemes belonging to the semantic classes listed in 3.2.2. I use the term expressive morphemes to refer to these, the non-loanword, lexical subset of the peripheral class.

3.2 Analysis

3.2.1 Fudge's approach, and the present one

One of the few previous attempts at a systematic study of the relationship between phonotactic constraints and semantics is Fudge (1970). The rest of this sub-section compares Fudge's approach with that used here. Though his work has shortcomings (see below), he nevertheless presents a strong case for a relationship between phonological

---

6 After analysing data from Balinese, I looked for the same phenomenon in English and independently reached the same conclusions as Fudge, before becoming aware of his article. I am aware of the 'phonaestheme' hypothesis (Householder 1946 and several subsequent studies, cf Jakobson and Waugh 1979, Palmer 1992); many of the phenomena accounted for here are not explained by that approach.
markedness and 'expressiveness' in English, while pointing out that then current theories of phonology had no means of dealing with it.

Fudge argues that in English the following sequences are regularly associated with 'expressive' meanings: (i) syllable-initial consonant clusters and (ii) syllables containing homorganic consonants in both onset and coda. I show below that in Balinese too these sequences are regularly associated with 'peripheral' semantic types. Fudge assigns the label 'expressive' to a given word if it belongs to any of the following types (1970:162):

(i) Onomatopoeic words
(ii) Movement words
(iii) Affective words [including 'words denoting intense reactions (horrible, marvellous), words with a pejorative connotation (grumble, sly etc) and words with a jocular or endearing connotation (plump)']
(iv) Hypercoristic [ie 'pet' names eg Bob, Tom
(v) Nursery words eg teddy, bib
(vi) Colloquial words and taboo words
(vii) Abbreviations eg perk, rep

For a given syllable structure, Fudge tallies two values (i) Usage (U) 'the proportion of possible syllables of the type in question which are actually used in words (166)' and (ii) Restriction (R): 'the proportion of actually used syllables which are restricted to 'expressive' words.(id.)' Fudge takes the Usage figure to be a measure of the 'peripherality' or otherwise of a given structure (161). An example of Fudge's findings is given in Table 3.1, which investigates the relationship between the presence of consonant clusters within the syllable, and expressiveness:

<table>
<thead>
<tr>
<th>Value of x and y</th>
<th>Possible</th>
<th>Actual</th>
<th>U (%)</th>
<th>Restrict'd R (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>x = 0, y = 0 (no clusters)</td>
<td>3588</td>
<td>1280</td>
<td>36</td>
<td>207</td>
</tr>
<tr>
<td>x ≠ 0, y = 0 (Onset cluster only)</td>
<td>3864</td>
<td>574</td>
<td>15</td>
<td>172</td>
</tr>
<tr>
<td>x = 0, y ≠ 0 (Coda cluster only)</td>
<td>6240</td>
<td>623</td>
<td>10</td>
<td>112</td>
</tr>
<tr>
<td>x ≠ 0, y ≠ 0 (Both Onset &amp; Coda)</td>
<td>6720</td>
<td>141</td>
<td>2</td>
<td>62</td>
</tr>
</tbody>
</table>

Table 3.1 Usage and Restriction of syllables with short vowels, of the form CxVyC2. (Adapted from Fudge 1970:169)

Clearly, Fudge's figures show that as the number of clusters in a syllable increases, the value of U drops dramatically, from 36% to 2%. At the same time, the value of R increases from 16% to 44% (though a Coda cluster alone has little effect on the value of R).

Fudge presents a series of tables similar to this one, investigating various aspects of syllable structure. However, while he gives abundant statistics to support his case, there is virtually nil exemplification - to the extent that I assume he has made a deliberate decision to avoid giving examples. This, while not invalidating the methodology, makes the article less readable than it might have been. There are also no statistical tests to show that the distributions he obtains could not have been due to chance (though the repeated pattern of a decrease in U being associated with an increase in R is so regular,
from table to table, that chance effects seem ruled out). Perhaps because of these and
other (mainly minor) shortcomings (see also below), his study appears to have been
largely overlooked. I will thus illustrate the phenomenon in English, before going on to
compare the approach I have used with his.

Fudge's claim that complex syllable onsets in English are regularly associated with
morphemes with expressive meanings (1970:168) is easy to demonstrate. A good
candidate is the sequence /tw/.7 Two kinds of phonological evidence suggest that
syllable-onset /tw/ is a 'peripheral' structure in English:

1) it occurs before a much smaller set of possible nucleus vowels than other clusters
   (Gimson (1989:241)),

2) unlike most other initial CC clusters, it has no /sCC/ counterpart,

The following lists all (non-loan) words beginning with the sequence /tw/ listed in the
Macquarie dictionary (1982), and grouped by Fudge's semantic types:

(6)  bad/pejorative: twaddle, twat, twee, twerp, twinge
     onomatopoeic: twang, tweet, twitter
     movement: twaddle, twill, twiddle, twirl, twine, twist; twitch, tweak,
                twinge, tweezers
     colloquial: twig (understand)
     unclassified: twin, twice, twenty; twig, twitch (k o grass)

Note that most of the items listed in (6) as 'unclassified' by Fudge's categories fall into
semantic classes found in other studies to be regularly associated with peripheral
structure: grammatical morphemes and plant names.

Similarly, the following list of morphemes beginning with /b/ and containing a
tautomorphemic labial consonant, (taken from the Macquarie dictionary, and grouped
according to Fudge's semantic categories) clearly supports Fudge's claim that there is 'an
evident relationship between sameness of place of articulation and 'expressiveness"
(1970:172).8 The list is again exhaustive, with only loanwords excluded.

7 Another example is the onset structure /sCC/, which is highly marked cross-linguistically. Words
   beginning with /sCC/ onsets overwhelmingly fall into expressive semantic categories. For example,
   these are all the /skw/-initials in the Macquarie dictionary, classified with Fudge's categories:
   onomatopoeic: squall, squawk, squeak, squeal, squelch; pejorative: squander, squib; movement:
   squash, squat, squeegee, squeze, squirt, squash, squint, squirm; colloquial: squiggle, squitters,
   squiz; unclassified: squid, squirrel (both animal names).
8 Fudge looks only at tauto-syllabic consonants: I have included also tauto-morphemic labials in the
data. The result is the same, in fact even more categorical than Fudge's, since loanwords are
excluded in the counts.
(7) **onomatopoeics:** babble, bebop, beep, blabber, bleep, blob, blub(ber), boom, bow-wow, burble

**movement words:** bob, bobble, biff, buff ('slap'), bump

**nursery words:** bib, bub, babe

**colloquial/affective:** baffle, bamboozle, bloorp, blurt, bluff, boop ('dunce'; 'breast', 'prison'), booby, brummy, bum, bob, boffin, boomer, buff ('enthusiast'), bumm, bumble, bummer, bumkin, bumpkin, bumptious

**unclassified:** baboon, beaver, bobcat, booby (animals); bramble, broom (k.o. plant); blimp, baby (originally a diminutive of babe); beam, bloom, brim

Again, many of the items 'unclassified' by Fudge's categories fall into semantic classes regularly associated in other studies with peripheral structure. The names of plants and animals, for example, are regularly associated with aberrant phonological shape in both Balinese and the closely related Javanese (Uhlenbeck 1950).

The rest of this sub-section contrasts the approach I have taken in classifying morphemes semantically for the purposes of this analysis with that of Fudge.

Fudge discusses the difficulty of classifying morphemes in terms of expressiveness, arguing that 'expressiveness is probably a continuum, rather than a series of discrete steps' (1970:163). He relies on personal intuitions of 'expressiveness', and thus classifies slither as expressive, 'but not slip and slide'. Fudge's method, apart from being usable essentially only by analysts working on their own languages, suffers in that it establishes in advance the semantic types being looked for when analyzing phonologically marginal morphemes (items the researcher feels intuitively to be expressive), rather than examining the data, and seeing what categories are there. The latter is the approach that I have taken.9

Thus in analyzing Fudge's examples, slither, slip and slide, I would say that all three show a correlation between a marked phonological structure and a common semantic class which recurs in such structures: 'movement verbs', a subset of 'sense' terms (defined in 3.2.2).10 I would classify them all as 'expressive'. My use of this term is thus different from Fudge's, though related.

Clearly, the set of semantic types used in an analysis like this must be limited. A problem with perhaps all previous analyses is that the lists of expressive semantic types they propose are too open-ended. Fudge for example has the vague categories 'colloquial' and 'affective', while Uhlenbeck includes an 'affective-expressive' class, and Trubetzkoy (1969 [1939]:254), all "words with an 'expressive' coloration". If any

---

9 Fudge in an appendix does attempt to classify phonologically peripheral Hokkien lexemes for expressiveness, based on their English glosses. Several of the forms not classified by him as expressive nevertheless fall into cross-linguistically commonly recurring 'expressive' classes: they have glosses such as 'carry in front of bosom (SENSE), 'personal name', 'to be bruised by a fall (BAD)', 'plum flower, plum (PLANT)'.

10 /sl/-initial words in English are generally expressive: sloven (BAD), sly (BAD), slap, slog (TRUE PATIENT VERB), skir (BAD), slag (BAD), slouch (BAD/SENSE), sloth (BAD, ANIMAL), slip, slide, slurp (SENSE), slack, sleazy (BAD), and so on.
lexical item is potentially classifiable as peripheral, the analysis becomes unfalsifiable. I therefore have avoided vague classes such as 'colloquial', or the obviously circular 'expressive'/'affective'; this is particularly needed when the language being analysed is not the analyst's native tongue (not a problem in Fudge's case). And to avoid a proliferation of peripheral classes in statistical analyses (3.2.3), I first test for the numbers of morphemes classifiable into three classes only: SENSES, PLANT/ANIMAL, or BAD (defined in 3.2.2). These three classes recur in sufficiently high numbers in morphemes containing peripheral phonological structures for them to be given special diagnostic status.

A further difficulty with Fudge's analysis is that it does not distinguish loanwords or grammatical morphemes from other items in the data base: while noting that loanwords often are exceptional, Fudge includes them in the non-expressives in his totals. In my analysis, I include both these classes in the peripheral morpheme category, for the following reasons.

1. It is common, both in Balinese and cross-linguistically, for grammatical morphemes to be 'aberrant' phonologically. In Balinese, for example, there is a strong tendency for morphemes to be disyllabic: approximately 90% of the morphemes listed in Warna et al (1990) are disyllabic (6.1). At the same time, over 90% of grammatical morphemes are monosyllabic. A well known example in English is /ɔ/, which is 'illegal' morpheme-initially, except in grammatical morphemes. Initial /hw/, now archaic in most varieties, is similarly restricted.

2. It is also commonly recognized that, in the operation of phonological rules, languages regularly treat loanwords as a distinct subclass, or subclasses, in the lexicon (see section 3.4.1, also Chomsky and Halle 1968:373). The many loanwords in Balinese from Sanskrit and Javanese abound in instances of 'irregular' phonological structures.

Because of the various methodological decisions he made, Fudge's conclusions are expressed as tendencies, rather than as more tightly constrained rules. If for example Fudge had allowed for loanwords and grammatical morphemes in his calculations, his conclusions would have been more categorical. His study nevertheless remains a valuable one, since (except in allowing the vague 'colloquial' and 'affective' classes) it generally errs on the side of caution.

3.2.2 The 'expressive' semantic categories defined

A variety of semantic classes consistently recur in phonologically peripheral morphemes in Balinese. Of these, three - 'BAD', 'SENSES' 'PLANTS & ANIMALS' - are accorded special criterion status in statistical analyses (3.2.3). These and the other 'expressive' classes (3.1) are exemplified in Table 3.2, together with the particular well-formedness condition(s) 'broken' in each example. The classes are given working definitions in the paragraphs following the table.
### Table 3.2 Some recurrent semantic and phonological types associated with peripheral morphemes in Balinese

<table>
<thead>
<tr>
<th>Semantic class</th>
<th>example</th>
<th>Preferred rule 'broken'</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SENSES sound</strong></td>
<td>kedék 'laugh'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>smell</strong></td>
<td>milk 'fragrant'</td>
<td>*{CNV}</td>
</tr>
<tr>
<td><strong>taste</strong></td>
<td>manis 'sweet'</td>
<td>*{CNV}, *{Ca,Ca}</td>
</tr>
<tr>
<td><strong>touch</strong></td>
<td>dingin 'cold'</td>
<td>*{CNV}, *{Ca,Ca}</td>
</tr>
<tr>
<td><strong>sight</strong></td>
<td>-kebyar 'flash (v)'</td>
<td>*{CL}</td>
</tr>
<tr>
<td><strong>movements</strong></td>
<td>-kakeb 'lie face down'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>emotions</strong></td>
<td>gedeg 'angry'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>BAD</strong></td>
<td>beler 'behave badly'</td>
<td>*{VHI,V MID}, *{Ca,Ca}</td>
</tr>
<tr>
<td><strong>PLANT/ANIMAL</strong></td>
<td>coring 'k.o. fruit'</td>
<td>*{C,VH1,VM1D}, *{Ca,Ca}</td>
</tr>
<tr>
<td><strong>HI/HON voc</strong></td>
<td>jlati 'earthworm'</td>
<td>*{CL}, *{Ca,Ca}</td>
</tr>
<tr>
<td><strong>TRUE PAT VBS</strong></td>
<td>cingak 'see (HON)'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>GRAMMATICAL</strong></td>
<td>dados 'may (HI)'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>emotions</strong></td>
<td>gedig 'hit'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>BAD</strong></td>
<td>tigtig 'hit'</td>
<td>*{complex}</td>
</tr>
<tr>
<td><strong>PLANT/ANIMAL</strong></td>
<td>teken 'with'(prep)</td>
<td>*{Ca,Ca}, *{V}</td>
</tr>
<tr>
<td><strong>GRAMMATICAL</strong></td>
<td>lakar 'will' (tense)</td>
<td>*{Ca,Ca}, *{complex}</td>
</tr>
<tr>
<td><strong>NAMES pers. place</strong></td>
<td>Kényi, Kédí, Dodi</td>
<td>*{VHI,V MID}, *{Ca,Ca}</td>
</tr>
<tr>
<td><strong>ADDRESS terms</strong></td>
<td>Bébéting, Buléléng</td>
<td>*{VHI,V MID}, *{Ca,Ca}</td>
</tr>
<tr>
<td><strong>very</strong></td>
<td>Bapa 'father'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>BODY PARTS</strong></td>
<td>Kadék '2nd-born child'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>small: 1.CAKES, FOODS</strong></td>
<td>miekag 'very (smelly)</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>CORING</strong></td>
<td>cunguh 'nose'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>PIPI</strong></td>
<td>pipi 'cheek'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>LAKLAK</strong></td>
<td>laklak 'k.o. cake'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>APEM</strong></td>
<td>apem 'k.o. cake'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>CRAMCAM</strong></td>
<td>cramcam 'meat dish'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>PLECING</strong></td>
<td>plecing 'veg. dish'</td>
<td>*{Ca,Ca}</td>
</tr>
<tr>
<td><strong>BUBUK</strong></td>
<td>bubuk 'sleep'</td>
<td>*{complex}</td>
</tr>
<tr>
<td><strong>LOLOK</strong></td>
<td>lolok 'penis'</td>
<td>*{complex}</td>
</tr>
</tbody>
</table>

**Meanings of symbols used:**

- *{Ca,Ca}*: prefer morphemes in which two consonants of the same place of articulation do not cooccur (6.2.1)
- *{VHI,V MID}*: prefer lexical morphemes which are disyllabic (6.1)
- *{CNV}*: prefer grammatical morphemes which are monosyllabic (6.1)
- *{complex}*: prefer morphemes which do not contain nasal consonants in syllable-onsets (5.2.1)
- *{complex}*: prefer morphemes which are not complex (4.0, 6.4)
- *{CL}*: prefer morphemes which do not contain CL clusters (6.2.2)
- *{VHI,V MID}*: prefer morphemes where cooccurring [+ATR] vowels agree in height (6.3.1)

In the following paragraphs I give the working definitions I have used to classify morphemes into the semantic classes referred to in table Table 3.2, except where they seem self-evident. I stress that I do not consider these semantic classes to be semantic primitives in any sense. In section 3.4.2 I propose a continuum of related expressive
semantic elements; I assume that a given expressive morpheme will be associated with one of those elements, or a closely related one.

I take the classes **BODY PARTS**, (terms of) **ADDRESS**, (personal and place) and **NAMES** to be self-explanatory. Three classes are given special diagnostic status in statistical analyses. I define these classes as follows:

- **the PLANTS & ANIMALS** class refers to names of both plants and animals, and of parts of plants and animals.
- **the SENSE** class includes (a) any morpheme which lexicalizes vivid sense impressions in some way, be these impressions of sound (onomatopoeics), sight, touch, taste or smell.11 (b) morphemes lexicalizing human feelings and emotions.12
- **the BAD** class includes morphemes with either pejorative value, or else which refer to undesirable things (eg evil spirits) and states (eg mental, bodily, atmospheric).

Other classes whose semantics are not self-explanatory are:

- **TRUE PATIENT VERBS** are those in which the undergoer of the action is radically affected by the action concerned, often suffering some sort of physical damage, or else being in some sense overpowered. Where they are transitive, they are also signalled by the verb morphology: see 11.1.
- **VERY** indicates words with a meaning like 'extremely, really, as [verb] as possible', and particularly the class of intensifiers, each of which occurs in collocations with a single lexical item only. In these examples the order is (as is usual) head before intensifier:

  (8)  
  suung mungmung 'very quiet'  
  selem dekot 'very/jet black'  
  matah kelur 'very raw'  
  barak ngékak 'very red'  
  pait meklit 'very bitter'  
  manis nyer 'very sweet'
- **HI/HON** refers to the 'High' and Honorific lexical classes, part of the elaborate speech levels system borrowed virtually in toto from Javanese at least 400 years ago (Clynes 1989). These are loanwords, though I occasionally list them as a distinct semantic type. There are many parallels both in morpheme structure constraints, and in the function and pragmatics of the high/honorific lexis, in the two languages (see Uhlenbeck 1949).
- The class 'SMALL things (about which one typically has good feelings)' cross-linguistically is expressed with special markings, such as 'diminutive' derivational morphemes.13 It includes minor subclasses such as **CHILDISH TERMS**, names of **CAKES/FOODS**, names of woven temple offerings (**sampian**).

---

11 Sight' terms include verbs of movement, particularly those which lexicalize movements of the body, and of body parts, such as (in English) skipping, nodding, twitching, and so on.
12 Balinese *asa* translates variously as 'taste', 'feeling (physical, including touch)' and 'feeling (emotion)'; the emotions are thus lexicalised as a 'sixth sense' in Balinese.
13 See for example Wierzbicka's (1986) explication of the English diminutive *-ie*, in words like *birdie, horsie, mummy, daddy*: '[...] talking about it I am in a good mood as people are when they talk about small things for which they feel good feelings.'
These semantic categories closely match those found by Uhlenbeck (1978 [1950]) to be associated with peripheral morphemes in Javanese, as well as those listed by Trubetzkoy (1969 [1939]:254) and by Fudge (including those found in the studies of Finnish and African languages cited by Fudge); they are similar too to those Blust (ms.) associates with reflexes of his reconstructed proto-Austronesian expressive affix *qali-/kali-. 14

3.2.3 Statistical tests: central versus peripheral morphemes

The claim that morphemes which contain non-optimal phonological structures belong in unusually large numbers to a limited set of recurrent semantic (and other) classes is meaningless unless the frequencies of these classes in the phonologically regular portion of the lexicon are known. As a control, a sample of the semantic types of morphemes predicted to have the most 'neutral/optimal' possible structure in Balinese was made, again using Warna et al (1990) as database. This involved a sample of 209 morphemes of shape C1V(N)CVC3, where C1 was one of the three commonest initial consonants, /p/, /t/ and /s/, and C3 was one of the two commonest final consonants, /η/ and /h/. 15 Any morpheme containing a dispreferred phonotactic combination, or one suspected as such, was excluded. The numbers of morphemes belonging to the three 'criterion' semantic classes, the PLANTS & ANIMALS, SENSE and BAD classes (defined in 3.2.2), was then calculated, for this sample. (See appendix A for the detailed analysis). The 'peripherality' or otherwise of any other data set can then be established by comparing the numbers of morphemes in it belonging to the criterion semantic classes, with their levels in the control sample. Results of the control group analysis are given in Table 3.3, with a breakdown by expressive and other semantic/functional classes. The background level of morphemes belonging to the SENSE, PLANT/ANIMAL, and BAD classes is thus 33%:

<table>
<thead>
<tr>
<th>/C₁...C₃/</th>
<th>SENSE</th>
<th>PL/AN</th>
<th>BAD</th>
<th>Other LOANS</th>
<th>PERIPH.</th>
<th>UNCLSS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p...η/</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>21</td>
<td>40</td>
</tr>
<tr>
<td>/p...η/</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>14</td>
<td>37</td>
</tr>
<tr>
<td>/t...η/</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>20</td>
<td>35</td>
</tr>
<tr>
<td>/t...η/</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>35</td>
</tr>
<tr>
<td>/s...η/</td>
<td>4</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>3</td>
<td>11</td>
<td>33</td>
</tr>
<tr>
<td>/s...η/</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>12</td>
<td>29</td>
</tr>
<tr>
<td>totals</td>
<td>23</td>
<td>17</td>
<td>30</td>
<td>28</td>
<td>22</td>
<td>89</td>
<td>209</td>
</tr>
<tr>
<td>=%'age</td>
<td>11%</td>
<td>8%</td>
<td>14%</td>
<td>13%</td>
<td>11%</td>
<td>43%</td>
<td>100%</td>
</tr>
</tbody>
</table>

| Total PLANT/ANIMAL, BAD, SENSE | 70 (=33%) |
| Total Expressives & loans      | 120 (=57%) |
| Total Unclassified             | 89 (=43%)  |

Table 3.3 Semantic classes of morphemes in control sample

14 Semantic classes Blust associates with *qali/kali: 1. creepy-crawlies [insects, lizards, crabs]; 2. bats/birds; 3. striking natural phenomena [rainbows, echoes, storms, whirlpools, sunshowers]; 4. body parts; 5. muddled psychological states; 6. skin conditions [dandruff, rash]; 7. "unclustered", including: dust, feline quadruped, tree sp., restless, topsy-turvy, shy/timid, turbid; slingshot, rustle in wind, summit (Blust ms.:14-15)

15 Like Javanese (Uhlenbeck 1949), Balinese has around 90% disyllabic morphemes (6.1). The types CVCVC and CVNCVC together constitute around 50% of all disyllabics.
Compare these figures with those in Table 3.4. This gives the semantic classes of morphemes that, in different ways, break the constraint preferring no more than one consonant of a given place of articulation per morpheme, underlyingly (6.2.1):

<table>
<thead>
<tr>
<th>structure</th>
<th>SENSE</th>
<th>PL/AN</th>
<th>BAD</th>
<th>LOANS</th>
<th>OTHER</th>
<th>UNCLS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>/k... ...k/</td>
<td>71</td>
<td>18</td>
<td>17</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>125</td>
</tr>
<tr>
<td>/#gVg.../</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>/...tVt#/</td>
<td>15</td>
<td>11</td>
<td>8</td>
<td>12</td>
<td>6</td>
<td>6</td>
<td>58</td>
</tr>
<tr>
<td>/#tVd.../</td>
<td>5</td>
<td>-</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td>Totals:</td>
<td>98</td>
<td>31</td>
<td>35</td>
<td>24</td>
<td>20</td>
<td>18</td>
<td>226</td>
</tr>
</tbody>
</table>

- total SENSE, PLN/AN, BAD = 164 (=73%) p < .0000
- total Expressives & loans = 208 (=92%) p < .0000
- total Unclassified = 18 (=8%)

Table 3.4 Semantic classes of some morpheme-types breaking the Preferred Rule 'one C_place per morpheme' (data from Warna et al 1990).

The 'diagnostic' classes BAD, SENSES, PLANT/ANIMAL now constitute 73% of the total sample, versus only 33% in the control. A standard statistical test indicates that there is less than one chance in ten thousand that the difference between these figures could be due to random effects. The procedure used to determine this is described in the following paragraphs.

I take the data samples analysed in Tables 3.3 and 3.4 to derive from distinctly different statistical populations, since from the outset the samples have been selected as phonologically 'optimal' and phonologically 'non-optimal' respectively. The proportion of morphemes belonging to the classes BAD, SENSES, PLANT/ANIMAL in the control sample is 0.33. Given this, the probability of obtaining a proportion of 0.73 of morphemes belonging to those same classes in the second sample can be calculated using the Median Test (Spiegel 1972:171, Hatch & Lazaraton 1991:271):

\[ p(Z > t) = \Phi^{-1}(Z) \]

Using this formula, we obtain the result \( Z = 9.105, p(Z > 9.105) < .0000 \). That is, the probability of getting this difference between the two proportions in a purely random way is less than one in ten thousand. If we redo the analysis, including all the potentially 'peripheral' elements (57% of the control sample classifiable as 'expressives'

---

16 The /k... ...k/ data, and my analysis of it, is given in full in Appendix B. The /k... ...k/ and /#gVg.../ samples were the first 125 and 20 items respectively of the lists produced by a data base search, using Alsop (1990). The /...tVt#/ and /#tVd.../ samples include all the morphemes of those structures in the data base.

17 I am grateful to Yvonne Pittelkow of the RSPAS Computer Services Centre, Australian National University, for help with the statistical analysis used in this section.

18 Where \( Z = \frac{p_1 - p_2}{\sqrt{p_1(1-p_1)/N_1 + p_2(1-p_2)/N_2}} \), and \( p_x \) = the proportion of 'peripherals' in sample number 'x', \( q_x \) = the proportion of 'unclassified' elements in sample number 'x', and \( N_x \) = total number of items in sample 'x'.
or loans, versus 92% in the second sample), we obtain the result: $Z = 9.042$, $p(Z > 9.042) < .0000$. Again, a probability of less than one in ten thousand.

In this and later chapters a variety of non-optimal phonological structures are analysed in this way - see for example 4.2, 3.2.5, 3.2.6, 6.2.1, 6.2.2. Repeatedly, I find that a very high proportion of morphemes containing these structures belongs to the three criterion 'expressive' classes - a much higher proportion than is found with phonologically regular morphemes. In each case, the statistical test shows that the disparity cannot be plausibly attributed to random effects.\(^{19}\)

I conclude from the above analysis that in Balinese morphemes containing marginal phonological structures do regularly associate with a limited set of 'expressive' (and other) classes, and to a far greater extent than do phonologically regular morphemes. This difference cannot be plausibly explained away as due to random effects. I propose an explanation of the correlation in 3.4.2.

I take this correlation to constitute independent evidence for the phonological 'peripherality' of structures identified as such on purely phonological grounds. The correlation thus lends independent support to the phonological model developed by Smolensky & Prince, Goldsmith (1990) and Archangeli & Pulleyblank (1994) and others. In these recent models of phonology, phonological well-formedness is a matter of degree rather than an absolute. Similarly, analysts of syllable structure have long taken as self-evident the existence of rules operating in terms of preferences, rather than categorically (for example Fudge 1969, 1970, Itô 1988, Vennemann 1988).

Archangeli & Pulleyblank (1994:406, citing unpublished work by Prince & Smolensky) propose that less-than-optimally well-formed structures arise as the product of competing principles in the phonology and other components of linguistic structure. This is clearly the case in Balinese, where the regular correlation between otherwise dispreferred phonological structures and expressive semantics indicates 1) that linguistic structure recognises a distinct 'expressive' lexical class within the lexicon, and 2) that there is a principle or principles operating, whose net effect is to prefer less 'well-formed', less 'optimal', phonological structures in the creation of members of that class.

Independent evidence that Preferred Rules are active synchronically, and deliberately and productively broken to form expressives, is given in the next subsections, 3.2.4, and 3.2.5. A mechanism to account for this is proposed in 3.4.2.

---

\(^{19}\)Sample size can affect the reliability of the probability test used above. Occasionally the number of morphemes in my samples is a good deal smaller than the sample used above: 30 morphemes is normally the smallest sample used, except where a given phonological sequence occurs less frequently than that. When the necessary statistical corrections have been made, however, it turns out that even at these smaller sample sizes, the test remains reliable. A cut-off point of minimally 50% 'diagnostic peripherals' (BAD, SENSES, PLANT/ANIMAL) is valid for a probability of less than one in 1000, assuming that the proportion of 33% 'diagnostics' found for the control sample holds where the sample size is smaller. This is certainly the case for the data sub-sets in Table 3.3 (\(/p...g/\) etc).
3.2.4 Synchronic evidence for the productive interaction of phonology and semantics - I: Nicknames

It could be argued that the correlations described above between peripheral phonological structures and 'expressive' semantic classes are not the results of Preferred Rules (3.0), but rather reflect in some way diachronic constraints that are no longer part of the synchronic grammar of the language. Naming practices in Bali provide one important source of evidence that these rules do function synchronically.

As indicated in Table 3.2, terms of address, including personal names, are one of the semantic types regularly associated with marginal phonological structures. Most Balinese as children are given a deliberately inelegant and often meaningless, nickname, which for some is the only name they are known by throughout their lives. There can be no doubt that these names do have inelegant connotations. The generic term used to refer to them is adan jelé, 'bad name'. (Where they are meaningful, nicknames often poke fun at physical characteristics: gentong 'jar' [of a fat child], pékés 'squashed flat', pengah 'acrid smell', ribut (BI, 'rowdy'), or are onomatopoeic: geledag 'thump!' engkik 'squeak!', or in other ways undesirable: util 'snitch', kebo 'buffalo'). Not surprisingly, most people these days also have an alternative, usually Sanskrit-derived, 'good name' (adan melah), which others use when referring politely to them.

These nicknames invariably contain dispreferred phonological structures. Where they are meaningless, they provide good evidence that certain preferred rules do operate synchronically, since their inelegant connotations must derive directly from their irregular form. Examples of meaningless, one-off, and definitely inelegant names, are:

\[ (9) \quad \text{Glemuk, Cikadu, Cibuhuk, Nyaamprut, Sengkug, Gombios, Dodi, Ratep, Jedur, Uk, Cruit, Ngahnguh, Co'ét, Protokonyos} \]

Other (apparently) unique names, while not inelegant, are still phonologically marginal, such as these:

\[ (10) \quad \text{Kédi, Kéni, Maséni, Toti} \]

The name Cibuhuk breaks an Absolute Rule, having \(/h/\) in a syllable onset (5.2.1). Otherwise, the names in (9) and (10) variously 'break' one or more of the following Preferred Rules:

1. Prefer morphs in which two consonants of the same place of articulation do not cooccur (6.2.1): Glemuk, Sengkug, Ratep, Jedur
2. Prefer morphs which are disyllabic (6.1): Uk, Cikadu, Cibuhuk, Protokonyos
3. Prefer morphemes which do not contain nasal consonants in syllable-onsets (5.2.1): Ngahnguh, Nyaamprut, Glemuk, Protokonyos

\[ ^{20} \text{Some of these names are of Peguyangan people, others of people living nearby, but outside the Peguyangan area. I have not been able to get their approval for citing here, but my consultants doubt that they would take offense to this usage.} \]
Chapter 3

4. Prefer morphemes where cooccurring [+ATR] vowels agree in height (6.3.1): Kédi, Kéni, Maséni, Dodi
5. Prefer morphemes which do not contain CL clusters (5.2.2): Glemuk, Nyamprut, Gomblos, Crui
6. Prefer morphemes which are not complex (4.0): Ngahnguh, Dodi
7. Prefer morphemes which do not end in labial consonants (5.3.2): Ratep

I take this productive coining of 'inelegant' nicknames, using otherwise meaningless morphs containing dispreferred phonological structures, to constitute strong evidence for a synchronic interaction between semantics and phonology in Balinese. Based on evidence like this I argue in 3.4.2 that marginal phonological structures generate 'expressive' meaning components directly.

3.2.5 Synchronic evidence for the productive interaction of phonology and semantics - II: Dorsal-dorsal interactions

One particularly striking example of the productive way in which phonologically marginal structures are used to create expressive vocabulary was touched upon in 3.2.3: the figures in Table 3.4 suggested that the rule preferring no more than one consonant of a given place of articulation per morpheme is regularly broken to generate expressives. This is especially clear when morphemes containing more than one dorsal consonant are examined. The raw figures for morphemes of structure /#C_initial... ...C_final#/ in Warna et al 1990 are given in Table 3.5:

<table>
<thead>
<tr>
<th>p</th>
<th>b</th>
<th>m</th>
<th>t</th>
<th>d</th>
<th>n</th>
<th>l</th>
<th>r</th>
<th>s</th>
<th>k</th>
<th>g</th>
<th>nj</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>120</td>
<td>36</td>
<td>65</td>
<td>64</td>
<td>69</td>
<td>76</td>
<td>131</td>
<td>22</td>
<td>144</td>
<td>100</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>14</td>
<td>105</td>
<td>21</td>
<td>57</td>
<td>89</td>
<td>64</td>
<td>77</td>
<td>124</td>
<td>58</td>
<td>145</td>
<td>102</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>9</td>
<td>31</td>
<td>2</td>
<td>22</td>
<td>27</td>
<td>25</td>
<td>37</td>
<td>18</td>
<td>18</td>
<td>58</td>
<td>40</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>9</td>
<td>4</td>
<td>4</td>
<td>25</td>
<td>12</td>
<td>6</td>
<td>17</td>
<td>3</td>
<td>37</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>18</td>
<td>14</td>
<td>17</td>
<td>19</td>
<td>78</td>
<td>59</td>
<td>41</td>
<td>110</td>
<td>138</td>
<td>37</td>
<td>125</td>
<td>113</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>7</td>
<td>26</td>
<td>1</td>
<td>33</td>
<td>31</td>
<td>14</td>
<td>15</td>
<td>48</td>
<td>9</td>
<td>56</td>
<td>29</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>13</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>15</td>
<td>3</td>
<td>78</td>
<td>47</td>
<td>38</td>
<td>0</td>
<td>100</td>
<td>66</td>
<td>80</td>
<td>48</td>
<td>116</td>
<td>112</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>13</td>
<td>59</td>
<td>14</td>
<td>18</td>
<td>4</td>
<td>0</td>
<td>18</td>
<td>68</td>
<td>26</td>
<td>63</td>
<td>56</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>8</td>
<td>34</td>
<td>38</td>
<td>17</td>
<td>135</td>
<td>22</td>
<td>122</td>
<td>61</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>5</td>
<td>45</td>
<td>5</td>
<td>21</td>
<td>15</td>
<td>26</td>
<td>13</td>
<td>29</td>
<td>27</td>
<td>59</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>51</td>
<td>28</td>
<td>20</td>
<td>187</td>
<td>52</td>
<td>66</td>
<td>63</td>
<td>118</td>
<td>7</td>
<td>97</td>
<td>52</td>
<td>176</td>
<td>139</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>27</td>
<td>28</td>
<td>157</td>
<td>55</td>
<td>86</td>
<td>88</td>
<td>100</td>
<td>115</td>
<td>139</td>
<td>42</td>
<td>243</td>
<td>114</td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>24</td>
<td>101</td>
<td>6</td>
<td>18</td>
<td>64</td>
<td>89</td>
<td>66</td>
<td>93</td>
<td>80</td>
<td>188</td>
<td>87</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>4</td>
<td>10</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.5 Numbers of morphemes of structure /#C_initial... ...C_final#/ in Warna et al (1990) (column = C_initial, row = C_final)

The raw data alone suggest that there are restrictions dispreffering cooccurrence of more than one labial, apical stop, liquid (/I/ or /r/, which here apparently behave as a separate class), or laminal.21 However, these figures also suggest that there is no constraint dispreffering coooccurrence of consonants belonging to the remaining place category,

21Glides do not occur syllable-finally; /s/ is the only laminal to occur word-finally.
dorsal-dorsal. In fact there are more morphemes with dorsals in both initial and final positions, than with any other place combination.

It turns out, however, that morphemes with both initial and final dorsal consonants are almost exclusively expressive in meaning. This is shown in Table 3.6 below, which gives the results of an analysis of 153 morphemes of this structure (given in full in Appendix B). A random sample for each sub-type was taken. Clearly there is a very high correlation with expressiveness: 80% of the sample belongs to the three criterion semantic classes, with a big clustering in the SENSE category. The skew is again highly statistically significant, $Z = 10.247$, $p < 0.000$, using the methodology outlined in 3.2.3.

<table>
<thead>
<tr>
<th>structure</th>
<th>SENSE</th>
<th>PL/AN</th>
<th>BAD</th>
<th>LOANS</th>
<th>Other PERIPH</th>
<th>UNCLS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>/k... ...g/</td>
<td>27</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>/k/ ...g/</td>
<td>14</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>-</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>/g/ ...k/</td>
<td>18</td>
<td>4</td>
<td>2</td>
<td>-</td>
<td>4</td>
<td>5</td>
<td>33</td>
</tr>
<tr>
<td>/g/ ...g/</td>
<td>25</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>/g/ ...g/</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>-</td>
<td>2</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>16</td>
<td>15</td>
<td>3</td>
<td>6</td>
<td>21</td>
<td>153</td>
</tr>
<tr>
<td>total SENSE, PL/AN, BAD</td>
<td>123 (=80%)</td>
<td>p &lt; .0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total Expressives &amp; loans</td>
<td>132 (=86%)</td>
<td>p &lt; .0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.6 Semantic types of /#C_dorsal ... ...C_dorsal#/ morphemes

Table 3.6 does not include data for morphemes of structure /#k... ...k#/; this was given in Table 3.4, and is repeated below. For this combination the skew is particularly striking, with only 5% of morphemes not immediately classifiable into one of the expressive or other peripheral categories (see Appendix B).

<table>
<thead>
<tr>
<th>structure</th>
<th>SENSE</th>
<th>PL/AN</th>
<th>BAD</th>
<th>LOANS</th>
<th>Other PERIPH</th>
<th>UNCLS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>/k/ ...k/</td>
<td>71</td>
<td>18</td>
<td>17</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>125</td>
</tr>
<tr>
<td>total SENSE, PL/AN, BAD</td>
<td>106 (=85%)</td>
<td>p &lt; .0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total Expressives &amp; loans</td>
<td>119 (=95%)</td>
<td>p &lt; .0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.7 Analysis of a random sample of morphemes of shape /k... ...k/ (data from Warna et al 1990).

In short, and despite their occurrence in large numbers, the fact that morphemes with dorsal consonants in both initial and final position contain a very high proportion of expressives (and particularly so in the case of morphemes of shape /#k... ...k#/) leads to two conclusions:
1. The rule disprefering cooccurrence of consonants of the same place of articulation still holds, though in the non-expressive component of the lexicon only.
2. In the coining of expressive lexis this rule is productively 'broken', using the dorsal place of articulation in particular.
3.2.6 Diachronic evidence for the interaction of phonology and semantics: lenition of /r/

If certain morphemes are exempted from the application of synchronic rules because they belong to specific semantic classes, then we can expect that rules of sound change operating over time could also fail to apply for the same reason.²² I now present some evidence from Balinese that this is the case.

Texts dating back more than a thousand years show that in Old Balinese, a rule variously deleting /r/ in syllable onsets, and leniting it to /h/ in syllable-codas had already begun to take effect (Goris 1954, Clynes 1989). Examples of the application of the rule include:

(11) Old Balinese > Modern Balinese

<table>
<thead>
<tr>
<th>Old Balinese</th>
<th>Modern Balinese</th>
</tr>
</thead>
<tbody>
<tr>
<td>ratu 'ruler'</td>
<td>atu</td>
</tr>
<tr>
<td>rumah 'house'</td>
<td>umah</td>
</tr>
<tr>
<td>karu 'coconut shell'</td>
<td>kau</td>
</tr>
<tr>
<td>turun 'descend'</td>
<td>tuun</td>
</tr>
<tr>
<td>bayar 'pay'</td>
<td>bayah</td>
</tr>
<tr>
<td>lemar 'daytime'</td>
<td>lemah</td>
</tr>
</tbody>
</table>

Warna et al (1990) nonetheless lists about 600 r-initial morphemes, and 900 r-final ones, many of them not obviously loans. To test the hypothesis that exceptions to the lenition rule could have been semantically motivated, I analysed the semantic classes of a sample of 58 otherwise 'phonologically neutral' /r/-final morphemes (see Appendix C). Results are shown in Table 3.8.

<table>
<thead>
<tr>
<th>SENSE</th>
<th>BAD</th>
<th>PL/AN</th>
<th>LOANS</th>
<th>OTHER</th>
<th>UNCLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>8</td>
<td>3</td>
<td>20</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>33%</td>
<td>14%</td>
<td>5%</td>
<td>34%</td>
<td>-</td>
<td>14%</td>
</tr>
</tbody>
</table>

Table 3.8 Semantic (etc) classes of a sample of otherwise phonologically 'neutral' /r/-final morphemes.

The total number of morphemes belonging to the BAD, SENSE and PLANT/ANIMAL classes is then 30 - of which 27 belong to just two classes, BAD and SENSE. The proportion 30/58 'diagnostics' was compared with that for the modern Balinese 'neutral' control sample (§3.2.3), and analysed using the statistical procedure outlined there. The result, \( Z = 2.733, P < .005 \), again indicates a very low probability that the discrepancy between the two samples is due to random effects.

²²This view is implicit in the argument of Gonda (1947), that 'Indonesian' (i.e Austronesian) languages in general contain a very high proportion of 'primitive' (i.e expressive) phonological elements, and that the normal methods of comparative historical linguistics could give incorrect results if these elements are not taken into account.
That the non-loan residue of /r/-finals shows a strong statistical correlation with in fact just two of the three diagnostic 'expressive' semantic classes suggests that there was a semantic motivation for their not undergoing the rule of /r/-deletion. Clearly this is just one piece of evidence. It does however support the claim that 'exceptions' to phonological processes can be semantically motivated, both over time as well as synchronically. The explanation proposed in 3.4.2 gives a single account for both the synchronic and diachronic facts.

3.3 Identifying Preferred Rules

To demonstrate that a given structure violates a synchronically active Preferred Rule in Balinese (or any other language), I take it that two independent types of evidence are required. There must first be independent phonological evidence that the structure concerned is in some way 'dispreferred' in the language. Secondly, there must be semantic evidence: a statistically significant proportion of the morphemes containing that structure must belong to the criterion peripheral semantic classes (3.2.3). Without evidence of phonological marginality, semantic effects with a morphological or other explanation could be falsely analysed as phonologically motivated. And without the semantic evidence, a list of 'exceptions' to a phonological rule could be no more than that. In practice, when looking for potentially marginal phonological sequences, I have used five types of data as a starting point:

1. Structures which I have noticed in passing to be recurrently associated with expressive meanings (eg morphemes containing complex syllable onsets, particularly when these occur in final syllables).
2. Constraints which apply to some members of a phonological class, but not to others (eg there are near-absolute constraints barring cooccurrence of labial consonants tauto-morphemically, while dorsals cooccur frequently).
3. Structures which occur with low frequency (cf Fudge 1970).
5. Cross-linguistic evidence generally on marked and unmarked structures, and those tending to cooccur with 'expressive' morphemes.

3.4 Explaining the correlations

So far I have attempted to demonstrate that certain semantic or functional classes regularly recur in morphemes with phonologically peripheral structures. How can this correlation be explained? I explore two different explanations in the next subsections. The first is that the effect occurs where phonological rules apply to only a certain sub-
class, or sub-classes, of the lexicon (3.4.1). However, I take this explanation to be insufficient to account for all the observed sound-meaning interactions. A radically different explanation, that marginal phonological structures themselves generate meaning, is proposed in section 3.4.2.

3.4.1 Explaining the correlation 1: Phonological sub-classes

It appears that many so-called 'exceptions' to phonological rules are simply members of lexical classes not targeted by a particular rule. Take for example the strong tendency for morphemes to be disyllabic in Balinese: close on 90% of the morphemes listed in Warna et al (1990) are disyllabic (6.1). At the same time, over 90% of grammatical morphemes are monosyllabic. The correct explanation for these facts, is perhaps not that grammatical morphemes are exceptions to the Foot Well-formedness rule, which leads to disyllabicity (5.6), but rather that they are simply not targeted by it:

(12) Foot Well-formedness
The input to, and the output of, all phonological processes must be prosodically well-formed at the Foot level

Domain: Lexical morphemes

Other, more universal, principles might then account for the monosyllabicity of grammatical morphemes, given the strong cross-linguistic tendency for this class to be mono-syllabic. Alternatively, they are possibly targeted by a specific rule:

(13) Prefer monosyllabic morphemes

Domain: Grammatical morphemes

It is also commonly recognized that, in the operation of phonological rules, languages regularly treat loanwords as a distinct subclass (or subclasses) in the lexicon. In English, for example, Chomsky and Halle argue that:

' [...] it is necessary to classify many lexical items in terms of a feature which, roughly, distinguishes items of Germanic origin from other items; and for certain rules, such as Velar Softening, we need a further classification of the non-Germanic part of the vocabulary into items of Greek and Romance origin, roughly. Parallel instances may be cited from a great many languages (1968:373).'

23

Similarly, in Balinese a rule variously deleting /r/ or leniting it to /h/ was underway more than a thousand years ago (3.2.6). However, the many Sanskrit and Javanese loanwords being incorporated into the lexicon concurrently with the period of operation of that rule were never subject to it.

23 They go on to cite Turkish, where phonological rules behave differently with respect to the native Turkish and Arabic parts of the lexicon, and Russian, which has a three-way distinction between native Russian lexis, other Slavonic, and non-Slavonic elements.
So a possible explanation for the exceptionality of the classes functors and loanwords with respect to certain phonological rules is that either those rules do not target them, or else that other rules (such as (13)) do. Nevertheless, the existence of phonological sub-classes in the lexicon is probably a secondary phenomenon, deriving from a more fundamental source. I argue in 3.4.2 that the pragmatics of loanword usage parallels that of expressives; the phonological exceptionality of the two classes may thus have a common explanation. The peripherality of functors is not as easily explained in these terms.

3.4.2 Explaining the correlation 2: Marginal phonological structures generate meaning

One reason for apparent 'exceptions' to phonological rules could therefore be that they belong to lexical classes not in fact targeted by those rules (3.4.1). This may be the case for grammatical morphemes and loanwords, but it seems less likely that the exceptionality of the expressives can be explained in this way. For one thing, the group is too heterogeneous, both semantically (3.2.2) and in the variety of word classes (stative verbs, transitive verbs, nouns of different types) it contains. There seem instead to be two possible hypotheses for the correlation between phonological marginality and expressiveness:

1. A constant expressive semantic element is associated with marked phonological structures; they thus realize an expressive morpheme.

2. An expressive element (or elements) is associated with these structures, but in a way unlike other form-meaning associations.

Both hypotheses are compatible with the first main claim of this chapter, that peripheral morphemes associate regularly with peripheral phonological structures, and in so doing manifest the existence of Preferred Rules of phonology. Hypothesis 1 is problematic for a variety of reasons, discussed below; I will argue in favour of hypothesis 2. I propose an explanation which requires revision of Woodbury's (1987) Revised Double Articulation account of phonology-semantics interactions (3.4.3).

The first hypothesis, that a morpheme 'expressiveness' exists, seems unlikely, for (at least) five reasons, outlined in the following paragraphs.

(1) Firstly, it would differ radically in kind from other morphemes, in having an inordinate number of lexically conditioned realisations, both in Balinese, and cross-linguistically (Fudge 1970). In Balinese these would include monosyllabicity, tri- and multi-syllabicity, reduplication of different types, cooccurrence of consonants of the same place of articulation within a single morpheme, cooccurrence of high vowels with mid-, non-central vowels, occurrence of CL sequences, and so on.

\[24\] Itô and Mester (1995) argue against such an analysis of very similar phenomena in Japanese, proposing instead a core-periphery model of the lexicon, consistent with frameworks like Optimality Theory.
(2) A second difficulty is that the Double Articulation hypothesis (in both its strong and revised (Woodbury 1987) versions) assumes that the relationship linking a given meaning with a given underlying form is purely arbitrary. However, in language after language, expressiveness is consistently associated with a variety of (often similar) marginal phonological structures (Fudge 1970, Hinton et al 1994). This suggests that this latter type of form-meaning association is a non-arbitrary process, deriving from universal principles, ones quite distinct from the regular (language-specific and arbitrary) processes linking form and meaning.

(3) Another reason for doubting the existence of such an expressive morpheme, necessarily associated with a constant semantic element, is the variety of semantic classes included in the notion 'expressive' in Balinese: from place names to personal names (most with 'inelegant' associations, but not all), through to onomatopoeia. It seems implausible that place names could have a similar expressive element to onomatopoeics, or perhaps any inherent expressive element at all.

(4) Any explanation also needs to account for the apparently gradient nature of the expressiveness-phonology correlation. By this I refer to Fudge's view that 'expressiveness is probably a continuum, rather than a series of discrete steps' (1970:163), or indeed, than the single 'all or nothing' property that a morphemic solution would require. Fudge's view is supported by his repeated finding (one not investigated in this study) that, as the peripherality a given structure increases, there is a gradual increase in the proportion of morphemes containing it which have expressive meaning - in Fudge's terms, 'as \( U \) decreases, \( R \) increases' cf 1970:170, 172 and passim, see 3.2.1.

(5) Finally, it seems that the peripheral phonology-semantics correlation is a strong tendency, demonstrable statistically, but with exceptions that cannot be accounted for under the expressive morpheme hypothesis.

For all the above reasons, that approach is rejected here. I will now sketch an alternative account, one where the association of expressive meaning with peripheral phonological structures is not achieved by the usual lexical process of arbitrary form-meaning association, but by a distinct, ultimately non-linguistic, mechanism. I assume then that phonological well-formedness is a matter of degree rather than an absolute, and that more or less 'illformed' structures must be tolerated at times (Prince and Smolensky 1993). At some level, such illformed structures are necessarily perceived as 'marked' or even, in extreme cases, 'not good'. This in turn is part of a more general tendency whereby all marked or unusual perceptions impinge on our cognitive processes, either at an unconscious or a conscious level: we notice, at some level, departures from the norm. In more extreme cases, vivid perceptions cause people to feel emotions, and even thoughts ('a beautiful sight', 'a terrible feeling'). There is then a continuum of affectedness associated with our perception of the world, from barely perceived reactions, to the generation of feelings - of pleasure, discomfort, and so on - and then finally of thoughts about these feelings and perceptions.
I assume that our use of language is also subject to this (basically non-linguistic) mechanism: marked linguistic structures impinge on our perceptions in a continuum of affectedness, increasing in intensity with the degree of aberration of the phonological structure used, until they generate linguistic reactions. Some at least of the stages in this continuum can be expressed as follows (using the approach of Wierzbicka (1986, 1988)):

(14)  

a. using a word of this [marked] shape, 'A', 
people feel something.

b. using a word of this shape, 'B',
(i) people feel something.
(ii) people think about this word.

c. using a word of this shape, 'C',
(i) people feel something
(ii) people think about the shape of this word.

d. using a word of this shape, 'D',
(i) people feel something BAD
(ii) people think something BAD about the shape of this word.

In all cases, I assume that users' reactions include a non-linguistic reaction such as (14a). Reactions (14b-d) are more extreme in that the user's attention is drawn, in a linguistic response (components b(ii)-d(ii)), to the word in which the peripheral structure occurs, and the feelings its shape generates. A tentative explication of the linguistic components of (14b(ii)-d(ii)), again expressed using the general approach of Wierzbicka, is:

(15)  

b. Using a word like this,
I feel something,
Not because of anything else

c. Using a word of this shape,
I feel something
Not because of anything else

d. Using a word of this shape,
I feel something bad
Not because of anything else

Such reactions can be assumed to violate a general principle requiring 'invisibility' of phonological form during communication:

(16) Principle of Transparency of Phonological Form:

Phonological form should be invisible during communication: it should not impinge on the speaker's or listener's consciousness.

This principle is violated in onomatopoeics, and in rhetorical use of language, such as poetry, which can exploit the full range of reactions (14a-d). More generally, the
Chapter 3

focussing effect of (14b) is exploited in a way which parallels the use of *italic font* or Capital Letters in writing. Such usages draw the attention of the reader to the word(s) thus marked, and tell the reader that they are to be interpreted as in some way salient in the current context. NAMES have important communicative functions. One reason that they are consistently given marked shapes, I propose, is because of the importance of these functions, and the consequent need for them to be picked out in the stream of speech (see below for a different, more basic explanation). The unique shapes of some FUNCTIONALITY (see 3.4.1) may well derive from the same principle.

Reaction (14c) may account for the predominance of onomatopoeia and other SENSE terms in very marginal structures. (Though words from the BAD class are also common in this type.) The most extreme reactions, such as (14d), probably lead in many cases to the structure involved being 'repaired': eliminated from the phonology (see 3.2.6, on sound change). Such reactions are no doubt the reason why many adults learning to pronounce foreign languages experience discomfort.

A common element shared by members of the expressive class is then this: they are all used in pragmatically marked contexts, precisely those contexts where the user might feel, or want the listener to feel, one or more of the elements 'available' in (14). This accounts for problematic features of the observed form-meaning correlation, such as the fact that NAMES (place names and personal names) behave in a similar way to SENSE terms, while not being inherently expressive in meaning: like SENSE terms, PERSONAL NAMES in fact almost always are used to communicate 'pragmatic' meaning. One could for example address someone named 'Robert Johnson' in various ways - Bob, Rob, Robbie, Johnson, Mr Johnson, etc - but none of these choices would be unmarked. Each would clearly imply a distinct relationship between speaker and addressee (Wierzbicka 1989).

There seems no doubt also that, for the people whose language and culture they are primarily associated with, PLACE NAMES have multiple pragmatic associations, cultural, religious and personal. Names of PLANTS and ANIMALS are equally non-neutral semantically, as Wierzbicka (1985) has again demonstrated. There is moreover no doubt that in Balinese culture plants and animals are of central importance to people's religious, affective and physical needs. It can even be argued that elements similar to (14a) and (14b) are part of the pragmatics of LOANWORD usage, at least in the initial phase of borrowing, when such words are still seen as foreign, together with other connotations, such as 'smart', 'exotic' and so on. There are then perhaps grounds for including loanwords as a subclass of expressives. I have chosen not to do this, since loanwords perhaps lose affective semantic elements.

---

25 Compare Fudge: 'There is a tendency for [members of the expressive classes] to occur with certain "emphatic" intonations [...], and in certain types of situation ("familiar" as opposed to "formal"), where their effect is to strengthen social bonds between speaker and hearers' (1970:164).

26 See for example Covarrubias 1970 [1937], on the importance of animals and plants in religious offerings.
over time, familiarity breeding indifference: they nonetheless remain phonologically aberrant (see footnote 23).

I propose that the generation of semantic elements by marked structures has repercussions in three areas: 1) in synchronic word-formation processes, 2) in processes of sound change over time, and 3) in the 'filtering' of loanwords as they enter the language. These are discussed and exemplified in turn in the following paragraphs.

1) Word formation. The information in (14) could be seen as a series of lexical entries associated with particular marginal word shapes, rather than particular words. This information is clearly available in word formation, and is exploited productively, when and where appropriate. Specifically, peripheral phonological structures are used productively to create morphemes belonging to peripheral semantic classes.

Abundant evidence of this productivity has been presented in this chapter. A particularly striking example is the sequence /k......k/, which violates the preferred rule disfavouring more than one consonant of the same place of articulation within a morpheme (6.2.1). /k......k/ nevertheless occurs with the second-highest absolute frequency of any /C......C/ combination in Warna et al (1990) (193 tokens, see 3.2.5). At the same time, virtually every morpheme in a sample of 125 (121, or 97% of the sample) belonged to one of the expressive classes (see Table 3.7). Importantly, the PLANT/ANIMAL, BAD, SENSE classes made up 86% of the sample (p < .0000).

New morphemes/words then are not coined in a 'pragmatic' vacuum. Rather, the evidence shows that attitudinal meanings associated with the usage of particular semantically-based word classes (SENSE, BAD and so on), as well as those generated by aberrant phonological structures, are both available at this stage. Appropriate phonotactic structures are 'chosen' in word-formation, to ensure that these two layers of attitudinal meaning are compatible.

2) Sound change. Evidence was given in section 3.2.6 that rules of sound change operating over time can fail to apply to expressive morphemes because they belong to specific semantic classes. One can thus hypothesize that the process by which meaning is generated phonologically actually constitutes one mechanism by which the selective application of sound changes proceeds.

Where a peripheral structure occurs in a neutral/central morpheme, phonologically generated semantic elements such as those in (14) will be incompatible with the lexical semantics of the morpheme, and because of this the structure is 'repaired'. On the other hand where such a semantic element is compatible with the semantics/function of the morpheme, the sequence is likely to be retained. This would lead to a clustering of expressive semantic classes with peripheral phonological structures, over time, just as is observed with /r/-final morphemes (3.2.6).

It might be thought that this historical explanation could then explain all 'synchronic' correlations between expressive semantics and peripheral phonological structures, as simply the end product of historical processes no longer part of the synchronic grammar. This would require the ad hoc and
3) Filtering of loanwords. With loanwords too, it is probable that the generation of attitudinal elements like (14) by non-native or marginal phonological structures influences whether or not a given morpheme undergoes 'repair' mechanisms when it is borrowed. Where these attitudinal elements are compatible with the lexical semantics of the morpheme concerned, as well as with pragmatic elements associated with usage, we would expect that a peripheral structure would be more likely to be retained. Where there is no such compatibility, the generating structure is more likely to be 'repaired'.

Evidence that meaning is involved in this process is clearest where 'bad' reactions like (14c & d) are generated in morphemes with compatible semantics. A striking example from English is the borrowing of various 'ill-formed' words from Yiddish, like schlock, schlemiel, schmuck, schmo, schmaltz, schnozzle, schnorrer, and 'shm-reduplication': actor-shmactor, friend-shmend. The ill-formedness of the initial consonant clusters, not found in the native English lexis, is desirable, precisely because it generates attitudinal overlay which is in harmony with the 'BAD' lexical semantics of the morphemes concerned. (Potentially analogous examples from Balinese include Balinese raksasa ~ rangsasa 'ogre, demon', from Sanskrit raksasa, where the non-native /ks/-sequence is optionally retained, and Balinese rangda 'demon widow', from Skt randa 'slut, widow', where Javanese/Balinese rangda has acquired a marginal consonant cluster, in the process of borrowing.)

The few non-pejorative /ʃC/-initial borrowings tend to be regularized: snitzel ~ schnitzel, snorkle ( ~ ??schnorkle), snapper ( ~ ??schnapper), whereas the pejoratives from Yiddish are not regularized (*slack, *smo, *smuck, etc). This difference in treatment is exactly what the above analysis would predict.28

3.4.3 On revising 'revised double articulation'

Woodbury's theory of Revised Double Articulation (RDA) has as its key element:

(17) Optional rules of postlexical phonology can bear meaning (1987:690)

While this is correct, it is clearly inadequate to account for the all the patterns described in this chapter. Semantics has been shown to also interact productively with phonology in pre-lexical word-formation processes.

----

28 schnozzle does vary with snozzle, but this is probably due to the influence of the expressive 'nose-related' words beginning with /sn/: snout, sniff, snot. /ʃC/ is probably retained in schnitzel where the foreignness of the dish is salient.
The question is then, can the notion of double articulation be retained through a 'Revised Revised Double Articulation', or is it simply no longer tenable as an axiom of linguistic theory?

Alongside (17) (which now has no special theoretical significance), Woodbury proposes three 'RDA principles', which specify 'just what should and should not receive treatment as a meaning-bearing rule of postlexical phonology' (691):

(18) RDA principles: For whatever may be formulated as an optional process bearing linguistic meaning, treat it as a rule in the postlexical phonology of a generative grammar only if:

a. It refers crucially to features of phonology in its structural description.

b. Its structural description, structural change, and effects on meaning are regular, predictable, and general (i.e. referring to natural classes, rather than arbitrary lists of elements)

c. Its output is not referred to by any rules or distributional patterns in the syntax or lexicon.

Clearly, (17) is not an integral part of a revised theory of double articulation. On the other hand, Woodbury's account accurately predicts the characteristics of the pre-lexical phonological processes I have described: 1) they are clearly optional rules, 2) their phonological patterning, and effects on meaning, are regular, predictable, and general, and 3) they appear not to refer to distributional patterns in the syntax or lexicon in their output. Could one then 'save' RDA by simply promoting the three RDA principles above to central status in the theory, as determining 'just what should and should not receive treatment as a meaning-bearing rule' of phonology in general (the word 'postlexical' having of course been deleted from line two of (18)?

Woodbury however argues that these largely follow from (17), plus independent 'principles of Lexical Phonology' (690-691). It appears that the most important of the three principles, if a principled double articulation is to be maintained, is RDA Principle c, which could perhaps therefore be promoted to central status in a theory delimiting possible interactions between semantics and phonology. Assuming that 'double articulation' as a notion is no longer a useful term (see below), I rename RDA as follows:

(19) 'Theory of Meaningful Phonological Processes', version 1:

For whatever may be formulated as an optional process bearing linguistic meaning, treat it as a rule in the phonology of a generative grammar only if its output is not referred to by any rules or distributional patterns in the syntax or lexicon.

Still, it is unlikely that the processes described in this chapter are compatible even with (19). As argued in 3.2.3, the evidence suggests that linguistic structure recognises a distinct 'expressive' lexical class within the lexicon, to which special principles of 'reverse optimality' apply during word-formation. That being the case, the output of
reverse-optimality phonological principles would violate principle (19). A further revision to take this into account would be to simply excise the words 'or lexicon':

(20) Theory of Meaningful Phonological Processes (revised):

For whatever may be formulated as an optional process bearing linguistic meaning, treat it as a rule in the phonology of a generative grammar only if its output is not referred to by any rules or distributional patterns in the syntax.

A principle such as (20) is clearly needed in a theory which seeks to bar purely phonological treatments of the many word-formation rules which have specifically phonological input conditions (Woodbury 1987:688ff, also Spencer 1991:16ff). The crucial question is, does such a principle have any empirical validity, or is it simply establishing boundaries 'by fiat'.29 There can be no value in replacing one clearly incorrect prescriptive claim, viz. that there 'can be' no interaction at all between semantics and phonology (Archangeli & Pulleyblank, 1994:433), with another.

In any case, it is clear from the empirical evidence, be it from Balinese, English (Fudge 1970), or Central Alaskan Yupik (Woodbury 1987), that semantics and phonology do interact in regular, productive ways, both at the pre-lexical, and the post-lexical levels. Given this, the mainstream view of interactions between linguistic modules as represented for example by Archangeli and Pulleyblank (1994:5, with minor modifications):

(21)

```
phonology

\[ \rightleftharpoons \]

syntax \[ \rightleftharpoons \] morphology

\[ \rightleftharpoons \]

semantics
```

should be replaced by a theory in which semantics is recognised as interacting with all components of linguistic structure in regular, albeit complex and predetermined ways. A more appropriate view of the relationship between the various components of linguistic structure, with minimal changes to diagram (21), is, I believe, the following representation:

\[ \text{\footnotesize \textsuperscript{29}Woodbury uses this argument in rejecting a placeholder analysis of the phenomena he describes.} \]
3.5 Conclusions

In this chapter I have shown that in Balinese there is a strong statistical correlation between morphemes containing a variety of demonstrably less-than-well-formed phonological structures, and 'expressive' semantic classes (3.2.3). I take the correlation to lend support to the phonological model outlined by Goldsmith (1990), Prince and Smolensky (1993). Importantly, it also enriches it, by supplying a source of semantic evidence for the marginality, or otherwise, of phonological structures.

It follows then that semantic data, properly handled, can be an important source of evidence in phonological analyses, where peripheral phonological structures are to be identified. In the following two chapters in particular, I regularly use semantic evidence, though always in conjunction with other kinds of evidence, to support analyses.

To explain the correlation between peripheral phonotactic structures and expressive semantic classes, I have proposed that such structures themselves generate 'expressive' semantic elements (3.4.2). This in turn throws further doubt upon the traditional view of the 'double articulation' of linguistic structure, and even on a modified version, such as Woodbury's (1987) Revised Double Articulation hypothesis, which would allow that certain classes only of phonological rules can bear meaning (3.4.3).

The nature of the relationship between phonological form and expressive meaning components, as I have described it, is not the usual one of purely conventional, arbitrary, association - nor even one of similarity, the usual basis of Peircean diagrammatic iconicity (Haiman 1985:9). Rather, it is one of identity. Haiman asks, reasonably, 'what similarity could possibly exist between a sound on the one hand, and any non-auditory phenomenon on the other[?] (1985:3)'. The answer I have proposed is simple: a common, subjective, human response. The feelings generated in response to marginal
phonological shapes are, at least in part, identical to the ones we feel when we experience a vivid sense impression, or when we hold an as-yet-nameless baby in our arms. Word-formation processes take advantage of this very natural, ultimately pre-linguistic, correspondence.

A strength of the model proposed in 3.4.2 is that it provides a unified explanation of both synchronic and (certain) diachronic phonological processes. I do not claim however that all sound change can be explained in this way, nor all exceptional phonological structures. One possible alternative account is given in 3.4.1.

It appears from this study that a partial distinction must be drawn between the expressive and non-expressive (neutral) components of the lexicon. Peripheral phonological structures are used productively in the formation of expressive lexis, whereas they are dispreferred in the generation of the non-expressive component of the lexicon. This then requires a nuanced view of what constitutes a well-formed, or 'optimal' phonological structure. Clearly 'optimal' structures are truly so only in the non-expressive lexical component. The model proposed in 3.4.2 nevertheless requires that such optimal structures be seen as more fundamental, in that they constitute the cognitive norm: there are no restrictions on the semantic type of morphemes containing them; the evidence suggests that departures from this norm are only allowed in word-formation processes deriving expressive lexis, otherwise repair processes will eliminate them.

Like Fudge (1970), I have found that, in general, the lower the frequency of occurrence of a given phonological structure, the more likely it is that only expressive and other peripheral morphemes will occur with it. However, the reverse does not apply: in Balinese a structure can be phonologically marginal, yet occur very frequently, given that peripheral phonological structures are used productively in the creation of expressive vocabulary (see for example 3.2.5, also 3.2.3, 3.4.2). This fact complicates the analysis of Balinese phonology: raw data on numbers of occurrences of a given sequence are of little value in determining its phonological 'markedness', if the function and semantics of the morphemes having that structure are not taken into account. A full study of Balinese phonology would therefore, for a given structure, take into account the semantics of all morphemes containing that structure - clearly beyond the scope of this and most studies.

The fact that 'pragmatically generated' semantic elements associated, say, with the usage of names (3.4.2), in turn influence the phonological shapes we give to newly coined names (or mutatis mutandis, any expressive terminology) is striking evidence for the fundamental inseparability of (linguistic) pragmatics and semantics; throughout this thesis I therefore follow Wierzbicka (1989) in treating linguistic pragmatics as a sub-discipline of semantics.

The finding that distinct principles of word-formation apply to expressives reflects how human cultural values intimately interact with, and shape, even the phonological
component of linguistic structure. Take for example the finding of this and other studies, that the names of plants and animals regularly pattern as expressives. This evidence that human interaction with the natural world is reflected in the phonologies of languages is perhaps surprising; yet it shows the fundamental role plants and animals play in human culture and cognition. Information from phonological analysis is thus relevant not merely to other branches of theoretical linguistics, but to a broad range of social sciences. For similar comments, see Wierzbicka (1985:146).

30Andy Pawley suggests that this comment probably applies only to linguists, who have been taught to view the sound-meaning association as arbitrary; to most non-linguists, the view I have been arguing for is self-evident.
Complex phonological formants

4

Complex phonological formants

4.0 Introduction

In 2.0 I described a hierarchy of non-prosodic phonological units, from features to morphs:

(1)    morph
        |  segment
        |  feature

This chapter introduces the description of the highest level in that hierarchy, the morph. Paralleling and interacting with this hierarchy is the hierarchy of prosodic units, including syllable, metrical foot and phonological word, described in chapter 5.

Morphs then are configurations more complex than individual segments, which behave phonologically as distinct units. While they can typically be described in terms of their segmental composition, morphs are also subject to prosodic constraints, and as a result are generally disyllabic or at most trisyllabic (5.6). Morphs play an important part in the signalling of meaning in Balinese: a single morph normally realises a single lexical morpheme, a form-meaning association, (6.0). So for example the morpheme gedé 'big' is realised by the morph /gade/, which is linked lexically with the meaning 'big'.

Nevertheless, morphs are not always clearly meaningful in Balinese. They can occur in their primary form, as meaningless, purely phonological units. When they do so, they combine with other phonological formants, in compound-like structures, to form morphemes (6.0). The following are examples of morphemes which consist of two such concatenated morphs:

(2)    katilambung /kati-|lambug|/   'ricochet'
       ketumbah /ke-|tumbah|/   'coriander'

When morphs are used in this way, as purely phonological 'building blocks', I call them complex formants. Others use the term 'meaningless morpheme' or simply 'morpheme', eg Aronoff (1976).

Complex formants provide especially clear evidence for the morph as a phonological unit. I present some of that evidence in 4.1. In this chapter I also give evidence for the purely phonological functioning of processes like concatenation, and reduplication (4.2).

1I use the symbol '||' to indicate material presumed to function as a complex formant.
Chapters 5 and 6 describe prosodic and other constraints on the phonological shape of morphs, be they meaningful or not.

Morphs then may or may not be associated with meaning. A recurrent theme of this thesis will be that meaning can be assigned at any level of linguistic structure - from feature (eg 2.1.2, through to clause level (chapter 13)). A corollary of this is that - at the lower levels in particular, from feature through segment to morph - meaning is often not assigned to such linguistic units, but rather to the higher units they form part of (cf also 15.1, 15.2).

In the following subsection, I concentrate on demonstrating the reality of the morph, in its purely phonological function. It should not be forgotten that the term 'complex formant' is primarily a descriptive convenience, referring simply to one kind of morph, one which happens not to have an independent meaning-signalling function.

I also use the term 'complex formant' to refer to processes, such as inherent reduplication, which have a purely phonological function, cf 4.2. I thus see processes such as concatenation and reduplication also applying 'below' the level of the morpheme, when complex formants concatenate to produce complex morphemes. Perhaps even more than is the case in English (cf Aronoff), a great many Balinese morphemes contain complex formants. The following are examples of morphemes containing different types of reduplication as a complex formant:

(3)  
| gigi | 'tooth' | -selsel | 'regret (vi)' |
| kekupu | 'butterfly' | kupukupu | 'butterfly' |
| kedongdong | 'k.o. fruit' | kalibukbuk | 'place name' |
| crorot | 'k.o. cake' | buléleng | 'place name' |

Evidence for inherent, non-meaningful reduplication is given in 4.2, and 8.9.1.

The underlying representation of a morpheme like kupukupu can be thought of then as consisting of two complex formants, [RED] and [kupu]: a segmental formant /kupu/, associated with a reduplication 'diacritic', indicating that /kupu/ is to be reduplicated. In the notation /|RED|·|kupu|/, the symbol '·' indicates that the reduplicated copy is concatenated like a prefix onto the base (see below and 8.0). Similarly, I analyse morphemes like kedongdong and kalibukbuk, which contain a fully reduplicated final syllable, as concatenations of prefix-like complex formants [kə] and [kəli] (see 4.1), plus distinct monosyllabic formants, [dɔŋ] and [buk], which undergo full base reduplication (8.3).

As a descriptive convenience, I use the symbol '|...|' to indicate material presumed to function as a complex formant: thus kedongdong /|kə|·|RED|·|dɔŋ|/. Where there is no evidence to the contrary, I assume that the material concatenated to a complex formant is itself also a complex formant: the bases which undergo inherent reduplication are thus treated in this way.
I claim that complex formants are *meaningless*, in the sense that they are not associated with conventional lexical meaning, as are lexical morphemes. At the same time, however, the use of complex formants to create morphemes serves to create expressive lexis (see 4.1), 6.4). However, I take the expressive semantic component associated with complex *morphemes* to be generated by their phonologically aberrant status, rather than by a conventional derivational process (3.4.2).

### 4.1 Segmental complex formants

Abundant evidence is presented in 4.2 that in inherently reduplicated morphemes, reduplication is a meaningless, purely phonological formant. Not surprisingly, Balinese also has purely segmental complex formants similar to those Aronoff (1976) argues for in English. Candidates for this status include the prefix-like sequences *ke*, *c(l)e*, and *kali ~ kali ~ katu*. I take it that the *'base'*-like sequences they concatenate with are also complex formants. I have found no evidence for *'meaningless suffixes'*.

Morphemes listed in Warna et al (1990) which contain the initial sequences *ke*, *c(l)e*, and *kali ~ kali ~ katu* are almost exclusively classifiable as expressives. There are for example 109 /ka/-initial trisyllables listed in Warna, and 29 with initial /kati/ ~ /kali/ before a disyllabic sequence. (The latter make up most of the native, non-reduplicated, quadsyllabic morphemes in the language, see 6.1.1.) The following are typical examples:

(4) ketekok  'clucking sound of hen' -ketekéh 'grumble'
kejengit 'bare one's teeth (eg laughing)' ketepuk 'small drum'
kedropon 'hasty, careless' ketengkung 'song of dove'
kemombo 'large bud just about to open' ketumbah 'coriander'
kemaruk 'ravenously hungry' keslab 'startled'

(5) kalialah 'rainbow' kaliadrem 'k.o. cake'
katulayah 'lie stretched out, helpless' katilambung 'ricochet'
katulajeg 'stand erect, in rows' kaliasem 'k.o. tree'
katimumbulan 'have a disease (of the fingers)'
kalimayah 'many-legged insect, shines in dark'

Given that in Balinese around 90% of morphemes contain only two syllables (6.1), these initial /ka/ and /kati/ ~ /kali/ sequences are prima facie suspiciously like *'prefixes'* on disyllabic bases. Lexemes containing these sequences have the broad range of meanings associated with expressives: iconics, plant and animal names, *'bad'* things and actions, and so on (3.2.2), and so could conceivably be complex. As with inherent reduplication, there seems to be no recurrent semantic element in these forms (apart from their expressiveness) which could justify analysing /ka/ or /kali/ ~ /kati/ as morphemes. There is then a case for assigning *'meaningless complex formant'* status to them.
Data like that in (4) and (5) is nonetheless problematic. The expressive nature of these morphemes could derive from fact that they are complex (cf rule (57), section 6.4) containing a 'meaningless prefix'. On the other hand, it could derive simply from their multisyllabic structure, itself marked in Balinese (6.1).

However a good number of these /ka/-initial and /kali/-~/katV/-initial morphemes are clearly complex, since they contain reduplicated sequences. Where they occur morpheme-finally, the reduplicated elements carry two primary stresses, strong evidence of synchronic reduplication (see 8.8):

(6)  kebilbil 'ashamed'  kedukduk 'k.o. shrub'
    keppir 'k.o. bird'  kedongdong 'k.o. tree'
    ketoktok 'k.o shrub' -kedokdak 'boil, bubble (vi)'
    ketestes 'freshwater shrimp' -kediding 'scurry'
    kenyongnyong 'k.o. tree'  kecalcil 'k.o. small fruit'
    kemangmang 'k.o. dangerous spirit' kebetbet 'butterfly'
    kebabas 'lose one's way'

kalikukun 'k.o. tree'  kalimatmatan 'have blurred vision'
kalisasuan 'k.o. wasp' kalibukbuk '(place name)'
katibangbung 'k.o. brightly coloured beetle'
katiwawalan 'jackfruit bud which doesn’t become a fruit'

Moreover, these, together with lexemes containing the expressive bound morpheme of shape C\text{laminal}re-/C\text{laminal}le-, are virtually the only sequences listed in Warna et al (1990) as occurring before inherently reduplicated, closed syllables (those of structure CVC): 26 out of the 28 items are of these shapes. A full list is given in Appendix D. This limited distribution is strong independent evidence for 'recurrent partials' of shape /ka/ and /kali/-~/katV/. Extending this analysis to the data in (4) and (5) is thus plausible.

There are moreover many disyllabic morphemes with initial /ka/, and others with initial /ca/, which are both semantically expressive, and which show other phonological behaviour suggestive of complex status.

Evidence for a complex formant ce- includes the many onomatopoeic nominals occurring with an optional initial /ca/ sequence. Warna et al (1990) lists amongst others:

(7)  cebug, bug 'thudding sound'
    cegir, gir 'sound of a gong reverberating'
    cekoh, koh 'coughing sound'
    cekug, kug 'sound made by a falling object'
    cekut, kut 'hiccoughing sound'

\footnote{Warna et al (1990) lists several lexemes of shape C\text{laminal}re-/C\text{laminal}le- [ + reduplicated base]:
  cre-/jre-/jle/-fire/-firele/-are/-ale-; all have meanings involving 'adopt a bad (eg awkward) body position':
  firegaogag 'sleep with legs part', firegaogag 'sit in silence', flepapph 'slump down onto the ground'.
  The variation is dialectal: only jre-/jle- (~ fire/-firele-) occur in Peguyangan.
Given just this limited set of examples one would be tempted to conclude that the [ca] sequence does have meaningful status. I will argue against such an analysis in the rest of this section. The glosses of the verbalised forms derived from those above suggest that the /ca/ sequence is interpreted as constituting a distinct, apparently meaningless, element:

(8)  
me-cebug  'fall and make a 'bug' sound'
me-ceped  'make a 'ped' (slapping) sound'
me-cekug  'fall and make a 'kug' sound'
me-cedig  'fall and make a 'dig' sound'
me-cengér  'cry, making a 'ngér' sound'
me-cekut-an  'make a 'kut' (hiccough) sound'

Analogous /ka/-initial forms include

(9)  
me-kepot  'make a 'pot' (popping) sound'
me-keplek  'make a 'plék' sound'
me-kepret  'make a 'pré' sound'
me-kesieg  'make a 'sieg' sound'
ngering  'make a ring-ring noise (of a cricket)'
ngering  'make a ring-ring noise (of a cricket)'
kepuak-an  'bamboo clapper, which makes a 'puak' sound'

A related set of lexical roots apparently combines a meaningful monosyllabic element with either initial /ke/ (or, in some cases, /ca/) or reduplication to form expressive doublets, mostly iconic of movement:

(10)  
kebés, básbés  'rip, tear (vt)'
kebis, bísbis  'rip, tear (vt)'
kedet, detdet  'tug (vt)'
ketek, tektek  'chop up (vt)'
ketog, totong  'give a sharp blow to (vt)'
kepak, pakpak  'chew (vt)'
kepik, pikpik  'pick, snap off (vt)'
-ketél, -télél  'drip (vi)'
kepus, puspus  'pick, break off (vt)'
kesét, sétsét  'torn (vi), tear (vt)'
kesit, sitisit  'make small tear(s) in (vt)'
cegut, gutgut  'bite (vt)'
cekoh, kohkoh  'coughing sound (n); cough (v)'

There is a regular meaning contrast in the pairs in (10): the reduplicated forms indicate a repeated action, while the forms with initial /ke/ ~ /ca/ indicate a single action. Since reduplication is used meaningfully to mark durative/repetitive aspect (see (18c)), its occurrence in these lexemes could be considered to manifest a meaningful derivational process. The lexically conditioned /ke/ ~ /ca/ alternation might then be analysable as a morpheme/realisation of punctiliar aspect.

These conclusions are however problematic, and I prefer to analyse pairs such as those in (10) as not reflecting synchronic derivational processes in the language. This is for several reasons, outlined in the following paragraphs.
1) Some verbal bases contain both the sequence /ka/ and reduplication: -kecigcig 'scurry', kebibil 'ashamed', -keséksék 'twitch uncontrollably', -ketabtab 'move legs suddenly because startled', -ketaktak 'suddenly weep loudly'. We would expect the two elements to be incompatible, hence never cooccur, if they both were meaningful.

2) Elsewhere the sequences /ka/- and /ca/ occur on a wide range of lexemes which have other, non-punctiliar, but still 'expressive' meanings (see also examples (4)).

(11) cekel 'hold between thumb and first finger' \[\text{celeg} \text{ 'misshapen'}\]
     cekok 'having a dent, depression, curved inwards' \[\text{cenikih} \text{ 'k.o. snake'}\]
     cepud 'have no descendants' \[\text{cenés} \text{ 'smooth, neat'}\]

3) If this is a kind of aspectual marking, operating on (what would be) a bound morpheme to derive a lexical base, it is not productive, at least not in terms of the marking of punctual aspect. A great many lexemes thus occur only in the reduplicated form, with no /ka/- or /ca/-initial counterpart. Examples include:

(12) batbat 'insult (vt)' \[\text{nengneng} \text{ 'stare at'}\]
     jagjag 'approach' \[\text{déngdén} \text{ 'dry (vt)'}\]
     kikhikh 'grate (vt)' \[\text{congcong} \text{ 'dig (with)'}\]
     daldal 'beat' \[\text{sengseng} \text{ 'plug (vt)'}\]
     gulgul 'bother (vt)' \[\text{gésgés} \text{ 'scratch (vt)'}\]

And many others occur only with a /ka/- or /ca/-initial sequence, but no corresponding reduplicated form:

(13) celek 'poke finger into' \[\text{keleg} \text{ 'bend (vt)'}\]
     kecong 'come apart, slip off handle' \[\text{cepol} \text{ 'collapse'}\]
     cegik 'angrily rebuke' \[\text{kedeng} \text{ 'tug, pull (vt)'}\]
     cegus 'sniff' \[\text{kepéh} \text{ 'split on side (vi, vt)'}\]
     cegeg 'punch repeatedly' \[\text{kepung} \text{ 'pursue (vt)'}\]
     celep 'enter' \[\text{kerik} \text{ 'scrape (vt)'}\]

4) Lexical bases like kebés and bésbés, cegut, gutgut are analysed by native speakers as monomorphemic. They receive distinct listings in dictionaries, and to my knowledge no previous studies have pointed out the relationship between such pairs (see for example Bawa et al 1984, Djendra et al 1975.)

5) Apart from the derivational use of reduplication, aspect is always expressed by distinct grammatical words, rather than being marked on the verb itself: /ka/- (~ /ca/) does not occur elsewhere, say as a preverbal particle, to mark punctual aspect.

I conclude that while the process of reduplication forming bases like -télél 'drip' and gutgut 'bite' could be described in terms of the regular morphological processes in the language, forms like -ketél 'drip' and cegut 'bite' are better analysed as monomorphemic, with underlying structures such as the following:

(14) -ketél \[\% /\text{ka} / -tel/\%
     cegut \[\% /\text{ca} / -gut/\%

74
Such lexical bases thus consist of a meaningful bound morpheme, plus a meaningless initial syllable, as opposed to possibly bi-morphemic -télétél and gutgut.

It is plausible to extend the analysis and conclude that these complex formants are present in some or all of the many other disyllabic expressives which begin with /ka/ or /cé/, but do not have a reduplicated counterpart (see for example (13) above).

Further evidence for /ka|, |cé| and |kali| ~ |kati| ~ |katu| comes from 'final-syllable reduplication' (8.4), which, only targets bases beginning with these sequences. In combination with the prefix pe-, it targets expressives commencing with the sequences /ka/ and /cé/, to derive a 'plural' meaning. Examples include:

| (15) | ketél | 'drip' | peketélél | 'id. (pl.)' |
| -cebug | 'make a thudding noise' | pecebugbug | 'id. (pl.)' |
| -cebwah | 'be bright and shining' | pecebwahbwah | 'id. (pl.)' |
| -kebyar | 'flash' | pekebyarbyar | 'id. (pl.)' |
| -cepol | 'fall' | pecepolpol | 'id. (pl.)' |

While there is a lexical base télétél, in all other cases the reduplicated final syllable does not occur independently as a lexical base.

Final syllable reduplication also occurs in combination with the 'agentive' prefix, {N}, on a limited group of /ka/-initial verb bases. These bases independently occur with both the 'applied' suffix -ang (11.3), and with full reduplication. Compare the following related sets:

| (16) | ngebir, kebir-kebir, ngebirbir | 'move back and forward in the wind' |
| ngebet, kebet-kebet, ngebetbet | 'throb repeatedly with pain' |
| kesir-kesir | 'blow gently (of wind)' |
| ngesirsir | 'blow gently continually (wind)' |
| keteb-keteb | 'stamp feet repeatedly' |
| ngetebteb | 'stamp one's feet repeatedly and quickly' |
| (...) | ngémélang | 'move (lips, to speak)' |
| kemél-kemél | 'open and close (eg lips, wound)' |
| ngémélémél | 'continuously talk to oneself' |
| ngemigang | 'move (lips, to speak)', |
| kemig-kemig | 'open and close (mouth, to speak)' |
| ngemigmig | 'mutter to oneself' |
| ngemikang | 'move (lips, to speak)', |
| kemik-kemik | 'open and close (mouth, to speak)' |
| ngemikmik | 'mutter to oneself' |

It was also found in the inherently reduplicated forms discussed above, such as:

3Warna et al (1990) also lists -jegir 'stand out in an elevated position' pejegirgir 'stand up tall (pl.), eg of houses'.
The crucial thing about this morpheme-internal reduplication is that it only occurs on morphemes beginning with either /ka/, /ca/, or /kali/ ~ /katâ~ /katu'. This restriction is further evidence that these sequences have a special status in the language, which I take to be that of 'complex formant'.

In conclusion, a variety of evidence, both semantic and phonological, supports the assignment of complex formant status to certain occurrences of morpheme-initial /ka/, /ca/ and /kati ~ kali ~ katu' sequences. These, and the bases they attach to, are by no means the only such sequences having this status - see Appendix F, as well as 4.2.

4.2 Inherent reduplication as a purely phonological process

Balinese has many lexemes which, while apparently single morphemes, contain reduplicated elements, such as memê 'mother', pipi 'cheek', jirji 'finger', tektek 'chop up (vt), cruükuk 'k.o. bird'. I refer to these as 'inherently reduplicated' morphemes (8.8).

Prince (1987), using an interpretation of the notion 'morpheme' similar to that of Aronoff (1976), claims that reduplication always 'implies bimorphemic or derived status'. I give evidence here that in Balinese reduplication can have either a purely phonological function (inherent reduplication), or a morphological one. The two types show different patterns of both semantic and phonological behaviour. A variety of evidence, from semantics, and from phonological patterning, supports the view that reduplication should be seen as representing a purely phonological process, only where it occurs in inherently reduplicated lexemes.

Reduplication is regularly used in a morphological way, to mark categories such as plurality, durational aspect, and to derive words from others. For example, reduplication of nouns (optionally) marks plurality, (18a), reduplication of intransitive stative verbs marks either plurality of the subject noun (18b), or durative aspect (18c) (cf me-kecuh 'spit (once)').' Examples of words regularly derived from other words by reduplication include (18d, e & f):

(18) a. bikul-bikul 'RED-rat: rats'
b. bikul gedé-gedé 'rat RED-big: big rats'
c. suud-ja kecuh-kecuh 'stop-PART RED-spit' 'Stop spitting [DUR]/'
d. mobil-mobil-an 'RED-car-an: toy car'
e. aba-aba-an 'RED-carry-an: 1. way of carrying, 2. bearing'
f. je-jait-an 'RED-sew-an: (things made by) needlework'

The reduplicational processes illustrated in (18) assign meaning in regular, predictable ways, to independently occurring morphemes. The semantic evidence shows that inherent reduplication does not assign meaning in the way that morphological processes
do. Apart from being consistently 'expressive' (in the sense of 3.2.2), morphemes containing inherently reduplicated elements have no appreciable common meaning element. They include both nouns of different semantic types:

(19)  

<table>
<thead>
<tr>
<th>Morpheme</th>
<th>Meaning</th>
<th>Morpheme</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>klungkung</td>
<td>'place name'</td>
<td>sirit</td>
<td>'place name'</td>
</tr>
<tr>
<td>nyonyo</td>
<td>'breast'</td>
<td>gigi</td>
<td>'tooth'</td>
</tr>
<tr>
<td>canguku</td>
<td>'grandchild'</td>
<td>mémé</td>
<td>'mother'</td>
</tr>
<tr>
<td>kedongdong</td>
<td>'k.o. fruit tree'</td>
<td>cekcek</td>
<td>'house lizard'</td>
</tr>
</tbody>
</table>

and verbs, transitive and intransitive:

(20)  

<table>
<thead>
<tr>
<th>Morpheme</th>
<th>Meaning</th>
<th>Morpheme</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>kisiksi</td>
<td>'whisper'</td>
<td>fitrekongkong</td>
<td>'sit with legs apart'</td>
</tr>
<tr>
<td>nyahnyah</td>
<td>'dry fry (vt)'</td>
<td>caccomcacas</td>
<td>'criticise'</td>
</tr>
<tr>
<td>sepsep</td>
<td>'suck'</td>
<td>cakcak</td>
<td>'chop up, pulverise'</td>
</tr>
<tr>
<td>mrésmés</td>
<td>'messy (with food)'</td>
<td>nangdanangda</td>
<td>'dither'</td>
</tr>
</tbody>
</table>

The above examples typify the expressive semantic classes inherently reduplicated morphemes consistently belong to. As argued in chapter 3, this is to be expected, since they 'violate' a preferred rule (6.4) by their complex status. A morphemic analysis of expressive meaning assignment is considered and ruled out in 3.4.2.

Table 4.1 below gives the the semantic classes of 60 inherently reduplicated disyllabic morphemes, such as lumbalumba 'dolphin', randomly selected from Warna et al (1990). The complete data and analysis is given in Appendix E.

<table>
<thead>
<tr>
<th>Class</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLANT/ANIMAL</td>
<td>12</td>
</tr>
<tr>
<td>BAD</td>
<td>22</td>
</tr>
<tr>
<td>MOVEMENT (SENSE)</td>
<td>7</td>
</tr>
<tr>
<td>SOUND (SENSE)</td>
<td>3</td>
</tr>
<tr>
<td>Total 'criterion' expressives</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>73%, p &lt; .0000 (Median Test)</td>
</tr>
<tr>
<td>VERY</td>
<td>1</td>
</tr>
<tr>
<td>CAKE</td>
<td>3</td>
</tr>
<tr>
<td>SMALL</td>
<td>2</td>
</tr>
<tr>
<td>PLURAL</td>
<td>2</td>
</tr>
<tr>
<td>LOANS</td>
<td>3</td>
</tr>
<tr>
<td>Total 'other' expressives</td>
<td>11</td>
</tr>
<tr>
<td>Total expressives</td>
<td>55 (91.7%)</td>
</tr>
<tr>
<td>Unclassified</td>
<td>5     (8.3%)</td>
</tr>
</tbody>
</table>

Table 4.1 Semantic classes of inherently reduplicated morphemes with disyllabic bases, data from Warna et al (1990).

As the table indicates, a very high proportion (p < .0000, Median Test) of the sample falls into just the BAD, PLANT/ANIMAL and SENSE classes (defined in 3.2.2), following the 'typical' expressive pattern. The procedure for the statistical analysis is that described in 3.2.3.

The semantic classes of a similar random sample of monosyllabic reduplicated forms where the base is of shape CVC, such as laklak 'k.o. cake', are given in Table 4.2. Again, the complete data is given in Appendix E.
Table 4.2 Semantic classes of inherently reduplicated monosyllabic bases, data from Warna et al (1990)

The patterns above indicate that the only common meaning element shared by inherently reduplicated morphemes is their 'expressiveness'. In 3.4.2 I propose an explication of this expressive meaning component, and argue that it is generated directly by the phonological form, rather than being assigned lexically, as in the case of morphological processes.

Phonological evidence that lexemes containing inherent reduplications behave as a single morpheme, rather than as bi-morphemic, is of five kinds: 1) persistence of contour segments linking identical vowels after inherent reduplication; 2) illegality of derived geminate consonant sequences in inherent reduplications 3) regularity of inherently reduplicated forms with respect to L-slot constraints (6.2.2); 4) failure of inherently reduplicated morphemes to undergo foot-reduplication; 5) postponement of inherent reduplication of disyllabic bases until the postlexical phonology. These are discussed in turn in the rest of this section.

1. In reduplication plus vowel overwrite the base is copied, but the final vowel in the copy is replaced by /a/ (8.7):

\[
\begin{array}{lll}
\text{-kelip} & \text{twinkle} & \text{kelap-kelip} \\
\text{tolih} & \text{look back} & \text{tolah-tolih}
\end{array}
\]

'twinkle on and off' 'look back repeatedly'

When a base has adjacent identical vowels, both are obligatorily replaced by /a/ in the copy, suggesting they share a single feature specification:

\[
\begin{array}{lll}
\text{-gliling} & \text{roll (vi)} & \text{galiang-gliling} \\
\text{-sruyung} & \text{sway} & \text{srayang-sruyung}
\end{array}
\]

'roll back and forth' 'sway back and forth'

Inherently reduplicated bases behave like single morphemes when they undergo further reduplication plus overwrite - both vowels are replaced:
Complex phonological formants

Separate morpheme status implies non-sharing of phonological structure. If forms like *dugdug, -klidkid* constitute two morphemes, the vowels, though identical, will not share feature specifications: the non-occurring forms *dugdag-dugdug, *klidkad-klidkid* would then be produced by reduplication plus vowel overwrite. The fact that such forms are not produced indicates that bases like *dugdug* and *klidkid* containing inherently reduplicated sequences are also single morphemes: reduplication does not imply bimorphemic status.

2. A second characteristic of inherently reduplicated bases also supports their monomorphemic status. This is the constraint (see 6.2.5) barring derived geminate sequences in inherently reduplicated morphemes of shape \( C_1 V C_2 C_1 V C_2 \) (such as cadcad 'criticise'). The medial \( C_2 C_1 \) cluster can apparently consist of any two consonants, except two identical consonants:

\[
* C_{[\text{place}]} C_{[\text{place}]} 
\]

McCarthy (1986) describes how this 'antigemination' effect holds in a wide variety of languages, but only *morpheme-internally*.\(^4\) Across a morpheme boundary, the same languages typically allow sequences of identical (though non-linked) segments. Balinese, significantly, does this regularly, for example in the sequence *-n né* 'GEN-3' (discussed also in 2.2.3):

\[
(25) \quad \text{aji-n-né} \quad [\text{ajrnné}] \quad \text{'father.HON-GEN-3; his/her/their father(HON)'}
\]

\[
\text{mata-n-né} \quad [\text{matanné}] \quad \text{'eye-GEN-3; his/her/their eye(s)'}
\]

This combination of facts is again strong evidence for the monomorphemic status of inherently reduplicated morphemes in Balinese, and thus for the purely phonological function of reduplication, when it occurs in that context.

3. The third type of phonological evidence for the purely phonological function of inherent reduplication comes from constraints on the L-slot. In (6.2.2), the following constraint is described:

\[
(26) \quad \text{Prefer morphs which contain only one filled L slot}
\]

Where \( \text{L} \) is /l/ or /rl/, only three unambiguously monomorphemic (ie non-reduplicated) exceptions to the constraint are listed in Warna et al (1990), all loanwords (see 6.2.2). Similarly there are only four exceptions involving inherent reduplication of a monosyllable:

\(^4\)Antigemination thus applies to consonants, but not to vowels in Balinese (cf. the preceding discussion of forms like *dagdag-dugdug*). While that fact is curious, it is does not negate the validity of the arguments being made here.
In all other cases of inherent reduplication of a monosyllabic base containing an L-slot (of which there are many), only one L is filled in the surface form:

(28)  
| prungpung | % | RED | prun | % | 'chipped (of a pot)' |
| -closcos | % | RED | clos | % | 'hop about' |
| gruguh | % | RED | gruh | % | 'hand (crude term)' |
| klakah | % | RED | klah | % | 'roof of bamboo shingles' |

If reduplication implies bi-morphemic status, as argued by Prince, the apparent adherence of inherently reduplicated forms to the L-slot constraint is inexplicable. If reduplication here is purely phonological, their adherence to that constraint can be explained.

4. The fourth type of phonological evidence indicating that inherent reduplication is different in kind from reduplication as a derivational process comes from foot reduplication. Meaningful bases (morphemes) can undergo either full-base reduplication (8.3) or foot reduplication (8.5), whereas the meaningless bases targeted by inherent reduplication never undergo foot reduplication.

Foot reduplication copies a purely phonological unit, with a closed final syllable. However it reduplicates only those bases containing (at least) a morpheme, which occurs elsewhere in a non-reduplicated context. Foot reduplication never targets bases containing 'just' a complex formant. Thus sequences such as:

(29)  
| oka-n-né | paling jelé-na | ngedé-nang | beli-nin | 'child.HON-GEN-3; his child.HON' |
| (%) | gade-an | (%) | 'most bad-IRR; (the) worst' |
| (%) | 'AP-big-APP; become bigger' |
| (%) | 'buy-LOC; buy from' |

occur in two possible reduplicated forms, either undergoing full base reduplication:

(30)  
| oka | oka-oka-n-né | paling jelé-jelé-na | ngedé-ngedénang | beli-beli-nin | 'his children' |
| (%) | 'the) worst (PL)' |
| (%) | 'become bigger and bigger' |
| (%) | 'repeatedly buy from' |

or foot reduplication (with no meaning difference):

(31)  
| oka | okan-okan-né | paling jelén-jelén-a | ngedén-ngedén-ang | belin-belin-in |

On the other hand, inherent reduplication produces only one possible surface form in the same contexts. Morphemes like lumbalumba 'dolphin' and nangdanangda 'shilly-shally, dither' never have alternate forms produced by foot reduplication:
Complex phonological formants

(32) | Lumba | lumbalumba-n-ne  'his dolphin'
  *lumban-lumban-né
| Nangda | paling nangdanangda-na 'most dithering(adj.)-IRR'
  *paling nangdan-nangdan-a
  nge-nangdanangda-nang 'dither more & more'
  *nge-nangdan-nangdan-ang

This is despite the fact that, phonologically and syntactically, disyllabic complex formants like _lumba_ and _nangda_ are largely indistinguishable from true morphemes,5 undergoing processes such as cliticisation and lowering at exactly the same stage. Compare the derivation of a) inherently reduplicated _picépicé-n-ne_ ( %RED | pici | %n-ne), 'his shellfish' and b) derived _béle(n)-béle-n-ne_ 'his elder brothers' (from _béli_ 'elder brother' %béli-n-ne %, elder.brother-GEN-3):

(33) Input: a. _pici_ b. _béli_  'elder brother'
    Cliticisation: picinne belinne  'id.-GEN-3'
    Lowering: picenne belenne
    Root-Red: picepicenne belebale nne
    Foot-Red: - belebale nne

Foot reduplication thus fails to apply to the complex formant _pici_, where it does apply to the morpheme _béli_. This is the only difference in the derivations, and apparently derives purely from the difference in morphemic status.6

5. The fifth type of phonological evidence also comes from the reduplicational possibilities of disyllabic bases. Where these bases are themselves separately occurring morphemes, they have two possible phonetic realisations, suggesting two distinct mechanisms of reduplication, one postlexical, the other lexical (cf 8.9):

(34) _alit_ [alit]w  'small.HI'
_alit-alit-é_ [alitelte]w  'RED-small-DEF; children.HI'

However where the bases only ever occur in reduplicated form (in inherent reduplications), only the postlexical option is available:

(35) _anting-anting-e_ [antingante]w  'earring'
  *[anting]w[anting]w

Again there are very different possibilities, depending on whether reduplication serves a morphological or a purely phonological function. See 8.9.1 for further discussion and data.

5 The one difference in phonological behaviour is discussed in subsection (5), below.
6 It might be argued that, since inherently reduplicated forms are lexically specified as undergoing full-base reduplication, their failure to undergo foot reduplication is a natural consequence of the Elsewhere Condition (Kiparsky 1982). However this still does not account for the difference in distribution of the two reduplication types.
7 On the distinction between lexical and postlexical phonology, see 2.0.
Other evidence that inherent reduplication is not a morpheme comes from native speaker intuitions: except in the case of reduplication of a disyllabic base, reduplications of monosyllabic bases, such as the following, are not generally perceived of as involving reduplication:

<table>
<thead>
<tr>
<th>Base</th>
<th>Meaning</th>
<th>Base</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>cucu</td>
<td>'grandchild'</td>
<td>cakcak</td>
<td>'chop up'</td>
</tr>
<tr>
<td>bubuh</td>
<td>'rice porridge'</td>
<td>loloh</td>
<td>'herbal drink'</td>
</tr>
<tr>
<td>-ëlél</td>
<td>'gaping, wide open'</td>
<td>pipi</td>
<td>'cheek'</td>
</tr>
</tbody>
</table>

This applies even in previous linguistic analyses, like Jendra et al (1977:126) which discusses only the reduplication of disyllabic bases. There are in fact many morphemes derived from reduplicated mono-syllabic bases. A variety of evidence that reduplication of monosyllabic bases is a synchronic process is given in 8.8.

Although the above examples do not show it, the standard orthography recognises inherently reduplicated tri- and di-syllabic sequences by writing them with a hyphen -gumatat-gumitit 'creepy crawlies', dangap-dangap 'k.o. lizard' - reduplication of smaller sequences is not signalled by the standard orthography: again speaker intuitions do not recognise reduplication here. (In the modified orthography used in this thesis, I adopt a more consistent practice: I do not insert a hyphen between base and copy in inherently reduplicated morphemes, no matter what the length of the base.)

To sum up, there is a wide variety of evidence that lexemes containing inherent reduplications behave as single morphemes. On can thus make a principled distinction between the morphological function of reduplication (exemplified at (18)), and the purely phonological one found in inherent reduplications.

This conclusion is consistent with the finding in 4.1 that morphs, while usually directly realising morphemes, can also occur as purely phonological formants.
5 Prosodic structure

5.0 The prosodic hierarchy

McCarthy & Prince (1990) argue that the following hierarchical prosodic structure is relevant to word-formation processes:

(1) P-Word
    | Foot
    | syllable
    | mora

In Balinese morphemes (and complex formants) must meet constraints on prosodic well-formedness at least at the syllable and foot levels. I did not find evidence that the mora plays a distinctive role in the phonology (5.1.1). The P-word is the relevant domain for several processes, including syllabification and phoneme allophony. At a higher level, there is evidence from stress patterns (5.8) for the phonological phrase (eg Nespor and Vogel (1986)).

These prosodic units (rather than grammatical units such as morphemes and so on) are the domains of purely phonological rules such as stress assignment and syllabification. Other processes, such as Foot reduplication (8.5) require information from both morphological and prosodic structure.

5.1 Syllable structure

The syllable in Balinese is minimally a single vowel, and maximally a sequence of shape CLVC (where L is a liquid or a glide). As a descriptive convenience only, I use the following structural labels:

(2) σ
    / \ onset rhyme
    / \ / \ C_i L nucleus coda
    V C
Chapter 5

The only obligatory element is the vowel nucleus. Each vowel potentially represents a distinct syllable nucleus, though the [+HI] vocoids can surface as either vowels or (onset) consonants, depending on syllabification processes (cf 2.1.3):

(3) baang 'give' /ba.an/  
keet 'cut up' /ke.at/  
byaung 'k.o. grass' /baun %, /bya.un/  
cai 'you (2sg)' /ca.i/

Many phonotactic constraints on underlying representations are in fact constraints on well-formed syllables. I assume that in Balinese organization into syllables holds at all levels of phonological representation (see eg 5.3.2). Syllables of shape V, CV, VC, CVC, CLV, and CLVC all occur:

(4) V.V  
V.CV  
VC.CVC  
CLV.CVC  
CVC.CLVC

5.1.1 Contrastive syllable weight?

Here I review evidence as to whether Balinese has a syllable weight contrast, between light and heavy syllables. I conclude in the negative, though tentatively so. Such a contrast would indicate that the mora, a prosodic unit below the level of the syllable, is active in the phonology of Balinese (Hayes 1989). Going against this analysis is the lack of evidence in Balinese for the following features, all typical of quantity-sensitive languages (Hayes 1989:259):

1. Balinese lacks contrastive vowel length (2.2), said to be typical of such systems.
2. There are no geminate consonants (6.2.5).
3. High vowels and glides do not contrast underlyingly (2.1.3)
4. Processes of compensatory lengthening do not occur.

As well, stress, which is often influenced by syllable weight, falls regularly on the word-final syllable, independently of whether it is closed or not (5.8).

Possible evidence in favour of a syllable weight distinction includes evidence from reduplication processes and preferred morpheme shapes.

1. The minimal unit to which CV-RED partial reduplication applies is a base consisting of a monosyllable, with a filled coda (8.2.1). 'Lighter' syllables, of shape CV, always undergo full base reduplication.
2. Derivational reduplication sometimes targets a base non-coextensive with morpheme boundaries. I call this Foot reduplication (8.5):

\[1\text{It is possible that at a late stage in the postlexical phonology, re-syllabification ceases: see 8.9.}\]
(5)  a. gedé-nang  'big-APP; enlarge (vt); get bigger (vi)'

    base-red: gedé-gedénang
    Foot red: gedén-gedénang

b. bé-n-né  'fish-GEN-3; his/her/their fish'

    base-red: bé-bé-n-né
    Foot red: béné-bé-n-né

This could be seen as evidence that this type of reduplication involves a prosodically determined iambic foot template (following McCarthy and Prince 1990), which implies a heavy-light syllable distinction. However an alternative analysis, involving a preferred extrametrical final consonant in the reduplicative template, is not ruled out (5.3.2, 5.6).

3. Another source of possible evidence for a light/heavy syllable weight distinction comes from morpheme structure preferences. Only 28 lexical morphemes of shape CV are listed in Warna et al (1990), while more than 430 of shape CVC are found, 6.1. This could be seen as evidence for a dimoraic minimal limit to morph structure, given that disyllabic morphs make up around 90% of the lexicon (6.1.1: the CVC type would be non-exceptional if it were heavy. A dimoraic foot would then be the minimal word in the language, while the iambic foot (CVCVC) would be the preferred prosodic word shape (hence its use in foot reduplication, and in preferred morpheme structure, 6.1.

A problem for this analysis is the evidence that morph-final consonants are extrametrical underlyingly, 5.3.2, hence would not contribute a mora to syllable weight. In that case the monosyllabic morphemes are simply exceptions to the preferred disyllabic shape.

It appears from this review that there is no unambiguous evidence in favour of contrastive syllable weight in Balinese. I therefore use a neutral terminology in this description.

5.2 Onset-internal constraints

5.2.1 The onset slot

I assume that in Balinese the following rule holds:

(6)  Prefer syllables which have a filled onset (\(\text{\textgreek{C}}\) ) slot

Evidence that this constraint operates comes from frequencies of occurrence of morphemes of different shapes. The following figures, set out in 'minimal pairs', are representative:
### Table 5.1 Frequencies of occurrence of syllable-types in morpheme-initial position (data from Warna et al. (1990)).

For this sample the overall counts are 1149 V-initial syllables, versus 6351 C-initial ones, or a ratio of approximately 6 to 1 in favour of filled onsets. Similar results are obtained for pairs differing only in whether the second syllable has a filled onset or not, indicating that all these figures reflect a syllable- rather than morpheme-level constraint:

<table>
<thead>
<tr>
<th></th>
<th>(σ).V</th>
<th>(σ).CV</th>
<th>(σ).CVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>V.V</td>
<td>8</td>
<td>V.CV</td>
<td>108</td>
</tr>
<tr>
<td>CV.V</td>
<td>75</td>
<td>CV.CV</td>
<td>966</td>
</tr>
<tr>
<td></td>
<td>V.CVC</td>
<td>CV.CVC</td>
<td>4410</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(σ).VC</th>
<th>(σ).CLV</th>
<th>(σ).CLVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>V.VC</td>
<td>115</td>
<td>V.CLV</td>
<td>7</td>
</tr>
<tr>
<td>CV.VC</td>
<td>566</td>
<td>CV.CLV</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>V.CLVC</td>
<td>CV.CLVC</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>V.V</td>
<td>8</td>
<td>V.CV</td>
<td>108</td>
<td>CL.V</td>
<td>15</td>
</tr>
<tr>
<td>V.CV</td>
<td>108</td>
<td>CV.CV</td>
<td>966</td>
<td>CL.CV</td>
<td>185</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CL.VC</td>
<td>1259</td>
</tr>
<tr>
<td>V.VC</td>
<td>115</td>
<td>V.CLV</td>
<td>7</td>
<td>CL.VC</td>
<td>122</td>
</tr>
<tr>
<td>V.CVC</td>
<td>901</td>
<td>CV.CLV</td>
<td>4410</td>
<td>CL.V.CVC</td>
<td>811</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>98</td>
<td>1259</td>
<td>803</td>
<td>6122</td>
<td></td>
</tr>
</tbody>
</table>

### Table 5.2 Frequencies of occurrence of syllable-types in morpheme-final position (data from Warna et al. (1990)).

As for which consonants occur in onset position, the following figures give an idea of the respective frequencies of occurrence. (They in fact give figures only for occurrence morpheme-initially (Table 5.3), and initially in final closed syllables (Table 5.4), respectively.)

<table>
<thead>
<tr>
<th></th>
<th>k</th>
<th>s</th>
<th>p</th>
<th>b</th>
<th>t</th>
<th>l</th>
<th>g</th>
<th>c</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1444</td>
<td>1410</td>
<td>1067</td>
<td>1051</td>
<td>985</td>
<td>971</td>
<td>857</td>
<td>636</td>
<td>592</td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ny</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ny</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 5.3 Frequency of occurrence of consonants morpheme-initially (data from Warna et al. (1990))
The initial $C_i$ slot may then be filled by any consonant, except (in Peguyangan and most varieties) /h/. 3 I conclude that this negative constraint is in force:

$$(7) \quad \text{\textsuperscript{*}I_o h}$$

In Peguyangan/Singaraja dialect, /h/ is deleted when it becomes syllable-initial due to suffixing (5.5):

$$(8) \quad \text{bayah} = '\text{pay}', \quad \text{[bayah]}$$
$$(8) \quad \text{bayah-ang} = '\text{pay-APP; pay for (BEN)}', \quad \text{[bayaŋ]}$$

There is rough agreement in the above figures with Clements (1990) Sonority Dispersion Principle, which predicts that oral stops and other obstruents will be favoured in the onset slot over more sonorous consonants. There are some disagreements, for example Clements' theory predicts that nasal consonants will be preferred over the liquids /l/ and /r/ as onset consonants; in Balinese the reverse pattern is found. (Equally, Clements predicts that liquids will be preferred over nasals in the coda slot: only nasals occur in this slot in non-final syllables). In Balinese, sonorant consonants, particularly nasals and glides, occur syllable-initially with the lowest frequency. And when they do occur in this position the morpheme generally belongs to one of the expressive semantic classes (3.2.2) or is a loan.

$$(9) \quad \text{ngaap} = '\text{sting (of eyes, sore)}', \quad \text{bengu} = '\text{stink}'$$
$$(9) \quad \text{nyambu} = '\text{k. o. tree}', \quad \text{méong} = '\text{cat}'$$
$$(9) \quad \text{benyah} = '\text{smashed to pieces}', \quad \text{lemes} = '\text{weak}'$$

The occurrence of the glides /w/ and /y/ morpheme-initially in the native vocabulary is very marginal. The sequences [wu] and [yi] do not occur, except as the result of

---

2Totals do not include underlying monosyllables, including %C(L)VC% reduplications (eg cakcak, crukcuk), or %CV-CVC% reduplications (eg kekek); nor do they include tri-syllabics ending in /an/: two common suffixes, 'NOM' and the 'transitive stem' suffix have this shape.

3The few /h/-initials listed in Warna et al (1990) are loans. These are used only in very formal/literary style and always have a vowel-initial alternant, the form in general usage. In Mountain Balinese varieties (including some on Nusa Penida) /h/ does occur syllable-initially, though this /h/ has been lost in the cognate forms in most varieties.
optional postlexical processes. All other /wV/ and /yV/ sequences, except /wa/ and /ya/ are strongly dispreferred (see 5.4.2).

5.2.2 The L slot

A syllable has a complex onset if both the Ci and the L-slot positions are filled. The L position can only be filled if Ci is also occupied. It can be filled both by the liquids, /l/ and /r/:

\[
\begin{align*}
(10) & & \text{jirit} & \text{‘spin quickly’} & \text{-cograh} & \text{‘fight (of dogs)’} \\
& & \text{-jlawat} & \text{‘flash by’} & \text{omplong} & \text{‘water dipper’ [onomat.]} \\
\end{align*}
\]

and (less frequently) by the glides /w/ and /y/:

\[
\begin{align*}
(11) & & \text{kwéni} & \text{‘k.o. tree’} & \text{-byayag} & \text{‘scatter untidily’} \\
& & \text{kacwak} & \text{‘cockroach’} & \text{grembyong} & \text{‘very loose (clothes)’} \\
\end{align*}
\]

The following preferred rule holds:

(12) Disprefer syllables with complex onsets

The first evidence for this comes from productivity. Table 5.5 below shows that, for disyllabic morphemes of otherwise identical structure, morphemes with initial-syllable CL clusters occur much less frequently than those without complex onsets: the ratios are approximately five-to-one on favour of the latter type:

<table>
<thead>
<tr>
<th>CV.CV</th>
<th>966</th>
<th>CVC.CV</th>
<th>490</th>
<th>CV.CVC</th>
<th>4410</th>
<th>CVC.CVC</th>
<th>2070</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLV.CV</td>
<td>185</td>
<td>CLVC.CV</td>
<td>51</td>
<td>CL.VCVC</td>
<td>811</td>
<td>CLVC.CVC</td>
<td>398</td>
</tr>
</tbody>
</table>

Table 5.5 Numbers of morphemes with and without complex onsets in the initial syllable, in Warna et al 1990.

Semantically, morphemes containing Ci-L sequences are generally (perhaps all) expressives or loans. This is particularly apparent where Ci is a sonorant, or where the CL sequence occurs morpheme-medially rather than initially; in these cases, the morphemes almost invariably have iconic or ‘bad’ semantics - see also the above examples, and 6.2.2:

\[
\begin{align*}
(13) & & \text{mrekak} & \text{‘lie boastfully’} & \text{bongglak} & \text{‘half bald’} \\
& & \text{kecrés} & \text{‘squirting sound’} & \text{ngléwa} & \text{‘[do] carelessly’} \\
& & \text{rilékéték} & \text{‘lie stretched out lifeless’} \\
\end{align*}
\]

A semantic category where this sequence occurs frequently is in (deliberately) inelegant personal nicknames (3.2.4). Children's nicknames in Peguyangan included Nyamprut, Gomblos, Glemuk, Gablaran, and Cruit, and were considered obviously inelegant.

\[\text{4The standard orthography writes these glides as vowels. Arguments in favour of the glide analysis are given in 2.1.3.}\]
The combined evidence from productivity and semantics therefore indicates that complex onsets are dispreferred in Balinese.

5.2.3 \(C_i\)-L interactions

Frequencies of occurrence of different \(C_i\)-L combinations are given in Tables 5.6 (below) and 5.7 (following page), for morpheme-initial and morpheme-internal occurrences respectively.

<table>
<thead>
<tr>
<th>L:</th>
<th>l</th>
<th>r</th>
<th>w</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C_i):</td>
<td>b</td>
<td>117</td>
<td>178</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>m</td>
<td>18</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>w</td>
<td>6</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>t</td>
<td>71</td>
<td>65</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>d</td>
<td>14</td>
<td>17</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td>n</td>
<td>0</td>
<td>3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>l</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>r</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>c</td>
<td>89</td>
<td>80</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>j</td>
<td>24</td>
<td>20</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>ny</td>
<td>18</td>
<td>138</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>s</td>
<td>144</td>
<td>0</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>y</td>
<td>0</td>
<td>143</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>g</td>
<td>118</td>
<td>162</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>n</td>
<td>10</td>
<td>18</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>totals:</td>
<td>876</td>
<td>947</td>
<td>47</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 5.6 Frequencies of CL sequences, morpheme-initially (data from Warna et al (1990)).

The following generalisations can be made:

1. If \(C_i\) is a liquid or a glide, the L slot is not normally filled. Sequences /\(rL\)/ and /\(wL\)/ do not occur. A handful of morphemes containing /\(yL\)/ and /\(wL\)/ sequences occur, but all are literary loans, not in general usage, or belong to the high language, eg yakti 'truly. Hf' (from Skt) has alternant forms wiakti /wyakti/~ yuakti /ywakti/.

2. For all \(C_i\), the possibility of occurring before /l/ or /\(r\)/ is much reduced, or barred completely, morpheme-internally (after another syllable). On the other hand /\(w\)/ and /\(y\)/ in the L slot occur only marginally morpheme-initially, and more frequently in intervocalic onset clusters.

3. Liquids outnumber glides in L position by a large factor.

4. \(C_i\)-L combinations occur most freely where \(C_i\) is a labial or dorsal oral stop or /\(s\)/. Initial apicals and the laminals other than /\(s\)/ cooccur with liquids with lower frequency.

\(^{b}\)Includes 87 Sanskrit loans with initial sequence /pra/. 

89
Table 5.7 CL sequences, intervocically

5. Where Ci is a laminal, L is never the laminal glide /y/, except in the case of a few medial /sy/ sequences in expressives:

(14) me-kesyab  'startled'
     me-kesyar  'shine'
     me-kesyeng 'be wafted about (odour)'

On the other hand, all the labials occur with the back/labial glide /w/ (underlying %u%), though again in low numbers:

(15) /bwaya/  'crocodile'
     /klampwak/ 'k.o. tree'
     /mwani/  'male'

6. Nasal stops hardly occur in CL clusters. This may simply reflect the general low occurrence of nasals in syllable-initial position.

5.3 Rhyme-internal structure

5.3.1 The nucleus

The nucleus of a syllable is always a vowel. Any of the six vowel phonemes may fill the syllable nucleus underlyingly (2.2.5). (The restriction barring underlying /θ/ morpheme-finally is discussed in 6.3.2.)
Occasionally liquids and nasal stops are realised as syllabic, though underlyingly /a/ fills the syllable nucleus (see 2.2.5 on optional deletion of underlying /a/):

\[(16) \ /mr\text{a}sidaan/ \ 'be able to (Hl)' \quad /mla\text{ma}/ \ 'person' (Skt) \quad /\text{anto}/ \ 'that'\]

5.3.2 The syllable coda

Closed syllables are of two kinds, depending on their position in the morph.

1. Morph- and morpheme-finally, the coda can be filled by any consonant except /w/, /vl/, or a laminal stop:

\[(17) \ \begin{array}{ll}
\text{panak} & \ 'child' \\
\text{nyem} & \ 'tasteless' \\
\text{amah} & \ 'eat' \\
\text{babad} & \ 'chronicle' \\
\text{lawar} & \ 'k.o. meat dish' \\
\text{ng\text{é}ng\text{é}s} & \ 'snot'
\end{array}\]

In the case of the laminals this restriction is easily ascertained by their complete absence in the lexicon (6.2.4).

The case is not quite so straightforward for the glides. Morphemes with final sequences represented in the orthography as \(<...\text{Vi}>\) and \(<...\text{Vu}>\) are the only possible case where syllable-final glides could occur in Balinese. These are consistently syllabified as vowels, eg \text{biu} 'banana' is syllabified as /bi.u/, not as a single syllable */biw/, \text{bui} 'prison' as /bu.i/, not as */buy/. Moreover, morphemes of this shape listed in Warna et al (1990) are overwhelmingly disyllabic, if \(<i>\) and \(<u>\) here represent /i/ and /u/; they are overwhelmingly monosyllabic if they represent /Ny/ and /Nw/. The disyllabic solution conforms to the normal pattern (6.1) and is therefore preferable. The constraint barring syllable-final glides falls out from principles of syllabification (5.5), the glide/vowel contrast not being found underlyingly (2.1.3).

2. In non-final syllables in simple morphemes the coda consonant may only be a nasal homorganic with a following stop (or, before /s/, /\text{\v{r}}/):

\[(18) \ \begin{array}{ll}
\text{-em\text{.}pu} & \ 'look after (child)'
\text{in\text{.}tuk} & \ 'pound (v)'
\text{mang\text{.}ku} & \ 'k.o. priest'
\text{blim\text{.}bing} & \ 'starfruit'
\text{ifi\text{.}jin} & \ 'black sticky rice'
\text{mans\text{.}sit} & \ 'smell of urine (v)'
\end{array}\]

This nasal segment is best analysed as underlyingly not specified for place of articulation (cf 2.1.4). This restriction on non-final coda consonants applies to the native vocabulary only (see below).

Inherently reduplicated disyllabic morphemes show much freer possibilities: any consonant occurring morph-finally can also appear in non-final codas:
The discrepancy between the types of consonant filling the coda position of final- as opposed to non-final syllables can be accounted for if the following constraint on underlying syllable structure applies to the native lexis (cf Itô 1986):

\[(20) \quad {^*C}_C \quad \text{[PLACE]}\]

That is, syllable-final consonants cannot be specified for place of articulation underlyingly. (In Itô's approach, the syllable coda is not a 'licenser' of place of articulation in underlying representations.) This analysis depends on the assumption that morph-final consonants are extrametrical at the stage when this constraint applies (see Itô (1986)). This must be either in the lexical representation itself (see Mohanan (1991) for arguments in favour of this possibility), or at an early stage in the lexical phonology.⁶

At a later stage in the phonology, the coda constraint does not hold, hence morph-final consonants, specified for place, can be syllabified and surface in coda position. Similarly, the reduplication process applying to inherently reduplicated morphemes is ordered after the constraint ceases to hold, hence morphemes like those in (19) are well-formed. Homorganic nasal-obstruent clusters are well-formed underlyingly, the nasal gaining its place of articulation only by a late spreading rule.

Here are some sample derivations, of linggih 'position, status (HON)', (21a), and linggihin /lingiin/ 'id.-LOC; sit on (HON)', (21b):

\[(21) \quad (a) \quad (b)\]

\begin{align*}
\text{Initial syllabification:} & \quad \sigma & \sigma & \text{EX} \\
\text{lìN gìh} & \quad \text{lìN gìhìn} \\
\text{Deletion of non-syllabifiable elements:} & \quad \sigma & \sigma & \sigma & \text{EX} \\
\text{n/a} & \quad \text{lìN gì-hìn} \\
\text{Subsequent syllabification:} & \quad \sigma & \sigma \\
\text{lìN gìh} & \quad \text{lìN gìhìn}
\end{align*}

In the initial syllabification, the underspecified nasal N syllabifies, since the coda constraint (20) applies only to consonants specified for place of articulation. Then in (21b), the /h/-segment, which is blocked from filling syllable-initial position (5.2.1),

---

⁶On the distinction between lexical and postlexical phonology, see 2.0.
cannot syllabify and is deleted. On the other hand the word-final consonant is protected from deletion by its extrasyllabic status.

In loanwords any coda consonant occurring in final syllables may also occur in non-final syllables:

(22) ak.sara 'letter of the alphabet (Skt)
ap.sari 'angel (Skt)'
bak.ti 'homage (Skt)'
isi.tri 'female (HON, Skt)'
brah.ma 'Brahma (Skt)'

I conclude therefore that loanwords are not subject to (20). (See Mohanan (1989) for similar exceptional syllabification of loanwords in Malayalam, and 3.2.1 for other exceptional phonological patterning of loanwords in Balinese).

5.3.3 Nucleus-coda interactions

Table 5.8 gives raw data for the numbers of different VC combinations morpheme-finally in Warna et al (1990). There appear at first inspection to be no significant interactions between the vowel nucleus and the coda consonant. These figures do not however allow for borrowings or the expressive/neutral morpheme distinction (3.2.1), and for this reason a statistical analysis has not been done. A detailed study is beyond the scope of this thesis.

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>u</th>
<th>é</th>
<th>o</th>
<th>a</th>
<th>e</th>
<th>totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>37</td>
<td>40</td>
<td>14</td>
<td>13</td>
<td>101</td>
<td>96</td>
<td>301</td>
</tr>
<tr>
<td>b</td>
<td>14</td>
<td>39</td>
<td>7</td>
<td>13</td>
<td>82</td>
<td>73</td>
<td>228</td>
</tr>
<tr>
<td>m</td>
<td>21</td>
<td>37</td>
<td>9</td>
<td>9</td>
<td>36</td>
<td>151</td>
<td>263</td>
</tr>
<tr>
<td>t</td>
<td>254</td>
<td>236</td>
<td>163</td>
<td>142</td>
<td>240</td>
<td>246</td>
<td>1281</td>
</tr>
<tr>
<td>d</td>
<td>37</td>
<td>68</td>
<td>54</td>
<td>56</td>
<td>92</td>
<td>125</td>
<td>432</td>
</tr>
<tr>
<td>n</td>
<td>64</td>
<td>110</td>
<td>115</td>
<td>127</td>
<td>7581</td>
<td>105</td>
<td>1102</td>
</tr>
<tr>
<td>l</td>
<td>127</td>
<td>129</td>
<td>81</td>
<td>106</td>
<td>181</td>
<td>153</td>
<td>777</td>
</tr>
<tr>
<td>r</td>
<td>160</td>
<td>167</td>
<td>102</td>
<td>114</td>
<td>218</td>
<td>150</td>
<td>911</td>
</tr>
<tr>
<td>s</td>
<td>151</td>
<td>146</td>
<td>103</td>
<td>100</td>
<td>227</td>
<td>139</td>
<td>866</td>
</tr>
<tr>
<td>k</td>
<td>216</td>
<td>309</td>
<td>173</td>
<td>241</td>
<td>374</td>
<td>234</td>
<td>1547</td>
</tr>
<tr>
<td>g</td>
<td>90</td>
<td>110</td>
<td>52</td>
<td>104</td>
<td>178</td>
<td>155</td>
<td>689</td>
</tr>
<tr>
<td>ng</td>
<td>334</td>
<td>334</td>
<td>187</td>
<td>263</td>
<td>561</td>
<td>278</td>
<td>1957</td>
</tr>
<tr>
<td>h</td>
<td>168</td>
<td>300</td>
<td>124</td>
<td>98</td>
<td>480</td>
<td>196</td>
<td>1366</td>
</tr>
<tr>
<td>Totals</td>
<td>1673</td>
<td>2025</td>
<td>1184</td>
<td>1386</td>
<td>3351</td>
<td>2101</td>
<td>11720</td>
</tr>
</tbody>
</table>

Table 5.8 Morpheme-final VC sequences (data from Warna et al (1990))

7This is 5 times as many as any other /VN/ combination; many of these 581 presumably carry one of the nominal or verbal suffixes of this shape.
Chapter 5

5.4 Onset-rhyme interactions

5.4.1 C\textsubscript{i}-C\textsubscript{coda} restrictions

Only one significant interaction between the initial and coda consonants of syllables of shape C(L)VC has been noted; where such consonants are of the same place of articulation, the morpheme is in most cases expressive or a loan:

\[(23)\] besus 'short-tempered' mangkug 'stink (v)'
   bobab 'lie, deception' déngkik 'screech'
   cedut 'sound of spinning propellor' beler 'naughty, irresponsible'

This is better interpreted to reflect the general morpheme structure constraint dispreferring the cooccurrence of two consonants with the same place of articulation (6.2.5).

5.4.2 Onset-nucleus interactions

The analysis of Ci-V and L-V sequences presented here is limited in that it is based only on morpheme-initial data; onset-nucleus sequences occurring morpheme-internally have not been examined.

The main interactions between onset and vowel noted involve the sequences /yw\textsubscript{[-LO]}/ and /yw\textsubscript{[-LO]}/: that is, where the glides are followed by any vowel other than /a/. Of the 52 y-initial items in Warna et al (1990), perhaps only these are non-borrowed:

\[(24)\] yag 'common knowledge (of news)' yuyu 'crab'
   yah, yéh 'exclamation of surprise' yéh 'water'

Similarly, of the 339 w-initials perhaps only these forms (all but two with initial /wa/) listed in Warna et al (1990) are not loanwords:

\[(25)\] wadah 'container' wadih 'tired of'
   wai 'day' waluh 'gourd'
   wangsit 'sign, signal' wangwang 'challenge, oppose'
   wani 'k.o. fruit' wanti 'continuously do'
   wauh 'close (of friends)' wayah 'old'
   wani, wéni 'k.o. fruit' wayan 'title, 1st-born child'
   -wél 'scold, rebuke angrily'

The marginal status of initial /w/ and /y/ before non-low vowels can also be seen from the way loanwords are taken into the language. While there are significant numbers of /w/– and /y/-initial loans, most of them are literary terms not in general usage. In some loanwords the /yw\textsubscript{hi}/ and /yw\textsubscript{hi}/ sequences of the original have been simplified by deletion of the glide:

\[81\] do not count a few items listed in Warna et al (1990), eg yab 'person of same age', yung 'lull', yus 'luck, situation': in Peguyangan and most dialects these have initial /i/: /iab/, /iung/. A number of listed w-initials similarly have /u/.

94
or substitution:

(27) bidadari 'angel', from Jav widyadari, Skt vidyadhari

In several cases with initial /wi/, doublets have been created:

(28) wimba ~ imba 'eyebrow (HON)
wibuh ~ ébuh 'rich, prosperous (HON)
widuh ~ idu 'saliva (HON)
wikti ~ yakti 'truly (HI)
widadari ~ bidadari 'angel"

Finally, where the source language has /we/ this has typically gone to /o/ in everyday Balinese, though again doublets are sometimes listed in Warna et al (1990) (the /we/-initial alternants are limited to formal literary contexts):

(29) wedak ~ odak 'powder (HON)', Jav. wedhak (id.)
welas ~ olas 'pity', Jav. welas (id.)
weka ~ oka 'child (HON)', Old Jav. weka (id.)

The homorganic sequences /yi/ and /wu/ never occur in Peguyangan Balinese, nor in most other varieties. Similarly, as indicated in the above discussion, the sequences /yu/ and /wi/, occur only in a couple of expressives in the native lexis, and tend to be reduced to plain /u/ and /i/ in loanwords. The descriptive statement can be diagrammed as follows:

(30) Disprefer:

\[ \sigma \]

\[ O \quad N \]

\[ [+HI] \quad [+HI] \]

This dispreference in turn is due to an OCP violation, applying to strictly adjacent phonemes. (Identical vowels in adjacent syllables must apparently always be linked at the melody or phoneme tier cf 8.7). This constraint overrides the normal tendency to syllabify the high vocoids as glides, whenever the result will produce a disyllabic surface structure (5.6): tri-syllabics, containing /ii/ and /uu/, surface in this context:

(31) sengguu /səŋ.gu.u/ */səŋ.gwu/ 'type of priest'
beduu /bə.du.u/ */bə.dwu/ 'black cat'
jempilt /ʃəm.pi.it/ */ʃəm.piyt/ 'k.o. bat'
kepii /ko.pi.i/ */ko.pyi/ 'k.o. bug')

---

9 Warna et al 1990 list /yip/ as a variant of /iip/ 'part of sugar palm leaf'; only the latter form is possible in Peguyangan. Forms with /#wu/ and /#wo/ listed by Warna have structure /#uV/ in Peguyangan, and in most dialects, eg wong 'mushroom, fungus', wot 'rice husks' are /oong/ and /oot/ respectively; otherwise, they are borrowings from Sanskrit or Javanese.
Chapter 5

The sequences /yo, ye, ya/ and /wo, we, wa/ are also strongly disfavoured, though do occur, almost entirely in loanwords, plus some expressives. I assume that this derives from the same constraint as that dispreferring cooccurrence of high and mid vowels (repeated from (42), 6.3.1):

\[(32) \begin{align*}
\text{Prefer:} & \quad +\text{ATR} \\
& \quad / \ \ \\
& \quad \bullet \ \bullet \\
& \quad \alpha \text{HI}
\end{align*} \quad \text{Disprefer:}
\begin{align*}
& \quad +\text{ATR} \\
& \quad / \ \ \\
& \quad \bullet \ \bullet \\
& \quad \alpha \text{HI} - \alpha \text{HI}
\end{align*} \]

These constraints, as expected, apply whether the glide fills the Ci or the L slot, in fact they are even more marked in the latter case. The following data show the strong preference for onset glides to occur before /a/ only, with 219 out of 231 or 95% of the lexicon showing this pattern:

<table>
<thead>
<tr>
<th>Onset:</th>
</tr>
</thead>
<tbody>
<tr>
<td>#CwV</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Nucleus V:</td>
</tr>
<tr>
<td>a</td>
</tr>
<tr>
<td>i</td>
</tr>
<tr>
<td>u</td>
</tr>
<tr>
<td>e</td>
</tr>
<tr>
<td>o</td>
</tr>
<tr>
<td>e</td>
</tr>
<tr>
<td>Totals:</td>
</tr>
</tbody>
</table>

Table 5.9 CGV sequences (data from Warna et al (1990))

Clearly, from the above discussion, distinct phonotactic sequences are tolerated for loanwords in Balinese, many of which derive from the ritually powerful languages, Old Javanese and Sanskrit. Borrowings from these languages are less liable to be regularised to the norms constraining the native lexis. A full study of phonotactic restrictions would distinguish patterns obtaining in the native lexis from those in loanwords.

Such a study would no doubt reveal further restrictions. For example, a second minor restriction noted was that disfavouring the sequence /ni/ both initially, and morpheme internally. Although Warna et al (1990) list 30 ni-initial forms, these are almost all Sanskrit and Javanese loans. Indeed, a high proportion of all /n/-initials appears to be made up of loanwords. The need to distinguish loanwords from native lexis is further illustrated by the following two tables. Table 5.10 gives raw figures for onset-vowel combinations, in initial syllables of shape CV(...).
Prosodic structure

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>u</th>
<th>e</th>
<th>o</th>
<th>a</th>
<th>a</th>
<th>totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>158</td>
<td>187</td>
<td>78</td>
<td>80</td>
<td>321</td>
<td>243</td>
<td>1067</td>
</tr>
<tr>
<td>b</td>
<td>123</td>
<td>166</td>
<td>81</td>
<td>87</td>
<td>323</td>
<td>271</td>
<td>1051</td>
</tr>
<tr>
<td>m</td>
<td>48</td>
<td>74</td>
<td>40</td>
<td>58</td>
<td>158</td>
<td>96</td>
<td>474</td>
</tr>
<tr>
<td>w</td>
<td>89</td>
<td>16</td>
<td>26</td>
<td>8</td>
<td>148</td>
<td>52</td>
<td>339</td>
</tr>
<tr>
<td>t</td>
<td>146</td>
<td>192</td>
<td>80</td>
<td>65</td>
<td>296</td>
<td>206</td>
<td>985</td>
</tr>
<tr>
<td>d</td>
<td>62</td>
<td>103</td>
<td>51</td>
<td>47</td>
<td>148</td>
<td>101</td>
<td>512</td>
</tr>
<tr>
<td>n</td>
<td>30</td>
<td>14</td>
<td>12</td>
<td>11</td>
<td>44</td>
<td>19</td>
<td>130</td>
</tr>
<tr>
<td>l</td>
<td>144</td>
<td>124</td>
<td>103</td>
<td>87</td>
<td>279</td>
<td>234</td>
<td>971</td>
</tr>
<tr>
<td>r</td>
<td>67</td>
<td>92</td>
<td>60</td>
<td>71</td>
<td>171</td>
<td>131</td>
<td>592</td>
</tr>
<tr>
<td>c</td>
<td>45</td>
<td>75</td>
<td>46</td>
<td>86</td>
<td>181</td>
<td>203</td>
<td>636</td>
</tr>
<tr>
<td>j</td>
<td>18</td>
<td>59</td>
<td>12</td>
<td>29</td>
<td>152</td>
<td>126</td>
<td>396</td>
</tr>
</tbody>
</table>

Table 5.10 C1-V interactions (morpheme initially), data from Warna et al 1990. \( \chi^2 = 607, \ DF = 85, \ p < .001 \)

Table 5.11 below gives \( \chi^2 \) values for each cell in the distribution. The high overall value of \( \chi^2 \) indicates that there is strong non-random skewing. However, a table like this is of little value, if the presence of loanwords is not taken into account. For example, the observed number of morphemes beginning with the sequence /wi/ (89) is much higher than expected, as indicated by the large value of \( \chi^2 \) (52.66) for that cell in the following table. In fact, of these 89 tokens, all are either loans or variant forms of native morphemes of shape /#ui/; the actual number of native forms beginning with /wi/ is only one (see above), or far lower than expected.

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>u</th>
<th>e</th>
<th>o</th>
<th>a</th>
<th>a</th>
<th>totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>5.06</td>
<td>3.87</td>
<td>0.10</td>
<td>2.35</td>
<td>0.03</td>
<td>4.66</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>0.39</td>
<td>0.26</td>
<td>0.02</td>
<td>0.45</td>
<td>0.30</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>1.95</td>
<td>0.05</td>
<td>0.47</td>
<td>5.94</td>
<td>1.97</td>
<td>6.31</td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>52.66</td>
<td>24.43</td>
<td>0.004</td>
<td>16.28</td>
<td>21.80</td>
<td>15.16</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>4.73</td>
<td>12.08</td>
<td>0.39</td>
<td>5.85</td>
<td>0.02</td>
<td>10.34</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>0.03</td>
<td>8.22</td>
<td>3.84</td>
<td>0.05</td>
<td>0.14</td>
<td>8.08</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>12.0</td>
<td>1.66</td>
<td>0.47</td>
<td>0.03</td>
<td>0.71</td>
<td>6.62</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>4.69</td>
<td>3.71</td>
<td>11.78</td>
<td>0.01</td>
<td>0.38</td>
<td>1.57</td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>0.54</td>
<td>0.046</td>
<td>5.12</td>
<td>6.38</td>
<td>0.17</td>
<td>3.66</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>14.47</td>
<td>4.81</td>
<td>0.10</td>
<td>15.28</td>
<td>0.39</td>
<td>8.08</td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>19.65</td>
<td>0.02</td>
<td>10.80</td>
<td>1.10</td>
<td>9.76</td>
<td>4.86</td>
<td></td>
</tr>
<tr>
<td>ny</td>
<td>1.46</td>
<td>2.22</td>
<td>0.04</td>
<td>0.83</td>
<td>10.96</td>
<td>1.30</td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>13.96</td>
<td>6.74</td>
<td>1.31</td>
<td>7.88</td>
<td>4.89</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>y</td>
<td>4.60</td>
<td>3.31</td>
<td>0.22</td>
<td>2.45</td>
<td>5.82</td>
<td>9.89</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>20.69</td>
<td>3.89</td>
<td>7.90</td>
<td>0.05</td>
<td>8.78</td>
<td>84.24</td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>6.43</td>
<td>4.10</td>
<td>0.74</td>
<td>18.69</td>
<td>0.22</td>
<td>2.98</td>
<td></td>
</tr>
<tr>
<td>ng</td>
<td>1.61</td>
<td>0.01</td>
<td>0.34</td>
<td>0.38</td>
<td>1.16</td>
<td>0.14</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.11 Chi-square values for the preceding table. \( \chi^2 = 607.432, \ DF = 95 \)
Similarly, the low value (0.004) of $\chi^2$ for morphemes beginning with the sequence /we/ gives the impression that this sequence is a normal one in the language. Yet of the 26 forms in Warna et al (1990), almost all are loans, the couple of potentially native forms being expressives. Simple statistical tests thus do not necessarily bring out the true situation.

5.5 Syllabification rules

Evidence from allophony and other phonological processes shows that syllabification is a cyclical process, reapplying as more complex phonological units are produced. The main syllabification rules, following Itô (1988), are:

(33) a. Assign vocoids to a distinct syllable nucleus, right to left
b. Maximise onsets, rather than codas
c. Delete non-syllabified elements (except where this would lead to irrecoverability of the underlying morpheme structure)
d. Epenthesize schwa

Domain: Phonological word

See the discussion of example (40) below for evidence that the domain for syllabification is the phonological word. The rules in (33) are now exemplified and discussed in turn.

1. Assign each underlying vowel a distinct syllable, right to left (rule (33a)):

(34) 

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Syllabification</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ua</td>
<td>/u.a/</td>
<td>'parent’s elder sibling'</td>
</tr>
<tr>
<td>mai</td>
<td>/ma.i/</td>
<td>'come here'</td>
</tr>
<tr>
<td>buung</td>
<td>/bu.un/</td>
<td>'not go ahead with intention'</td>
</tr>
</tbody>
</table>

Evidence for this includes native speaker syllabifications, and allophony of {N} (see the discussion in 2.2, 2.1.3. Right-to-left syllabification results in non-occurrence of glides in word-final position, 6.2.4.

2. Maximise onsets (rule (33b)). This ensures that (i) an intervocalic consonant must syllabify as the onset of the following syllable, rather than as the coda of the preceding one, (ii) a sequence acceptable as a complex onset (CL) syllabifies as such, rather than in two different syllables.

(35) 

<table>
<thead>
<tr>
<th>Word</th>
<th>Syllabification</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>bekel</td>
<td>/ba.kal/</td>
<td>&quot;pocket money&quot;</td>
</tr>
<tr>
<td>bah-ang</td>
<td>/ba.han/</td>
<td>* /bak.al/</td>
</tr>
<tr>
<td>cogroh</td>
<td>/co.groh/</td>
<td>* /cog.roh/</td>
</tr>
</tbody>
</table>

\(^{10}\text{h} \text{ is in this position is subsequently deleted, see below.}\)
In the case of the high vowels, which can syllabify either as vowels or as glides, (33a) and (33b) can have conflicting results. Which of the two takes precedence is determined by other well-formedness constraints: for discussion see 5.6.

3. **Delete non-syllabifiable elements.** This rule is illustrated by the deletion of morpheme-final /h/ when it becomes intervocalic due to suffixing or cliticisation. Rule (33b) would require /h/ to fill the syllable-onset in that case, but it is blocked from doing this by this constraint (5.2.1):

\[(36) \quad *_{\sigma}[h]\]

Intervocalic h is therefore deleted by (33c):

\[(37) \quad \begin{array}{ll}
\text{bah} & \text{'fall down (of tree etc)'} \\
\text{bah-ang} & \text{'fall down-APP'} \\
\text{belah} & \text{'shatter, break'} \\
\text{belah=a} & \text{'break = 3, broken by him'} \\
\text{pasih} & \text{'sea'} \\
\text{pasih-e} & \text{'sea-DEF'}
\end{array}\]

In very formal speech, rule (36) does not apply; /h/ occurs in the onsets of a few literary loanwords (sahasa, 'with great force' everyday form saasa) and /h/ is not deleted in derived forms like those in (37) above. Otherwise, /h/-deletion is widespread in the Buleleng district, and elsewhere, though it is not found in all dialects.\(^{11}\)

4. **Epenthesize schwa.** Rule (33c) does not apply where competing principles override it. One such principle is, I propose, the following semantically based constraint on repair, and other, processes:

\[(38) \quad \text{Phonological processes must leave the underlying morpheme identifiable}\]

Another is Foot Well-formedness (5.6), the requirement that the output of lexical phonological processes be prosodically well-formed at the Foot level.

The requirements of either of the above two constraints can lead to schwa epenthesis (33d), and with it the creation of an extra syllable. For example, both constraints influence the allomorphy of /N/, the 'agentive prefix' on verbs, of shape /N/ (a nasal consonant not specified for place, 7.2). Usually, where /N/ prefixes to an consonant-initial root, it assimilates to the place of articulation of that consonant: with jemak 'take', this produces the intermediate form *[njemak]. This contains an illegal onset cluster (5.2.1), which rule (33c) repairs by deletion, giving the surface form nyemak [nema?]. Schwa epenthesis 'intervenes' in /N/-prefixing in two different cases.

1. A version of the Foot Well-formedness rule (5.6) also applies, requiring the input and

\(^{11}\text{Karangasem dialects for example generally retain morpheme-final /h/ before vowel-initial suffixes and clitics. Even so, there and in most dialects, /h/ cannot occur anywhere but morpheme-finally in underlying forms. Exceptions are some mountain dialects.}\)
output of /NI/-prefixing to be an Optimal (disyllabic) Foot. As a result, with monosyllabic roots, such as cêt 'paint', the output of of /NI/-prefixing is not the ill-formed *nyêt *(net), but rather ngecêt [gecet], schwa-epenthesis (actually formally a prefix here) having applied to the base, before /NI/-prefixing.

2. Where the root begins with a nasal sonorant, eg mati-ang 'die-APP: kill', assimilation of /NI/ to the place of articulation of the sonorant is blocked (7.2). One could expect the output to be either ngmtiàng or ngatìang ngatìang is ruled out by principle (38) above, and ngmtiàng contains an ill-formed syllable-onset. /a/-epenthesis (see p102) and re-syllabification therefore produce ngmtiàng (/ŋmatiàng/).

That the domain of syllabification rules (33) is the phonological word is also indicated by the realisation of /h/, which only occurs in syllable codas, never in onsets. In compounds, word-final /h/ does not syllabify as onset (and hence undergo automatic deletion, cf (36)) where there is a following vowel-initial - hence onsetless - word: yé[h] anak-an 'water person-NOM: 'spring water'. Similarly, the NP ipah i ketut-é 'Ketut's sibling-in-law' consists of two phonological words, hence final /h/ is not deleted:

(39) ipah i ketut-é [ipah] [i (...) ] * [ipai (...) ] 'Ketut's child'

The cooccurrence patterns of the genitive morpheme, {n}, also indicate that syllabification is restricted to the phonological word. Genitive -n behaves like a suffix to its preceding head noun, and is always followed by a P-word boundary. This affects vowel allophony, only P-word-final high vowels being laxed. These examples show that {n} does not syllabify with following vowels:

(40) beli-n ipun [belin] [ipun] 'elder.brother-n-3.HI, his elder brother' 
   * [belinipun]

cucu-n ipun [cucun] [ipun] 'grandchild-n-3.HI, his grandchild'
   * [cucunipun]

Since syllabification applies only within a P-word, /n/ cannot syllabify with a following vowel-initial nominal, and is therefore deleted when it occurs immediately after a consonant:

(41) kulit ipun *kulitnipun 'skin 3.HI, his skin'
    pekak ipun *pekaknipun 'grandfather-n-3.HI, his grandfather'

5.6 The Foot

I follow MacCarthy & Prince 1990 in assuming that the prosodic unit immediately 'dominating' the syllable is the Foot, a preferably disyllabic unit which also constitutes the Minimal, or better, Optimal Word. Around 90% of morphs are disyllabic (6.1). Other formants are either monosyllabic (3%), or maximally tri-syllabic (probably a similar percentage) - see 6.1.1.
Morphemes larger than this - the maximum in the native lexis is four syllables - always consist of at least two complex formants, 6.1.1. There is thus evidence for the following hierarchy of foot templates (cf McCarthy & Prince ms., cited in Kenstowicz 1994:650):

(42) \begin{align*}
\text{Optimal Foot:} & \quad \sigma \sigma \\
\text{Maximal Foot:} & \quad \sigma \sigma \sigma \\
\text{Minimal Foot:} & \quad \sigma
\end{align*}

Morphs then must be prosodically well-formed, not just at the syllable level. They must parse to one of the Foot templates, in a process analogous to syllabification which I will refer to as the Foot Well-formedness rule:

(43) Foot Well-formedness
Morphs must be prosodically well-formed at the Foot level, throughout the lexical phonology

Morphs as either the input to, or the output of, phonological processes, must satisfy (43). In meeting the requirements of Foot Well-formedness, the Optimal Foot is imposed whenever possible. Foot Well-formedness, like all phonological rules, is only active in the phonology until 'switched off'. In the later postlexical phonology it no longer applies (8.9).\(^{13}\)

One result of Foot Well-formedness is a preference for disyllabic morphs (those which parse to the Optimal Foot template). This in turn leads to a preponderance of disyllabic morphemes in the lexic, 6.1. In the terms of McCarthy & Prince (1990), the Optimal Foot constitutes the 'Minimal Word' template in Balinese. I will use instead the expression Optimal Word, since Foot Well-formedness rejects output which is overly long as much as it disprefers overly short input: there are no underlyingly four-syllable morphs, and they may not be produced by inherent reduplication processes 8.9.)

The effects of (43) can be seen clearly in the realisation of underlying %i% and %u% as either vowels or glides, according to the prosodic context. In almost all cases, %i% and %u% are syllabified so as to generate a disyllabic structure:

(44) \begin{align*}
\text{wayan} & \quad \text{'first born'} & /\text{wayan}/ & */\text{uaian/} \\
\text{satwa} & \quad \text{'fable'} & /\text{sa.twa}/ & */\text{satua/} \\
\text{byasa} & \quad \text{'usual'} & /\text{bya.sa}/ & */\text{biasa/} \\
\text{tuak} & \quad \text{'palm wine'} & /\text{tu.ak}/ & */\text{twak/}
\end{align*}

These are generated by the Foot-Wellformedness requirement, which imposes Optimal word structure, giving precedence variously to syllabification rule (33b) (\text{wayan, satwa, byasa}) or (33a) (\text{tuak}). Evidence for this analysis is given in 2.1.3. This

---

\(^{12}\)I use the term Maximal Foot for expository purposes. McCarthy & Prince claim that the maximal Foot size is the disyllable (1995:321). The three-syllable unit is then a Foot plus extrametrical final syllable - their 'loosely interpreted' P-word template. Evidence in favour of this comes from the constraint on the position of NC clusters in the morph, see (28) in 6.2.3.

\(^{13}\)On the distinction between lexical and postlexical phonology, see 2.0.
disyllabic parse also means that monosyllabic formants of shape CGVC do not occur in the lexicon, 6.2.2.

Other evidence for a disyllabic Foot includes various phonological processes which either create or target disyllabic units:

1. Before most monosyllabic words, an epenthetic /ə/ can be optionally inserted:

   (45)  
   ling ~ eling 'weeping (n)'
   nu ~ enu 'still (aspectual)'
   bé ~ ebé 'meat; fish'
   bom ~ ebom 'bomb'

   Speakers disagree as to which surface form of such doublets represents the most basic one; Warna et al (1990) lists both forms, but gives priority to the monosyllabic root. The initial schwa does not surface when such morphemes bear suffixes or occur in compounds:

   (46)  
   ling-ang dum=a bé siap
   [liɲan] [dum3] [besiap] "cry-APP; make cry"
   *[elɨɲan] *[edum3] *[əbesiap] 'share-3ACT; shared by him'
   ng-
   "share out (v)"
   "AP-id."

   No other vowels are 'deleted' in this context, suggesting that these morphemes are underlyingly monosyllabic, with initial /ə/ functioning to produce a surface disyllabic unit.

2. {N-} 'ACTOR PIVOT' is realised as /ŋə-/ before monosyllabic roots (even where they are part of longer stems), adding a syllable to the surface form:

   (47)  
   bel dum séh-in
   'bell' 'share' 'change-LOC'
   ngebél ngedum ngeséhin
   'AP-bell; sound bell, ring (v)'
   'share out (v)'
   'AP-id.'

   Before disyllabic and longer roots an extra syllable is not added (7.2), suggesting again that the /ŋə/ realisation functions partly to create a disyllabic surface form.

3. Inherently reduplicated morphemes undergo reduplication at different stages, according to whether they are underlyingly monosyllabic or longer (see 8.9). Underlying monosyllables undergo reduplication before the application of many phonological rules such as lowering of final high vowels (7.4), vowel allophony rules, or raising of final %a% (7.5). On the other hand, inherent reduplication of underlying disyllabics (or longer) is virtually the last of all phonological rules to apply. Both of these can be explained in terms of (43). The earlier application of reduplication to underlying monosyllables creates morphs which meet the disyllabic optimal word shape of (43), while the 'blocking' of lexical reduplication of disyllabic bases results from a failure of the multi-syllabic output to satisfy the same constraint (see also 8.9).
4. Partial reduplications target morphs of no more than two syllables (8.2). The output has three syllables, the Maximal Foot.

The three-hundred-plus monosyllabic lexemes, listed in Warna et al (1990), as well as those tri-syllables and multi-syllables which contain a tri-syllabic formant, are exceptions to rule (43), just as semantically they are overwhelmingly 'peripheral' - either expressives or grammatical morphemes. Of a sample of 169 monosyllabic morphemes, only 24 (14.7%) did not fall into the peripheral/expressive categories listed in (3.2.2). See appendix G.

A further well-formedness constraint on the Foot results in a preference for morphemes which end in an (initially extrametrical, cf 5.3.2) consonant, rather than a vowel. For example, morphemes formed from the two most preferred morph structures in the language, CVCVC and CVCCVC, together constitute about 45% of roots in Warna et al (1990). Those based on their vowel-final counterpart structures, CV.CV and CVC.CV make up only 10% of the total, see 6.2.4.) I assume that this derives from a constraint on prosodic well-formedness, applying to the 'right edge' of the foot. Given that morphs must parse to a di- or tri-syllabic Foot template, this in tum implies a preferred Foot structure, formulated perhaps as:

(48) Prefer a Foot template which includes an extrametrical consonant at the right edge

From the evidence of NC cluster occurrence (see discussion of (28) 6.2.3), I conclude that (48) is better formulated as:

(49) Prefer a Foot template which includes an extrametrical consonant specified for place of articulation at the right edge

Rule (49) applies to all three foot types in (42), and so is independent of the Optimal word rule (43). It can be seen as a variant of McCarthy & Prince's (1990:240) 'obligatory extrametricality'. They cite cases in which an extrametrical final constituent must be present for a prosodic operation to apply to a base. In Balinese the presence of this constituent is a requirement only in Foot-reduplication (8.5). For Foot Well-formedness applying to all lexical formants (6.2.4), it is simply a preference.

5.7 The Phonological word

McCarthy & Prince (1990) claim that the most complex prosodic unit relevant to word-formation rules is the P-word (prosodic or phonological word). A P-word must contain at least one Foot, as must a lexical morpheme (by rule (43). In Balinese the P-word is the domain of a variety of phonological processes, including:

14 Many of course consist of smaller formants, eg kekupu 'butterfly', RED kekupu, 4.0.
15 The edge of the foot as a category relevant to prosodic operations in Balinese is found also in C-RED, which only targets left-most consonants of disyllabic formants in Balinese 8.2.2).
Chapter 5

1. Syllabification (5.5),
2. Consonant and vowel allophony rules (2.1, 2.2),
3. /o/ ~ /o/ alternations (7.3).

The following characteristics of the P-word in Balinese follow Cohn's (1989:200-201) description of the P-word in Indonesian:

(50) The Phonological Word:
1. The domain created between left edges of stems
2. Suffixes become part of the string-adjacent Phonological Word
3. Prefixes are not part of the stem

The P-word is thus defined by reference to morphological bracketing: the stem is a root morpheme, plus any suffix(es). Inherent reduplications like kupukupu 'butterfly', where reduplication serves a purely phonological purpose, constitute only a single P-word: their phonological behaviour supports this conclusion (see discussion of examples (74) and (75), section 8.9). Equally, segmental complex formants (4.1) do not constitute separate P-words.

That suffixes form part of the same prosodic unit as the roots to which they attach is seen in syllabification effects. Thus in the examples in (51), root-final /hi/ is deleted when it becomes syllable-initial (5.5):

(51) lalah lalah-an lalah-ang amah amah-in
lalah-an lalah-ang amah amah-in
[lalah] [lalaan] [lalaan] [amah] [amain]
'hot (spicy)' 'hot-CMP: hotter' 'hot-APP: make hot' 'eat.CRÚDE' 'eat-LOC: run (of colours)'

While formants are no more than 3 syllables long, P-words can be longer than this. Parsing to P-word units does not 'split' formants down the middle. Hence the following surface realisations of quadri-syllabic grammatical words:

(52) sabat-in=na slampar-ang=lālah-ang gesges-ang=ka
'spawned at by him' 'thrown by him' 'made spicy by him'
[sa.ba.tin3] [slam.pəraŋ3] [la.la.əŋ3]
'thrown.AT by him' 'thrown by him' 'made spicy by him'
[w3] [w] [w]
'spawn.AT' 'thrown' 'made spicy'

Words like those in (52) have structures as in (53). I assume then that lexemes are parsed, into feet, from left to right, with stray final syllables being incorporated into 'loosely interpreted' P-words, e.g. (53b) (McCarthy & Prince, ms. cited in Kenstowicz
1994). Extrametrical Foot-final consonants (5.6) such as final /t/ in the morph -sabat 'throw', become part of the following Foot when they can syllabify as onsets (53a):

(53) a. PWord b. PWd
\[
\begin{array}{cccc}
F & F & \sigma & \sigma \\
\sigma & \sigma & \sigma & \sigma \\
\sigma & \sigma & \sigma & \sigma \\
sa & ba & ti & na \\
\end{array}
\]

The term 'suffix' for the purposes of the definition in (50) includes not only verbal and nominal derivational affixes, but also the following, the last three of which may concatenate after (other) suffixes, and behave phonologically like them, except for stress assignment (5.8):

(54) -n 'genitive', linking two NP's to form a compound
-é 'definite marker' on nominals
-a 'irrealis'
-a 'third person actor'

This is seen for example from vowel and consonant allophony:

(55) bulu-n alis \[\text{bulun}_{\text{w}} \text{alis}_{\text{w}}\] 'hair of eyebrow'
*\[\text{bulunalis}_{\text{w}}\]

kuluk \[\text{kulu?}\] 'dog'
kuluke \[\text{kuluke}\] 'the dog'
pégpég \[\text{pegpeg}\] 'thickset'
(paling) pégpéga \[\text{pegpeg}_3\] '(most) thickset-IRR'
cekuk \[\text{cekuk}\] 'throttle (vt)
cekuka \[\text{cekuk}_3\] 'throttled by him'

Other cliticised pronominals (all disyllabics or longer) do not form part of the same phonological word as their host:

(56) kuluk ipun \[\text{kulu?}_{\text{w}} \text{ipun}_{\text{w}}\] 'puppy-3.HI, his puppy'
*\[\text{kulkupun}_{\text{w}}\]
kuluk ida \[\text{kulu?}_{\text{w}} \text{id}_{\text{3w}}\] 'puppy-3.HON, his puppy'
*\[\text{kulkuid}_{\text{3w}}\]

The definition in (50) specifies that, again as in Cohn's (1989) analysis of Indonesian, prefixes are not part of the same P-word as that associated with the stem to which they attach. Evidence for this is the failure of the prefix vowel to undergo the kinds of vowel alternations which occur at stem-suffix boundaries (see 7.6). They also do not behave like unambiguously P-word-final vowels (see 6.3.2). I therefore follow Cohn in concluding that prefixes do not constitute distinct P-words of their own (a conclusion which, as Cohn points out, violates the Strict Layer Hypothesis (Selkirk 1986)). I leave the status of prefixes undetermined in this analysis: they perhaps are extrametrical elements, or else are directly dominated by a higher metrical unit.
Chapter 5

Compounds, as predicted by rule (50), consist of (at least) two P-words in Balinese. This can be seen for example from vowel allophony:

(57) tai-n sampi 'dung-GEN-horse' \( \gets \text{tam l}_\text{w} \text{l sampi l}_\text{w} \)
mula padi 'plant(v) rice' \( \gets \text{mul l}_\text{w} \text{l padi l}_\text{w} \)

and the failure of h-deletion:

(58) bubuh iñjin 'porridge-black.rice' \( \gets \text{bubuh l}_\text{w} \text{iñjin l}_\text{w} \)

5.8 Stress patterns and interaction with prosodic structure

Previous descriptions, where they mention stress, at best simply state that 'the final syllable' in the word is stressed (Tinggen 1984).\(^7\) The discussion below shows that this is oversimplified. The stress rule for words uttered in isolation is generally simple: stress the final (rightmost) syllable of the (phonological) word. I could not discern, nor elicit, secondary stresses.\(^8\)

(59) mata \( \rightarrow [\text{ma'\text{t3}}] \) 'eye'
gedeg \( \rightarrow [\text{ge'deg}] \) 'angry'
takeh \( \rightarrow [\text{ta'kax}] \) '[I] thought [as much]!'
matah \( \rightarrow [\text{ma'tax}] \) 'pathetic!
katiwawalan \( \rightarrow [\text{katiwawa'lan}] \) 'undeveloped jackfruit bud'

Where a word is inherently reduplicated, it carries two stresses; stress is therefore assigned before reduplication, in the lexical phonology:

(60) gigi \( \rightarrow [\text{gi'gi}] \) 'tooth'
gigit \( \rightarrow [\text{gi'titt}] \) 'place name'
crukcuk \( \rightarrow [\text{cruk'cuk}] \) 'bird sp.'
lumbalumba \( \rightarrow [\text{lumbalumba}] \) 'dolphin'

This is true in forms with complex-formant 'prefixes', where a 'heavy' syllable is reduplicated, either inherently, or due to derivation:

(61) kedongdong-cebwah \( \rightarrow [kowsd'\text{d3d3}] \) 'k.o. fruit'
pe-cebwahbwah \( \rightarrow [pacs'bwah'bwah] \) 'id.PL'

---

\(^7\)Generic objects behave as though they are incorporated into the verb, forming a single compound verb (9.5.2).

\(^8\)The brief description of stress in Kersten (1984:14-15) apart from being internally inconsistent, makes unfounded claims, for example that (as in Malay), syllables containing a cannot be stressed. This is repeated in Barber (1978).

\(^9\)Of all the areas studied, eliciting stress data was the most difficult. I preferred to use speakers' own judgements of stress patterning, rather than my own, though they often had difficulty in discerning any stress. This presumably follows from the non-distinctive distribution patterns. I invite native (Singaraja/Buleleng) analysts to confirm, or correct, the present description, particularly with respect to secondary stress.
Prosodic structure

The doubling of stress in inherently reduplicated forms suggests that perhaps some unit other than the P-word is the unit of stress assignment in the lexical phonology. However, complex morphemes like *katiwawalan* (4.1), or *dumalada* (6.1.1), carry only a single stress. I conclude that the only relevant stress-bearing unit is the rightmost syllable in the phonological word. In 8.9 I argue that the phonetic environment of the copy (as opposed to the base) in inherent reduplications cannot be considered to be phonologically distinctive from that in the base itself. Hence the presence of a stress in the copy does not mean that there is an 'extra' stress in inherent reduplications, from a phonological point of view. It is merely part of the 'phonetic snapshot' which is the copy. In derivational reduplication, on the other hand, the extra stress can signify separate P-word status (8.9.1).

The addition of suffixes shifts the stress to the right:\textsuperscript{19}

(62) jagur \[ja'gur\] 'punch'
jagur-ang \[ja'gu'ran\] 'punch-APP; punch for'
me-jagur-an \[me'ja'gu'ran\] 'me-punch-an; be punched (vi, PL)'

Where reduplicated forms carry suffixes, stress shifts to that suffix:

(63) nengneng \[ne'ng'ne\]  'stare at (vt)'
me-nengneng-nengneng-an \[me'ne'ng'ne'ng'ne'\an\]  'stare at.DUR/RECIP'
lengeh \[le'ng\]  'intoxicated'
lengeh-lengeh-an \[le'ng'de\an\]  'dull, dopey'
crukcuk-é \[kurk'cu'ke\]  'bird sp.-DEF'
lumbalumba-né \[lumbalumba'ne\]  'dolphin-DEF'

The morphemes -é 'definite marker', -a 'irrealis', and -a 'third person actor' show ambiguous behaviour with respect to the P-word unit. By the criteria in 5.7 they behave as though they constitute a single P-word with the stem to which they affix. However they only attract stress optionally: sometimes they are stressed, sometimes not:

(64) mata-né \[ma'tane\]~\[ma'tane\]  'eye-DEF'
paling-a \[pali'ng\]~\[pali'ng\]  'steal-3; stolen by him'
paling matah-a \[ma'ta\3\]~\[ma'ta\3\]  'most pathetic!

(65) crukcuk-é \[kurk'cu'ke\]~\[kurk'cu'ke\]  'bird sp.'
lumbalumba-né \[lumbalumba'ne\]~\[lumbalumba'ne\]  'dolphin'

When they don't carry stress those morphemes behave like 'true' phonological clitics, which never take stress. Included in this group is the enclitic -né, third person pronominal:

(66) i\[i\] kadek, sing gedé bayo-n-né?
DET K. NEG big energy-GEN-3

*That Kadek, hasn't he got a lot of energy!*

\textsuperscript{19}This is not the case in Denpasar/Badung dialects, and has led to the loss of distinctions there. For example where Peguyangan/Singaraja varieties require a given verb to carry a derivational suffix, that same verb occurs suffixless in (low) Denpasar Balinese.
Other enclitics include particles such as séh 'question particle', ja 'imperative particle', suba ~ ba [ba] 'already', when used in commands, ti 'very' (stressed in its full disyllabic form, gati), and ké 'question particle':

(67)  

a. nyén = séh?  
who?  
b. eda = ja!  
DON'T-PART  
don't!  
c. jemak = 'ba!  
take-PART  
take it!  
d. gedeg = 'ti!  
angry-VERY  
very angry!  
e. oo kéné = ké gumi-né jani  
PART like.this-PART world-DEF NOW  
Oh, so the world is like this now?

Proclitics include aspectuals like lakar ~ kar ~ ka' [kæ] 'FUT, suba ~ ba[bə] 'already', the determiners used before human (and 'humanised') nouns (usually names), gender-neutral i (66), and ni (feminine): i madé, ni nyoman.

Enclitics are always monosyllabic. They can be considered to fill an extrametrical syllable at the edge of a higher prosodic unit, the P-PHRASE. The P-phrase carries a single prominent phrase stress, on the last P-word. Individual P-words earlier in the P-phrase are not perceptibly stressed:

(68) tai-n sampi  
dung-GEN-horse, horse dung'  
[ta'im]l_w[l'sam'pi]l_w|pp  
*[^ta'insi'mam'pi]  
mula padi  
'plant(v) rice'  
[mu'la]l_w[pa'di]l_w|pp  
*[^mu'lepa'di]  

As in English (Ladefoged 1993), then, words in isolation are stressed differently from when they occur in phrase: only the phrase-final word carrying stress. Is this two phenomena, or a single one: does a word spoken in isolation constitute a P-phrase, so that stress necessarily falls on the final syllable of that word? The dual (phonetic) stress of inherent reduplicates indicate that stress is assigned early in the lexical phonology: stress is therefore both a word-level as well as a phrase-level phenomenon. Imposition of phrase-level stress therefore involves the deletion of word-level stress everywhere except in final position.

Cliticised elements (as opposed to true clitics) attract stress: they therefore belong within the P-phrase. They include disyllabic pronominals, and nonreferential nominal agents which cliticise to UP verbs:

(69) beli-n-ipun lakar jemak-tiang  
'brother-GEN-3.DEP; his brother'  
amah = bikul!  
'eat.CRUDE-rat; (it) was eaten by a rat'  

The parsing of P-Phrases appears to generally correspond to syntactic bracketing, though I did not study this in detail.
Prosodic structure

(70)  \[ \textit{engkah-né} \_np \textit{ngebah-nge-bah-ang} \_vp \textit{témbo}k\_vp \]
halitosis-3 RED-AP-collapse-APP wall

\textit{His bad breath knocks down stone walls} \\

(71)  \[ \textit{ampun} \_vp \textit{m-aduk-an} \_vp \_ [di nyuh-é} \textit{nika} \_pp, \_vp \textit{kembali-ang pulang} \_vp \]
already VBL-mix-an LOC coconut-DEF that.HI return.BI-AP put.in

\textit{(When it) has been mixed into the coconut, put it in (the pot) again} \\

A second pattern found in Singaraja (but not to my knowledge in South Bali) stresses the \textit{penultimate} syllable, usually lengthening it considerably. This seems a marked alternative, only used to conveying pragmatic meaning:

(72)  \textit{eda} \_k[le:]t-ang-a
DON'T like.that.-APP-3
\textit{Don't [do it] like that!} \\

(73)  A: \textit{anak bali?} 
person Bali
A Balinese?

B: \textit{sing, anak} \_ja:wa, \_kala kurenan-né \_anak bali.
NEG Java but spouse
\textit{No, a Javanese, though her husband's a Balinese} \\

(74)  A: \textit{demen?} 
happy?

B: \textit{beih! demen} \_ja:ti! 
PART very
\textit{Oooh, I'm very happy} \\

109
6

Constraints on morph
and morpheme structure

6.0 Introduction: terminology

In 4.1 evidence was given for the existence of the morph as a purely phonological unit, involved in word-formation processes. The present chapter discusses in turn morph- and morpheme-level constraints relating to (i) prosodic constituents (ii) consonants (iii) vowels (iv) simple versus complex morphemes.

Certain structural restrictions then are conveniently expressed in terms of the morph and the morpheme (defined below), as phonological units, rather than in terms of prosodic units such as the syllable, or the phonological word. In fact, since morphs must also parse as Feet during the lexical phonology (5.6), many of these constraints appear equally, and sometimes better expressible in terms of the Foot. I generally give priority to the morph, as realising a purely segmental structure as well as a prosodic one.

These constraints include (i) restrictions relating to a particular position in a morph (eg restrictions on which consonants may occur morph-finally 6.2.4), and (ii) restrictions applying across the whole morph (eg consonant-harmony constraints, 6.2). Most of the latter can be attributed to constraints against adjacent like segments in phonological representations (Odden 1994, Clements 1995).

I reserve the term MORPHEME to refer to an entity which has both an underlying phonological specification and lexical meaning associated with it: melah /malah/ 'good', mémé %RED|me| 'mother'. A morpheme which contains no phonological elements more complex than a phoneme I refer to as a SIMPLE MORPHEME: melah /malah/ 'good', anak /anak/ 'person'. A morpheme containing a concatenation of complex formants (4.0) I call a COMPLEX MORPHEME: mémé %RED|-me| 'mother', katibangbung %kati|RED|-bun| 'k.o. brightly coloured beetle'.

The term MORPH is used in this description to refer to a phonological unit. It refers indiscriminately to either (i) non-prosodic complex formants (ie excluding processes, such as inherent reduplication) eg |kati| in katibangbung above, (ii) the output of inherent reduplication (/meme/, /lumbalumba/ 'dolphin') or (iii) to the phonological material constituting a simple morpheme /malah/, /anak/. A morpheme can therefore occasionally consist of more than one morph: for example, the morpheme katibangbung

1Strictly speaking, this applies only to lexical morphemes. Grammatical morphemes like affixes produce lexemes of predictable meaning; I assume however that that meaning is assigned by the derivation process as a whole (15.1).

111
% |kati| - |RED| - |bung| % 'k.o. brightly coloured beetle' consists of two disyllabic morphs (/kati/ and /banbung/).

The terms 'morpheme', 'morph' and 'complex formant' all convey information about the semantic status of formants (morphemes are meaningful units, the others are non-meaningful). Rather than multiply terminology, I follow general practice in using 'morpheme' in a second way, to refer just to the phonological material realising a given 'morpheme' (in the first sense). The definitions 'simple morpheme' and 'complex morpheme' above assume this second understanding of the term. I also sometimes use the cover term 'morpheme structure constraints', even though the majority of these are in fact constraints on the structure of morphs, which may or may not realise meaningful morphemes. Others, such as (57) in 6.4, do refer to the notion 'morpheme' in its primary sense, a linking of both form and meaning.

6.1 Preferred morph and morpheme structure: number of syllables

The Balinese lexicon contains roughly 3% monosyllabic morphemes and roughly 10% multi-syllabics (of three or more syllables). Disyllabics make up the bulk (around 87%) of the lexicon. A count using Fiesta software (Alsop 1990) gave the following figures, for a database of 14557 items taken from Warna et al (1990): 2

<table>
<thead>
<tr>
<th></th>
<th>%'age of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>monosyllabics</td>
<td>475</td>
</tr>
<tr>
<td>disyllabics</td>
<td>12629</td>
</tr>
</tbody>
</table>

These figures agree roughly with those found by Beratha (1992). 3 The great majority of disyllabic headwords in Warna et al (1990) each contain only a single, simple, morpheme, hence a single morph. However the situation is more complex with longer headwords: many - perhaps all containing four or five syllables - are complex (6.1.1). A separate search of Warna et al (1990) found 229 headwords containing four syllables (221 tokens) or five syllables (7 tokens): by elimination that leaves 1224 trisyllabic headwords (but not necessarily simple morphemes), in the 14557 item database.

A detailed semantic analysis of non-disyllabic morphemes is given in Appendix G. More than 90% of the samples consisted of expressives (3.2.2), loans or, if monosyllabic, functors. This parallels the claims, not supported by statistical analysis, made by Uhlenbeck for Javanese.

Once the morph-morpheme distinction is made (6.0), it is clear that well over 90% of the Balinese lexicon is made up of disyllabic morphs (see also 6.1.1). Both the

---

2 Warna et al (1990) contains more than 15,000 headwords. Certain items, such as names of literary figures and terms restricted to Old Javanese literary texts were not included in the database.

3 Beratha gives the ratios as 90% disyllables, 3% monosyllables, 7% 'trisyllables', as though this is the largest size; no actual figures are given, and the syllable analysis used is not clear.
frequency data and the semantic data then indicate a strong preference for disyllabic morphs. I take this to result from the Foot Well-formedness rule, (43), 5.6), repeated here:

(2) Foot Well-formedness

Morphs must be prosodically well-formed at the Foot level, throughout the lexical phonology

This, together with the rule preferring morphemes which are not complex (rule (57)), ensures that disyllabic morphemes predominate in the lexicon.

6.1.1 Maximum size of morphs, and of morphemes

Evidence is given here that the maximum size of a morph in Balinese (simple morpheme or complex formant, 6.0) is three syllables. This confirms that, to be well-formed, morphs must be parsable as Feet (6.1). Evidence is also given that the absolute maximum size of a (non-borrowed, complex) morpheme is four syllables.

The first evidence for a three-syllable limit to morph size comes from inherently reduplicated morphemes. Those which undergo full-base reduplication never have bases of more than three syllables:

(3) base morpheme
| [gumitt] | gumatagumit | 'creepy-crawlies' |
| [gedeber] | gedebargetedeb | 'body movements of fat person' |
| [kebir] | kebaratkebirit | 'run about wildly' |

A number of native morphemes of four syllables are listed in Warna et al (1990). In 4.1 evidence is given to show that morphemes such as the following reduce to concatenations of no more than di-syllabic formants:

(4) kalialah 'rainbow' kaliadrem 'k.o. cake'
katulayah 'lie stretched out, helpless' kaliambung 'ricochet'
katulajeg 'stand erect, in rows' kaliasem 'k.o. tree'
kalimayah 'many-legged insect, shines in dark'

Others consist of a disyllabic formant, plus a reduplicated monosyllable:

(5) katibangbuk %|kati|-[RED]-|bun|] % 'kind of insect'
kalibukbuk %|kali|-[RED]-|bunk|] % 'place name'

Still other apparently monomorphic quadri-syllabics like dumalada consist of two P-Words, and thus contain a morpheme-boundary, or are analysed as though they do (see Appendix F):
An analysis of all 229 headwords of four and five syllables in Warna et al (1990) identified only four out of that total which were not analysable (on currently available information) as concatenations of mono-, di-, or trisyllabic formants, such as the above, or as loanwords: see Appendix F.

I conclude from this evidence that while the preferred morph shape is a disyllabic foot (see previous sub-section), the maximal size of a morph (simple morpheme or complex formant) is three syllables. Anything bigger must consist of a concatenation of formants: either (meaningless) morphs or bound morphemes. The disyllabic-trisyllabic range corresponds respectively to the preferred and maximal sizes of the Foot, indicating that morphs have to satisfy Foot Well-formedness (5.6).

A corollary is that a simple morpheme is necessarily no more than three syllables long. I now give evidence to show that the maximal size of any (non-borrowed, complex) morpheme is four syllables.

The longest native headwords in Warna et al (1990) are five syllables long. There are just seven of these. They are treated as monomorphemic in some descriptions (eg Reoni et al 1987, Antara 1989), however they all end in the sequence /an/. Only 4% (581) of the 14557 headwords in Warna et al (1990) end in the /an/ sequence: cf Table 5.8. This indicates that the five-syllable lexemes bear one of the suffixes of this shape, either nominal or verbal.

I assume that five of the seven carry the nominalising suffix -an:

(7)  
katiwawalan % | kati | -[-REDCV | - wal |]-an% 'jackfruit bud'
katiwéwélán % | kati | -[-REDCV | - wel |]-an% 'jackfruit bud'
kalisasauan % | kali | -[-REDCV | - su |]-an% 'k.o. wasp'
tambullíngan % | tambu | -[-REDCV | - ling |]-an% 'bumblebee'
tamburérétan % | tambu | -[-REDCV | - rét |]-an% 'k.o. cicada'

while the others carry the -an suffix which derives 'locative adversative' intransitive verbs:

(8)  
katímbulan % | káti | -[-mumbul | -an% 'have a disease (of fingers)'
kalímatmatan % | kali | -[-RED | - mat |]-an% 'suffer from blurred vision'

Compare:
koréng-an 'skin disease-LOC' 'have (skin disease)'
kutu-an 'lice-LOC' 'lice-infested'
lumut-an 'moss-LOC' 'overgrown with moss'
oong-an 'fungus-LOC' 'mouldy'

Apart from the semantic and formal fit, there is phonological evidence for this analysis, from reduplication patterns. The final -an sequence is not reduplicated, just as suffixes never reduplicate (8.6):

4Similarly, of the 78 non-loan quadri-syllabics, 39 or exactly half end in the /an/ sequence.
There is then no evidence for a morpheme of five syllables (or longer) in Balinese: the maximal surface size that a complex morpheme in Balinese could possibly be is four syllables. (The quadri-syllabic sequences preceding the -an suffix in (7) and (8) are clearly composed of recurrent complex formants: I have no evidence for or against assigning them (complex) bound morpheme status.

To sum up this section, I have given evidence for the following limits to formant size:

(i) maximum morph size: three syllables,
(ii) maximum size of a simple morpheme: three syllables
(iii) maximum possible size of a complex morpheme: four syllables

6.2 Morph-level constraints on consonants

6.2.1 The 'one per place per morph' constraint

Balinese is similar to languages such as Arabic (Greenberg 1960) and Javanese (Yip 1989) in dispreferring the occurrence of more than one consonant of a given place of articulation within a morph. The following preferred rule applies:

(10) Prefer morphs in which two consonants sharing the same place of articulation do not cooccur

The evidence from semantics presented in Appendix B indicates that, at the very least, rule (10) applies to surface phonological structure: morphemes which violate it belong consistently to the expressive semantic classes (3.2.2). McCarthy (1986, 1994) argues that that similar effects in Arabic are due to constraints on underlying representations. This appears to be the case in Balinese too: the main class of potential exceptions is homorganic nasal-stop sequences. These are created by assimilation processes and involve a single underlying place specification (2.1.4) so would not violate (10) if it applies underlyingly; there is moreover no evidence for expressiveness being associated with homorganic nasal-stop sequences. All other structural types where more than one consonant of a given place of articulation occurs in a morph, including inherent reduplications, appear to violate (10) at a surface level.

Being a preferred rule, (10) has exceptions, but these consistently occur in expressive and other peripheral morpheme types. Indeed, one highly productive means of coinining

---

5 McCarthy (1994), and references there, argues for a universal 'No-Branching Condition', prohibiting underlying representations that contain a single Place node branching to two Root nodes: such structures are only created by assimilation processes, he claims.
expressives in Balinese is through 'violations' of (10), for example by using more than one dorsal consonant in a morph (3.2.5).

Only one absolute constraint on consonants of identical place of articulation is found. Strictly-adjacent sequences of identical consonants are never found, morph-internally: reduplicated forms such as *guggug, *tuttut do not occur. This is despite the facts that (i) monosyllables of shape C1VC1 do occur, (ii) there are apparently no other constraints on medial clusters in inherently reduplicated monosyllables (see 6.2.5 for further discussion). The following constraint, a more restricted variant of the Obligatory Contour Principle (Leben 1973, cited eg in McCarthy 1986), applies during the lexical phonology only: 7

(11) Root-node-adjacent identical elements are prohibited (Odden 1994)

Sequences in apparent violation of constraint (11) are created by postlexical processes - see 6.2.5, 8.9.

There are many exceptions to (10), but they are almost invariably expressives and other peripherals. This is demonstrated in Appendix B, where lexical and statistical data is examined in detail. In this section I firstly present raw statistical data, then discuss and exemplify.

Tables 6.1, 6.2 and 6.3 present the raw figures for cooccurrence of different combinations of consonants, in respectively initial (/C1VC2.../), final (/...CVC#/), and 'peripheral' (#C... C#) positions. The data base used is again Warna et al (1990). 8

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>b</th>
<th>w</th>
<th>t</th>
<th>d</th>
<th>l</th>
<th>r</th>
<th>c</th>
<th>j</th>
<th>s</th>
<th>y</th>
<th>k</th>
<th>g</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>45</td>
<td>1</td>
<td>6</td>
<td>79</td>
<td>56</td>
<td>109</td>
<td>101</td>
<td>45</td>
<td>27</td>
<td>67</td>
<td>12</td>
<td>39</td>
<td>31</td>
</tr>
<tr>
<td>b</td>
<td>3</td>
<td>45</td>
<td>5</td>
<td>45</td>
<td>74</td>
<td>30</td>
<td>88</td>
<td>26</td>
<td>24</td>
<td>47</td>
<td>20</td>
<td>55</td>
<td>31</td>
</tr>
<tr>
<td>m</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>21</td>
<td>25</td>
<td>41</td>
<td>46</td>
<td>8</td>
<td>11</td>
<td>25</td>
<td>6</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>w</td>
<td>1</td>
<td>5</td>
<td>14</td>
<td>19</td>
<td>31</td>
<td>31</td>
<td>47</td>
<td>9</td>
<td>10</td>
<td>23</td>
<td>7</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>t</td>
<td>42</td>
<td>45</td>
<td>17</td>
<td>45</td>
<td>23</td>
<td>69</td>
<td>67</td>
<td>0</td>
<td>18</td>
<td>8</td>
<td>8</td>
<td>58</td>
<td>46</td>
</tr>
<tr>
<td>d</td>
<td>18</td>
<td>6</td>
<td>15</td>
<td>11</td>
<td>36</td>
<td>40</td>
<td>43</td>
<td>0</td>
<td>3</td>
<td>17</td>
<td>6</td>
<td>29</td>
<td>17</td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>10</td>
<td>4</td>
<td>5</td>
<td>22</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>l</td>
<td>40</td>
<td>38</td>
<td>18</td>
<td>31</td>
<td>28</td>
<td>54</td>
<td>20</td>
<td>26</td>
<td>19</td>
<td>34</td>
<td>18</td>
<td>53</td>
<td>42</td>
</tr>
<tr>
<td>r</td>
<td>30</td>
<td>44</td>
<td>8</td>
<td>17</td>
<td>15</td>
<td>0</td>
<td>45</td>
<td>26</td>
<td>23</td>
<td>9</td>
<td>31</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>49</td>
<td>33</td>
<td>8</td>
<td>37</td>
<td>29</td>
<td>47</td>
<td>39</td>
<td>36</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>70</td>
<td>28</td>
</tr>
<tr>
<td>j</td>
<td>11</td>
<td>22</td>
<td>4</td>
<td>14</td>
<td>17</td>
<td>32</td>
<td>28</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td>27</td>
</tr>
<tr>
<td>h</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>19</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>s</td>
<td>74</td>
<td>52</td>
<td>35</td>
<td>51</td>
<td>79</td>
<td>106</td>
<td>110</td>
<td>4</td>
<td>12</td>
<td>46</td>
<td>17</td>
<td>75</td>
<td>58</td>
</tr>
<tr>
<td>y</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>k</td>
<td>124</td>
<td>86</td>
<td>19</td>
<td>87</td>
<td>65</td>
<td>107</td>
<td>120</td>
<td>97</td>
<td>43</td>
<td>21</td>
<td>14</td>
<td>55</td>
<td>4</td>
</tr>
<tr>
<td>g</td>
<td>16</td>
<td>55</td>
<td>5</td>
<td>39</td>
<td>49</td>
<td>73</td>
<td>64</td>
<td>12</td>
<td>9</td>
<td>21</td>
<td>16</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>j</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>30</td>
<td>10</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.1 C1-C2 combinations of Balinese (C2 = non-nasal) (data from Warna et al (1990), column = C1, row = C2).

---

6Inherently reduplicated monosyllables constitute a single morph, see 6.1.
7On the distinction between lexical and postlexical phonology, see 2.0.
8I have not used chi-square or other statistical tests on this raw data, for the reasons given in the discussion of table 5.10, 5.4.2.
Table 6.1 gives raw data for morphemes of shape \(C_1 VC_2(...)\), column = \(C_1\), row = \(C_2\).\(^9\) For \(labial-labial\), \(laminal-laminal\) and \(dorsal-dorsal\) combinations, the pattern is clear: these combinations are strongly avoided, as (10) predicts. This is so, except where \(C_1\) is identical to \(C_2\). In this latter case, analysis of the following vowels shows that these are inherently reduplicated: see 8.8. The apparent violation is thus a 'surface' one only, due to reduplication processes.

With the coronals the picture in table 6.1 is more complex. There are quite a number of morphemes containing \(apical-apical\) obstruent combinations; these in fact consist largely of expressives (see Appendix B). The \(liquids\) /\(l/\) and /\(r/\), appear to behave as though they constitute a distinct place of articulation; the expected avoidance pattern is found for /\(\text{rV}VI.../\) sequences, but there is an unexpectedly large number of /\(\text{rV}VI.../\) (a pattern repeated in the other tables). Large numbers of /\(\text{rV}VI.../\) and /\(\text{rVI}.../\) are found: on inspection these again turn out to be reduplications (Appendix B). Finally, where \(C_1\) is an \(apical\) and \(C_2\) a \(laminal\) the avoidance pattern is strong, though this is not the case when the order is reversed. See Appendix B for data and further discussion.

Table 6.2 gives figures for consonant cooccurrence in final closed syllables (\(l...CVCC\)). The patterns are similar to those for table 6.1: there is a strong avoidance of \(labial-labial\) combinations.\(^10\) Laminal-laminal combinations in table 6.2 are also avoided (\(\text{/sl/}\) is the only laminal occurring finally).

<table>
<thead>
<tr>
<th></th>
<th>p</th>
<th>b</th>
<th>m</th>
<th>t</th>
<th>d</th>
<th>n'</th>
<th>l</th>
<th>r</th>
<th>s</th>
<th>k</th>
<th>g</th>
<th>n</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>21</td>
<td>30</td>
<td>26</td>
<td>22</td>
<td>18</td>
<td>14</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>b</td>
<td>5</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>22</td>
<td>31</td>
<td>27</td>
<td>23</td>
<td>19</td>
<td>15</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>c</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>23</td>
<td>32</td>
<td>28</td>
<td>24</td>
<td>20</td>
<td>16</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>d</td>
<td>7</td>
<td>9</td>
<td>11</td>
<td>13</td>
<td>15</td>
<td>24</td>
<td>33</td>
<td>29</td>
<td>25</td>
<td>21</td>
<td>17</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>e</td>
<td>8</td>
<td>10</td>
<td>12</td>
<td>14</td>
<td>16</td>
<td>25</td>
<td>34</td>
<td>30</td>
<td>26</td>
<td>22</td>
<td>18</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 6.2 \(...CVCC\) interactions (column = \(C_1\), row = \(C_{\text{final}}\)).\(^12\)

With the \(liquids\), there is a strong pattern of avoidance, except that, again, large numbers of the sequence /\(\text{rV}VI.../\) are unexpectedly found. For the \(apical\) obstruents, only

\(^{9}\)For ease of presentation, \(C_2\) = non-nasals only; data where \(C_2\) is a nasal (typically before a homorganic stop) shows the same patterning as in table 6.1. Data and analysis is available from the author. Very low numbers of morphemes are found where \(C_1\) is a glide or a nasal, no matter what the place of articulation of the following consonant is. These moreover are consistently expressive in meaning, or are loanwords: they are therefore not included in the present discussion.

\(^{10}\)This is so, even taking into account the general avoidance of final labials.

\(^{11}\)Multi-syllabics (ie tri-syllabic or more) ending in /\(\text{an}/\) are here considered to bear the suffix -\(\text{an}\), and are not counted.

\(^{12}\)Totals do not include underlying monosyllables, including \%CV\%CV\% reduplications (eg cakcak, cruksuk), or \%CV-CV\% reduplications (eg kekeb); nor do they include tri-syllabics ending in /\(\text{an}/\); They DO include \%CV\% reduplications, such as ubub.
/...tVd/ and /...dVd/ seem to be avoided. There is also an apparent avoidance of the sequence /...C_laminalVd/, but many examples of the sequence apical-laminal. There appears to be free dorsal-dorsal combination; in fact morphs with this sequence turn out to occur consistently in expressive morphemes (Appendix B).

Table 6.3 gives data for 'long range' interactions, of consonants in morph-initial and -final positions. Again the pattern of avoidance is clearly found for labial-labial, laminal-laminal and liquid-liquid combinations (with, again, the unexpected exception of /l...l/ sequences, which however turn out to occur in consistently expressive morphemes).

<table>
<thead>
<tr>
<th>p</th>
<th>b</th>
<th>m</th>
<th>t</th>
<th>d</th>
<th>n</th>
<th>l</th>
<th>r</th>
<th>s</th>
<th>k</th>
<th>g</th>
<th>η</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>120</td>
<td>36</td>
<td>85</td>
<td>64</td>
<td>69</td>
<td>76</td>
<td>131</td>
<td>22</td>
<td>144</td>
</tr>
<tr>
<td>b</td>
<td>3</td>
<td>0</td>
<td>14</td>
<td>105</td>
<td>21</td>
<td>57</td>
<td>88</td>
<td>64</td>
<td>77</td>
<td>124</td>
<td>58</td>
<td>146</td>
</tr>
<tr>
<td>m</td>
<td>1</td>
<td>9</td>
<td>31</td>
<td>2</td>
<td>22</td>
<td>27</td>
<td>25</td>
<td>37</td>
<td>18</td>
<td>56</td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>l</td>
<td>21</td>
<td>18</td>
<td>14</td>
<td>17</td>
<td>19</td>
<td>78</td>
<td>69</td>
<td>41</td>
<td>110</td>
<td>138</td>
<td>27</td>
<td>126</td>
</tr>
<tr>
<td>d</td>
<td>12</td>
<td>0</td>
<td>7</td>
<td>26</td>
<td>1</td>
<td>33</td>
<td>31</td>
<td>14</td>
<td>15</td>
<td>48</td>
<td>9</td>
<td>56</td>
</tr>
<tr>
<td>n</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>r</td>
<td>11</td>
<td>10</td>
<td>13</td>
<td>59</td>
<td>14</td>
<td>18</td>
<td>4</td>
<td>0</td>
<td>18</td>
<td>58</td>
<td>26</td>
<td>63</td>
</tr>
<tr>
<td>c</td>
<td>9</td>
<td>7</td>
<td>10</td>
<td>65</td>
<td>10</td>
<td>8</td>
<td>34</td>
<td>38</td>
<td>17</td>
<td>135</td>
<td>22</td>
<td>122</td>
</tr>
<tr>
<td>l</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td>45</td>
<td>5</td>
<td>21</td>
<td>15</td>
<td>26</td>
<td>13</td>
<td>29</td>
<td>27</td>
<td>59</td>
</tr>
<tr>
<td>h</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>9</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>s</td>
<td>51</td>
<td>28</td>
<td>20</td>
<td>157</td>
<td>56</td>
<td>86</td>
<td>68</td>
<td>100</td>
<td>115</td>
<td>193</td>
<td>42</td>
<td>243</td>
</tr>
<tr>
<td>y</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>g</td>
<td>22</td>
<td>3</td>
<td>24</td>
<td>101</td>
<td>6</td>
<td>18</td>
<td>64</td>
<td>89</td>
<td>66</td>
<td>93</td>
<td>80</td>
<td>188</td>
</tr>
<tr>
<td>η</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>9</td>
<td>4</td>
<td>10</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 6.3 /C_initial... C_final/ combinations (column = C_initial, row = C_final).

With apical-apical obstruct sequences there also seems to be avoidance, when figures for other 'place'-apical combinations are considered. With laminal/apical pairs, in either order, there is perhaps absolute avoidance, but of certain sequences only. Importantly, there appears to be no interaction whatsoever with dorsal-dorsal (or back­
back, including /h/) combinations, given the high numbers of morphemes containing these sequences. However, this is illusory. Dorsal-dorsal combinations are in fact highly preferred, but in the productive formation of expressives only (3.2.5, Appendix B).

On the basis of the three tables, one could postulate three independent constraints, barring cooccurrence in each of the structures singled out there. This would be to miss the obvious single generalisation (10), which accounts for all the avoidance and expressivity patterns behind the raw data presented in Tables 6.1, 6.2 and 6.3, and Appendix B. It also explains related facts, such as the overwhelming tendency in morpheme-internal sequences of structure /...(C_laminal V).../ i.e containing identical consonants in adjacent syllable onsets, for the vowels immediately following them to also be identical, eg buléning, seririt (both place names). Such structures are clearly dispreferred, Except where produced by processes of reduplication (see 8.4), such structures are clearly dispreferred.
In the rest of this section I give some examples of morphemes which contain structures 'violating' (10), and sum up the main qualifications that need to be made. Again, for a fuller discussion, see Appendix B. Examples of exceptional morphemes include the following, all expressives (data from Warna et al (1990), expressive class indicated):

(12) \{#C_o···C_o#\}
- bengep 'shameless' BAD
- mésem 'smile sneeringly' BAD
- cedit 'sound of coconut falling' SENSE
- -gedag 'fall heavily, with a thump' SENSE

(13) \{...C_oVC_o#\}
- bintit 'take to one's heels in fear' SENSE
- kecos 'suddenly jump (eg frog)' SENSE
- mangkug 'putrid (smell)' BAD
- degag 'impudent' BAD

(14) \{#C_oVC_o...\}
- tadah 'boastingly claim [to be...]’ BAD
- césécéstan 'k.o. flying fish' ANIMAL
- kagok 'struck speechless' SENSE

(15) \{...C_oVC_oV...\}
- glilik 'roll, spinning rapidly' SENSE
- kediding 'scamper, bolt' SENSE
- beduda 'dung beetle' ANIMAL
- becica 'k.o. songbird' ANIMAL

Several qualifications to rule (10) are needed: (see Appendix B for details):

(i) Significant interactions were not observed for the sequence /t... ...C_apical/, but all other combinations of two apical consonants in a morph, in any order including /C_apical··· ...t/, were expressive.

Evidence from semantics thus suggests that neither coronal nor any other place of articulation is unmarked underlyingly in Balinese (compare Paradis & Prunet 1991).

(ii) The liquids /l/ and /r/ interact more strongly with each other than they do with the other apicals, such that they clearly form a distinct sub-class.\(^{13}\)

(iii) Just as Uhlenbeck (1949) found for Javanese, apical and laminal consonants interact. They share the same general place of articulation, the front of the tongue (Sagey (1990)'s coronal class feature).

(iv) There are very few labial 'exceptions' to rule (10), but many which involve dorsals; these however virtually all belong to the SENSE or BAD semantic classes. A favoured way of coining expressive lexis is by using cooccurring dorsal consonants (3.2.5).

\(^{13}\)McCarthy (1994:226) describes a cross-linguistic split in behaviour between coronal obstruents and coronal sonorants, with respect to OCP constraints. In Balinese /n/ appears to interact with obstruents.
(iv) For morphs of shape /...C_{dorsal}Vh#//, the semantic evidence (Appendix B) supports the grouping of /h/ with the dorsal consonants in a single 'back' place of articulation (2.1).

Rule (10) is not broken by homorganic NC clusters (which are not associated with expressive semantics), since these share a single place specification (2.1.4).

Rule (10) is broken by morphs containing inherent reduplications, such as susuk 'change (n)' /REDcv susuk/, bubu 'fishtrap' /REDbu/, kupukupu 'butterfly' /REDkupu/, even though their underlying representations do not violate the rule.

6.2.2 Morph-level constraints on the L slot

Complex onsets, containing a CL sequence (where L can be /l, r, w/ or /y/) occur overwhelmingly in semantically peripheral morphemes. This is presumably because the structure violates a preferred rule of syllable-structure (5.2.2). A related preferred rule applies at the morph level. The rule here is formulated informally:

(16) Where a morph contains a CL cluster, prefer the structure where it occurs in the left-most syllable.

All other things being equal, a morpheme such as clempung 'make a splashing noise when falling into water' has a more preferred structure than that of its less common variant form cemplung. This applies whether a morph has two, three or more syllables. In practice, since the bulk of the lexis is disyllabic, (16) most commonly disfavours CL clusters in the right-most syllable. The first evidence that these onsets are disfavoured comes from productivity.

<table>
<thead>
<tr>
<th>CV.CV</th>
<th>966</th>
<th>CVC.CV</th>
<th>490</th>
<th>CV.CVC</th>
<th>4410</th>
<th>CVC.CVC</th>
<th>2070</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLV.CV</td>
<td>185</td>
<td>CLVC.CV</td>
<td>51</td>
<td>CLV.CVC</td>
<td>811</td>
<td>CLVC.CVC</td>
<td>398</td>
</tr>
<tr>
<td>CV.CLV</td>
<td>92</td>
<td>CVC.CLV</td>
<td>23</td>
<td>CV.CLVC</td>
<td>242</td>
<td>CVC.CLVC</td>
<td>179</td>
</tr>
</tbody>
</table>

Table 6.4 Occurrence of complex onsets in morphs of comparable structure: without complex onsets (row 1); with complex onsets in initial syllable (row 2); and in final syllable (row 3), in Warna et al. 1990.

Table 6.4 shows that, for disyllabic morphs of otherwise identical structure, while morphs with initial-syllable CL clusters occur much less frequently than those without any complex onsets, CL clusters are even less favoured in final syllables. All other things being equal, there should be four times as many morphemes of shape CVCLVC as there are of shape CVCVC, since the L slot can be occupied by 4 different phonemes. The actual realization figures show the reverse tendency: there are nearly 20 times as many morphemes of shape CVCVC as there are of shape CVCLVC. This pattern recurs consistently for disyllabic morphemes of otherwise identical structure: there are from ten to twenty times as many morphemes without final complex onsets as there are with such onsets.
This indicates a continuum of markedness, with initial-syllable-onset CL disfavoured, but with final-syllable complex onsets least favoured of all.

When the semantic classes of a random sample of 120 morphemes of shape CV(C)CLVC were tallied, 91 (76%) belonged to just 3 expressive semantic classes: BAD, SOUND, MOVEMENT \((p < 0.0000)\),\(^{14}\) and only 8 (7%) were tentatively unclassified as expressive (see Appendix H). Typical examples include:

\begin{verbatim}
(17)  damprat 'yell furiously at'
caplok 'eat piggishly, gobble'
bunglun 'burnt outside, raw inside'
singlad 'awkward, clumsy'
kecret, kacrit, kecrit 'sound of spurtting water'
gedeblés, keblés, keblos 'suddenly emerge, pop out'
\end{verbatim}

Evidence from both productivity and semantics then supports the existence of rule (16). Further evidence for it comes from morphs consisting of a reduplicated CLVC sequence: the right-hand onset is almost always 'missing': crućuk (*crućruk, *cukcruk) 'k.o. bird', mresmès (*mrèsmrès, *mèsmrès) 'messy (when eating rice)' (see below for exceptions). That it is the right-most L which does not surface suggests again that complex onsets in the final-syllable are least preferred.

In tri- and other multi-syllabics filled L slots occur mainly in loans (eg brahmacari 'abstinence', anugraha 'grant', both from Sanskrit, mutyara 'pearl' (of Indian origin, cf Gonda 1973)).

If the orthography correctly reflects underlying structure, Warna et al (1990) lists only five monosyllabic morphemes with a filled L-slot, all expressives:

\begin{verbatim}
(18) plang 'obstacle (Hl)'
plek 'exclamation of surprise'
plès 'very in nguda--very young'
ngrek 'remain silent and motionless'
\end{verbatim}

It is likely that underlying forms of shape CLVC surface as /CeLVC/, undergoing schwa epenthesis (2.2.6) due to the requirements of Foot Wellformedness (5.6). Otherwise, the following surface constraint applies to the L slot:

\begin{verbatim}
(19) L is not filled in mono-syllabic morphs.
\end{verbatim}

This is only a surface constraint. L can occur (filled only by /l/ or /r/) in underlying monosyllabic complex formants, as in:\(^{15}\)

\begin{verbatim}
(20) prungpung
    -closcos % |RED| - |prug| %
    'chipped (of a pot)
(21) gruguh
    klakah % |RED| - |gruh| %
    'hand (crude term)'
    % |RED| - |klah| %
    'roof of bamboo shingles'
\end{verbatim}

\(^{14}\) The latter two semantic categories are subsets of the SENSE class, defined in 3.2.2. Procedures used in statistical analyses are described in 3.2.3.

\(^{15}\) I assume that sequences such as /cruk/ in crućuk 'k.o. bird' do not violate (19) because they do not yet constitute morphs, ie the morph is what is produced after reduplication. See also 8.9.
Chapter 6

Examples (20) and (21) also illustrate the effect of the following constraint:

(22) Prefer morphs which contain only one filled L slot

The morphemes in (20) and (21) show that the 'right-most' L is deleted after inherent reduplication, to satisfy (22). The only exceptions found so far to (22), where both L's are liquids, are three loanwords sléndro 'musical scale' (Jav.), clídra 'jealous' (Skt), bludru 'velvet' (source unknown, but also found in Malay), as well as these reduplicated expressives:

(23) -plokplok, -plukpluk 'sound made by pigs eating'
    -plongplog 'quick to grow and put on weight (of a baby)'
    (pe-)dladlad 'unevenly cut (of hair)'

Inherent reduplications of disyllabics appear to break (22): slégaksleguk 'sob', kebrinkebrin 'tousled (hair)', klingengklingeng 'look about, in confusion'. For disyllabic bases, reduplication does not occur until late in the postlexical phonology (8.9): presumably after (22) has ceased to apply.

A different reduplication process copies just the edge-most consonant of the base, to produce forms such as the following (see 8.2.2):

(24) -blańja 'go shopping' be-blańja-an 'purchases (n)'
    -slampar 'throw' me-se-slampar-an 'throw repeatedly'

One reason for preferring that reduplication analysis is because it enables a simple statement of L-deletion processes.

The constraint in (22) could be seen as simply a sub-case of the more general constraint, (10), dispreffering cooccurrence of two consonants of the same place of articulation within a morph. However it also applies in inherent reduplications, every other consonant of which continues to violate it. Other evidence that (22) is needed independently is that it also rules out cooccurring L slots with a liquid and a glide respectively. There are only a handful of morphemes of this shape, and again they belong to the 'expressive' semantic classes, or are loans:

(25) srambyah 'thick, dangling (of foliage, hair)'
    grombyong 'too big (clothes)'
    krebwak 'sound of falling branch'
    klampwak 'k.o. tree'
    tragya 'at the ready'
    mretyu 'Death' Skt

6.2.3 Morph-level constraints on the N slot

The following exceptionless constraint applies to the N slot (ie the structural slot occupied by a nasal consonant homorganic to a following obstruent):

16An analysis sometimes suggested, where /LI/ is infixed after reduplication (eg |RR|·|pung| > *pungpung > p-r-ungpung) is not motivated: there is no independent evidence for infixation as a process in Balinese.
Morph and morpheme structure

(26) No more than one N slot may be filled per morph.

There are thus no trisyllabics containing sequences such as /...VNCVNC.../. It does not seem that this rule can be explained in terms of an OCP violation (as can the non-occurence of derived 'geminate' sequences, 6.2.5). Morphemes containing nasals of different places of articulation, one in an N slot, the other, say, in final coda position, are common. For example, in Warna et al (1990) more than 280 morphemes end in either /(...).ntVN/ (panting 'wash (vt)', kantun 'still.HI', etc), or /(...).ndVN/ (lindung 'eel', pandan 'pandan'). Interactions with expressive semantics were not investigated.

While rules (22) and (26) have parallel effects, they do not bar both L and N slots being filled simultaneously (though in expressives only):

(27) brintug-an  'bumpy (of road)'     damprat     'scold, rebuke'
    grembwang  'slamming noise'     brengga     'cumbersome'

A third constraint is a preferred rule:

(28) Where a morph contains a filled N-slot, prefer the structure where that slot occurs in the leftmost syllable of the morph.

This constraint means, for example, that a structure like CVNCVVC is preferable to CVCNVC. In a sample of 3168 morphemes containing NC sequences only 85, or less than 4%, were surface exceptions to (28). Most of these however were either loans (themselves often historically complex):

(29) acintya    'one of the names of God' (Skt)
    narénda    'royal title' (Skt)
    olanda     'Holland' (Dutch)

or are morphemes containing recurrent complex formants:

(30) clekontong  'name of small container'   %|cla|-kontonj|%
    clebingkah  'potsherds'          %|cla|-binkah|%
    katilambung 'ricochet (v)'       %|kati|-lambung|%
    kalimanteg  'k.o.tree'           %|kali|-manteg|%
    ketumbah    'coriander'           %|ko|-tumbah|%
    kedongdong  'k.o.tree'           %|ko|-RED|doŋ|%
    ketengkung  'sound of dove'       %|ko|-teŋkung|%

Constraint (28) applies to tri-syllabic formants, and vacuously to disyllabics; the few longer items in the database containing NC sequences, such as katilambung, kalagaɲjur, katimumbulan are all complex (6.1.1); the disyllabic formants making them up are perfectly regular with respect to (28).

17This is assuming the analysis of orthographic <CL...> sequences as /CL/, not /CaL/ (cf 2.2.6). The /CaL/ analysis would give significantly higher figures for 'non-leftmost' NC sequences, but then these would be limited almost entirely to this environment, which would in turn require explaining.
These patterns appear to have their explanation in prosodic structure: what I refer to in 5.6 as the 'maximal (three-syllable) foot' is perhaps a foot plus an extrametrical final syllable - McCarthy & Prince's 'loosely interpreted P-word' template has this structure (cited in Kenstowicz 1994). Constraint (28) falls out from this foot-plus-extra-metrical-syllable morph structure, plus with the 'obligatory extrametrical consonant' requirement, formulated as in (49, section 5.6), and repeated here:

(31) Prefer a Foot template which includes a consonant specified for place of articulation at the right edge

6.2.4 Constraints on morph-final consonants

Morphs which end in a consonant are preferred to those which end in a vowel. This is clear when the numbers of morphs of these respective shapes are compared. Representative figures for a variety of disyllabic 'minimal pairs' are given in Table 6.5:

<table>
<thead>
<tr>
<th>V(σ)</th>
<th>CV(σ)</th>
<th>CVC(σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V.V</td>
<td>8 75</td>
<td>CVC.V</td>
</tr>
<tr>
<td>V.VC</td>
<td>115</td>
<td>CVC.VC</td>
</tr>
<tr>
<td>V.CV</td>
<td>108</td>
<td>CVC.CV 490</td>
</tr>
<tr>
<td>V.CVC</td>
<td>901</td>
<td>CVC.CVC 2070</td>
</tr>
<tr>
<td>VC(σ)</td>
<td>CLV(σ)</td>
<td>CLVC(σ)</td>
</tr>
<tr>
<td>VC.V</td>
<td>- 15</td>
<td>CLVC.V</td>
</tr>
<tr>
<td>VC.VC</td>
<td>- 122</td>
<td>CLVC.VC</td>
</tr>
<tr>
<td>VC.CV</td>
<td>56 185</td>
<td>CLVC.CV 51</td>
</tr>
<tr>
<td>VC.CVC</td>
<td>457 CLVC.CVC 398</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.5: Occurrence of disyllabic morph types - with and without final consonants in Warna et al (1990)

Each cell in Table 6.5 contains a 'minimal skeletal pair'. For a given pair, the number of tokens with a final consonant easily outnumbers those without - generally by a factor of five to one, and ranging up to more than ten to one. There are for example 966 tokens of structure CV.CV, versus 4410 of CV.CVC. There is then a clear preference for morphs with a closed final syllable.¹⁸

At the same time, for the non-final syllable, the preference is reversed: generally around twice as many morphs have an open penultimate syllable, as those with an otherwise identical CV structure, but with a closed penultimate (eg 966 of structure CV.CV versus

¹⁸An initial inspection suggests that a very high proportion of vowel-final morphemes are loanwords from Indian languages. If this is correct, the preference for final consonants is even greater than these figures suggest.
This latter fact confirms that the 'filled-final coda' preference is not a syllable-level preference, but rather a morph/morpheme level one.

The figures for monosyllabic morphs, show an equally strong preference for a syllable-final consonant:

<table>
<thead>
<tr>
<th>Shape</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>3</td>
</tr>
<tr>
<td>CV</td>
<td>28</td>
</tr>
<tr>
<td>[CLV]</td>
<td>4</td>
</tr>
<tr>
<td>VC</td>
<td>8</td>
</tr>
<tr>
<td>CVC</td>
<td>433</td>
</tr>
<tr>
<td>[CLVC]</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 6.6 Shapes of monosyllabic morphemes in Balinese**

Balinese then resembles Arabic, in preferring that lexical morphemes end in a consonant; they moreover both require that that consonant be extrametrical (see below, also McCarthy & Prince 1990:240). In Balinese this is due ultimately to a constraint on syllable structure. Morph-final consonants differ from all other consonants in being extrametrical with respect to lexical syllabification processes (5.3.2). The analysis assuming extrametricality explains why a far wider range of consonants can occur in morpheme-final position than in other potential syllable codas. The relative frequency of occurrence of different consonants in morpheme-final position is shown in the following figures from Warna et al (1990), for 11720 lexemes with a filled final coda:

<table>
<thead>
<tr>
<th>Consonant</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>1957</td>
</tr>
<tr>
<td>k</td>
<td>1547</td>
</tr>
<tr>
<td>h</td>
<td>1366</td>
</tr>
<tr>
<td>t</td>
<td>1281</td>
</tr>
<tr>
<td>l</td>
<td>777</td>
</tr>
<tr>
<td>s</td>
<td>866</td>
</tr>
<tr>
<td>r</td>
<td>911</td>
</tr>
<tr>
<td>g</td>
<td>689</td>
</tr>
</tbody>
</table>

**Table 6.7 Occurrence of morpheme-final consonants in Warna et al (1990)**

The totals in Table 6.7 show that morphemes never end in glides, nor in laminal stops, while the back consonants (/ŋ/, /k/ and /h/), as well as the apicals are favoured in final position. The laminal fricative /s/ also occurs with high frequency.

---

19 Warna lists pra-, tri-, dwi-, swa- all from Sanskrit, and of dubious synchronic status as independent derivational affixes.
20 The expressives plang, plek, plés, ngrek are listed in Warna; these should probably be analysed as underlying disyllabics CaLVC, cf 2.2.5.
21 Including 881 forms ending in /an/, not analysed for morphological complexity. See footnote 11.
The non-occurrence of the glides /w/ and /y/ in final position is the product of directional right-to-left syllabification processes (5.5), rather than a distinct structural constraint. I assume that the non-occurrence of non-continuant laminals morpheme-finally (/p/ does surface syllable-finally in homorganic clusters, 5.3.2) is due to a P-word-level negative constraint, applying throughout the phonology:

\[(32) \quad [\text{LAMINAL}]\]
\[\quad \text{*C} \text{P-word} \]
\[\quad [\text{-cont}]\]

Labials are disfavoured in morpheme-final position, occurring with the lowest absolute frequencies of any place of articulation, and then generally in expressives and loans. I assume a constraint such as the following:

\[(33) \quad \text{Disprefer} \]
\[\quad \text{C} \text{P-word} \]
\[\quad [\text{LABIAL}]\]

Similarly, the voiced member of a stop pair always occurs less frequently in morpheme-final position than its voiceless counterpart, eg 1547 /k/-finals compared to 689 /g/-finals. Morphemes ending in /g/ are almost always expressives: with either SENSE or BAD semantics.

\[(34) \quad \text{bug} \quad \text{‘thudding sound’} \quad \text{empug} \quad \text{‘break open eg coconut, bottle’}\]
\[\quad \text{cegceg} \quad \text{‘loom’ (onomat.)} \quad \text{gedeg} \quad \text{‘angry’}\]
\[\quad \text{jadig} \quad \text{‘rude (speech)’} \quad \text{beseg} \quad \text{‘water-logged, sopping wet’}\]

This preference for voiceless stops morpheme-finally probably reflects a natural phonetic tendency, rather than a language-specific rule. The related languages Malay and Javanese have lost voiced stops morpheme-finally, Javanese only relatively recently.

### 6.2.5 Cross-syllable consonant clusters

As with syllable coda constraints (5.3.2), a distinction must be made between constraints applying to cross-syllable consonant clusters in the native lexis, and those applying to loanwords. Cross-syllable clusters in the native lexis are extremely limited, and are the result of constraints on syllable structure (5.1): only homorganic nasal-obstruent sequences occur in simple morphemes:

\[(35) \quad \text{sampun} \quad \text{‘already.HI’} \quad \text{m-ambuh} \quad \text{‘shampoo (v)’}\]
\[\quad \text{untu} \quad \text{‘tooth.HON’} \quad \text{kendi} \quad \text{‘drinking vessel’}\]
\[\quad \text{a\text{\textsc{f}}cuk} \quad \text{‘prod’} \quad \text{i\text{\textsc{f}}jin} \quad \text{‘glutinous rice’}\]
\[\quad \text{angkuk} \quad \text{‘thrust with hips’} \quad \text{engging} \quad \text{‘protruding (buttocks)’}\]
Morph and morpheme structure

I argue in 2.1.4 for an analysis where the nasal consonant gains its place of articulation from the following obstruent, except in the case of /Ns/ sequences, which surface as /ns/: Sangsit (place name), dongsok 'shove', engsap 'forget'.

Only five exceptions to the above pattern have been noted in the native lexis, they are of course expressives:

(36)  
- nangdanangda  'dither, shilly-shally'
- pakpung  'have a bath (used speaking to infants, onomat.)'
- nungtit  'sound of wooden drum'
- nongcrét, crétnong  'cicada (onomat.)'

Inherently reduplicated monosyllables of shape C1VC2C1VC2, are necessarily complex, and show different surface behaviour. Constraints on medial C2C1 clusters are limited mainly by constraints on morph-final position (C2) and syllable-initial (C1) positions:

(37)  
- gudgud  'bent with age'
- butbut  'pull out'
- léklék  'eat.CRD'
- ngéknog  'have sex.CRD'

The one derived C2C1 combination which never occurs is a 'geminate' sequence: no forms such as *sus-sus, or kak-kak occur. This is despite of the facts that (i) many morphemes of shape /C1... C1/ occur (6.2), and (ii) there are at least 14 monosyllabic morphemes of form C1VC1 listed in Warna et al (1990), including:

(38)  
- ngeng  'shy, reluctant'
- mem  'soak (vt)'
- beb  'fat and short'
- gog  'playing card (king, queen, or jack)'

This striking absence, together with the existence of morphemes like the pair 'tut' 'follow', tutut 'follow, obey', as well as tatit 'lightening', -kakak 'laugh loudly', -kékék 'id.', indicates that there is an absolute constraint barring adjacent identical consonants morph-internally, even where they result from reduplication processes. This can be explained as an OCP violation (Leben 1973, cited eg in MacCarthy 1986), applying only to segments adjacent at the root node:

(39)  
Root-node-adjacent identical elements are prohibited (Odden 1994)

Where the reduplicated base is disyllabic, this constraint does not hold: kayak-kayak 'lie spreadeagled & helpless', kisik-kisik 'move slightly', see 8.9 for a suggested explanation. (With vowels, in contrast, identical sequences can be set up by phonological processes, 8.7.)

---

22Intervocalic CL clusters such as those in bunglon 'k.o. lizard', caplir 'miss (target)' are analysed as syllable-onset internal, see 5.2.1.
23The semantic marginality of this group is clear: several are used para-linguistically, in commands given to animals: kek 'command used in ploughing', sus 'sound used to drive away animals', sis 'turn right! (to cattle, when ploughing)
Loanwords allow a much wider range of cross-syllabic cluster types than occur in the native lexis:

<table>
<thead>
<tr>
<th>Loanword</th>
<th>Meaning</th>
<th>Loanword</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>mangda</td>
<td>'in order that' Jav.</td>
<td>wangledé</td>
<td>'cancel' Jav.</td>
</tr>
<tr>
<td>amerta</td>
<td>'water of life' Skt</td>
<td>orta</td>
<td>'news' Skt</td>
</tr>
<tr>
<td>sakti</td>
<td>'magically powerful'</td>
<td>aksara</td>
<td>'letter' Skt</td>
</tr>
<tr>
<td>brahma</td>
<td>'Brahma' Skt</td>
<td>wisnu</td>
<td>'Wisnu' Skt</td>
</tr>
<tr>
<td>awignam astu</td>
<td>'let there be no obstacle!'</td>
<td>atma</td>
<td>'soul' Skt</td>
</tr>
<tr>
<td>bioskup</td>
<td>'cinema' Dutch</td>
<td>saniscara</td>
<td>'Tuesday' Skt</td>
</tr>
</tbody>
</table>

6.3 Morph-level constraints on vowels

Structural restrictions discussed here are again ones which must be stated in terms of the morph, rather than the syllable. They include ones applying across the morpheme generally (6.3.1), and others applying to morpheme-final syllables (6.3.2).

6.3.1 Constraints on vowels in adjacent syllables

While a detailed study of vowel cooccurrence has not been done, some data for cooccurrences of vowels in adjacent syllables was gathered, using a Fiesta (Alsop 1990) search of the Warna et al (1990) database. Table 6.8 gives totals for adjacent $V_1V_2$ sequences (with no other intervening segment), while Table 6.10 gives frequencies of $V_1CV_2$ sequences, in adjacent syllables separated by a single onset consonant. Since nearly 90% of that data base is disyllabic, the figures give a good first approximation to vowel cooccurrence patterns in the lexicon. However, since the data base also contains morphemes of three and more syllables, as well as disyllables of shapes not picked out by these two search patterns, the figures in this table should be treated as preliminary.

Each table is followed by a further table, giving the results of chi-square analysis. In each case the figures are distributed in a highly non-random way. Nevertheless, as the investigation of onset-nucleus interactions demonstrates, (5.4.2), it is unwise to conclude too much from statistical analyses, unless the presence of loanwords and expressives has been taken into account, which is not the case in either of tables 6.8 or 6.10. The following generalisations are offered with those caveats.
Table 6.8 \( V_1V_2 \ldots \) sequences in Warna et al (1990).

<table>
<thead>
<tr>
<th>V1</th>
<th>i</th>
<th>u</th>
<th>é</th>
<th>o</th>
<th>e</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>36</td>
<td>74</td>
<td>5</td>
<td>32</td>
<td>47</td>
<td>169</td>
</tr>
<tr>
<td>u</td>
<td>46</td>
<td>66</td>
<td>32</td>
<td>5</td>
<td>36</td>
<td>148</td>
</tr>
<tr>
<td>é</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>33</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>o</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>21</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>e</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>57</td>
<td>0</td>
</tr>
<tr>
<td>a</td>
<td>66</td>
<td>104</td>
<td>8</td>
<td>27</td>
<td>24(30)</td>
<td>79</td>
</tr>
</tbody>
</table>

Table 6.9 \( V V \ldots \) sequences in Warna et al (1990). (o-e) /e for each combination.

<table>
<thead>
<tr>
<th>V1</th>
<th>i</th>
<th>u</th>
<th>é</th>
<th>o</th>
<th>e</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>1.69</td>
<td>0.00</td>
<td>15.81</td>
<td>0.38</td>
<td>0.38</td>
<td>10.05</td>
</tr>
<tr>
<td>u</td>
<td>0.59</td>
<td>0.04</td>
<td>3.75</td>
<td>23.51</td>
<td>2.65</td>
<td>5.76</td>
</tr>
<tr>
<td>é</td>
<td>8.51</td>
<td>14.03</td>
<td>49.47</td>
<td>101.18</td>
<td>9.78</td>
<td>3.35</td>
</tr>
<tr>
<td>o</td>
<td>7.9</td>
<td>13.01</td>
<td>36.50</td>
<td>34.40</td>
<td>9.07</td>
<td>0.30</td>
</tr>
<tr>
<td>e</td>
<td>7.03</td>
<td>11.59</td>
<td>3.90</td>
<td>5.61</td>
<td>296.18</td>
<td>20.81</td>
</tr>
<tr>
<td>a</td>
<td>19.20</td>
<td>25.25</td>
<td>8.44</td>
<td>0.49</td>
<td>(62.50)</td>
<td>480.61</td>
</tr>
</tbody>
</table>

From tables 6.8 and 6.9 above it appears that where two vowels are immediately adjacent, with no intervening consonant, the following occur with much higher than expected frequency:

(i) Except for the high vowels, sequences of identical vowels are strongly favoured.

(ii) The mid vowels /e/ and /o/ occur with high frequency with each other, in either order.

The following occur with lower than expected frequencies:

(i) the mid vowels /e/ and /o/ never occur immediately before a following high vowel. The reverse sequences occur, but in low numbers in the case of /uo/ and /ie/ (but not /ue/ or /io/ - though morphemes of these shapes shows are consistently expressive, or loanwords: compare Appendix I.

(ii) /a/ never occurs immediately before another vowel, except itself.\(^{25}\) On the other hand, there are many morphemes with /aCV/ sequences, see below.

(iii) there are lower than expected occurrences of all vowels, except /i/, before /a/. In Singaraja Balinese, the sequence /aa/ no longer occurs morpheme-internally: thus South Bali mael 'expensive', jaen 'delicious', Singaraja maal, jaan

(iv) The sequence /ae/, is also disfavoured, as well perhaps, as /ao/ and /ea/.

\(^{24}\)In Singaraja Balinese this sequence does not occur at all, being replaced by /aa/.

\(^{25}\)The one exception to this is an expressive, belhl a 'strongly emphatic particle'.

129
Chapter 6

The constraint barring /aVi/ sequences applies both morpheme-internally, and across a morpheme-boundary, when that boundary is between a root and a suffix (7.3). It does not apply across a prefix-root boundary. I conclude that it is a P-word level restriction:

\[ *aV \]p-word \[ where V is a vowel other than a \]

Table 6.10 gives figures for vowels in adjacent syllables, where a consonant intervenes. This is a much more common pattern, due to the preference for filled syllable onsets (5.2.1).

<table>
<thead>
<tr>
<th>V1</th>
<th>i</th>
<th>u</th>
<th>e</th>
<th>o</th>
<th>a</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>394</td>
<td>183</td>
<td>28</td>
<td>30</td>
<td>174</td>
<td>506</td>
</tr>
<tr>
<td>u</td>
<td>259</td>
<td>513</td>
<td>43</td>
<td>17</td>
<td>148</td>
<td>454</td>
</tr>
<tr>
<td>e</td>
<td>37</td>
<td>19</td>
<td>321</td>
<td>113</td>
<td>34</td>
<td>195</td>
</tr>
<tr>
<td>o</td>
<td>48</td>
<td>2</td>
<td>135</td>
<td>350</td>
<td>40</td>
<td>213</td>
</tr>
<tr>
<td>a</td>
<td>459</td>
<td>385</td>
<td>233</td>
<td>231</td>
<td>715</td>
<td>606</td>
</tr>
<tr>
<td></td>
<td>1793</td>
<td>1577</td>
<td>926</td>
<td>910</td>
<td>1364</td>
<td>3228</td>
</tr>
</tbody>
</table>

Table 6.10 ...V1CV2... combinations in Warna et al (1990)

Table 6.11 gives the results of a chi-square analysis applied to table 6.10. The figures are distributed in a highly non-random way (p < .0000).

<table>
<thead>
<tr>
<th>V1</th>
<th>i</th>
<th>u</th>
<th>e</th>
<th>o</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>12.2</td>
<td>-2.0</td>
<td>-9.6</td>
<td>-9.2</td>
<td>-0.4</td>
</tr>
<tr>
<td>u</td>
<td>0.2</td>
<td>22.5</td>
<td>-8.8</td>
<td>-11.2</td>
<td>-3.9</td>
</tr>
<tr>
<td>e</td>
<td>-9.3</td>
<td>-10.0</td>
<td>34.0</td>
<td>6.4</td>
<td>-7.2</td>
</tr>
<tr>
<td>o</td>
<td>-9.0</td>
<td>-12.4</td>
<td>8.0</td>
<td>35.9</td>
<td>-7.3</td>
</tr>
<tr>
<td>a</td>
<td>-0.7</td>
<td>-1.8</td>
<td>-0.8</td>
<td>-0.6</td>
<td>23.6</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>-1.1</td>
<td>-9.2</td>
<td>-8.7</td>
<td>-10.9</td>
</tr>
</tbody>
</table>

Table 6.11 Sample of V1CV2 combinations in Warna et al (1990)

\[ DF = 25, X^2 = 4146.03 \ P < .0000 \]

*italics* = lower frequency than predicted

*underline* = higher frequency than predicted

Combinations occurring with higher than expected frequency include:
(i) Again, morphemes where V1 and V2 are identical. These are very strongly preferred.
(ii) Again, morphemes where V1 and V2 are /o/ and /e/, in either order.
(iii) Morphemes where V1 and V2 are /a/ and /a/, in that order only - the reverse order being strongly disfavoured. (Both orders were dispreferred in VV sequences).

Combinations occurring with much lower than expected frequency include:
Morph and morpheme structure

(i) Again, morphemes where either /o/ or /el/ cooccurs with either /u/ or /i/, in either order.

This pattern is described in detail in 3.1, as well as in Appendix I. It is also found in Javanese (Uhlenbeck 1949) and Malay (Adelaar 1990). Evidence from loanword 'regularisation' and semantics (3.1), as well as glide-vowel cooccurrence (5.4.2) supports the existence of a preferred rule such as the following:

(42)  
\[
\begin{array}{c|c}
\text{Prefer:} & \text{Disprefer:} \\
\hline
+ \text{ATR} & + \text{ATR} \\
\text{\}/ /} & \text{\}/ /} \\
\text{-} & \text{-} \\
\alpha \text{HL} & \alpha \text{HL} \quad - \alpha \text{HL}
\end{array}
\]

That is, 'prefer morphs where cooccurring [+ATR] vocoids agree for the feature value [HL], either both [+HL], or both [-HL].'

(ii) also of low frequency are morphemes where /a/ occurs in a syllable preceding any of the non-high vowels, /el/, /al/ or /ol/. This is true also for immediately adjacent VV sequences (above). This suggests, though this would need to be confirmed by other evidence, the existence of a constraint such as:

(43)  
\[
\begin{array}{c|c}
\text{a. Prefer:} & \text{b. Disprefer:} \\
\hline
- \text{HL} & - \text{HL} \\
\text{\}/ /} & \text{\}/ /} \\
\text{-} & \text{-} \\
\alpha \text{LO} & \alpha \text{LO} \quad + \text{LO} \\
\end{array}
\]

That is, 'prefer morphs where cooccurring [-HL] vowels agree for the feature value [LO], either both [+LO], or both [-LO].' The mirror image of (43b) appears to hold as well, except where V1 is /a/: the sequence /aCa/ appears strongly favoured.

(iii) As in VV sequences, /a/ is dispreferred after other vowels, except itself and /il/.

Without a semantic analysis of the data, strong conclusions about preferred rules cannot be reached: again, see 5.4.2 for a demonstration of this fact.

6.3.2 A constraint on morpheme-final vowels: non-occurrence of /a/ finally.

Evidence is given in this section that the following negative constraint applies to underlying representations:

(44)  
\[\text{\^a]}_{\text{lexical morpheme}}\]
Chapter 6

As well, this constraint applies to surface realisations:

(45) \[^{\text{V}}\]P-word

[-ATR]

I also give reasons for the decision to treat P-word-final [\textipa{3}] as a surface allophone of an intermediate /\textipa{a}/, rather than of /\textipa{a}/ (cf 2.2.4).

At the phonetic level, the phone [\textipa{a}] never occurs phonological-word-finally. Orthographic final <\textipa{a}> is always realised as [\textipa{3}] P-word-finally, and in final <\textipa{aa}> geminate sequences:

(46) mata [mat\textipa{̃}] 'eye'
    aba [ab\textipa{̃}] 'carry'
    paa [p\textipa{̃}] 'thigh'

However, when such forms carry suffixes or enclitics, the phonetic realisation is always [\textipa{a}], where it occurs lexical-morpheme-finally:

(47) mata-né [mat\textipa{̃}e] 'eye-DEF'
    aba-ang [ab\textipa{̃}aŋ] 'carry-APP'
    paa-n-né [p\textipa{̃}ae] 'thigh-GEN-3; his thigh'

This alternation suggests two possible underlying forms: morpheme-final <\textipa{a}> could represent either %\textipa{a}% or %\textipa{e}% underlingly, [\textipa{3}] being phonetically closest to the vowel /\textipa{a}/. For Peguyangan/Singaraja dialect at least, the simplest analysis is that it represents underlying %\textipa{a}%, thus mata %mata%, aba %aba%, and so on.

A corollary is that in the underlying representation of lexical morphemes, %\textipa{a}% never occurs in final position. Evidence for this comes from /\textipa{a}/-to-/o/ alternation. In that process, unambiguous cases of underlying /\textipa{a}/ regularly alternate with /\textipa{o}/ when they come into pre-vocalic position due to suffixing/cliticisation and /h/-deletion (cf (7.3)):

(48) keneh %k\textipa{en}eh% 'thought'
    pe-keneh-in %p\textipa{e}k\textipa{en}eh-in% 'think about'
    CAUS-thought-LOC
    t\textipa{e}geh %t\textipa{e}geh% 'high'
    t\textipa{e}geh-an %t\textipa{e}geh-an% 'high-CMP, higher'

This contrasts with alternations such as:

(49) kena %k\textipa{en}a% 'be hit (vi)'
    kena-in %k\textipa{en}ain% 'hit (vt)'
    aba %aba% 'carry'
    aba-ang %aba-ang% 'carry for (vt)'

If aba, kena and other morphemes of shape <(...)a> are underlyingly %(...a)% their failure, without exception, to undergo the schwa-to-o alternation when suffixed cannot
be plausibly motivated. If, as assumed here, such forms underlyingly have final □a□, there is no difficulty.

As well if paa is □pae □, we might expect paa-n-ne 'thigh-gen-3' to be realised as the non-occurring □paa-ne □, cf (47); this would parallel lowering in forms such as kepil-n-ne /kapienne/, 'his bug' (7.4).

All lexical-morpheme-final [3] then reflect underlying □a□, since there are no lexical morphemes which both end in [3] and do not alternate when affixed, of the hypothetical type □mate □mate-ne □mate-ne □matene □. I conclude that this negative constraint applies to underlying representations:

(50)

□a□ | lexical morpheme

The domain is specified as 'lexical morpheme', since prefixes can apparently be of shape □...a□ underlyingly, cf 7.6.

There is a parallel lack of occurrence of □a□ P-word-finally, with two possible explanations:

Either (i) as well as (50), there is a rule of alternation, /a/ always going to /a/ (phonetic [3]) P-word finally:

(51)

□a□ > □a □Pword

Or (ii) There is no alternation of □a□ to /a/; rather, a simple allophony rule realises phoneme /a/ as [3], P-word finally.

A strong case can be made for both analyses, so that a choice between the two is difficult. The first analysis, to have surface [...3] derive from intermediate /a/, has these advantages: (i) it gives a straightforward relationship between the phonemic form and the unmarked surface realisation, [3] being closer phonetically to /a/ than to /a/; (ii) in non-standard spellings, final [3] is often written as < e >, indicating that it is perceived to be /a/.

Evidence in favour of the second solution comes from reduplication, when it is followed by [a]-overwrite, 8.7. Where the final syllable of the base is open, [a]-overwrite occurs with bases ending in all vowels except [3] - compare examples in (52a) and (52b). This is despite that fact that were overwrite to procede in (52b) a surface alternation paralleling that obtained with other vowels would be obtained:

(52) a. ngrépé
    grupa-grépé
    bulu
    bala-bulu
    ng-loso
    lasa-loso |
    □%r-grepe □ |
    [grapagrepe] |
    'gropes'
    'gripe about'
    'body hair, fur, feather'
    'animals of the gods, of many kinds'
    'slither'
    'slither about'
A simple explanation for the failure of overwrite to apply is that, phonemically, [3] is still /a/. This would account for the lack of occurrence of overwrite just in this context. This analysis has two other advantages: (i) it does away with the need for a separate rule of /a/-to-/a/ alternation, word-finally, and (ii) it allows the underlying phonological form to correspond to the simplest surface realisation, mata, /mata/, [mat3]

Disadvantages of the 'final-/a/'-analysis are that (i) it requires /a/ to have a surface allophone which more naturally realises /a/ (ii) in another, non-final, environment [3] does realise /a/: keet [k3t], 'cut off' (2.2.4), (iii) evidence from non-standard spellings appears to favour final /a/.

I therefore have adopted the 'final-/a/' solution, while noting that the 'final-/a/' solution is phonologically appealing (i) in retaining a single phonemic shape at all levels of derivation); (ii) in allowing a more parsimonious rule inventory, and (iii) in having greater explanatory power (viz, the non-occurrence of overwrite in one environment only, which therefore remains unexplained).

With the exception of the clitic né '3rd person' (7.4), [-ATR] realisations of vowels never occur finally. Either of the above solutions is therefore compatible with the view that the alternation (51) is a repair strategy, for a violation of the following surface constraint:

(53)  *V[I-ATR]

In the 'final-/a/' solution, there is simple phonetic adjustment:

(54)  V[I-ATR, +LOW] > [+ATR, -LOW]

In the 'final-/a/' solution, %a% is replaced /a/, which then automatically goes to the [+ATR] allophone in this environment.

### 6.3.3 A constraint on identical vowel sequences

In virtually all simple morphemes of more than two syllables, syllables containing identical vowels must occupy a continuous sequence from the right-most syllable:

(55)  sengguu 'kind of priest'  kepii  'k.o. seashell'
tengkeem 'clear one's throat'  jempiring  'gardenia'
jengkuung 'large concrete pipe'  temisi  'fresh-water snail'
Simple morphemes with structure \(^{*}V_{\alpha}-V_{\alpha}-V_{\beta}\) or \(^{*}V_{\alpha}-V_{\beta}-V_{\alpha}\), such as the following hypothetical forms, are not found in the native lexis:

\[(56) \quad *suunggeng \quad *kitipeh; \quad *suenggung \quad *kitepih\]

Exceptions are either loanwords, such as \textit{tumuli} 'then, next (Old Java.)', \textit{kalpataru} 'tree in the Ramayana epic', or are morphologically complex, or both.\(^{26}\)

A search of the data base found only one morpheme where the left-most vowel in structures like this was not /ə/, \textit{tembaang} 'forbid', (from [?tan] \textit{baang} '[NEG]-allow?'). This predominance perhaps simply reflects the strong tendency for the antepenultimate vowel to be schwa.

### 6.4 Complex morphemes

A simple morpheme contains no phonological elements more complex than the phoneme, while a \textsc{complex morpheme} contains two or more morphs/complex formants (6.0). The following structural condition holds:

\[(57) \quad \text{Prefer simple morphemes}\]

In Balinese, then, complex morphemes are expressives, with few exceptions. Evidence for this is given in 4.1 and 4.2. I take the expressive nature of complex morphemes to derive from their phonologically aberrant status, rather than from a meaningful derivational process (3.4.2).

\(^{26}\)These patterns are similar to those reported by McCarthy (1986, 1995) for consonants in Arabic, and can thus be explained in the same way: in an autosegmental phonological model, assuming exhaustive one-to-one, left-to-right association of melody elements and skeletal slots, the non-occurring structures set up OCP violations.
7.0 Morphophonological processes

This chapter treats phonological processes associated with morphology. As well as these, various automatic processes of deletion and epenthesis occur. The latter can be seen as repair processes, induced to satisfy well-formedness requirements - see for example sections 5.6, 7.2 and 8.2.2, on schwa epenthesis; 5.5 on deletion induced by syllabification processes; and 6.2.2 on deletions of consonants in the L-slot. Reduplication processes are described in a separate chapter (chapter 8), because of their complexity, and because of their inherent theoretical interest.

7.1 Concatenation

Concatenation is the phonological process whereby two phonological units link sequentially to create a more complex unit. I assume that this process is involved in affixing, as well as in cliticisation. The units linked in this way are:

a) complex formants (to form a complex morpheme) (4.0):

(1) Seririt
    clebingkah
    kedongdong
    'Place name'
    'pieces of broken pottery'
    'k.o fruit'

b) complex formants with morphemes, to form a lexeme (4.1):

(2) -ketél
    cegut
    % | kə | - tel |
    % | cə | - gut |

c) two or more morphemes (forming a derived stem or lexeme):

(3) pelut-ang
    me-daar
    nge-mati-ang
    'peel-APP'
    'VBL-eat'
    'AP-die-APP, kill'

d) two grammatical words, to form a single phonological word:

(4) pelut-ang = a kedongdong-é
    'peel-APP-3; peeled by him/her/them'
    'k.o fruit-DEF, the kedongdong'
7.2 Nasal assimilation, and other processes associated with \{N\}

The agentive verbal prefix \{N\} has different realisations, depending on the number of syllables in the verb root, and the initial segment of that root.\(^1\)

1. As /ŋ/ before roots of two syllables or more beginning with a continuant sonorant (vowels, liquids, and glides):

\[(5)\]

\(\%N\)-igal\% /ŋigal/ 'AP.dance'
\(\%N\)-empu\% /ŋempu/ 'AP.look after child'
\(\%N\)-amah\% /ŋamah/ 'AP.eat (of animals)'
\(\%N\)-rah-in\% /ŋorain/ 'AP.say-LOC, tell'
\(\%N\)-emes\-in\% /ŋemesin/ 'AP.flatter-LOC'
\(\%N\)-orges\% /ŋorges/ 'AP.scratch'
\(\%N\)-wayan\% /ŋwayan/ 'AP.perform wayang'
\(\%N\)-yasa-an\% /ŋyasaan/ 'AP.good.deed-APP'

2. As a nasal consonant homorganic to, and replacing, the initial obstruent of roots of two syllables or more:

\[(6)\]

\(\%N\)-pačin\% /ŋpačin/ 'AP.angle, fish (vi)\(^2\)
\(\%N\)-baañ\% /ŋbaañ/ 'AP.give'
\(\%N\)-tagak\% /ŋtagak/ 'AP.sit'
\(\%N\)-aar\% /ŋaar/ 'AP.eat'
\(\%N\)-ccak\% /ŋccak/ 'AP.chop'
\(\%N\)-jak\% /ŋjak/ 'AP.cook rice'
\(\%N\)-surud\% /ŋsurud/ 'AP.take used offering'
\(\%N\)-konkoñ\% /ŋkonkoñ/ 'AP.bark (of dog)'
\(\%N\)-gae\% /ŋgae/ 'AP.do/make'

3. /ŋ/ before all roots beginning with nasal consonants:

\[(7)\]

\(\%N\)-mati-an\% /ŋmati-an\% 'AP-dead.APP: kill'
\(\%N\)-nahnañ\% /ŋnahnañ\% 'AP-dry.fry'
\(\%N\)-naññæg\% /ŋnaññæg/ 'AP-stare.at'

and before all \textit{mono}-syllabic root morphemes:

\[(8)\]

\(\%N\)-bel\% /ŋbel/ 'AP.car.horn'
\(\%N\)-dot-in\% /ŋdotin/ 'AP.want-LOC'
\(\%N\)-ñiñ\% /ŋñiñ/ 'AP.cry'
\(\%N\)-ñæg\% /ŋñæg/ 'AP.lean (vi)'
\(\%N\)-ñæjin\% /ŋñæjin/ 'AP-summon(HI)-LOC'
\(\%N\)-yein\% /ŋyein/ 'AP.water-LOC'

As examples such as \(\%N\)-ñæg\% and \(\%N\)-ñæg-in\% in (8) show, it is the size of the root, not that of the derived stem, which determines the allomorphy of \{N\}. This requires that information about the internal structure of complex stems is retained in derived forms, at

---

\(^1\)In the following examples, the underlying form is written as \%N\% to simplify exposition, see discussion below

\(^2\)\{N\} has a free variant allophone, /mæŋ/, only occurring before certain roots with an initial labial, e.g. mæncing ~ memæncing, ñætæng 'night' > (me)mætæng 'be, leave in darkness', budæh 'crazy' > memudæh 'act in a crazy way'; see 12.2.
least at the stage when \{N\}-affixation applies. The alternative account, that \{N\} affixation occurs before stem formation appears highly implausible.

There are two plausible candidates for the underlying representation of \{N\}: /y/ or /C[+nasal]/ (a nasal consonant unspecified for place of articulation). Both occur independently in the phonology (2.1).

Assuming the underlying form is /C[+nasal]/, the assimilation of \{N\} to the place of articulation of the stem initial-obstruent in the examples in (6), as well as its failure to do so in (5) and (7) can be accounted for with the following rule:

\[
(9) \quad [\emptyset \text{place}] \quad [\text{place}] \\
\quad \bullet \quad C_1 \quad C_2 \\
\quad \bullet \quad [+\text{nasal}] \quad [-\text{son}]
\]

An assumed condition on assimilation (following Hankamer and Aissen (1974) as cited in Itô (1988:66)) is that it is determined partly by a sonority difference, with the nasal assimilating to an obstruent, but not to segments of equal or greater sonority: nasals, liquids, or glides (hence the realisations in (5).

Again assuming that the underlying phonological representation of \{N\} is /C[+nasal]/ (represented as /N/ in the following paragraphs), the following stages can be posited to take place in the phonological derivation of the surface form. I assume the unit composed of prefix plus root is the target for a morphologically-conditioned variant of Foot Well-formedness (5.6), applying to the input and output of all stages in the derivation:

a. Where the root is disyllabic or longer (and the output of concatenation of /N/ therefore automatically satisfies Foot Well-formedness) one of the following,

(i) \textit{where the root begins with an obstruent:}

- concatenation of /N/ to the root (eg */N-pańcin/), then
- spreading of the place of articulation of the root initial consonant by rule (9) (eg */Npańcin/ > */mpańcin/) then
- deletion of the root-initial consonant to meet syllable-onset requirements, 5.1: */mpańcin/ > /mańcin/

(ii) \textit{where the root begins with a non-nasal sonorant:}

- simple concatenation, then, since by rule (9) assimilation is blocked,
- resyllabification of root-initial liquids and glides to the L-slot of the new

---

3If there is a process of 'Bracket Erasure' or 'Tier-Conflation' (McCarthy 1986) which removes internal morphological information from the derived stem (root plus suffix), it must occur after (N)-prefixing.
complex syllable onset (5.5)
- /N/ surfaces as /ŋ/ by the default rule (10).

(iii) where the root begins with a nasal sonorant:
- concatenation (e.g. */Nmatiŋ/) then, since assimilation is again blocked,
- /ə/-epenthesis (/Nmatiŋ/) to enable resyllabification.4
- /N/ surfaces as /ŋ/ by the default rule (10).

b. Where the root is monosyllabic I assume:
- /ə/-epenthesis (e.g. /abəl/), since /bel/ does not satisfy Foot Well-formedness,
- concatenation of /N/ (/Nəbel/), then
- /N/ surfaces as /ŋ/ by the default rule (10).

In all cases, the surface form generated is the one which best meets prosodic well-formedness, at either syllable or Foot level.

A default rule is required, realising the underspecified C[+nasal] segment as /ŋ/ when it does not gain its place of articulation by spreading:

(10) \[ C[+nasal] > /ŋ/ \]

Such a rule is needed independently in any case, to account for /ŋs/ clusters morpheme-internally (see 2.1.4). If the underlying representation is /ŋ/, the derivational steps would be essentially as above, with minor modifications. The spreading process in rule (9) is motivated on purely phonological grounds if /C[+nasal]/ is the underlying phonological form of {N}.

The following filter was posited to account for the occurrence of medial /ŋs/, rather than the expected /ns/, in 'homorganic' NC clusters, dongsok 'elbow (vt)', blangsah 'nosebleed' (see 2.1.4):

(11) \*[laminal] /
\[ \begin{array}{c}
C_1 \\
C_2 \\
\end{array} \]
\[ [+cont] \]

I take it that where C₂ is /s/, this constraint rejects the output of (9) in both {N}-assimilation and in medial /Ns/ clusters. Assuming that different repair strategies apply in each context, these would produce /ŋs/ clusters morpheme-internally (via the default rule (10)), and /ŋ/ initially (via the processes independently needed to repair all unsyllabifiable /N/-obstruent sequences, cf a(i) above).

---

4 I take it that deletion of the root-initial nasal consonant is blocked by the principle blocking repair processes when the underlying morpheme(s) would become irrecoverable as a result, cf 5.5.
7.3 Alternation of /a/ and /o/

In Peguyangan, speakers below the age of around 50 have a rule producing morphophonemic alternations such as these (affected phonemes underlined, on right):

\[(12)\]
- kayeh 'bathe (vi)' kayeh/
- kayehang 'bathe-APP (vt)' kayohan/
- eficeh 'urine' eficeh/
- eficehin 'urine-LOC (v)' eficehin/
- keneh 'thought' keneh/
- keneh-é 'thought-DEF' keneh-é/
- tegeh 'high' tegeh/
- tegeh-an 'high-COMP; higher' tegeh-an/
- paling tegeh-a 'most high-IRR' paling tegeh-a/

That is, in the process of suffixing or cliticisation, after final /h/ is regularly deleted before a vowel (5.5), /a/ is then replaced by /o/, where it would otherwise occur immediately before a following vowel. Rule (13) is an automatic process, repairing the illformed /aV/ sequence (6.3.1):

\[(13)\]
\[ V[-low], [+back] > [+round] / _ V[=low][^back] \]

**Phonological-word-internally**

I use the notation 'V[=low][^back]' in the environment of the rule to indicate that this vowel is independently specified for values of those features. The rule does not then block the surfacing of /aa/ sequences (/səb/, */səb/ 'peer') since they share a single feature specification. Nor does it apply to /aV/ sequences across a prefix-root boundary: me-adan [meadan] 'VBL-name; named', ke-amah [kaamah] 'UP-eat.CRUDE; eaten'; prefixes are external to the P-word which contains roots and suffixes only (5.7).

Some Peguyangan speakers aged in their sixties or older do not have this rule, retaining /a/ always:

\[(14)\]
- kayeh kayeh/
- kayehang kayohan/
- eficeh eficeh/
- eficehin eficehin/
- keneh keneh/
- keneh-é keneh-é/

For these speakers, the restriction barring /aV/ sequences applies only morpheme-internally.

---

5 Inkelas and Cho (1993) argue against Hayes' (1986) linking constraint as an explanation of this and other examples of 'geminate inalterability'.

---

141
Chapter 7

7.4 Lowering of high vowels to mid, in the 'genitive' construction

With younger Peguyangan speakers, the high vowels /i/ and /u/ are lowered to /el/ and /lo/ respectively when they occur in the final syllable of a noun occurring before either the 'genitive' clitic -n, or the third person nominal clitic -né.6 This is always in complex noun phrases of structure NP(n)NP; where -né occurs, it fills the second NP slot.

Examples, with the affected vowel underlined are:

(15) %tilu-n-kupiq % /tilonkupiq/ [tilonkupiq]
    wax-n-ear    'ear wax'

    %raos bibi-n-ne% /raos bibgnne/ [ra òsbibgnne]
    speech aunt-n-3 'her aunt's words'

The alternation takes place even where the targetted vowel is in a closed syllable:

(16) %kuluk-ne i nanot % /kł offender i nanot/ [kł offender inanot]
    dog -3 DET N. 'Nanot's dog'

    %kc~it-ne% /kc~itne/ [kc~itne]
    piglet-3 'his piglet'

Since it is triggered by either -n or nominal -né, the rule requires a disjunction in the (morphologically conditioned) environment:

(17) \( V_{[+\text{HI}]} > [-\text{HI}] \)

\( \bigwedge \{n\} \bigwedge (C)\{{}\}

I assume that this is a morphological process, rather than an automatic, purely phonological one, since it is associated with this genitive construction only. Moreover, the morphemes -n 'GEN' and -né '3rd person' both only occur in the above syntactic structure. This suggests that it is only one of the two which is causing lowering. -né is ruled out, since lowering also occurs with 1st- and 2nd-person possessors: eg beli-n cai /belencai/ 'your brother'. Two alternative hypotheses assuming that -n alone causes lowering - both unsatisfactory - are:

(i) -n is always inserted underlingly in NP-n-NP sequences, inducing lowering (before it is then deleted in the environment C_{p-word due to failure to syllabify). Hence /kłok-ne/ 'his dog', from %kłok% and %ne% would derive from an intermediary */kłok-n-ne/. This hypothesis is unsatisfactory, as it predicts non-occurring lowering, such as */kłok-tia\_e/ 'dog-lexc-DEF (from *%kłok-n-tia\_e%) or */buboh-baas 'porridge-rice: rice porridge'. In fact only the non-lowered forms occur: /kłok-tia\_e/, /bubuh baas/, as in all complex NP's where the two component NP's are conjoined without {n}.

---

6 né 'third person', of invariant shape, should not be confused with -n6, the definiteness marker, realised as -né after a vowel-final head, -e after a consonant-final one. See also 5.7 and Appendix L.
(ii) -n is always inserted, but induces lowering only where it is not deleted. The form /kulokne/ 'his dog' would then be accounted for if the third person pronoun had an allomorph /e/, only occurring after -n thus /kulokne/, */kulok-n-e/. This analysis is ruled out by the fact that, in unambiguous cases, -n always forms a part of the P-word to the left (5.7); it cannot syllabify with a following vowel, as it would have to, to get forms like /kulokne/. This analysis too would predict non-occurring forms such as */buboh-n-injin/ 'black rice porridge' from %bubuh-n-injin%, */buboh-n-tian-e/ 'my porridge', and so on.

It appears then that the disjunction in (17) is required

There remains a related analytical problem, the realisation of the morpheme, né, as [né]. This is the only case where a [-ATR] vowel occurs word-finally in Balinese: other word-final enclitics of shape CV occur with [+ATR] allophones (ké [ke] question particle; ko [kol] 'emphatic' particle, -ti [ti] 'very'). I assume that, rather than representing a distinct phoneme of extremely limited distribution, final [e] here is the product of the above morphological process. 7

Where the left-hand nominal ends in a sequence of two identical high vowels, only the rightmost is lowered:

(18) %luu-n-ne% rubbish-n-3 /lug-n-ne/ [luonne] 'his rubbish'
%kopi-n-ne% k.o.bug-n-3 /kopiagne/ [kopiennie] 'his bug'

and similarly where the nominal contains two identical high vowels separated by a consonant:

(19) %kutu-n-ne% louse-n-3 /kutonne/ [kutonne] 'his lice'
%bibi-n-ne% aunt-n-3 /bibenne/ [bibenne] 'his aunt'

Elsewhere sequences of identical vowels behave as though they are linked at the phoneme/melody tier (8.7). The apparent non-operation of the Obligatory Contour Principle in both (18) and (19) can be accounted for if there is tier-collapsing (McCarthy 1986) and subsequent delinking prior to the formation of the complex NP.

This lowering was frowned upon by KMP (aged in his sixties) as 'not Peguyangan' speech, though it occurs freely in the speech of younger speakers. I heard it often used

7I propose that one of the exponents realising the genitive construction as a whole is the floating element [-HI], which attaches to the final vowel in the P-word complex (including the final clitic). This, plus a morphologically conditioned rule spreading [-HI] just to the syllable to the left (Odden 1994), would account for high vowel lowering.
Chapter 7

by speakers aged in their forties: a well educated woman in her twenties described the avoidance of lowering as 'putting on fancy manners' (bergaya BL).

Although this rule has not to my knowledge been previously described, it (or versions of it) is widespread. I have heard it used by speakers from Banjar and Kubutambahan villages (20km to the west and east of Singaraja respectively), as well as others from Bangli and Klungkung districts in south Bali. A speaker from Tangkup in the Klungkung district had a more restricted variant of the rule, lowering in final open, but not closed, syllables.

7.5 Alternation of /a/ and /ø/ in lexical morphemes

In lexical morphemes, morpheme-final %a% is always realised as [3] phonological-word-finally, and as [ø] elsewhere.

(20) mata %mata% 'eye' [matʃ]
mata-né %mata-ne% 'eye-DEF' [matane]
teka %teka% 'come' [tekʃ]
teka-ang 'come-APP; make come' [tekaan]

I take this to be due to the following alternation:

(21) a > ø /_[word]

Other possibilities, (i) that the underlying element is /ø/, going to /a/ when non-final, and (ii) that [3] realises %a% directly, are considered and rejected in 6.3.2.

7.6 Alternations in the shape of prefixes and prepositions: Ca ~ Cø ~ C

Prefixes are generally of shape /(...)CV/. Most such prefixes are realised as /(...)Cø/ before a consonant-initial stem.8

(22) me-daar 'VBL-eat; eat (vi)'
pe-daar-in 'CAUS-eat-LOC; feed'
ke-daar 'UP-eat; be eaten'
seke-dasa-ang 'seke-ten-APP; do in tens'

However, the actual realisation of the prefix vowel is influenced by both stylistic and phonological factors.

8The only exception to this is the nominal prefix pi-, which can be considered a formal/high stylistic alternant of nominal pe- (~ pe- ~ p-), viz. pi-anak 'child.HI' p-anak 'child'; pinath (the most numerous clan in Peguyangan), everyday form penath.
In very formal or literary style, it may optionally be realised as /(...)/Ca/, as in these examples, the first taken from a shadow puppet performance, the second two from an oral folk tale narration:

(23) a. pacang k[a]-séda-yang
    will.HI UP-die.HON-APP 'He(HON) will be killed.'

b. k[a]-crita, di tengah alas-é, ada lutung
    UP-story LOC middle forest-DEF, exist monkey
    'It is told, in the middle of the forest was a monkey.'

c. i kekua nawang unduk p[a]-teka-n i lutung-é
    DET turtle know fact NML-come-GEN DET monkey-DEF
    'The turtle knew about the coming of the monkey'

In all three cases in (23), the alternative /a/ realisation would be perfectly acceptable.

Before vowel-initial roots the same prefixes may have up to three different realisations:
1. Before /i/ /u/ /ø/ and /e/ (the 'non-central vowels'), underlyingly monosyllabic prefixes are realised as any one of /Ca/ ~ /Ca/ ~ /C/. 9

(24) me-intuk [maintu?] ~ [meintu?] ~ [mintu?]
    'me-pound, pounded'
ke-icen [kaicen] ~ [kaicen] ~ [kicen]
    'UP-grant'
me-orah-an [maoraan] ~ [meoraan] ~ [moraan]
    'tell s.o. sthng'

If the prefix is disyllabic, the final vowel is not lost:

(25) seke-ulung-dasa-ang ~ seka-ulung-dasa-ang
    'seke-eight-ten-APP,
*sek-ulung-dasa-ang
    (do) in lots of eighty'
meke-ukud ~ maka-ukud
    me-ukud '*me-body; (do) to the whole body'
*mek-ukud

2. Before /a/- and /ø/-initial roots, only two forms, /Ca/ and /C/ occur in my recorded texts, or are accepted in elicitation:

(26) me-empug ~ m-empug,
    *ma-empug
    'me-open, opened'
me-adan ~ m-adan,
    *ma-adan
    'me-name, named'

And while informants only use and accept [a] realisations when 'a-final' lexical morphemes are suffixed (mata [matæ] 'eye', mata-né [matane] *matane] 'the eye'), in elicitation they hesitate between [a] and [æ] when prefixes and proclitics carry the same suffixes:

(27) da-ang [daæn ~ deæn] ~ 'DON'T-APP; 'say 'don't' (to U)'
    me-né [mane ~ mene] ~ 'me-DEF; 'the (prefix) me-'~

9Phonologically, there may be no distinction between prepositions and prefixes. The only preposition containing final /a/ ~ /ø/ is ke 'to'. Occurring before uma 'ricefields' it has the same range of realisations as the prefixes in (24): ka uma [kaumæ] ~ ke uma [kaum3] ~ kuma [kum3], 'to (the) ricefields'. In elicitation however only the /ka/ form was accepted before other /i/- and /u/-initial nouns: * ka Ubud, *ka Indra, etc. See also 5. 7 on the problematic status of prefixes.
On this data alone, there seem two plausible candidates for the underlying representation of such forms: either /Ca/ or /Ca/.

Evidence that /a/ is synchronically perceived as the underlying vowel here comes from spelling reform proposals. In the standard orthography, vowels in prefixes are always written as <a>: ma-daar, pa-daar-in and so on. Recently, however, there have been proposals that <e> should replace this <a>, to reflect the commonest pronunciation (Udara Narayana et al 1991). Popular orthography (in letters, messages on T-shirts and so on) similarly prefers <e> in prefixes.

Morpheme-internally, /a/ never occurs immediately before another vowel. If the underlying vowel in the prefix is /a/, the alternate realisations before a vowel-initial stem (/Ca/ and /C/) could be seen as being phonologically motivated by, in this case not an outright ban on, but a related avoidance of, /aV/ sequences across morpheme boundaries.

The non-occurrence of the allomorph /ma/ before /a/- and /a/-initial stems (as in (26)) could perhaps be seen as further evidence in favour of /ma/ and against /ma/ as the underlying form. However, the reason for the unacceptability of /ma/ here is unclear: in this variety morpheme-internally /aa/ sequences do occur, though /aa/ sequences do not.

On the other hand, evidence in favour of the /Ca/ form as underlying is its occurrence even before C-initial stems in formal speech, as in (23). A rule of vowel raising could then account for the /a/ realisation.

Historically at least, the prefix vowel may have been /Ca/ invariably. Evidence for this comes from both of the standard orthographies (aksara and latin). They each use a single grapheme to represent both /a/ and the morpheme-final segment which varies as in (24). In the aksara, the unmarked interpretation of the same grapheme (ie in all other contexts) is /a/.

Whichever form is underlying, the differing behaviour of the vowel when it occurs prefix-finally and when it occurs lexical-morpheme-finally must be accounted for. If the underlying form is /Ca/, this /a/ fails to undergo the otherwise common alternation with /o/ when it becomes root-final before a vowel due to deletion of /h/ (see 7.3):

(28) enceh 'urine' enco-in 'urine-LOC, urinate on'

And if on the other hand the underlying form is /Ca/, this /a/ behaves differently from root-final /a/ in optionally raising to /a/ before a vowel (examples (24), (26)); root-finally /a/ goes to /a/ only when it is also word-final (7.5):

---

10Note that I have adopted this proposal in the orthography used here, though I use Ca-, and C- where these are the surface forms.
The simplest analysis can be obtained if the underlying vowel in prefixes is /a/. This requires that the constraint against /a/ occurring morpheme-finally underlyingly (6.3.2) be limited to lexical morphemes only. This is not problematic however, since cross-linguistically grammatical morphemes commonly have distinct phonological shapes and constraints from lexical ones (cf 3.2.2).

The differences in the behaviour of prefix-final and root-final vowels can then be seen to be due to the presence of a different underlying vowel in each case. Moreover, if prefixes do not form part of the P-word formed by the stem and suffixes (5.7) the prosodic environments of these vowels are also quite distinct: root-final vowels belong to the same P-word as any following suffix-initial vowel, while prefix-final vowels never belong to the same P-word as that of a following root-initial vowel.

The rule raising underlying /a/ to /al/ exemplified in (29) then does not apply to the prefix vowel because it is not /a/ underlyingly. And the rule backing underlying /a/ to /o/ exemplified in (28) does not apply to the prefix /a/ because it only applies where /a/ occurs before another vowel within the same P-word domain, not one across a P-word boundary.

I have adopted this solution in this thesis, while noting that the evidence is ambiguous, and that it rests upon several other assumptions. Some further requirements which either analysis would entail are now discussed.

A rule optionally deleting the prefix vowel is needed in either case, given forms such as m-intuk, k-icén, m-adan. This can be seen as producing a prosodically preferred sequence: either a disyllabic Optimal Foot,\(^{11}\) or else, to account also for forms like m-ubet-an 'VBL-close-an, closed' (~ meubetan, maubetan), perhaps a structure in which all syllable onsets are filled, cf 5.2.1.

Underlying /a/ would require a rule lowering it to /a/, conditioned partly phonologically (cf (24)):\(^{12}\)

(30) \(a > a / V[ + ATR]\)

Optional rule

---

\(^{11}\) Though this would require the prefix forming part of the same P-word as the base, something I have just argued against.

\(^{12}\) Assuming this rule applies only to underlying /a/, during the lexical phonology, and cuts out before surface /a/ is produced when underlying /a/ occurs P-word-finally (7.5), no further environment specifications are needed. Alternatively, the domain of the rule could be a higher prosodic domain such as 'clitic-group'-internal (cf Nespor & Vogel 1986, Cohn 1989). This would block the application of the rule to P-word-final underlying /a/.
and partly stylistically (in formal speech). Underlying /a/ on the other hand would require a rule raising it to /æ/ in most contexts, say:

(31)  a  >  æ  / clitic[C[P-word]V[+ATR]]

Optional rule

The environment for the latter rule is more complex, and so perhaps less plausible. I do not argue specifically for the prosodic unit 'clitic-group' in this thesis, as there is no unambiguous evidence for it.
8

Reduplication

8.0 Reduplication, introductory overview

Evidence from reduplication processes has played an important part in the development of phonological and morphological theory in the eighties and nineties (for example, in the work of Marantz (1982), Steriade (1988), McCarthy & Prince (1990)). I have tried therefore to describe reduplication processes in Balinese in some detail, drawing on the general approach of McCarthy & Prince (1990).

This chapter aims firstly to give a clear description of the various formal reduplication types. Amongst other things 1) it discusses two previously undescribed reduplication types in Balinese: Foot-reduplication and 'internal Reduplication' (a variant of CV-RED); 2) it argues that one previously identified type, 'final syllable' reduplication, is a surface variant of Full Base reduplication; 3) it demonstrates that morphemes like gigi 'tooth' are produced by synchronic reduplication processes (despite previous claims that there is no evidence for this); 4) it describes theoretically problematical reduplication processes (cf Steriade 1988) where non-distinctive material is transferred, and offers an explanation for them in terms of the Prosodic Morphology theory of McCarthy and Prince; 5) in doing so, it offers further evidence (in addition to that presented in 4.2) that one can distinguish in a principled way between a purely phonological function of reduplication, in pre-lexical morpheme-formation processes, and a morphological one, in derivation processes; this in turn is evidence for the existence of complex phonological formants, and concatenation processes, below as well as above the level of morpheme (4.0).

There are five surface types of reduplication, and four underlyingly, with two of the latter analysable as variants of a single process. They are:

(i) and (ii) Two types of partial reduplication, 8.2, the two surface realisations depending on whether the base is mono- (1a) or di-syllabic (1b):

(1) a. duduk 'pick up'
    pipis 'money'
    b. kekupu 'butterfly'
    je-jaian 'RED-sew-NOM; sewing (n)'

(iii) Full-base reduplication, where the entire phonological material of the base is copied, 8.3\(^1\). The base may be either an independently occurring morpheme:

---

\(^1\)I use the term base to refer to the element which is the target of reduplication, and copy to the element created by this process.
Chapter 8

(2) gedé 'big' gedé-gedé 'id. PL'

luh 'female' luh-luh 'id. PL'

or a complex formant, which often only occurs in reduplicated form (inherent reduplication):

(3) ongol-ongol 'k.o. cake' cakcak 'chop up'
udud 'clear with hoe' cucu 'grandchild'

(iv) 'Final syllable' reduplication, 8.4. Here all or part of the final syllable of a disyllabic base is copied:

(4) -cebwah 'flash (v)' pecebwahbwan 'id., PL'
-kétél 'drip' peketétél 'id., PL'
-kebeng 'open (of lips)' ngebébéng 'grimace with lips open'
kedongdong 'k.o. fruit' Buléléng '(place)'

Although final syllable reduplication appears to differ radically from other reduplication types, it is not in fact a separate process: I argue that all instances are better treated as either full-base- or partial (CV-RED) reduplication, applying to a meaningless complex formant.

(v) Foot reduplication, 8.5. Here the base must contain a morpheme or morpheme-sequence, never just a complex formant; it also necessarily ends in a vowel. The copied sequence consists of the morpheme(s), plus a 'stray' following consonant, which serves to close the final syllable of the copy (underlined in these examples):

(5) bé-n-né 'fish-GEN-3; his fish' bén-bé-n-né 'id. PL'
ng-ajin 'AP-cost-LOC; charge (vt)' ng-aiin-ajin 'id. PL'
pe-gedé-nin 'CAUS-big-LOC; make bigger' pegedén-pegedé-nin 'id. PL'

I assume that reduplication involves these stages:

1) A parsing process, which determines the contents of the base formant: its prosodic structure, its right and left edges and so on; parsing criteria may be morphological (full-base reduplication) or phonological (partial reduplication), or a combination of these (foot reduplication).

2) A process which delimits the element to be copied from the base parsed in 1) (McCarthy & Prince 1990). That process may involve phonological criteria (partial reduplication, foot reduplication) or a combination of morphological and phonological criteria (full-base reduplication).

3) A separate operation COPY (McCarthy & Prince 1990, Archangeli 1991), where the delimited material is copied.

4) The concatenation of the copied material like a prefix to the base.

The delimitation and prefixing of the Copy can occur very late in the phonological derivation of a word, after many postlexical processes, 8.9. Because it is so late, non-distinctive phonological material is typically transferred in reduplication (see
Steriade 1988). This is true even in inherent reduplications, where reduplication is part of the underlying phonological representation of a morpheme. I propose an explanation for this in 8.9.

Reduplication is assumed to be always a prefixing, rather than a suffixing, process (though see 8.1). I assume also that the parsing stage of reduplication, like syllabification (5.5) and association of melody elements (6.3.3), is directional, parsing only material to the right (Itô 1988). This removes the need for some ordering specifications. Square-bracket notation, as in the following examples, marks this directionality of application; it does not indicate an extrinsic ordering of reduplication before concatenation with other formants:

(6)  
\begin{align*}
\underline{kati} & \underline{-l} | \underline{RED} | \underline{-bun} | \\
\underline{ka} & \underline{-l} | \underline{RED} | \underline{-dong} |
\end{align*}

Of the four logically possible combinations of copy with empty prosodic affixes (Archangeli, 1991:261), I analyse Balinese to have: (i) copy without prosodic affixes: consonant reduplication C-RED; (ii) copy with prosodic affixes: foot reduplication, and perhaps CV-RED; (iii) no copy, no prosodic affixes: segmental affixes such as me- 'VBL', -ang 'APP'. The fourth type, prosodic affixes without copy, is possibly exemplified by CV-RED, which could be analysed as a case of the imposition of a Foot template, 5.6, on an underlying CVC sequence, and involving spreading of features rather than reduplicative copying. I will use the term 'reduplication' to include this possible analysis of CV-RED.

Reduplication serves two distinct functions in Balinese: purely phonological (in inherent reduplications) and derivational. Of the four underlying types above, CV-RED partial reduplication occurs only in inherent reduplications, while foot reduplication has a derivational function only. The other types, C-RED and full-base, have both functions.

Although cross-linguistically inherent reduplications are sometimes assumed to be fossilised forms, there is good evidence that they derive from synchronic processes in Balinese (8.8). Previous descriptions (Jendra 1975, Bawa et al 1984) mention only inherent reduplication of disyllabic bases:

(7)  
\begin{align*}
lumbalumba & \text{‘dolphin’} \\
dangapdangap & \text{‘k.o. lizard’}
\end{align*}

Bawa et al (1984) moreover suggest that there is no evidence for these as synchronic reduplications. I give a variety evidence in 8.4 and 8.8 that not only those, but also morphemes of shapes such as the following, contain synchronic reduplications:

(8)  
\begin{align*}
gigi & \text{‘tooth’} \\
agag & \text{‘gaping, wide open’} \\
cakcak & \text{‘chop up’} \\

\text{bubuh} & \text{‘rice porridge’} \\
\text{kedongdong} & \text{‘k.o. fruit’} \\
\text{Buléléng} & \text{‘place name’}
\end{align*}
Chapter 8

8.1 The direction of reduplication

The direction of reduplication in all cases is assumed to be a prefixing of the copy to the left of the base, thus: COPY-BASE (‘right-to-left reduplication’). The evidence for this is mixed, with some evidence for a BASE-COPY analysis. Evidence in favour of prefixing reduplication is:

1. In reduplication plus vowel overwrite, 8.7, it is the left-most formant which is overwritten, the right hand formant retains the shape of the base morpheme:

   (9)  gelur-gelur  ‘yell PL’  bading-bading  ‘turn over PL’
        gelar-gelur  ‘id., in a varied way’  budang-bading  ‘id. variously’

   grépé-grépé  ‘grope about’
   grapa-grépé  ‘grope about, back and forth’

2. In C-RED partial reduplication, the copy is to the left of the base:

   (10)  -sangi  ‘oath’  -daar  ‘eat’
        me-se-sangi  ‘make an oath’  de-daar-an  ‘food’

3. Inherently reduplicated morphemes like kupukupu ‘butterfly’, crućuk ‘k.o. bird’, can be further reduplicated: kupukupu-kupukupu ‘butterfly, PL’, crućuk-crućuk ‘k.o. bird, PL’. Suffixes are otherwise never reduplicated in Balinese:

   (11)  minum-an  ‘drink (n)’  takon-ang  ‘ask about U’
        minum-minum-an  ‘drink (PL)’  takon-takon-ang  ‘id. PL’
        minuman-minuman  ‘drink (PL)’  takon-ang-takonang

If the order of formants in the non-reduplicated morpheme is BASE-COPY, this would be the only case where reduplication of a suffixed element occurs.

Evidence in favour of the opposite, left-to-right, analysis is:

1. The shape of inherently reduplicated monosyllables with a complex syllable onset:

   (12)  jriji  ‘finger, toe’  plispis  ‘fall bit by bit (rice)’
        mrésmés  ‘messy (when eating rice)’

Here it is the right-most syllable which has apparently undergone simplification, suggesting it is a copy of the base, simplified in the copying process (Steriade 1988).

2. In final syllable reduplication, only the right-most syllable of the base is copied, again suggesting that it is suffixed to the base:

   (13)  -ketél  ‘drip’  -kemél  ‘move (of lips)’
        pe-ketél-tél  ‘drip, PL’  ngemél-mél  ‘mutter to oneself’

152
Rather than posit two distinct reduplication processes, one prefixing, the other suffixing, I assume that in all cases reduplication is prefixing. I analyse data such as that in (12) as being due to the independent phonotactic constraint favouring occurrence of the L segment in the 'left-most' position of a morph, 6.2.2. And I take it that in forms like (13) the copy is inserted on the left: pe-ketetel, ngemelmel, 8.4.

8.2 Partial reduplication

Partial reduplication targets bases consisting maximally of a disyllable, and minimally of a monosyllable with a filled coda. It therefore applies after a parsing process has established the size of the base (McCarthy & Prince 1990).

It has two main surface realisations, depending on the length of that base. If the base is a mono-syllable (of shape %C(L)VC%), it undergoes CV-reduplication, with both the initial consonant and the vowel copied: pipis 'money', gruguh 'hand (crude)'. If the base is disyllabic, the prefixed element is of shape /C₁/, where C₁ copies the initial consonant of the base. I refer to this variant as C-RED: kekupu 'butterfly', se-sangi 'vow', 8.2.2. Bases longer than two syllables, as well as any base beginning with a vowel, do not undergo partial reduplication.

The following sub-sections treat CV-RED and C-RED in turn. I attempt to present the descriptive facts in a reasonably theory-neutral way; the two reduplication types can nonetheless be well described in terms of the prosodic circumscription theory of McCarthy & Prince (1990).

8.2.1 CV-reduplication

In CV-reduplication, a copy of a light syllable appears to be affixed to a monosyllabic base. The base must end in a consonant, and perhaps have a filled onset (see below):

```
(14)     base       reduplication
    duk  duduk  'pick up'
    buh  bubuh  'porridge'
    keb  kekeb  'lid'
```

Where the base is of shape CLVC, I assume that the L is copied across in reduplication; its surface absence in the right-most syllable is due to independent principles (see 5.2.2).

```
(15)     base       reduplication
    Blair  blabar  'flood'
    gruh   gruguh  'hand (CRUDE)'
    klah   klakah  'bamboo shingle'
```

2Alternatively, CV-RED could be seen to involve the imposition of a disyllabic template on an underlying CVC segment sequence, rather than reduplicative copying (8.0).
Chapter 8

To undergo CV-RED reduplication the base must have a filled syllable coda. There is good evidence for this: morphemes of shape C₁V₂C₁V₂ always carry equal stress on each syllable, whereas those like dudukor blabarbear a single, final, stress. Bases of shape CV therefore always undergo full-base reduplication, 8.3, rather than CV-RED. It is unclear whether a filled onset is also required for CV-RED. Warna et al (1990) lists 46 morphemes of shape /V₁V₁C/, such as iis 'slice (v)', oon 'weary', ééd 'example, model'. It may be that these morphemes are also produced from underlying %VC% sequences by partial reduplication. The other reduplication type where the copied element is a light syllable is C-reduplication. C-reduplication, described in the next sub-section, applies to disyllabic bases only - it is therefore in complementary distribution with CV-reduplication; C-reduplication clearly never applies to vowel-initial bases. This is the only evidence I have found to indicate that partial reduplication is not found in morphemes of shape /VVC/.³ In this description I assume that forms like iis are simple morphemes which just happen to lack an onset consonant.

CV-RED is also the only reduplication type functioning just to derive morphemes from complex formants, never to derive lexemes from pre-existing morphemes. All other types have both functions. For these reasons, it could be wondered if CV-reduplications are in fact fossilised forms, rather than products of a synchronic process. Evidence in favour of CV-RED as synchronic is given in 8.8, as well as in 8.4. I take it that CV-RED, functions to derive morphemes of Optimal Foot shape (5.6) from monosyllabic bases.

8.2.2 C-reduplication

The second type of partial reduplication is 'C-reduplication', found only with di-syllabic bases which begin with a consonant. These bases may be meaningless formants:

(16) base
   baï    bébéï    'k.o. malevolent spirit'
   kupu   kekupu   'butterfly'
   tani   tetani   'termite'
   géndong  gegéndong   'tramp (n)'

or independent morphemes:

(17) jagur  'punch'  me-je-jagur-an  'punch (pl.)'
    kedas  'clean'  me-ke-kedas   'tidy up'

Where the base begins with a CLV sequence, only the initial consonant, and not the L-slot segment, surfaces in the copy:

(18) blakas  'chopper'  be-blakas-an  '(do) in an open, frank way'
    slampar  'throw'  me-se-slampar-an  'throw (pl.)'

³Diachronically, some of these morphemes derive from non-reduplicated morphemes: compare eg Bali. iis, Mal. iris 'slice'; Bali. aap, Jav. arep 'front part.'
Reduplication

As the name suggests, I analyse C-reduplication as reduplication of the initial consonant of the base (cf Archangeli (1991)), followed by the insertion of an epenthetic /a/ to satisfy syllabification requirements. The alternative would be to assume a single CV-copying process (CV-RED, see previous section), plus, where the base is disyllabic, either a rule of /a/ overwrite (8.7.1), or else one laxing the copied vowel to /a/, after reduplication. I assume a distinct C-reduplication, since it gives a simple account of the surface patterning of L segments, 5.2.2.5

Where a base is disyllabic but begins with a vowel, only full base reduplication is used. Partial reduplication, of the initial V only, is never found. Compare these deverbal nominalisation doublets:

(19) a. aba 'carry.LOW'          aba-aba-an 'thing carried (LOW)'
     bakta 'carry.HI'          *a-ba-an     be-bakta-an 'thing carried (HI)'

b. ambil 'take.HI'          ambil-ambil-an 'way of taking (HI)
     jemak 'take.LOW'          *a-ambil-an je-jemak-an 'way of taking (LOW)'

The non-occurrence of forms like *aabaan and *aambilan is compatible with both the C-RED analysis, and the alternative where the initial syllable is copied. If the process did involve copy of the first syllable, plus vowel overwrite/laxing, the forms expected after this process would be *aabaan and *aambilan, both with initial /ea/, an 'illegal' vowel sequence in Balinese, 6.3.1.

8.3 Full-base reduplication

Full-base reduplication copies the entire phonemic structure of the base to which it applies. Bases of any length can undergo this type of reduplication. They can be either simple morphemes:

---

4 The strongest independent evidence a rule of pre-penultimate laxing comes from doublets such as the following, all heard in Peguyangan:

Buléláng ~ B(elléláng) '(place name)'  Pinatih ~ Penatih 'personal (clan) name'
pi-takén ~ pe-takén 'NML-ask.HI; question(HI)'  pe-tulung ~ pi-tulung 'NML-help'
bængin ~ bænging 'banyan tree'  kucalcil ~ kecalcil 'sugar palm fruit'

However, if these forms are produced by a synchronic process, it is sporadic and lexically limited, since other tri-syllabics occur in only one form: sumangah 'k.o. ant', cialongan 'k.o. bird', bukakak 'roast chicken', and so on. (Cross-dialectally, though, the 'missing' alternants of all these forms occur, see Warna et al (1990). The sporadic nature of this laxing makes its existence as a synchronic rule unclear.

5 If forms like be-blañja-an and me-se-slampar-an involve intermediate forms like *biV-blañja-an and *me-sIV-slampar-an a more complex, and less plausible account of L-deletion seems required, with a rule such as: 'Where reduplication results in two adjacent syllables containing filled L-slots, prefer the representation where the L segment (/l/ or /n/) occurs in the penultimate syllable of the derived sequence (copy plus base).'

---

155
A morpheme formed by phonological full-base-reduplication may undergo further
derivational full-base-reduplication.

In 8.8 evidence is given that full-base reduplication of complex formants is a synchronic
process. Forms like gigi appear to be equally describable in terms of
light-syllable-reduplication (cf gigis 'soft, low (sound)', from \%RED\CV|-|GIS \%).
However, they differ from partial reduplications in that they have two stresses, not one,
thus patterning rather with forms like cakcak

8.4 'Final syllable reduplication'

With 'final-syllable reduplication' all or part of the final syllable of a disyllabic base is
copied. This formal type is found with both derivational (23) and inherent (24)
reduplication functions:

(23)  -cepol  'collapse'  pecepolpol  'id., PL'
      -kebyar  'flash (vi)'  pekebyarbyar  'id., PL'
      -kebet  'throb (eg of sore)'  ngebetbet  'id., PL'

(24)  sedédég  'lean against stg (vi)'
      giuluk  'roll about'
      ñirekongkong  'sit, thighs visible'
      buléléng  'place name'
      sengangal  'gape open'
      kedongdong  'k.o. fruit'
Final syllable reduplication is not analysed here as a distinct formal type of reduplication, but as a realisation of either full-base- or partial reduplication, applying to a meaningless complex formant. A good deal of evidence is given in 4.1 that morphemes like ce-pol, kebyar and kedongdong while monomorphemic, nevertheless contain concatenations or 'prefix'-like and 'root'-like complex formants:

(25)  
-cepol 'collapse'  | ce | | pol |
-kebyar 'flash (v)'  | ke | | byar |
kebilbil 'timid'  | ke | | RED | | bil |

I conclude therefore that in forms like pecepolpol and ñrekeñongkong derivational reduplication targets the right-most (root-like) complex formant. This parallels regular derivation of lexemes like me-laib-laib and me-se-slampar-an where the root morph(eme) is reduplicated, but not the prefix material to its left:

(26)  
ñrekeñongkong %ñre{-|RED|-|køŋ|}%
pecepolpol %ñø{-|ce|-|RED|-|pol|}%
me-slampar-an 'VBL-throw-an; throw (vi)'
me-se-slampar-an %me{-[RED_v-slampar]-an}% 'throw.PL'
me-laib 'VBL-run; run (vi)'
me-laib-laib %me{-[RED-laib]}% 'run.PL'

Where the reduplicated final formant is of shape CVC (full-base reduplication), the initial syllable is from a very restricted set, either ke-, ce-, kati/kali/katu, ñire-/ñile-and no others - see 4.1. This, with the other evidence presented in 4.1 and 6.1.1 for the independent status of the formants ke-, ce-, kati, etc, I take to be strong evidence for the full-base reduplication model proposed here.

An alternative analysis, after McCarthy & Prince 1990, would be to posit a simple morpheme base, say *%købil%, in the case of kebilbil 'timid', the final syllable of which is first prosodically circumscribed, and then subject to a copying operation, mapping it to a monosyllabic prefix template. However, if that were the process used, we would expect a far greater range of formants to undergo it than is the case: the limitation to just those morphemes with initial ke-, ce-, kati/kali/katu, ñire-/ñile-would be a mystery. It would also require the addition of a further reduplicational type to those posited. I have opted for the more economical, and more explanatory, analysis.

On the other hand, where 'final syllable partial reduplication' applies to a meaningless complex formant, the shape of the initial formant is much more variable, as in these examples:

(27)  
sréré 'glance sideways'
buléléng 'k.o. millet/place name'
ngèbèbèng 'pull faces'
tamblélé 'trunk (elephant)'
tambullilingan 'bumblebee'
-brarak 'in a mess'
bluluk 'palm seeds'
plilit 'entwine, plait'
tembréré-an 'cicada'
wlélé 'infertile hen'
Given the apparently much greater variability in the forms of the base, the 'prosodic circumscription' analysis appears appropriate in the case of partial reduplications, and is compatible with either of the analyses of CV-RED discussed in 8.2.

In the case of full-base reduplication, there is a variety of evidence indicating that reduplication is synchronic, including both the evidence of stress patterns, and the limited set of prefix-like complex formants occurring with it.

On the other hand partial reduplications like those in (27) have not, to my knowledge, been identified as such in previous descriptions of Balinese. The evidence is therefore presented in the rest of this section. I exemplify using data involving /...rV1rV2.../ and /...IV1IV2.../ sequences, though any consonant type would do as well. There is good evidence that the constraint dispreferring more than one consonant of the same place of articulation morph-internally, 6.2, holds particularly strongly for the liquids. With the exception of the sequence /...HV1IV2.../rV1rV2..., occurrence of two liquids in a simple morpheme is strongly dispreferred. For example, no morphemes at all of the shapes /...l/ and /r... ...r/ are found. Only four morphemes of shape /...IV1IV2/ occur, and only one of shape /...rV1rV2/. In all five cases, the vowels are identical, indicating reduplication of an underlying monosyllable:

(28) alal 'long (time)' elel 'sell badly'
edel 'gaping, wide open' olol 'swallow.CRUDE'
mirir 'gentle (of breeze)'

For morphemes with non-initial /...(...)IV1IV2.../ and /...(...)rV1rV2.../ sequences, such as those in (27), the figures are as follows:

<table>
<thead>
<tr>
<th>Shape</th>
<th>Example</th>
<th>Total Tokens</th>
<th>$V_1 = V_2$</th>
<th>%'age $V_1 = V_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>...IV1IV2...</td>
<td>plilit</td>
<td>27</td>
<td>19</td>
<td>73%</td>
</tr>
<tr>
<td>...rV1rV2...</td>
<td>sréré</td>
<td>32</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td>59</td>
<td>43</td>
<td>73%</td>
</tr>
</tbody>
</table>

Table 8.1: Numbers of morphemes containing medial /(...IV1IV2...)/ and /...(...)rV1rV2.../ sequences (data from Warna et al (1990))

As with other cases, see table 8.3, section 8.8 and discussion there, there is a large, highly statistically significant discrepancy ($p < .0000$), between these figures, and the 'background' level in the lexicon of less than 25% of morphemes where $V_1 = V_2$ (6.3.1). I conclude that this indicates a synchronic reduplication process. The full data supporting this analysis is given in Appendix J. Exceptions (where $V_1 \neq V_2$) can be accounted for by assuming lexically specified overwrite after reduplication, which is needed independently (8.7.1).

---

$^6$Excluding loanwords and morphologically complex items, see appendix J.
To sum up, rather than assume a distinct reduplication type, cases of 'final syllable
reduplication' can be analysed as either full-base reduplication, applying to a complex
formant, or CV-RED 8.2 applying to a prosodically circumscribed light syllable. With
inherently reduplicated morphemes, this implies that purely phonological reduplication
and 'affixing' occur below the level of the morpheme; other evidence for this is
presented in 4.1 and 4.2.

8.5 Foot reduplication

Foot reduplication applies only to morphemes (never to complex formants) or
morpheme-sequences ending in a vowel, and followed by a consonant-initial suffix - it
therefore requires both grammatical and phonological information about the stem to
proceed. The copied sequence consists of the stem morpheme(s), plus the initial
consonant from the following suffix. The effect of Foot reduplication is to close the
final syllable of the copy, (underlined in these examples):

(29)  
\[
\begin{align*}
\text{bé-n-né} & \quad \text{'fish-GEN-3; his/her/their fish'} & \text{bé-n-né} & \quad \text{'id. PL'} \\
\text{oka-n-né} & \quad \text{'child.HON-GEN-3; his (etc) child.HON'} & \text{oka-n-né} & \quad \text{'id. PL'} \\
\text{ng-liu-nang} & \quad \text{'AP-many-APP; multiply (vi)'} & \text{ng-liu-nang} & \quad \text{'id. DUR'}
\end{align*}
\]

Certain prefixes, such as all those of shape pe- (causative verbal prefix, nominal prefix),
me- 'verbal prefix', and the nasal segment realising {N}, the 'Agentive prefix', can be
optionally included in the copied sequence:

(30)  
\[
\begin{align*}
\text{ng-aji-nin} & \quad \text{'AP.price.LOC; charge (vt)'} & \text{ng-aji-nin} & \quad \text{'id. PL'} \\
\text{pe-gedé-nin} & \quad \text{'CAUS.big.LOC; make bigger'} & \text{pe-gedé-nin} & \quad \text{'id. PL'}
\end{align*}
\]

The final consonant of the copy is always /n/, since all consonant-initial suffixes begin
with that phoneme. Foot reduplication is optional. The same base may undergo either
this or full-base reduplication:

(31)  
\[
\begin{align*}
\text{ngedénang} & \quad \%N\text{-gade-naŋ} \% \quad \text{'AP.big-APP; enlarge (vt); get bigger (vi)'} & \text{ngedénang} & \quad \text{'id. PL'} \\
\text{beli-n-né} & \quad \text{'elder brother-GEN-3; 'his/her/their elder brother'} & \text{beli-n-né} & \quad \text{'id. PL'}
\end{align*}
\]

I assume that Foot reduplication involves the following stages:  
1. The stem is identified - where the 'stem' is the root morpheme, plus optional prefixes,
   eg \text{p-liu-nang} 'AP-many-APP; 'become many';
2. A sequence is parsed, from the left edge of the stem, to the right; that sequence is the
smallest possible Foot Template which will include all the material in the stem (cf (42), 5.6). So monosyllabic stems are parsed to a monosyllabic Foot, and so on.\(^7\) The parsed template satisfies Preferred Foot structure (repeated from 5.6):

(32) Prefer a Foot template which includes an extrametrical consonant at the right edge

Thus, in the example: \([\text{ñojú̃}]_{\text{Foot}}\).  

4. Finally, the parsed 'Preferred Foot' is copied and prefixed to the base: \(\text{ñojú̃-n-ú̃-ñ}\).

Since filled onsets are also the unmarked preference (5.2.1), it might be thought that the optional reduplication of the \(\{N\}\) prefix before sonorant-initial roots is part of the same process:

(33) a. \(\text{ng-ali-h} \quad \{\text{N}-\text{alih-in}\} \quad \text{AP-look.for-LOC: search about for}^\prime\)

\(\text{full-base red: ngalih-alihín}\)

\(\text{Reduplication of \{N\}: ngalih-ngalihín}\)

b. \(\text{ng-omong} \quad \{\text{N}-\text{omong}\} \quad \text{AP-speech}^\prime\)

\(\text{full-base red: ngomong-omong}\)

\(\text{Reduplication of \{N\}: ngomong-ngomong}\)

However, it appears that \(\{N\}\)-reduplication applies independently of foot reduplication. In elicitation, all of the following reduplications of \(\text{ng-aji-nin} \quad \{\text{N}-\text{price-LOC, 'charge'}\) were accepted:

(34) \(\begin{array}{l}
\text{ng-aji-aji-nin} \quad \{\text{N} \text{-root only reduplicated}\} \\
\text{ng-aji-in} \quad \{\text{N} + \text{root + initial C of suffix}\} \\
\text{ng-aji-ngaji-nin} \quad \{\text{N} + \text{root reduplicated, minus initial C of suffix}\} \\
\text{ng-aji-in} \quad \{\text{N} + \text{root + initial C of suffix}\}
\end{array}\)

The reduplication in \(\text{ng-aji-ngaji-nin}\) of the filled onset, minus the initial C of suffix, indicates a process distinct from Foot reduplication: see 8.6. Equally, forms like \(\text{ng-aji-aji-nin}\) suggests that foot reduplication does not require concurrent reduplication of a filled onset.

It might also be hypothesised that full-base reduplication (8.3) is not a distinct reduplicational type, but is rather a simple variant of Foot reduplication - one where the Foot may optionally circumscribe a base exactly coextensive with the stem, without the (now optional) extra consonant. That this is \textit{not} the case is indicated by the fact that inherent reduplications surface only as full-base reduplications, never as Foot reduplications: \(\text{kupo-kupo-n-né} \quad \text{butterfly-GEN-3; his butterfly}, \quad *\text{kupon-kupo-n-né}\) and so on. If the two were variants of the same process, we would expect both to occur in inherent reduplications.

\(^7\)The maximal Foot is three syllables. I have no example of Foot-reduplication applying to a base longer than this.
8.6 Reduplication across morpheme-boundaries

Where a lexeme consists of more than one morpheme (root plus affixes), typically only the root morpheme reduplicates, or the root plus prefix(es). Suffixes do not reduplicate along with the root:

\[(35)\] gedé-nang \(\rightarrow\) 'big-APP; make bigger'
gedé-gedé-nang \(\rightarrow\) 'make/grow bigger and bigger'
*gedé-nang-gedé-nang

beli-nin \(\rightarrow\) 'buy-LOC, buy from'
beli-beli-nin \(\rightarrow\) 'id. (repeatedly)'
*beli-nin-beli-nin

parek-an \(\rightarrow\) 'close-NML; servant'
parek-parek-an \(\rightarrow\) 'id. PL'
*parekan-parekan

The only apparent exception to this is the seeming optional reduplication of the 'genitive' suffix \(\{n\}\). This is in fact due to a distinct process, foot reduplication, 8.5:

\[(36)\] oka-n-né \(\rightarrow\) 'child-GEN-3' oka(n)-oka-n-né \(\rightarrow\) 'id. (pl.)'

Prefixes which optionally reduplicate include the verbal and nominal prefixes of shape pe-, the 'actor prefix' \(\{N-\}\), only before sonorant-initial roots; and, occasionally, me-:

\[(37)\] pe-dingo-ang \(\rightarrow\) 'CAUS-hear-APP; listen to'
penyero-an \(\rightarrow\) 'NML-inside-\(N\)-lan; servant'
%peN-jaro-an%

ng-waas-ang \(\rightarrow\) 'AP-healed-APP; heal (vi)'
ng-lais-ang \(\rightarrow\) 'AP-sell-APP; sell well (inch.)'
me-daar \(\rightarrow\) 'VBL-eat'
m-ampak-ang \(\rightarrow\) 'VBL-open-AN'

The occurrence of forms like pe-dingo-pe-dingo-ang, where the copy is a tri-syllable, or ng-waas-ng-waas-ang, where it contains a complex onset, suggests that this optional inclusion of prefixes is not phonologically motivated: the reduplicated copies are non-preferred structures. I conclude then that the prefix may optionally be included in the morphological material constituting the stem. This is indicated also in that foot reduplication, which always targets bases defined (at least in part) in morphological terms, allows optional reduplication of prefixes:

\[(38)\] pe-gedé-nang \(\rightarrow\) CAUS-big-APP; 'enlarge'
pe-gedén-pe-gedén-ang

It might be conjectured that doublets like ng-waas-waas-ang \(\rightarrow\) ng-waas-ng-waas-ang (37), result from reduplication occurring (i) before and (ii) after prefixing. This can't be the case: non-occurring forms like *malihballihin, and the others in (39) would be found:
(39) malih-in %N-balih-in% 'AP-look.at-LOC; look at, check'
malihmalihin *malihbalihin

ngedé-nang %N-gedé-nang% 'AP-make bigger; become bigger'
ngedén-ngedénang *ngedén-gedénang (foot red)
ngedé-ngedénang *ngedé-gedénang (full-base red)

Their non-occurrence indicates that %N% affixation must always occur before either full-base- or foot reduplication.

While Full-base reduplication (like Foot reduplication) is sensitive to morphological boundaries, it is, by contrast, oblivious of complex formant boundaries. Where a base contains a single morpheme, even though that morpheme contains concatenated complex formants, the entire morpheme is copied: kupukupu-kupukupu 'butterfly PL', katibangbung-katibangbung 'beetle sp. PL'.

8.7 Reduplication plus /a/ 'overwrite'

A process occurring only in association with reduplication is vowel alternation or 'overwrite'. This involves replacement of the final vowel of the copy by another vowel, normally /a/:

(40) gelur-gelur
    gelar-gelur
    kiud-kiud
    kiad-kiud
    -kejengit
    kejengat-kejengit
    -klidkid
    kladkad-klidkid
    kipu-kipu
    kipa-kipu

'yell repeatedly'
'id., in a varied way'
'strech, bending one's back (vi)'
'id., in a varied way'
'bare one's teeth'
'repeatedly bare one's teeth'
'look for a hiding place'
'id., here and there'
'wallow (pl.)'
'wallow about'

This can be expressed informally by the following rule:

(41) **Overwrite:**
    In the vocalic melodeme of the final syllable of the copy, change the feature specification to [±low].

In inherent reduplication, vowels other than /a/ may be lexically specified in the copy (8.7.1).

---

8I follow McCarthy & Prince (1990) in assuming 'overwrite', as a distinct process, rather than a range of reduplicative templates pre-specified for the final vowel.
Reduplication

As with the examples in (40), overwrite (REDv) may have a derivational function (adding meanings such as 'do in a varied or haphazard way; do back and forth, etc'). It also occurs in inherently reduplicated (hence expressive) morphemes, often with similar meaning components:

(42) kesyarkesyur  'anxious (feelings) because of fear'
    ngahngahngéngéh  'whimper pleadingly'
    ngiahngiuh  'sigh anxiously'
    ketekahketekéh  'curse; grumble'
    lendalendi  'k.o. evil spirit'
    kucalcil  'k.o. small fruit'

Where the final syllable of the base is open, /a/-overwrite proceeds. The /a/ segment in the copy is not subject to the usual rule raising final /a/ to /a/:

(43) grapa-grépé  [grapagrepe]  'grop about'
    lendalendi  [lendalendi]  'k.o. evil spirit'
    srara-sréré  [srarasrere]  'glance quickly from side to side'
    bala-bulu  [balabulu]  'animals of the gods, of many kinds'
    lasa-loso  [lasaloso]  'wriggle like an eel'

This suggests that a new P-word-boundary is not inserted between base and copy when reduplication occurs; alternatively, it could indicate simply that overwrite is a later process than /a/-to-/a/ alternation, and subsequent allophony. The account of 'non-distinctive transfer' (8.9) assumes the correctness of both hypotheses. An explanation in terms of 'geminate inalterability' is inadequate, see discussion of examples (18), 7.4.

Overwrite does not occur when the final open syllable contains surface /a/, suggesting that the latter is 'still' underlyingly %a% (6.3.2):

(44) teka  /taka/  'arrive'
    teka-teka  /takateka/  'id., (durative)'

Where the final and penultimate vowels of the base are identical, both are replaced with /a/:

(45) saab-seeb  'peer about'
    daas-diis  'sigh repeatedly'
    gradag-grudug  'go around and about in a group, noisily'
    jangkak-jongkok  'squat about'
    crakcak-crékcék  'drip water all over the place'
    srandang-sréndéng  'sway back and forth'
    ketapas-ketépéss  'make a repeated slapping motion'
    kecagcág-kecigcig  'run about busily, to and fro'
    gumatatgumitit  'creepy-crawlies'

Forms where only the final vowel is replaced are ungrammatical:
Both of these observations can be accounted for in the formalism of autosegmental phonology by assuming that the identical vowels of the base are linked at the melody level. \(^9\) In the process of overwriting, a single set of feature specifications is changed. Since the melody element is linked to two skeletal slots, both vowels are replaced in the surface form. The only exceptions to this pattern are certain morphemes where the base contains two /a/ vowels, at least on the surface:

(47)
-**dlekep** 'press one's face against a surface'  
-**celap** 'go inside'  
-**jeleg** 'stand up tall'  
-**gremek** 'mumble'  
**keblat-keblet** 'flicker on and off'

I suggest in 2.2.5 that this exceptionality indicates that /a/ and an underspecified vowel contrast in underlying representations. Forms patterning like gramang-gremeng (-gremeng 'obscured by darkness') are specified for two /a/'s underlyingly, while those in (47) contain an underspecified vowel in the penultimate syllable.

The overwriting of the final vowel may set up a sequence of two /a/'s in the copy, even where the vowels in the base are not identical: \(^10\)

(48)
-**jagur** 'punch'  
-**sangkol** 'carry'  
-**balik** 'return'

This last type of overwrite has not been reported in previous studies of Balinese, hence is possibly not found in other varieties. It seems to be accepted as a minor secondary form, speakers sometimes disagreeing about the acceptability of its application to particular bases. For example, KPM, while accepting some such forms, rejected badang-bading, from -bading 'turn over', while others accepted it. Another indication that it is a minor form is that where the vocalism of the base is a - i, a second, preferred, surface realisation is found. This, the most common and 'usual' form, my teachers emphasized, is where both vowels are replaced, the first by /u/:

(49)
-**balik** 'repeat, return'  
-**badang** 'turn over'  
-**pantig** 'throw to ground'  
-**mait**

These facts suggest that forms such as badang-bading set up an OCP violation (McCarthy 1986) of two identical adjacent melody elements: \(^11\)

---

9. This assumes a model where vowels and consonants occupy separate tiers in Balinese.
10. Contrast consonant behaviour: sequences of (strictly adjacent) identical morpheme-internal consonants cannot be created in inherent reduplications of monosyllabics 6.2.5.
11. The C & V symbols are for expository convenience only.
This is resolved normally by what can be seen as a subsequent dissimilation rule raising the first /a/ segment to /u/ (hence the forms in (49)). If the leftmost /a/ of the copy is specified as [+low, +back] at the time of overwrite, the simplest dissimilation process would be one which changes a single feature specification:

\[
[+\text{low}] > [-\text{low}] /\_ [+]\text{low}
\]

This produces either the observed /u/ vowel, or else /o/, the replacive vowel found in some varieties (cf entries in Warna et al (1990)).

This rule only applies where the vowel overwritten in the copy is /i/ (and not where it is u, ê, o or ø), suggesting a more complex model of dissimilation in which the feature [+high] of the /i/ vowel is somehow involved.\(^{12}\)

Forms such as jagar-jagur (48) suggest a second, less preferred way of resolving the OCP clash - either by fusing the two adjacent autosegments to produce a linked segment, or (if the OCP is not a universal requirement of underlying representations, Odden (1988)) by simply tolerating the violation.

### 8.7.1 Lexically specified vowel overwrite

In many morphemes formed by inherent reduplication, the vowels in the copy and the base differ. Figures are given in Table 8.2:

<table>
<thead>
<tr>
<th>shape</th>
<th>example</th>
<th>total tokens</th>
<th>total ( V_1 \neq V_2 )</th>
<th>%'age ( V_1 \neq V_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_a V_1 C_b C_a V_2 C_b )</td>
<td>dugdug</td>
<td>481</td>
<td>40</td>
<td>8.4%</td>
</tr>
<tr>
<td>( C_a L V_1 C_a V_2 )</td>
<td>bilbis</td>
<td>138</td>
<td>13</td>
<td>9.4%</td>
</tr>
<tr>
<td>( \sigma\sigma )</td>
<td>kupukupu</td>
<td>115</td>
<td>13</td>
<td>11.3%</td>
</tr>
<tr>
<td>( C_a V_1 C_a V_2 )</td>
<td>lalah</td>
<td>413</td>
<td>60</td>
<td>14.5%</td>
</tr>
<tr>
<td>( V_1 C_a V_2 C_a )</td>
<td>udud</td>
<td>41</td>
<td>13</td>
<td>31.7%</td>
</tr>
<tr>
<td>( C_a V_1 C_a V_2 )</td>
<td>gigi</td>
<td>55</td>
<td>20</td>
<td>36.4%</td>
</tr>
</tbody>
</table>

Table 8.2: Inherently reduplicated morphemes where vowels in copy and base differ, by shape (data from Warna et al (1990))\(^{13}\)

---

\(^{12}\)The fact that /i/ and /a/ have 'mirror-image' feature specifications (2.21 has perhaps something to do with it. The resulting sequence of u - a - a - i involves just those vowels at the farthest corners of the vowel space.

\(^{13}\)For each shape, the sample is all lexical entries of that structure listed in Warna et al (1990), with the exception of reduplicated disyllabic bases, for which the sample is all entries listed in sections L-V of Warna et al (1990), or about half the book.
In many cases, the vowel patterning differs from that found in derivational overwrite in that the overwrite vowel is not /a/:

(52) sungsang 'upside down (eg in a breech birth)'
lélor 'weak from emotion'

No clear patterns of vocalism seem to occur. For example, of the 40 morphemes listed in Warna et al (1990) of shape RED-C₁VC₂ and with non-identical vowels (sogsag, léglog and so on), only 11 have /a/ in the copy syllable, while 20 have it in the base syllable. The full figures for the different V₁ - V₂ combinations are:

(53) o-a 8 o-e 4 a-i 2
    u-a 6 e-o 4 e-a 2
    a-u 5 a-e 2 i-u 1
    i-a 4 a-e 2

Similar figures are found for the other inherent reduplication types. Since there is good evidence that such forms, like other inherent reduplications, are synchronically reduplicated (Table 8.2 above, and section 8.8), I conclude that they should be analysed as undergoing lexically specified vowel overwrite, with the appropriate variants of the rule proposed at (41).

8.8 Inherently reduplicated morphemes as synchronic reduplications

Balinese has many forms which look as though they contain reduplicated sequences. The sequence (either monosyllabic, or longer) may appear to be wholly reduplicated:

(54) gigi 'tooth'
    bubu 'fish trap'
    paskak 'chew'
    lumbalumba 'porpoise'
    nyonyo 'breast'
    pipi 'cheek'
    laklak 'k. o. cake'
    kupukupu 'butterfly'

or else seem partly reduplicated in different ways:

(55) pipis 'money'
    kekeh 'stiff'
    sasak 'tease (hair)'
    kekupu 'butterfly'
    sesapi 'swallow (bird)'
    dedalu 'flying termite larvae'

These forms, which I refer to as being 'inherently reduplicated', may show apparent interaction with other processes, such as affixation (4.0):
Reduplication

(57) Buléléng 'place name'
kedongdong 'k.o. fruit'
férekongkong 'sit with legs apart'
kepupungan 'wake up disoriented'
katiwawalan 'young jackfruit bud'

deletion of L in a complex onset (6.2.2):

(58) cruclcuk 'k.o. bird'
Klungkung 'place name'
blabar 'flood'
riji 'finger'

and vowel alternation (8.7.1):

(59) sosgag 'fight (v)'
mingmang 'figure-of-8'
lélor 'weak, from desire'
linglung 'confused, from love'

or more than one of these:

(60) katibangbung 'k.o. insect'
kucalcil 'k.o. small fruit'

All these forms are analysed here as complex morphemes, undergoing synchronic reduplication as part of 'pre-lexical' word formation. So for example the morpheme gigi underlyingly consists of two complex formants in association, \(|\text{g}i|\) and \(|\text{RED}|\), \(|\text{RED}||\text{g}i|\), rather than a simple \(*%\text{g}i\text{g}i%\) sequence. Evidence for synchronic reduplication is of two kinds: 1) phonological and 2) morphological.

Phonological evidence shows that both inherent and derivational reduplication occur after both stress assignment and certain phonological rules have applied.

1) Primary stress normally falls on the last syllable of a word, other syllables being equally unstressed (5.8):

(61) gedé 'big' [ge'de]
gédé-nang 'enlarge' [gade 'nal] kurenan 'spouse' [kure'nan] katiwawalan 'jackfruit bud' [katiwawa'lan]

Where a formant is wholly reduplicated, it receives double stress, indicating reduplication occurs after stress assignment:

(62) gigi 'tooth' ['gi'gi] bubu 'fish trap' ['bu'bu] cakcak 'chop up' ['cak'cak] Gitgit '(place name)' ['git'gtt] lumbalumba 'porpoise' [lum'b]lum'b] kisikisi 'whisper' [ki'siki'si]

14Except in the case of CV-RED, 8.2.1.
Such forms are the only lexical items to receive double primary stress.

2) Final orthographic <a> is realised as [a] by a rule raising /a/ phonological word finally (7.5). With inherently reduplicated disyllabics, alternations like the following occur: 15

(63) lumbalumba [lumbəlumbə] 'porpoise'
lumbalumba-ne [lumbəlumbəne] 'the porpoise'
kawakawa [kawa kawa] 'spider'
kawakawa-ne [kawakawane] 'the spider'

A form like lumbalumba [lumbəlumbə] (never *[lumbəlumbə]) requires that reduplication occur after the raising rule has applied. This parallels the stress behaviour of these forms.

3) The rule lowering high vowels to mid (cf 7.4) is also ordered before inherent reduplication of disyllabic bases: 16

(64) kupukupu [kupəkupənnə] 'butterfly'
kupukupo-ne [kupəkupənnə] 'his butterfly'
picipici [picəpicənnə] 'k.o. seashell; group of kids'
picepicenne [picəpicənnə] 'his seashells/kids'

4) Inherently reduplicated peripheral morphemes of form \( C_1 VhC_1 Vh \), such as cakcak 'chop up', have two possible surface realisations, \( C_1 V[k]C_1 V[k] \) (the usual realisation) or less-commonly, \( C_1 V[k]C_1 V[k] \) (2.1):

(65) cakcak [cakcak]~[caʔcaʔ] 'chop up'
    *[caʔcak], *[caʔcak]
cekcek [cakcek]~[caʔcaʔ] 'house lizard'
    *[caʔcak], *[caʔcak]

With vowel-initial enclitics, only the [k] realisation is allowed:

(66) cakcak =a [cakcakə] 'chop up-3; chopped up by him'
    *[caʔcakə]
cekcek-é [cekceke] 'house.lizard-DEF'
    *[ceʔcake]

Similarly, morphemes of shape \( C_1 VhC_1 Vh \) are realised as [CVCV] before a following vowel-initial suffix or clitic:

---

15 When the reduplicated element is a monosyllable, raising of %a% does not apply before reduplication: eg dada 'chest (Hi)' %RED-da%, [dadə], *[dadə]; see 8.9 for an explanation.
16 Again, when the reduplicated element is a monosyllable, reduplication occurs first: gigi-n-ne [gigenne], *[gegennə].
Reduplication

Examples (66) and (67) show that reduplication occurs after suffixing/cliticisation and syllabification, which then determines both /k/ allophony, and deletion of /h/ (see also 5.5).

The second kind of evidence for synchronic reduplication comes from the following descriptive fact about morpheme structure, which derives from the 'dislike' of morphs which contain more than one consonant of a given place of articulation, underlingly (6.2).

(68) In a morpheme of shape $C_1 (L)V_1 (C)C_2 V_2 (...)$, where $C_1$ is identical to $C_2$, then $V_1$ and $V_2$ will also be identical. 17

Figures showing the percentages of morphemes obeying the patterns in (68), as well as percentages for similar patterns, are given in the following table:

<table>
<thead>
<tr>
<th>shape</th>
<th>example</th>
<th>total tokens</th>
<th>$V_1 = V_2$</th>
<th>$V_1 = V_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_v V_1 C_v V_1 V_2 C_v$</td>
<td>dugdug</td>
<td>481</td>
<td>441</td>
<td>91.6%</td>
</tr>
<tr>
<td>$C_v LV_1 V_2 V_1 C_v$</td>
<td>blibis</td>
<td>138</td>
<td>125</td>
<td>90.6%</td>
</tr>
<tr>
<td>$C_v V_1 C_v V_2 C_v$</td>
<td>lalah</td>
<td>413</td>
<td>353</td>
<td>85.5%</td>
</tr>
<tr>
<td>$C_v C_v V_1 V_2$</td>
<td>kupukupu</td>
<td>115</td>
<td>102</td>
<td>88.7%</td>
</tr>
<tr>
<td>$V_1 C_v V_2 C_v$</td>
<td>ubud</td>
<td>41</td>
<td>28</td>
<td>68.3%</td>
</tr>
<tr>
<td>$C_v V_1 C_v V_2$</td>
<td>gigi</td>
<td>55</td>
<td>35</td>
<td>63.6%</td>
</tr>
<tr>
<td>Overall total</td>
<td></td>
<td>1243</td>
<td>1084</td>
<td>87.2%</td>
</tr>
</tbody>
</table>

Table 8.3: Numbers of morphemes containing $C_v - C_v$ sequences, by shape (data from Warna et al (1990))

If vowel patterns in disyllabic morphemes are purely random, given six vowel phonemes, the probability that a morpheme will contain two identical vowels is one in six. That is, without reduplication, we would expect to find around 17% of morphemes where $V_1 = V_2$. The actual figures in table 8.3 are much higher, and are highly significant ($Z = 5.72, P < .0000$) using the Median Test), even for the 'lowest' percentage of 63.6%.

17 The only exception being trisyllabic morphemes of shape $C_v e C_v (ee)$ such as gegitik 'ticklish spot', bebiki 'mischievous', kekupu 'butterfly', which I analyse as undergoing inherent C-RED (8.2.2).

18 For each shape, the sample is all lexical entries of that structure listed in Warna et al (1990), with the exception of reduplicated disyllabic bases, for which the sample is all entries listed in sections L-Y of Warna et al (1990), or about half the book.

169
The logical explanation for the these figures, together with the evidence from phonological processes such as stress assignment and h-deletion, is that reduplication is synchronic in all cases.

It might be argued that the figures in table 8.3 show that such forms were formed by reduplication in the past, but are now merely fossilised forms, not undergoing synchronic reduplication. This would account for the fact that for each structural type in the table there are exceptions to the descriptive generalisation (68). **Diachronic** evidence comes from the 130 morphemes of structure CᵥVCᵥV(C) reconstructed for either Proto Austronesian, Proto-Malayo-Polynesian or Proto-Western-Malayo-Polynesian by Dempwolff (1938) and Blust (1980, 1983, 1986) (see Appendix K). Of these, 110 or 85% have identical vowels in both positions: a figure very close to that found in the synchronic Balinese data. This suggests that the Balinese data reflects a *stable* situation over time, and so supports the conclusion that morphemes of the shapes like those in table 8.3 are synchronically reduplicated.

The small residue of exceptions to the requirement that vowels be identical in reduplicated forms can be accounted for by either of the ways suggested in 8.7.1.

A theory-dependent argument in favour of inherent synchronic reduplication is that it accounts for the constraint that C₁ must be identical to C₂ if they are of the same place of articulation in terms of currently accepted models of feature geometry (eg Sagey 1986, 1990), rather than requiring the radically different one proposed by Mester (1988) to account for parallel facts in Javanese. Mester's feature geometry requires manner features to be dependent on place features, which is not the case in models such as Sagey (1986, 1990), Clements (1985). Mester's account also requires right-to-left spreading of association lines for consonants, undesirable in Balinese (and probably Javanese) since vowel association clearly needs the reverse direction of association, see 6.3.3.

### 8.9 Ordering: Lexical versus Postlexical Reduplication

Just when reduplication occurs in the phonology depends on 1) the size of the base being copied 2) whether reduplication is purely phonological ("inherent"), or morphological in function. If the base is monosyllabic, inherent reduplication occurs before a variety of postlexical processes have applied.¹⁹

However, if the base is disyllabic (or longer) reduplication - at the least the copy and prefixing sub-stages (8.0) - is delayed until the postlexical phonology: paradoxically, then, word-formation is 'postlexical' for this group. This is the case both for inherent reduplications - even though such morphemes are specified underlyingly as undergoing reduplication - and optionally (see 8.9.1) for derivational reduplication.

---

¹⁹On the distinction between lexical and postlexical phonology, see 2.0.
Reduplication

Not only is most reduplication carried out postlexically - more than 90% of bases are disyllabic or longer - it occurs after most other processes, including allophony rules, have ceased to apply. In the process, non-distinctive phonological material created by postlexical allophony rules is copied from the base to the copy (cf Steriade 1988, on a parallel situation in the neighbouring languages, Javanese and Madurese, and the difficulty this poses to theories of reduplication).

The hiatus between 'prelexical' marking of (inherent) reduplication, and its postlexical realisation represents a challenge for an organising principle of the Lexical Phonology framework, the Strong Domain Hypothesis (Kiparsky 1984, cited in Rice 1990). In this section I will present the evidence for these differing patterns (lexical reduplication of monosyllables, versus postlexical elsewhere), and propose an explanation compatible with the Strong Domain Hypothesis.

I focus the discussion on inherently reduplicated morphemes, since for these the difference in behaviour is particularly striking, given that they are all equally underlyingly specified as undergoing reduplication, whether they are formed from monosyllabic bases, like kuku, 'fingernail', gudgud, 'extremely (old)', or from longer sequences, like dangapdangap, 'lizard. sp.' and kebaratkebrit, 'run about wildly'. Examples involving derivational reduplication are also provided (see also 8.9.1).

I present the evidence in two parts: first I show that inherently reduplicated monosyllabic bases pattern in a variety of ways like simple morphs, and very differently in these respects from disyllabics. I then go on to describe the contrasting ways in which reduplication of disyllables and monosyllables, respectively, is ordered with respect to various phonological processes. The explanation then proposed accounts for all the differences between the two types.

1. Monosyllabic bases. Almost the only difference between inherently reduplicated monosyllables and simple morphemes is that, like all reduplications except CV-RED, they carry two stresses:

(69)    
<table>
<thead>
<tr>
<th>Base</th>
<th>Copy</th>
<th>Phonemic form</th>
</tr>
</thead>
<tbody>
<tr>
<td>cucu</td>
<td>'grandchild'</td>
<td>['cu' 'cu']</td>
</tr>
<tr>
<td>gésgés</td>
<td>'scratch (vt)'</td>
<td>[ges'ges]</td>
</tr>
<tr>
<td>crukcuk</td>
<td>'k.o. bird'</td>
<td>['cruk' 'cuk']</td>
</tr>
</tbody>
</table>

Otherwise, reduplicated monosyllables behave largely like simple morphs.

(i) There is no evidence of any boundary between base and copy which might influence allophony:20

(70)    
<table>
<thead>
<tr>
<th>Base</th>
<th>Phonemic form</th>
<th>Morphological form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gitgit</td>
<td>[gitg1t]</td>
<td>*[gitg1t]</td>
<td>'place name'</td>
</tr>
<tr>
<td>gudgud</td>
<td>[gudgud]</td>
<td>*[gudgud]</td>
<td>'extremely (old)'</td>
</tr>
<tr>
<td>prunpung</td>
<td>[prunpun]</td>
<td>*[prunpun]</td>
<td>'chipped'</td>
</tr>
</tbody>
</table>

---

20The asterisked (non-occurring) forms would occur if there was a morpheme-boundary between base and copy.
(ii) like simple morphemes, they only allow one filled L-slot in the surface morph (base plus copy), (71a)), whereas disyllabics allow two (71b), 6.2.2:

\[(71)\]
\[
\begin{array}{ll}
\text{a. } & \text{blabar} \quad [\text{RED}_{CV} \text{b}| | \text{lar} ] \quad \text{‘flood’} \\
& \text{bréngbéng} \quad [\text{RED}\text{b}| | \text{bren} ] \quad \text{‘dilapidated’}
\end{array}
\]

\[
\begin{array}{ll}
\text{b. } & \text{silaksiuik} \quad [\text{RED}_{a}\text{a}| | \text{liuik} ] \quad \text{‘bob up and down’} \\
& \text{sumbrangsambring} \quad [\text{RED}_{a}\text{a}| | \text{sambriq} ] \quad \text{‘dishevelled (clothing)’}
\end{array}
\]

(iii) Reduplicated monosyllables cannot contain adjacent identical consonants at the copy-base boundary, \(*C_{1}VC_{1}-C_{1}VC_{1}\) (72a), even though the same sequence can occur in inherently reduplicated disyllabics, (72b), (6.2.5):

\[(72)\]
\[
\begin{array}{ll}
\text{a. non-occurring hypothetical forms:} \\
& *\text{kakkak} \\
& *\text{ngungung} \\
& *\text{kikkik}
\end{array}
\]

\[
\begin{array}{ll}
\text{b. kayakkayak} & \text{‘lie stretched out, helpless’} \\
& \text{nguengngueng} \quad \text{‘sound of a bumblebee’} \\
& \text{kikkikik} \quad \text{‘sound of chicks, ducklings’} \\
& \text{kisikkisik} \quad \text{‘move slightly’}
\end{array}
\]

To sum up, in a variety of ways, reduplicated monosyllables behave like simple morphs, while reduplicated disyllabics (and longer) appear to behave like two distinct morphs. I now go on to contrast ordering behaviours with respect to other phonological processes.

2. Where the reduplicated base is a disyllable, or longer, the copy shows the results of many processes, which necessarily have applied to the base before reduplication. These include (i) realisation of final \(%a%\) as \([3] (7.5):

\[(73)\]
\[
\text{lumbalumba} \quad %|\text{RED}|-|\text{lumba} | % \quad [lumb3lumb3]
\]

(With reduplicated monosyllables, \(%a%\) raising takes place after reduplication: \text{dada} \quad \text{‘chest’} \quad [d3d3], not \*\([d3d3]\).)

(ii) lowering of high vowels (7.4):

\[(74)\]
\[
\begin{array}{ll}
\text{kapukapu} & \text{‘k.o. water plant’} \\
\text{kapokapo-n-né} & \text{‘id.-GEN-3; his kapukapu’} \quad [\text{kapòkapònè}] \\
\text{undurundur} & \text{‘any small flying insect’} \\
\text{unduurndor-n-nee} & \text{‘id.-GEN-3; his undurundur’} \quad [\text{undòundònònè}]
\end{array}
\]

(Again, with reduplicated monosyllables, high-vowel lowering takes place after reduplication: \text{gige-n-nê} \quad \text{‘tooth’} \quad [\text{gigennè}], not \*\([\text{gègennè}\].)

The realisation of \text{kapokapo-n-né} in (74) also shows that a new P-word boundary is not inserted between the copy and the base: \*\([\text{kapo}]_{w}[\text{kapo-n-né}]_{w} \quad *[\text{kapò}][\text{kapòn}]_{w}\). The copy is an ‘empty’ reflection of the allophony of the base. This is confirmed by examples such as these:
At first glance, vowel pairs like [u] and [u] (75a&b), or [i] and [i] (75c&d) seem to occur in contrasting environments, when they occur in the copy only, whereas elsewhere their occurrence is purely predictable. The correct analysis is not to double the inventory of vowel phonemes, based only on the evidence of disyllabic reduplications like these. Clearly, reduplication has occurred late in the postlexical phonology, after allophony processes, and this has clouded the picture. The phonetic content of the copy is 'conditioned' by whatever conditions the shape of the base, and as such is no more than an 'empty' reflection of that base; the copy therefore cannot be considered to constitute an environment distinct from that of the base.

Similarly, the realisation of (75a) gives the impression that a P-word boundary has been inserted between base and copy, triggering the usual [-ATR] vowel allophony in the copy (rule (47), section 2.2.2). However this is clearly not the case, given (75b), which does not surface as *[undurundure], and (75f), not as *[cólócolónne]. Inherent reduplication of disyllabic bases thus does not trigger parsing of a new P-word between base and copy.

Even derivational reduplications often show exactly the same, clearly postlexical, reduplication, transferring the results of allophony to the copy, with no evidence for a P-word boundary separating copy and base:

This data suggests then that the P-word parse, presumably with all other cyclic prosodic parses, has ceased to apply when postlexical reduplication occurs.21 (On problems with derivational reduplication and P-word bracketing, see 8.9.1.)

This leads directly, I believe, to the explanation for the patterns observed above. The explanation hinges on the following, largely independently motivated, assumptions:

21If failure to insert a P-word boundary was limited to inherent reduplications, this could be explained differently: inherent reduplication is a purely phonological process (4.2), rather than a derivational one: the prefixed copy is not a separate morpheme, and so cannot be parsed as a distinct P-word. This account fails for the derivational reduplications, such as those in (76).
Chapter 8

1. All phonological rules apply, or attempt to apply, continuously from the first level of the lexicon, until they are 'switched off' (the Strong Domain Hypothesis, Kiparsky 1984, cited in Rice 1990).

2. The input to, and the output of, all phonological processes must satisfy Foot Well-formedness (5.6), as long as it is active in the phonology. All well-formedness constraints based directly or ultimately on the foot (such as L-slot constraints) must also be met.

3. With inherent (purely phonological) reduplications, first the base, then later the copy-base sequence both constitute a single morph for the purposes of Foot Well-formedness.

4. Foot Well-formedness ceases to apply late in the postlexical phonology, before reduplication ceases to apply.

5. Reduplication is a four-part process (8.0): 1) delimit the base 2) impose a reduplicative template on the base 3) copy the phonological material within that template 4) prefix the copied material to the base.

6. The 'copy' stage of reduplication copies totally whatever is in the phonological material previously delimited by the reduplicative template, at the time that it (copy) applies: hence copying of non-contrastive material in postlexical reduplication.

Assumptions (1) to (4) above account for most of the differences between mono- and disyllabic bases, outlined above: both with respect to reduplication, and to well-formedness constraints generally. Inherently reduplicated monosyllabic bases undergo the four-part process (cf 8.0) of reduplication during the lexical phonology, before many other phonological rules apply. The disyllabic morph created by the copy+base output passes the Foot Well-formedness requirement vacuously.

On the other hand, while assumption (1) requires that all bases - mono-, di- or tri-syllabic - undergo the initial reduplication processes of parsing the base and determining the nature of the prefixed copy, in the lexical phonology, Foot Well-formedness rejects the output of reduplication of disyllabics, since that output would be a quadrisyllabic, larger than the maximal Foot (5.6). Only after Foot Well-formedness ceases to apply can the final stage in the reduplication of disyllabics, prefixing of the copy, take place.

There are two possible ways in which the blocking process, and the subsequent resumption of reduplication, could proceed: either (i) the first three stages of reduplication are completed, then the final PREFIX stage simply stalls, or applies iteratively, failing each time, until Foot Well-formedness ceases to apply or (ii)

---

22 In the Lexical Phonology framework, this can be seen as a Structure Preservation effect, applying at the prosodic level.

23 Assume that inherently reduplicated monosyllabics are underlyingly specified as undergoing reduplication, rather than automatically undergoing it due to failure to satisfy Foot Well-formedness; the latter account would have to explain the existence of 400-plus monosyllabic major class morphemes.
Reduplication

Reduplication fails totally once PREFIX has been blocked; it must begin again, from the first parse stage, iteratively, until satisfied. The evidence from copying of non-distinctive features indicates that the second model is the correct one. The transfer of non-distinctive features means that the process of COPY too must be delayed until after allophony rules cease to apply. Since there is no reason why COPY, as opposed to PREFIX, should be blocked at an earlier stage, I conclude that the whole process of reduplication is reiterated, and finally effected in the postlexical phonology, once Foot Well-formedness has gone.

Postlexical reduplication of disyllabics then can be seen to be compatible with the Strong Domain Hypothesis (Kiparsky 1984, cited in Rice 1990), which requires that if a phonological rule is assigned to more than one level within the phonology, then those levels must form a continuum. Rather than being arbitrarily assigned to two widely separated levels, the reduplication process is present from the early lexical phonology to the late postlexical phonology; within that continuum it continues to try to apply until the necessary conditions for its realisation have been created. The alternative, a 'placeholder' analysis of inherent reduplication (discussed in Woodbury 1987), would necessarily violate the SDH hypothesis.

Assumption (1) also accounts in a straightforward way for the unexpected presence of double stress in reduplicated monosyllables (examples (69), which otherwise pattern as simple morphemes; this requires of course that stress is assigned early in the lexical phonology, before reduplication. Stress patterns clearly cannot be used as a diagnostic for P-word, or other prosodic, boundaries.
Part 1 - Concluding remarks

When I began work on this thesis, I assumed that the section on phonology would be the shortest, and the most straightforward to write. Instead, it grew and grew, as I found more and more things which I had not seen discussed elsewhere, and whose description was not at all straightforward. The following elements in the phonology of Balinese presented particular analytical challenges:

1) The interactions of phonology and semantics demonstrated in chapter 3. These influence not only allophony (2.1.2, 2.2.8), but also prelexical processes determining the shapes of morphemes (for example, 6.2, 6.2.2, 6.3.1). The claims made here, as well as presenting a strong challenge to the 'standard theory' of non-interaction between the phonology and syntax components of linguistic structure (eg Archangeli & Pulleyblank 1994:5), enrich the notion of optimal phonological structure (Prince & Smolensky 1993), given that marginal phonological structures are preferred in the creation of expressive lexis in Balinese.

2) The 'long-distance constraints' on morph structure (eg 6.2, 8.7), which recall those operating in Semitic languages.

3) The important role of the Foot as a higher prosodic unit active in the phonology (chapters 6 and 8, also eg 2.1.3, 5.5).1

4) The existence of purely phonological formants, more complex than the phoneme, and described in chapters 4 to 6, and 8. The evidence from Balinese is that these are available for prelexical phonological processes of morpheme building. In chapter 4 I give evidence for a principled distinction between the purely phonological use of reduplication (and of the purely segmental complex formants identified), and a derivational/morphological one.

Those then are aspects of Balinese phonology I see as being of particular interest to theoreticians of phonology and morphology. Other things, such as the account of morphophonemic processes not previously described in a Balinese dialect (eg 7.3, 7.4), and the cataloguing of a variety of reduplication types, including types not identified in other descriptions of Balinese (8.5), will I hope be of interest to students of Balinese, and of phonology in general.

---

1 The role of the Foot in the phonologies of Western Indonesian languages in general is discussed in Clynes 1996.
PART II

MORPHOSYNTAX
9

Overview of Morphosyntax

9.0 Introduction

This chapter provides a brief survey of the principal features of Balinese verbal morphology and of some basic syntactic structures. The main points made here are exemplified in greater detail, and justified, in later chapters. In this introductory sketch, I occasionally use terms such as 'Subject', and 'transitive', without defining them, anticipating the definitions to be found in later sections.

9.1 Typological sketch

Balinese is a morphologically agglutinating language, with primarily SV(O) word order (though V(OS) is also common).

As with other Austronesian languages (see eg Cumming and Wouk 1987, Kroeger 1991a), there is debate as to whether in its syntactic orientation it is classifiable as nominative-accusative, ergative-absolutive, or neither of these (see also 14.2). Balinese is a head-initial language, with these basic orderings:

Verb > Object/Complement NP
Noun > Stative verb
Noun > Genitive
Noun > Relative clause
Adposition > Noun phrase

Examples of each of these orderings are given in (1):

(1) a. Verb > Object/Complement (see footnote 1)

\[
\begin{array}{cccc}
\text{id}a & \text{be}t\text{ar}a & \text{Brah}ma & \text{ng-}r\text{auh-in} \\
3\text{HON} & \text{god} & \text{B.} & \text{AP-}c\text{ome-LOC} \\
\text{Sang} & \text{Gana} & \text{3HON} & \text{(title)}
\end{array}
\]

The God Brahma visited Sang Gana

b. Noun > Stative verb\(^2\)

\[
\begin{array}{cccc}
\text{an}ak & \text{lacur} & \text{biu} & \text{me-}l\text{ablab} \\
\text{person} & \text{poor} & \text{banana} & \text{me-boil} \\
[\text{a}] \text{poor person} & & \text{boiled banana}
\end{array}
\]

---

\(^1\)Evidence for the Object grammatical relation is in fact very limited, see 9.5.2.

\(^2\)There is no evidence for a separate word class 'adjective' in Balinese.
9.2 Simple clause structure

The simplest clauses in Balinese consist of just a Subject noun phrase (NP) and a Predicate. The notion ‘subject’ is defined in 9.5.1. The predicate can consist of an NP, such as anak Bali in the following example:

(2) tiang anak Bali
1exc person B.
I am a Balinese

or a prepositional phrase (PP):

(3) i Madé uli Bona
DET M. LOC B.
Madé is from Bona

(4) i Ketut ke peken
DET K. to market
Ketut has gone/is going to the market

More commonly it will consist of a verb phrase (VP). In simple clauses this might be filled by an intransitive verb:

(5) i Wayan gedeg
DET W. angry
Wayan is angry

3The genitive suffix occurs only after nouns ending in a vowel.
Overview of morphosyntax

(6) i Komang ngeling
DET K. AP-CRY
Komang is crying

(7) kuluk-é me-laib
dog-DEF me-run
The dog is running

(8) m-endep koné masih i Bawang
me-be.silent PART TMA DET B.
Bawang remained silent

or by an intransitive verb with an adjunct NP in a prepositional phrase:

(9) demen ajak poh Om Tut?
happy with mango 'uncle'.T.
Does 'uncle' Tut like mangoes?

Alternatively, it will be filled by an Actor-pivot transitive verb (9.4), and one or more NPs:

(10) ia nyemuh padi
3 AP-dry rice
she dried rice

(11) tiang lkar ng-lamar ia
1exc will AP-apply.for[BI] 3
I will propose to her
[From 'Nang Kepod' by Ketut Aryana, in Bagus and Ginarsa (1978)]

(12) unduk-é totonan banget ngawé bendu ida Hyang Gana
fact-DEF DEM very AP-make angry 3.HON H. G.
That fact really made the god Gana angry

(13) i Putu naar lawar
DET P. AP-eat (dish)
Putu eats lawar

At clause level, both Subject - Predicate and Predicate - Subject word orders are common:

(14) kuluk-é medem
dog-DEF AP-sleep
The dog is sleeping

(15) medem, kuluk-é
The dog is sleeping
Subject-initial order involves a single (falling) intonational curve in unmarked utterances; predicate-initial order can contain two such tone groups, and as such is probably the more marked order. As indicated above, the verb complement or 'object' NP (see 9.5.2) almost always immediately follows the verb, regardless of the position of the Subject NP:

(16) kuluk-é medem-in tikhe(-é).
dog-DEF AP.sleep-LOC mat(-DEF)

The dog is sleeping on the mat

(17) sing ngelah pipis, tiang
NEG possess money 1 exc

I've got no money

As in English, it may very occasionally be fronted for pragmatic effect:

(18) pipis sing ngelah
money NEG possess

Money [I] don't have

(19) [ané kéto-kéto]⁴ tusing nyidaang kola makmak
REL RED-like.that NEG able.to 1 sg chew

Things like that, I can't 'chew' [i.e. deal with].
[from Cupak, by Gdé Dharna, Sukasada, Buléleng: BP-7 Dat II 1990]

9.2.1 Ellipsis of Noun Phrases

Ellipsis of topical noun phrase arguments is very common. In example (20) the subject NPs of two successive 'Undergoer Pivot/passive' verbs are deleted, while in (21) the understood object of an 'Actor pivot/active verb' is omitted (in B's response):⁵

(20) tumbén [ ] jemak=a ajak niné-n-né, [ ] efico-in=a
first.time pick.up=3 with grandma-GEN-3 urine-LOC=3

the first time [he] was picked up by his grandmother, [she] was urinated on by him.

(21) A: dija i Kadék?
where DET K.

where is Kadék

B: i Ketut ng-ajak [ ]
DET K. AP-with

Ketut is accompanying [him]

⁴This is a noun phrase, consisting of a 'headless' relative clause.
⁵The brackets indicate places where the ellipsed NP's could occur. Other positions are also possible. On my understanding of the terms Undergoer and Actor see 9.3.1, for the Undergoer Pivot and Actor Pivot constructions, see 9.4.
Example (20) also illustrates how contextual/pragmatic information is often important in establishing the referent when such ellipsis occurs.

A different context in which 'zero NPs' occur is in anaphoric reference to non-humans. Pronouns only encode human referents, and some anthropomorphic uses with particular animals. Otherwise, anaphoric reference to animals and inanmites necessarily consists either in the repetition of the entire NP, or (most commonly) simple 'zero anaphora':

(22)\[\text{dija} \ [ \ ] \ \text{meli} \ [ \ ]\]
where buy
Where did [you] buy [it]?

This last example also illustrates the way 1st- and 2nd-person pronouns are very commonly ellipsed. Chapter 13 investigates the way pronominal avoidance has influenced the use and interpretation of Undergoer Pivot constructions.

9.3 Basic verbal morphology

Non-derived verbs (those which are inherently verbs) can be divided into two main classes: transitive and intransitive (9.3.4). These in turn can be sub-classified in terms of the affixes they carry: none, one or more than one. Morphology in Balinese is agglutinating in the sense that boundaries between affixes and roots are generally easily defined. The maximal verb structure usually recognised is: 6

Prefix-Prefix-ROOT-Suffix

Examples of the different possible root-affix combinations are given in the following table:

<table>
<thead>
<tr>
<th>structure</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROOT Prefix-ROOT Root</td>
<td>tua 'old'</td>
</tr>
<tr>
<td>Root-Suffix</td>
<td>me-daar 'eat'</td>
</tr>
<tr>
<td>Prefix-ROOT-Root-Suffix</td>
<td>tegak-in 'sit on (UP)'</td>
</tr>
<tr>
<td>Prefix-ROOT-Suffix</td>
<td>negak-in (%N-tegak-in%) 'sit on (AP)'</td>
</tr>
<tr>
<td>Prefix-Pfx-Pfx-ROOT-Suffix</td>
<td>me-tegak-an 'be caused to sit'</td>
</tr>
<tr>
<td></td>
<td>me-gedé-nang (%N-pa-gade-naŋ%) 'make bigger (AP)'</td>
</tr>
<tr>
<td></td>
<td>me-se-sambatan (%ma-RED-sambat-an%) 'call out (DUR)'</td>
</tr>
<tr>
<td></td>
<td>me-pe-gaé-nan 'be married off'</td>
</tr>
</tbody>
</table>

Table 9.1 Examples of different root-affix combinations in Balinese

6In the Buleleng variety, verbs with two suffixes occasionally occur: tegak-in-ang 'sit-LOC-APP; 'sit on U for someone [eg, minding their seat at the cinema]'; ngaja-nang-ang %N-kaja-arj-arj% AP-mountainwards-APP-APP, 'be going more and more mountainwards'. These have not to my knowledge been mentioned in previous descriptions.
9.3.1 Terminology 1: Actor and Undergoer semantic roles

In the analysis of Balinese verbal morphology presented here, the notions Actor and Undergoer recur constantly. Foley & Van Valin (1984:20) define these terms thus:

'Provisionally we may characterize the Actor as the argument of a predicate which expresses the participant which performs, effects, instigates or controls the situation denoted by the predicate, and the Undergoer as the argument which expresses the participant which does not perform, initiate or control any situation but rather is affected by it in some way.'

This definition is useful as an initial 'working' approximation. However, to fully account for the data presented in the following chapters, one important qualification is necessary. Foley & Van Valin's definition refers explicitly to NPs realised as syntactic arguments ('the argument of a predicate which expresses [...]'). In Balinese, Actor and Undergoer are primarily semantic notions. The evidence from both morphology and syntactic patterning (considered in detail in 15.3) shows that a verb may refer to both Actor and Undergoer participants in its semantic structure, yet only involve a single independent NP argument in the surface syntax. This is the case for example with:

- Actor Pivot verbs where the generic Undergoer is incorporated (9.5.2, 15.3),
- Formally Actorless 'Bare stem-' and 'ke-' Undergoer Pivot verbs (9.4, 15.3),
- Verbs where reference to an undergoer is encoded in the verb suffixes, but not permitted in the syntax, eg ng-lengheh-in 'be intoxicating' (11.2.2.2), and nge-mokoh-ang 'become fatter' (11.3.2.1, 15.3).

In applying these notions to Balinese then, at the very least a modified version of Foley & Van Valin's definition is needed, one which refers not to arguments in the syntax, but to participants in the semantic/conceptual structure.

(23) Actor:
'the participant which performs, effects, instigates or controls the situation denoted by the predicate'

Undergoer:
'the participant which does not perform, initiate or control any situation but rather is affected by it in some way'.

For Foley & Van Valin (1984:29-36), Actor and Undergoer are 'macroroles', each representing a wide range of possible semantic roles. Thus an Actor could bear any one of a range of 'etic' semantic roles, such as Agent, Instrument, Experiencer, or Source: all will be encoded in the syntax in the same way, in contexts where the language is sensitive only to the Actor-Undergoer opposition. Similarly an Undergoer could be describable etically as Patient, Locative, Source, or otherwise: again where the language
is sensitive only to the Actor-Undergoer opposition, these will all receive the same morpho-syntactic expression.⁷

The Actor-Undergoer opposition is prominent in Balinese morphosyntax. It gives rise to a 'split-S' patterning in the morphology of intransitive verbs (12.3), and informs the opposition between the Actor Pivot and Undergoer Pivot voice diatheses (9.4).

The use of the semantic categories Actor and Undergoer, in combination with the understanding of the term 'transitive stem/verb' adopted here (15.3) enables generalisations to be stated more concisely and accurately than, say, using a grammatical-function based approach.

For example, one simple generalisation that can be made is that the suffix on the transitive stem signals further information about the semantic role of the Undergoer (15.3). Using a non-semantic approach this must be expressed as a disjunction, something like 'the suffix on the transitive stem gives further information about the semantic role of either the Object NP of an AP verb or the Subject NP of a UP verb'. It fails completely to account for intransitive verbs which carry this suffix (eg 11.3.2.1). The latter description also requires the problematic assumption that the Undergoer NP associated with an Actor Pivot verb always bears the grammatical function 'Object' (9.5.2); it moreover fails to signal the special status of the Undergoer NP as the one NP which is obligatorily subcategorised in both Actor Pivot and Undergoer Pivot voices.

9.3.2 Verbal prefixing

Most commonly, verbs carry only a single prefix, or no prefix at all. Paradigmatic contrasts in the prefix position give information about the Subject NP (9.5.1). They are illustrated in the following table (the glosses are simplifications for expository purposes, see chapter 12 for analysis):

<table>
<thead>
<tr>
<th>prefix</th>
<th>approximate gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>'∅'</td>
<td>no prefix ('zero prefix')</td>
</tr>
<tr>
<td>/N/</td>
<td>'agentive prefix'</td>
</tr>
<tr>
<td>ke-</td>
<td>'passive' on transitive verbs⁸</td>
</tr>
<tr>
<td>me-</td>
<td>1. 'agentive prefix', certain non-derived intransitives</td>
</tr>
<tr>
<td></td>
<td>2. 'passive' on derived intransitives</td>
</tr>
<tr>
<td>pe-</td>
<td>'causative'</td>
</tr>
</tbody>
</table>

Table 9.2 Paradigmatic contrasts in the prefix slot

⁷A different view (Givon 1990, Wierzbicka 1992) which I will adopt in 15.3, is that Actor and Undergoer, like the related 'transitivity', are prototypic notions: participants in individual predicates will (or will not) be realised in the morphosyntax in the way that prototypical Actors and Undergoers are, according to the degree to which they conform to, or depart from, the prototype.

⁸Clearly, I am using a special understanding of the term 'transitive verb'. See 9.3.4.
I will now exemplify these prefixes in turn:

(i) Verbs with no prefix have an Undergoer subject (12.1). They may be either intransitive:

$$\begin{align*}
\text{pules} & \quad \text{'sleep'} \\
\text{ulung} & \quad \text{'fall'} \\
\text{gedeg} & \quad \text{'be angry'}
\end{align*}$$

or else realise the passive-like (Bare Stem) Undergoer Pivot (9.4) diathesis of transitive verbs, e.g. ingsap-in in this example:

$$\begin{align*}
\text{mani} & \quad \text{déén} & \quad \text{ingsap-in} & = a \quad \text{`ba} \\
\text{tomorrow} & \quad \text{just} & \quad \text{forget-LOC} & = 3 \quad \text{already}
\end{align*}$$

By tomorrow, even, he’ll have forgotten about [it]! (lit. [lit] will have been forgotten by him)

(ii) The 'agentive' prefix, %N%, is also found both on intransitives:

$$\begin{align*}
\text{negak} & \quad \text{(}{%N\text{-}}\text{tagak}{\%}) \\
\text{ng-øyong} \\
\text{ngwél} & \quad \text{%N-wel{\%}} \\
\text{ng-ijeng}
\end{align*}$$

and on Actor Pivot (9.4) transitive verbs:

$$\begin{align*}
\text{ida} & \quad \text{nyumpéna} & \quad \text{mangghihin} & \quad \text{anak} & \quad \text{lingsir} \\
\text{N-sumpena} & \quad \text{N-pangghih-in} & \quad \text{3.HON AP-dream} & \quad \text{AP-meet.HON-LOC} & \quad \text{person old.HON}
\end{align*}$$

He (the king) dreamed he met an old person (Kirtya 2070)

In both cases, the semantic role of the subject NP is that of Actor.

(iii) The prefix ke- occurs on a second passive-like verb form, which I call the ke-Undergoer Pivot. This is described in 9.4.3, and in more detail in chapter 13.

$$\begin{align*}
\text{napi} & \quad \text{ke-juang?} \\
\text{what.HI} & \quad \text{ke-take}
\end{align*}$$

What was taken ([Understood:] by you)?

(iv) Verbs carrying the prefix me- are always intransitive. Where the root morpheme is also a verb, there are two sub-types, firstly where the subject NP is an Actor (29a), and secondly, where it is an Undergoer (29b):

---

9 Purely due to time constraints, pe-, the relatively marginal causative prefix, is not analysed in this thesis; nor are reduplication as a derivational process, and one or two minor derivational types listed in reference grammars. I have no reason to believe that the broad conclusions reached about the functioning of affixes (15.1, 15.2, 15.3) are affected by this.

10 Some writers confuse me- and certain realisations of the homophonous sequence, %N-pə-% 'AP-CAUS-'.

186
Overview of morphosyntax

(29) a. me-daar 'eat'
    me-laib 'run'
    me-jujuk 'stand up'
    me-jagur 'be in a state of having been punched'
    m-adep 'be ( ...) sold'
    me-goréng 'be ( ...) fried'

I argue in 12.4.3 that Actor-subject 'me-' verbs like those in (29a) are inherently intransitive verbs, not derived from another source. On the other hand, the Undergoer-subject me-verbs like those in (29b) are always derived from transitive verbs (12.4.1).

Undergoer-subject me- verbs constitute the third passive-like form. They are both formally and semantically actorless:

(30) jani lakar me-geseng cai Rama
    now will me-burn 2 R.
    now you are going to be burnt black, Rama
    [From a wayang performance by JDP, Lovina beach March 1991]

They cannot therefore cooccur with an Actor NP, even in an oblique PP:

(31) suba me-mem, baas-é, (*tekén i Madé)
    already me-soak rice-DEF (by DET M.)
    The rice is already soaked (*by Madé)

Finally, me-prefixed verbs are also formed from nominal roots. The subject is always in some sense a Locative; pragmatic factors then determine whether that subject is further interpreted to be Actor or Undergoer (12.4.4):

(32) suba me-suah, i Komang?
    already me-comb DET K.
    Has Komang combed [her hair]?/Has Komang's hair been combed?

To sum up verbal prefixing patterns, assuming that the word class of verbs with prefix me- is known, prefixing contrasts in Balinese give information about the semantic role of the Subject NP:

(i) Undergoer subjects:
    Zero prefix verbs, plus detransitive verbs with prefix me-

(ii) Actor subjects:
    Verbs with the 'agentive' prefix %N%, plus underlyingly intransitive verbs with prefix me-

Balinese is then a 'split-S' language (Dixon 1994:70): it obligatorily signals the semantic macrorole (Actor or Undergoer) of the subject NP of both intransitive and transitive verbs (see 12.3 for discussion).

11On the analysis of bound root morphemes, see 10.4.
9.3.3 Verbal suffixing

If verbal prefixes give information about NP's bearing the grammatical relation 'Subject', verbal suffixes give purely semantic information, about the precise role of the Undergoer of the predicated event. Suffixing is particularly important in the formation of transitive verbs. These are formed on a transitive stem, which has a three-way meaningful contrast in suffixing possibilities (including '-∅, the absence of a suffix):

<table>
<thead>
<tr>
<th>suffix</th>
<th>approximate gloss</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>-∅</td>
<td>no suffix ('zero suffix')</td>
<td>jagur 'punch'</td>
</tr>
<tr>
<td>-(n)in</td>
<td>'Locative'</td>
<td>jagur-in 'punch PL'</td>
</tr>
<tr>
<td>-(n)ang</td>
<td>'Applicative'</td>
<td>jagur-ang 'punch for'</td>
</tr>
</tbody>
</table>

Table 9.3 Paradigmatic contrasts in the suffix slot, transitive verbs

Here are examples of the transitive stem, with the typical Undergoer semantic roles occurring with each suffixing possibility:

(33)  'Suffix' Type: Typical Undergoer Semantics:

a. ROOT-∅ true patient; thing sensed
   jagur 'hit'
   cakcak 'chop up'
   adek 'smell (vt)'
   adep 'sell'

b. ROOT-in locative roles: goal, source, recipient
   tegak-in 'sit-LOC: sit on'
   beli-in 'buy-LOC: buy from'
   kauk-in 'call-LOC: call to'
   demen-in 'happy-LOC: like, love'

c. ROOT-ang other peripheral roles: causee, recipient, moved theme, etc
   pules-ang 'sleep-APP: put to sleep'
   tegak-ang 'sit-APP: sit someone down'
   beli-ang 'buy-APP: buy for (RECIP), 'buy with (THEME)'

The Actor Pivot and (two, distinct) Undergoer Pivot verb forms are formed on the transitive stem:

12 That undergoer is not necessarily realised in the syntax; hence my use of the term 'transitive verb' in a purely semantic sense (9.3.4, 15.3).
The stems listed in (33) thus also realise the bare-stem Undergoer pivot form without further affixation (see 9.4).

The patterns in (33) apply only to transitive verbs. Only one suffix occurs exclusively on intransitive verbs, -(n)an. Intransitives carrying -(n)an always derive from other lexemes, and have subjects with 'peripheral' semantic roles (Locative, Theme, Source, etc):

(35) SOURCE LEXEME DERIVED VERB

<table>
<thead>
<tr>
<th>SOURCE LEXEME</th>
<th>DERIVED VERB</th>
</tr>
</thead>
<tbody>
<tr>
<td>kutu</td>
<td>'lice-ridden'</td>
</tr>
<tr>
<td>bading-ang</td>
<td>'U be turned over'</td>
</tr>
<tr>
<td>duduk-in</td>
<td>'U be collected from'</td>
</tr>
</tbody>
</table>

One cannot speak of a paradigmatic contrast between -(n)an and absence of a suffix in the case of intransitive verbs; suffixless inherently intransitive verbs with both actor and undergoer subjects occur (cf examples (24) and (26)).

Where intransitives derive from verbs of shapes ROOT-in or ROOT-ang, they have shape me-ROOT-an; the -in/-ang suffixing contrast is therefore neutralised in these derived verbs (11.4, 12.4.1):

(36) SOURCE LEXEME DERIVED VERB

<table>
<thead>
<tr>
<th>SOURCE LEXEME</th>
<th>DERIVED VERB</th>
</tr>
</thead>
<tbody>
<tr>
<td>tulis-in</td>
<td>'U be written on'</td>
</tr>
<tr>
<td>tulis-ang</td>
<td>'U be written with'</td>
</tr>
</tbody>
</table>

From 'ROOT-Ø' transitive, intransitive verbs of shape me-ROOT-Ø are derived:

(37) SOURCE LEXEME DERIVED VERB

<table>
<thead>
<tr>
<th>SOURCE LEXEME</th>
<th>DERIVED VERB</th>
</tr>
</thead>
<tbody>
<tr>
<td>jagur</td>
<td>'U be hit'</td>
</tr>
<tr>
<td>cak-cak</td>
<td>'U be chopped up'</td>
</tr>
<tr>
<td>tulis</td>
<td>'U be written'</td>
</tr>
</tbody>
</table>

9.3.4 Terminology 2: A note on the use of the term 'transitive'

In the previous sub-section, as throughout this thesis, I refer to both Actor Pivot ('active-like') and Undergoer Pivot ('passive-like') verb forms as 'transitive'. This is despite the fact that both forms may at times occur in syntactically intransitive clauses: UP always (14.2), and AP sometimes (9.5.2). I am in fact using the term 'transitive' in a semantic sense, to refer to verbs which refer to both an Actor and an Undergoer in their semantic structure. There is good evidence, from both verb morphology, and from the syntax, to support this practice (see 15.3). For essentially the same reasons, the terms 'Actor' and 'Undergoer' are used in this thesis to refer to semantic entities (9.3.1 and 15.3).

I also use the term 'transitive' in the traditional syntactic sense, to refer to verbs which occur with a direct object, or to clauses containing such a verb (see eg Crystal
1991:360). Unless used specifically about clause-level syntax, the term 'transitive' should be understood to have the first, semantic, sense.

9.4 Actor Pivot/Undergoer Pivot diatheses

Transitive verbs occur in two basic voice oppositions, 'passive(-like)', or Undergoer Pivot (UP) and 'active(-like)' or Actor Pivot (AP), according to which participant is realised as the grammatical subject. I follow the generally accepted Austronesianist practice of avoiding the terms 'passive' and 'active', given the differences between typical active/passive systems and the voice systems found in Austronesian languages (see 14.1)).

I discuss in turn first the single Actor Pivot construction (9.4.1), then the two Undergoer Pivot constructions, the Bare Stem UP (9.4.2) and the ke-UP (9.4.3). The third passive-like construction (cf Kersten 1984:93-99), that where the verb carries the prefix me-, is discussed in 12.4.1.

9.4.1 Actor Pivot

There is then a single (morphologically complex) Actor Pivot construction. The word order is typically subject-initial, with the verb carrying the agentive prefix, %N%, indicating that the subject is an Actor:

(38) kuluk-é medem-in tikeh
    N-pedem-in
    dog-DEF AP-sleep-LOC mat

The dog is sleeping on a mat

This is the best candidate for the 'basic transitive construction' of Balinese (pace Artawa (1992) and Blake (1993), who give that title to the Bare Stem Undergoer Pivot construction (9.4.2 below)). I show in 9.5.2 and 14.2 that the Actor Pivot construction is the only one which can (at times) be characterised as 'transitive', being the only one in which two core NPs can occur. Where the Undergoer has indefinite reference, as in the above example, the indications are that the clause is syntactically intransitive (9.5.2).

9.4.2 Bare stem Undergoer Pivot

The first of the Undergoer Pivot constructions is the unmarked Undergoer Pivot, which I will refer to as a 'Bare Stem' Undergoer Pivot. It consists of the bare verb stem (including suffixes); the Undergoer NP is subject, while the verb obligatorily carries a
clitic actor (pro)nominal, either a pronominal (ida in (39)), or a noun phrase with generic reference (kuluk in (40)).

(39) sirsa-né engkeb-ang = ida di tengah angga-né
head.HON-DEF hide-APP= 3.HON LOC inside body.HON-DEF
The head was hidden by him inside his body ['he' = the god Brahma, revealing only four of his five heads. from Hyang Kumara, by I Nyoman Antasa, Banyuning, Singaraja, KVlb 5975]

(40) tikeh-é pedem-in = kuluk
mat-DEF sleep-LOC = dog
The mat was slept on by a dog

Such cliticised actors form a single phonological word with the verb, with a single main stress on the final syllable: [pədəminkɿu'kɪ], [əŋkebənɿi'd3]. The verb-Actor unit can never be interrupted by other constituents:

(41) !*sirsa-né engkeb-ang di tengah angga-né ida
(42) !*pedem-in tikeh-é kuluk

The third person (non-honorific) pronominal =a is unique in its phonological behaviour: it is the only monosyllabic pronominal, and the only one which is a true clitic, not attracting stress (eg [ja'mak3], ['pak'pak3] in example (43) following). It only occurs on Bare Stem UP verbs (never on ke-UP verbs):

(43) umbe-n-né jemak =a laut pakpak =a
tuber-GEN-3 pick.up= 3 then chew= 3
the tuber was taken by him, then chewed by him
[From 'Satua I Lutung Ian I Kekua', by drs Ida Kadé Suarioka, Banjar, broadcast on RRI, March 1991]

If the actor is realised by a noun (phrase), in the 3rd person, and with definite reference, it must occur in a separate prepositional phrase (44b); the verb still carries the 3rd-person pronominal clitic, ' =a'.

(44) a. jemak =a bangké-né laut tegen =a tekén pengalu-né
pick.up= 3 corpse-DEF then carry =3 with peddler-DEF
the corpse was picked up, then carried by the peddlers
[KVlb 1317/4, Satua Pengalu Mati Tepén Buah]

---

13 If the Actor is 3rd person, the verb must carry an obligatory 3rd person clitic pronoun. 1st and 2nd person Actors of Bare Stem Undergoer Pivot constructions can be ellipsed in many contexts, however such formally Actorless clauses will always be assumed to have a 1st- or 2nd-person Actor (see chapter 13).
14 Artawa's (1992:15-16) claim that =a is a 'passive' suffix, rather than a pronominal is clearly wrong: it would have to derive passives which only ever have non-honorific third-person Actors. At the same time Artawa is forced to posit a distinct 'basic verb construction' which occurs with everything but non-honorific 3rd person Actors. Clearly there is only a single construction, the Bare Stem UP, with =a being a pronominal, not a passive marker.
15 In relative clauses the actor can occur in an NP rather than a PP, see for example (53a).
9.4.3 \textit{ke}-Undergoer Pivot

The second Undergoer Pivot form is the \textit{ke}-Undergoer Pivot. Morphologically it is identical to the Bare Stem UP, except that it carries the prefix \textit{ke-}. It differs syntactically in that an Actor NP is not obligatory, and where present, usually occurs in a separate prepositional phrase, particularly in formal high styles:

(45) \textit{indik-é punika ke-takén-ang ring ida sang sura bisama}
\hspace{1cm} \text{matter-DEF that ke-ask.HI-APP PREP.HI 3.HON S. S. B.}
\hspace{1cm} \text{the matter was asked about by Sang Surya Bisama}
\[\text{[From a wayang performance by JDP, Lovina beach March 1991]}\]

With third person agents the \textit{ke-UP} verb never carries the clitic \textit{=a}, in high or low styles:

(46) \text{*indik-é punika ke-takén-ang =a ring ida sang sura bisama}

In both low and high high style, formally agentless \textit{ke-UP} constructions are common:

(47) \text{ida ndauh-in\textsuperscript{16} i patih,}
\hspace{1cm} \text{3HON AP-call-LOC DET official}
\hspace{1cm} \text{His highness called his adviser,}
\hspace{1cm} \text{[and the advisor] was ordered to assemble the vassals}
\hspace{1cm} \text{[from Kirtya 2070 'I Sudamala', by I Giret, Bratan, 11/10/1940]}

However, in Singaraja Balinese the \textit{ke-UP} construction resembles the Bare Stem UP in that it too occurs with (apparently cliticised) generic and pronominal agents:\textsuperscript{17}

(48) \text{désa-n déwék-é k-entas-in = blabar agung}
\hspace{1cm} \text{village-GEN body-DEF (=1PRO) ke-pass-LOC=flood great}
\hspace{1cm} \text{our village was traversed by a great flood}
\hspace{1cm} \text{[From 'Nang Kepod' by Ketut Aryana, in Bagus and Ginarsa (1978)]}

(49) \text{jaga ke-pe-kidih-ang = titiang sampun,}
\hspace{1cm} \text{will.HI ke-CAUS-askfor-APP = 1exc.DFR PART}
\hspace{1cm} \text{[she] will be given away by me now, my child Sri}\textsuperscript{18}

In low informal styles, clauses with formally actorless \textit{ke-UP} constructions are always interpreted to have 1st- or 2nd-person actors (see chapter 13). This is not the case in formal styles, where they can have 3rd-person actor readings.

Though they have some differences then in syntactic behaviour, and in pragmatic value (chapter 13), the Bare Stem and \textit{ke}-Undergoer Pivot verb forms share many overlapping functions. Artawa (1992) calls the \textit{ke-UP} a 'passive', with the implication that it derives

\textsuperscript{16}This verb carries an unexpected, literary, form of the \textit{AP} prefix, influenced by Javanese texts.
\textsuperscript{17}See also 13.3.4.1, on the differing syntax of UP constructions in relative clauses. Kersten (1984:98) incorrectly claims that \textit{ke-UP} verbs bearing the suffixes \textit{-in} and \textit{-ang} can only occur with a third-person actor.
\textsuperscript{18}Said in a formal speech during a pre-wedding meeting of the families concerned, Peguyangan, May 1991.
Overview of morphosyntax

from the Actor Pivot form. For the Singaraja dialect at least, there is no evidence to warrant this assumption. ke-Undergoer Pivot verbs more plausibly derive from the much more similar Bare Stem Undergoer Pivot verb. Apart from its simpler morphology, evidence from the behaviour of certain verb classes indicates that the Bare Stem UP is more basic than the ke-UP (13.3.2).

9.5 On Subjects, Objects, and Core NP's

In this section, I describe some distinctive characteristics of NPs referred to respectively as 'subject NP', 'object NP', 'core NP', and 'non-core NP'.

9.5.1 The Subject NP

I use the terms 'subject' and 'pivot' interchangeably; both are common in the western Austronesian literature. Characteristics of noun phrases which can be considered to bear the subject grammatical relation in Balinese include the following:

(i) Subjects are the only obligatory (though ellipsable) referential NPs which are not part of the VP syntactic unit, and which do not occur in a prepositional phrase. (ii) Subjects must normally be topical or 'definite', with uniquely identifiable referents (including generic reference to all members of a given class):

(50) anak baas-é telah
PART rice-DEF used.up
The rice is all gone

Where subjects are not definite, a special presentational construction, with ada 'exist', must be used:

(51) di pesisi kangin ada anak mamondok padidi-in =a pesan LOC shore east EXIST person live.in.hut do.alone = 3 very
On the east coast there was a person who lived all alone [Kirtya K 1937]

(iii) As in many Austronesian languages, only subjects (or possessors of subjects) can be relativised, hence the grammaticality of (52a&b), but not (52c):

(52) a. pipis ané jemak =a
money REL take = 3
Money is what was taken by him

b. ia ané nyemak pipis
N-jemak
3 REL AP-take money
He's [the one] who took money

c. *pipis ané ia nyemak
Money is what he took
Requirement (iii) subsumes two other tests for subjects: 'only subjects can occur in Cleft constructions', and 'only subjects can be the target of 'Focussed Questions', since both of these constructions involve relative clause structures in Balinese. I will illustrate only focussed questions. If a question-word is sentence initial (hence pragmatically focussed), it must question the (deleted) subject of the accompanying relative clause. The relative clause marker ané is often omitted: ¹⁹

(53)

a. apa (ané) beli-na i Ketut di peken?
   what REL buy=3 DET K. LOC market
   What is [the thing which] was bought by Ketut at the market?

b. *apa (ané) i Ketut meli di peken?
   what REL DET K. AP-buy LOC market
   What is [the thing which] Ketut bought at the market?

9.5.2 The Object NP

Tests for the grammatical relation Object are more difficult to find, given that Actor Pivot/active sentences plausibly derive from Undergoer Pivot/passive in Balinese (14.1): the commonest test for objecthood, the ability to become subject of a passive verb (Andrews 1985), requires the reverse direction of derivation, which does not occur in Balinese.

I will take it that one test for objecthood in Balinese is the ability of the Undergoer NP occurring as complement to an Actor Pivot verb (underlined in the following examples) to be formally marked for definiteness, and yet still not occur in a prepositional phrase:

(54)

Mén Paluk [...] mulih nyuun kacang-é ento
   M.P. go.home AP-carry.on.head bean-DEF THAT
   Men Paluk went home, carrying those beans
   [Kirtya 1316/4, 'Satua Mén Paluk', from Banjar Dangin Peken, Singaraja]

(55)

aji lakar nyembah raga-n cening-é
   father N-sambah body-GEN child-DEF
   1 will pay.homage 2PRO
   [I, your] father will pay homage to your person

This is the only test I have found for Object status. Evidence for this relation is then not nearly as robust as that for Subject. Nevertheless, if it is valid, one can draw a single generalisation for NPs bearing either of the two core grammatical relations. Just as Subject NPs in simple clauses must be definite (see discussion of (51)), so it appears, must Objects:

¹⁹Such questions involve two NPs; the focussed question pronominal fills one, and is functionally the predicate; the headless relative clause fills the other, and is functionally the subject. The requirement that subject NPs be referential/definite is therefore not broken by focussed questions.
Definiteness requirement:
NP's bearing a core grammatical relation must be formally and semantically definite.

I argue in 15.3 that transitivity is primarily a semantic notion in Balinese. It is likely that the 'definiteness requirement' is a secondary effect deriving from the special salient status of the Actor and Undergoer participants in semantic structure.

In Singaraja Balinese, four kinds of evidence suggest that (semantically) transitive verbs behave as though they are syntactically intransitive when the complement/object NP is indefinite. For simplicity, here I give evidence from imperative constructions. Entirely parallel claims can be made about interrogative constructions, see 14.1.

1. *Form of the transitive verb in imperative clauses.* Imperatives containing transitive verbs have two forms. Where the undergoer argument is definite/referential, the bare transitive stem is used:

(57) baca ja buku-né!
read PART book-DEF
Read the book!

However where the undergoer argument is indefinite/generic, the verb stem carries the {N} affix:

(58) maca buku ja!
read book PART
Read books/do book-reading!

In this the verb in (58) behaves like the following agentive intransitives, which always retain the {N} affix in imperatives:

(59) negak!
%N-tagak % 'sit down (IMP)' *tegak!
nyakan!
%N-jakan% 'cook rice (IMP)' *jakan!
ngoyong!
%N-ooyo% 'stay(IMP)' *oyong!

2. *Position of ja in imperatives*

The 'emphatic particle' ja appears to behave like a 'second position clitic' in coming immediately after the first constituent in the sentence, as in these possibilities:

(60) baca ja buku-né!
buku-né ja baca!

Compare these unacceptable sequences:

(61) *baca buku-né ja!
*buku-né baca ja!

---

20I assume that this form of the imperative is simply a Bare Stem UP, see also 14.1; this analysis is not crucial to the argument being made here.
However when the Undergoer is indefinite, as in (58), the particle must come after the entire verb-complement sequence.

(62)  maca  
    buku  
    ja!  

??maca  
    ja  
    buku!

These patterns strongly suggest that the Undergoer is incorporated into the verb when generic, as in in (62), forming a single intransitive unit. The standard orthography clearly does not reflect the syntactic grouping in this case.

3. Non-possibility of fronting indefinite Undergoers in imperatives

While definite Undergoers can be fronted (63a), this is not acceptable for indefinite Undergoers (63b):

(63)  a.  
    buku-ne  
    ja  
    baca!  
    book-DEF   
    PART    
    read  
    Read the book!

b.  
    *buku  
    maca  
    ja!  
    *buku  
    ja  
    maca!

Again, this suggests that where the Undergoer is indefinite, the Verb-Undergoer sequence constitutes a single constituent, an intransitive verb with an incorporated undergoer.

4. Evidence from stress placement. The evidence from stress placement supports that from other tests.

(64)  
    baca  
    ja  
    buku-né!  
    [ba'c3j3buku'ne]  
    Read the book!

(65)  
    maca  
    buku  
    ja!  
    [macabu'ku3]  
    Read books/do book-reading

Whereas baca in (64) carries a separate stress [ba'c3], and the clause contains two distinct (stressed) phonological words, in (65) maca buku ja behaves as a single P-word in only allowing a single stress on the final syllable: [macabu'ku].

There are thus strong grounds for the following conclusion:

(66)  Indefinite nominal complements of Actor Pivot verbs are incorporated into the verb to form a syntactically intransitive unit.

A second conclusion, which complements the previous one, is this:

\[^{21}\text{The clitic ja optionally attracts stress, see section 5.8. This does not affect the argument made here.}\]
NP complements of AP verbs must be topical/definite for them to carry the grammatical relation 'Object'.

The latter conclusion is however not as 'solid' as the previous one, in that it is dependent on a single test for Object status.

9.5.3 Core and Non-core NPs

In the preceding two subsections, I gave evidence that only NPs which bear either the Subject or Object relations have the ability to be marked for definiteness in Balinese, without having to occur in a Prepositional Phrase. Put another way, only topical NPs can bear either of the Subject and Object relations. I shall refer to these NPs as core NPs.

In contrast, unambiguously non-core arguments must occur in a prepositional phrase if they are marked as definite (as with karung and talenan in these examples):

(68) a. ia ng-wadah-in baas-é aji karung(-é)
3 AP-container-LOC rice-DEF PREP(INST) sack(DEF)

b. dapak-né tektek-ang =a di talenan(-é)
chopper-3 chop-APP = 3 PREP(Loc) choppingboard

He stored the rice 'with/using' the sack

his chopping knife was chopped by him onto the board

Non-core NPs may occur not within a PP - in fact many only have this latter possibility - but only if they are formally and semantically indefinite, (again, see karung and talenan in these examples):

(69) a. ia ng-wadah-in baas-é karung(*-é)
3 AP-container-LOC rice-DEF sack

b. dapak-né tektek-ang =a talenan(*-é)

His chopping knife was chopped by him onto a board.

Other examples of such obligatorily indefinite NPs, which can't occur in a PP, include the [underlined] Theme arguments in these examples:

22 Clause-external topics may also be marked for definiteness (9.5.4). Beratha (1992) reserves the term 'topic' for what I refer to as the Subject/pivot NP. This usage is problematic, given the topicality requirement for Objects. She uses the term 'focus' in a novel way, to refer to VP-internal NPs, including clitic actors of Bare Stem passives. While at times she refers to these nominals as bearing a 'pragmatic role', in fact she defines this 'focussed' nominal in purely syntactic terms (1992:136-142).

23 wadah-in is a transitive verb formed from a noun root, wadah 'container'. The derived verb has the regular meaning in this case: the Undergoer (the rice) is conceived of as a 'location' which is affected by a container, manipulated by the Actor (11.2.1). The karung is the instrument (cf the preposition aji) with which this is done.
And in contrast to core NPs (see eg examples (15), (17), (18) and (19), in 9.2), these NPs cannot be freely fronted or otherwise moved. In the following example, karung can occur in only two possible positions - either immediately after the Object NP, or between it and the verb:

(72)  
-a. ia ng-wadah-in baas-é karung  
  b. ia ng-wadah-in karung baas-é  
  c. *ia karung ng-wadah-in baas-é  

  *He stored the rice 'with/using' a sack

It cannot be fronted, even in the corresponding Undergoer Pivot sentence:

(73)  
  baas-é wadah-in =a karung  
  *karung baas-é wadah-in =a  

  *The rice was stored by him in a sack

The same is true for the non-core NPs in the UP counterparts of (70) and (71):

(74)  
  *jaja pekak meli-ang i Putu  
      *Grandpa bought Putu a cake

(75)  
  *poh tiang pelut-ang=a tekén i Madé  
  1exc peel-ang =a 3 mango(-DEF) with DET M.  
      *I was peeled a mango by Made

Indefinite non-core NPs thus appear to be part of a Verb Phrase (VP) constituent, with only limited movement possibilities, strictly within the VP.

I show in 14.2 that the actor nominals in Bare Stem Undergoer Pivot constructions behave like non-core NPs, despite claims that that construction is the 'basic transitive construction' of Balinese (eg Blake 1993).

9.5.4 The External Topic NP

In addition to core NPs, 'external topic' NPs can be marked for definiteness, without having to occur in a prepositional phrase. The external topic NP occurs in sentence-initial position. As the name indicates, I assume that the external topic NP, which is always optional, is external to the clause (cf Andrews 1985:82-3):

(76)  
  ibok-né, basang-né bawak  
  mother-3 stomach-3 short  

  [As for] her mother, her stomach is short [i.e. she has a quick temper]

198
(77) 

i ₩ Bawang, mémé-n-né mati
DET B. mother-GEN-3 dead

[As for] Bawang, her mother was dead.
[from 'Bawang Kesuna' told by Ibu Ketut Supani, Peguyangan, 3/1991]

(78) 

I Gedé, sing taén ngenah mata-n-né
DET G. NEG ever visible eye-GEN-3

As for Gedé, his eyes are never visible [i.e. he's never here]
[From 'Nang Kepod' by Ketut Aryana, in Bagus and Ginarsa (1978)]
10

Verbal morphology - approach

10.0 Introduction

Here I discuss my treatment of various categories often referred to in analyses of verbal morphology. The first section, 10.1, briefly describes the method of data collection, itself of importance to the subsequent analysis. The distinction between free and bound morphemes is described in 10.2, and that between derivational and inflectional morphology in 10.3. In 10.4 I give my reasons for rejecting the existence of a 'precategorial' lexical class in Balinese.

In chapter 15 I draw some general conclusions about the function of verbal morphology in Balinese.

10.1 Collection of data

In collecting data on verbal morphology I followed the approach used by Dixon in describing Boumaa Fijian (1988:205): a sample of roots was made, covering a range of semantic areas. This sample included (Dixon's categories): verbs of motion and stance; of giving and carrying; verbs describing an activity which affects an object (destroying, throwing, cutting, touching, burning, making etc); verbs of speech; verbs of mental activity (thinking, knowing, reading etc); verbs of bodily activity; verbs of sensory perception and stative verbs/adjectives'. The sample included around 100 lexical roots. For each of these I tested their ability to co-occur with the principal verbal affixes of Balinese: \{N\}, me-, pe-, reduplication (partial and full), the suffixes -in and -ang, as well as '-Ø' (no suffix), and various combinations of these affixes. I also tested the ability of some nominal roots, and various 'particles' to occur with these affixes. Much data was collected through elicitation, as well as from texts, oral and written. In this way, correlations between combinatorial possibilities and factors such as the semantic type, and the transitivity of each root, and of hundreds of derived lexemes, were explored.

The use of Dixon's methodology was a crucial factor that enabled many of the generalisations offered in this analysis to be subsequently uncovered.

The elicitation sessions were time consuming and not always interesting for my teachers. It was however important to get this data, since although there are several dictionaries of Balinese available (eg Panitia Penyusunan Kamus Bali 1990, Kersten 1984), none of them gives anything like a complete listing of the derivational
possibilities for any given lexical base: non-inclusion of a base-affix combination in the
dictionaries is not an indication that such a combination is not meaningful, or even not
in common use, at least in the Peguyangan variety of Balinese.

10.2 Zero derivation, bound morphemes, and free morphemes

In Balinese some morphemes occur unaffixed, as independent lexical items: batu 'stone
(n)', ingsap 'forget (vi)', inggih 'yes.HI'. Some of these appear to have more than one
syntactic function: uug 'in ruins (vi); demolish (vt)', gambar 'picture (n), make picture
(v)', pules 'sleep (vi), (n)'. These cases can be described as due either to a process of
zero derivation (or 'conversion') from one syntactic class to the other, or to polysemy.

Other morphemes are inherently bound: they never occur alone, but must 'surface'
concatenated with another morpheme. There are two broad types of bound morphemes
in Balinese, affixes and bound roots.

1. Affixes productively concatenate with a variety of bases to form lexemes. The
resultant lexemes typically belong to a single word class: an example is the prefix me-,
which always forms intransitive verbs, from a variety of bases:

(1)  
| keneh  | 'thought' |
| sepatu | 'shoe'    |
| -gedi  | 'leave (vi)' |
| -gaang | 'crawl (vi)' |
| jagur  | 'hit (UP)' |
| gorêng | 'fry (UP)' |
| me-keneh | 'think' |
| me-sepatu | 'wear shoes' |
| me-gedi | 'leave (vi)' |
| me-gaang | 'crawl (vi)' |
| me-jagur | 'be hit (stative)' |
| me-gorêng | 'fried (stative)' |

2. Bound roots can be assumed to carry lexical semantic content (see 10.4), for example
-tegak 'sit (vi)', -peju 'defecate (vi)' and -jagjag'approach (vi)'.

(2)  
| a. tegak  |
| negak (%) | 'sit (vi)' |
| negak-an | 'seat (n)' |
| negak-in | '(A) sit on U' |
| me-negak-an | 'be seated' |
| b. peju |
| meju (%) | 'defecate' |
| peju-peju | 'defecate uncontrollably' |
| peju-in | 'defecate on U' |
| c. jagjag |
| nyagjag (%) | 'approach (vi)' |
| jagjag-in | 'approach U (vt)' |
| jagjag-ang | 'make U approach (vt)' |
| jagjag-jagjag-an | 'way of approaching (n)' |

Lexemes then can be realised as:
(i) free morphemes (either simple morphemes, or containing complex formants, cf 5.0):

(3)  

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>batu</td>
<td>'stone (n)'</td>
</tr>
<tr>
<td>kupukupu</td>
<td>%REDkupu% 'butterfly'</td>
</tr>
<tr>
<td>inggih</td>
<td>'yes. Hi'</td>
</tr>
<tr>
<td>katibangbung</td>
<td>%katIRECVbung% 'k.o. beetle'</td>
</tr>
</tbody>
</table>

(ii) concatenations of two or more bound morphemes:

(4)  

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>negak</td>
<td>%N-tagak% 'sit'</td>
</tr>
<tr>
<td>srombot-an</td>
<td>'k.o. dish'</td>
</tr>
<tr>
<td>me-takon</td>
<td>'ask (vi)'</td>
</tr>
</tbody>
</table>

(iii) combinations of bound morphemes with those elsewhere occurring as free morphemes:

(5)  

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>jagur</td>
<td>'hit (vt, UP)'</td>
</tr>
<tr>
<td>jagur-jagur</td>
<td>(%RED-jagur%) 'repeatedly punch (vi)'</td>
</tr>
<tr>
<td>me-jagur-an</td>
<td>'be repeatedly punched (vi)'</td>
</tr>
</tbody>
</table>

10.3 Derivation and Inflection

The traditional distinction between a derivational and inflectional function of morphology is not an important, or very useful, one in Balinese. The affixes discussed in Part II of this thesis are often used as part of derivations creating a new lexeme from 'elsewhere'. Nevertheless, I argue in (15.2) that the primary function of affixation in Balinese is the signalling of semantic content inherent in lexemes, independently of whether that lexeme is derived or not. This function then is not described by either of the terms 'derivation' or 'inflection'.

Perhaps the only case of an inflectional process is the use of reduplication to mark plurality; this is classifiable as inflectional on purely semantic grounds. Reduplication of nouns (optionally) marks plurality, (6a), reduplication of intransitive stative verbs marks either plurality of the subject noun (6b), or durative aspect (6c):

(6)  

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>toris-toris</td>
<td>'RED-Westerner: Westerners'</td>
</tr>
<tr>
<td>anak luh-luh</td>
<td>'person RED-female: women'</td>
</tr>
<tr>
<td>sing teka-teka</td>
<td>'NEG RED-come'  [He] hasn't arrived! [over an extended period of time]'</td>
</tr>
</tbody>
</table>

10.4 Are bound morphemes 'precategorial'?

It is a common practice in Western Austronesian descriptive linguistics to refer to inherently bound lexical morphemes as 'precategorial' (see for example Prentice 1987:920, Kridalaksana 1990:36, Foley 1991:passim, Artawa 1992:13), implying that they don't acquire word category status until they undergo processes such as affixing. I
reject this approach, both for the description of Balinese and generally, for the reasons outlined in the following paragraphs.

First, I know of no rule which treats bound lexical morphemes differently from others. In their combinatorial possibilities with affixes, and the semantics of the resulting verbs, bound bases behave like free ones. For example, many bound morphemes show identical combinatorial behavior to free intransitive verbs. The bound bases in (7a) behave in a parallel fashion to the free intransitive bases in (7b), and quite differently from the nominal bases in (7c):

(7)  

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Derived Base</th>
<th>Derived Meaning</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>-tegak '?'sit, vi'</td>
<td>tegak-in</td>
<td>'sit on U'</td>
<td>['put a seat on U']</td>
</tr>
<tr>
<td></td>
<td>-siram '?bathe.HI, vi'</td>
<td>siram-in</td>
<td>'bathe in U'</td>
<td>['put a bath on U']</td>
</tr>
<tr>
<td>b.</td>
<td>pules 'sleep'</td>
<td>pules-in</td>
<td>'sleep on U'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ulung 'fall'</td>
<td>ulung-in</td>
<td>'fall on U'</td>
<td></td>
</tr>
<tr>
<td>c.</td>
<td>sepatu 'shoe'</td>
<td>sepatu-in</td>
<td>'put shoe on U'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pupur 'powder (Bl)'</td>
<td>pupur-in</td>
<td>'put powder on U'</td>
<td></td>
</tr>
</tbody>
</table>

When occurring with other verbal affixes, such as -ang (11.3), bound morphemes again show parallels to free intransitives (8b), and not to free transitives (8c) or nouns (8d):

(8)  

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Derived Base</th>
<th>Derived Meaning</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>-tegak 'sit, vi'</td>
<td>tegak-ang</td>
<td>'cause to sit'</td>
<td>'cause to bathe'</td>
</tr>
<tr>
<td></td>
<td>-siram 'bathe.HI'</td>
<td>siram-ang</td>
<td>'cause to bathe'</td>
<td>'cause to bathe'</td>
</tr>
<tr>
<td>b.</td>
<td>pules 'sleep'</td>
<td>pules-ang</td>
<td>'cause to sleep'</td>
<td>'cause to fall'</td>
</tr>
<tr>
<td></td>
<td>ulung 'fall'</td>
<td>ulung-ang</td>
<td>'cause to sleep'</td>
<td>'cause to fall'</td>
</tr>
<tr>
<td>c.</td>
<td>jagur 'punch'</td>
<td>jagur-ang</td>
<td>'hit (U patient) for someone else'</td>
<td></td>
</tr>
<tr>
<td>d.</td>
<td>sepatu 'shoe'</td>
<td>*sepatu-ang</td>
<td>(doesn't occur)</td>
<td>(doesn't occur)</td>
</tr>
<tr>
<td></td>
<td>pupur 'powder (Bl)'</td>
<td>*pupur-ang</td>
<td>(doesn't occur)</td>
<td>(doesn't occur)</td>
</tr>
</tbody>
</table>

The same remarks apply when they occur with the nominal affix -an:

(9)  

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Derived Base</th>
<th>Derived Meaning</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>pules 'sleep'</td>
<td>pules-an</td>
<td>'thing slept on'</td>
<td>'thing which falls'</td>
</tr>
<tr>
<td></td>
<td>ulung 'fall'</td>
<td>ulung-an</td>
<td>'thing slept on'</td>
<td>'thing which falls'</td>
</tr>
<tr>
<td>b.</td>
<td>-tegak 'sit, vi'</td>
<td>tegak-an</td>
<td>'thing sat on'</td>
<td>'place bathed at'</td>
</tr>
<tr>
<td></td>
<td>-siram 'bathe.HI, vi'</td>
<td>siram-an</td>
<td>'thing sat on'</td>
<td>'place bathed at'</td>
</tr>
<tr>
<td>c.</td>
<td>sepatu 'shoe'</td>
<td>*sepatu-an</td>
<td>(doesn't occur)</td>
<td>(doesn't occur)</td>
</tr>
<tr>
<td></td>
<td>pupur 'powder (Bl)'</td>
<td>*pupur-an</td>
<td>(doesn't occur)</td>
<td>(doesn't occur)</td>
</tr>
</tbody>
</table>

On the basis of these parallels in combinatorial behaviour, it is possible to conclude that -tegak and -siram, and many other verbs like them, are underlyingly intransitive verbs (as I have tentatively glossed them above), and not 'simply' precategorial. This analysis obviates the need to introduce a further analytical category and so should be preferred.
As well, assuming precategoriality introduces immense difficulties into any theory of semantics, and I would argue, should not be assumed in a linguistic description unless that description incorporates such a theory.\textsuperscript{1} Since the assumption that such bound morphemes are underlyingly assigned a syntactic class accounts for the data equally well, and does not require the introduction of the dubious category 'precategorial', in this description I refer simply to either \textit{bound} or \textit{free} morphemes. I assign individual bound morphemes glosses (and hence category status) based on such combinatorial behaviour: in fact, the majority of bound bases are underlyingly verbs - see 15.2.\textsuperscript{2} Further evidence for the position argued for here is given in 15.2, as well as evidence for a simple explanation for the existence of bound morphemes in Balinese: Balinese obligatorily signals the presence of NP's bearing the Actor and/or Undergoer semantic roles, via verb affixation; as a result, certain verbs never 'surface' without such an affix.

### 10.4.1 Inherent and derived transitivity values

Most monomorphemic verbs can be assumed to be either inherently semantically intransitive or transitive. The two cases are dealt with below. (I argue in 10.4 and 15.2 that some morphologically complex verbs also have an inherent transitivity value.) English has many verbs which can be used either intransitively or transitively: \textit{cook, eat, walk, march, sing, soak}, and so on. In Balinese this is not the case: nearly all monomorphemic verbs can be used either \textit{only} intransitively:

\begin{align*}
\text{(10)} \quad & \text{cai } \text{tusing } \text{lakar } \text{mati} \\
& \text{you } \text{NEG } \text{FUT } \text{die/dead} \\
& \text{You're not going to die}
\end{align*}

or \textit{only} transitively:

\begin{align*}
\text{(11)} \quad & \text{mem } \text{malu } \text{baas-é} \\
& \text{soak } \text{first } \text{rice-DEF} \\
& \text{First, soak the rice!}
\end{align*}

Derivational processes must apply to create new lexemes, If these root morphemes are to be used with a transitivity value opposite to that which they bear inherently (15.2).

\begin{align*}
\text{(12)} \quad & \text{i } \text{raga } \text{mati-ang=a nyanan} \\
& \text{[DET body]} \\
& \text{we } \text{die-APP=3 IRR} \\
& \text{We'll be killed by him! He'll kill us!}
\end{align*}

\begin{align*}
\text{(13)} \quad & \text{suud me-mem, ingsah} \\
& \text{finish me-soak wash} \\
& \text{After [it] has been soaked, wash [it]}
\end{align*}

\textsuperscript{1}Precategoriality would require that a lexical base have associated with it all necessary semantic information compatible with, say, both noun and verb status, and yet only receive final categorial information on affixing. This in turn would require a precategorial system of semantic units, one which human metalanguages would be incapable of reproducing. For arguments against this see Goddard (1994).

\textsuperscript{2}Examples of bound nominals are given in (87), 11.3.2.1.
Chapter 10

A very few monomorphemic verbs do occur with two argument frames, one intransitive, the other transitive. The examples in my data were: bah 'fall over (vi), fell (vt); belah 'break (vi, vt); eleg 'bent (vi), bend (vt); uug 'in ruins (vi), demolish (vt).

Monomorphemic intransitive verbs can be distinguished from monomorphemic UP transitives using the following criteria:

(i) enclitic pronominals do not occur on intransitives like pules 'sleep' (14a), whereas they mark the agent of UP verbs such as jagur 'hit' (14b):

(14)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| a | pules, ia | (*pules-a)  
  | sleep | 3  
  | He/she/they is/are asleep |
| b | jagur—a | (*jagur, ia)  
  | hit(UP) | 3  
  | (X) was hit by him/her/them. |

(ii) the negative imperatives of intransitives never carry the irrealis morpheme found on UP transitives in the same context:

(15)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| da | (Madé) | pules(*-a)  
  | NEG | (M.) | sleep-IRR |
  | Don't sleep, Madé |
| da | (Madé) | jagur*(-a)  
  | NEG | (M.) | hit-IRR |
  | Don't hit Madé |

and (iii) there is generally no {N} affixed alternate form of intransitives (16a, also 12.3), whereas there is always one for UP transitives (16b):

(16)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| a | teka | 'arrive'  
  | *neka  
  | mokoh | '(be) fat'  
  | *ngemokoh  
  | kayeh | 'bathe'  
  | *ngayeh |
| b | jagur | 'hit, UP'  
  | nyagur | 'hit, AP'  
  | daar | 'eat, UP'  
  | naar | 'eat, AP'  
  | baca | 'read, UP'  
  | maca | 'read, AP' |

---

3The one near-exception is the irregular verb laku 'go', which behaves formally as a transitive, but occurs only in UP forms, never with AP. The actor is usually an incorporated 3rd-person clitic, on a Bare Stem UP: kija laku-na 'to where did s/he go?'; ke-UP forms also occur: ke duur kop ke-laku 'to the bonnet (of the car) I went'. It appears that the PP bears the subject relation (cf Bresnan 1994).

4I have given this morpheme its usual gloss of 'irreals' (Clynes 1989); an alternative analysis is that this is the use of a Bare Stem UP, with =a a 3rd-person clitic actor, understood to have 2nd-person reference in an avoidance context, cf chapter 13, Appendix L).
11 Verbal Suffixes

11.0 The Transitive stem, 3 forms: '-∅, -in, -ang

The transitive stem has a three-way paradigmatic contrast in formal suffixing possibilities. It may be a simple monomorphemic root, with no suffix ('-∅):

(1) jagur 'punch'
daar 'eat'
beli 'buy'
dingeh 'hear'

or else it may consist of a root plus one of two suffixes, either -in, the 'Locative' suffix:

(2) tegak-in 'sit-LOC: sit on'
daar-in 'eat-LOC: eat a little of'
beli-nin 'buy-LOC: buy from'
kauk-in 'call-LOC, call to'
jagur-in 'punch-LOC; punch repeatedly'

or -ang, the 'Applicative':

(3) tegak-ang 'sit-APP: make sit down'
dingeh-ang 'listen-APP: listen to'
beli-ang 'buy-APP: buy for U, buy with U (U = money)'

Purely as a descriptive convenience, and following Pawley and Reid (1979), I refer to the first type as carrying '-∅, a 'zero suffix'. A given root may occur in combination with all three, two, or only one, of these suffixes to form transitive stems.

The suffix encodes purely semantic information: a further 3-way distinction in the semantics of the Undergoer participant of transitive stems. Broadly, the Undergoers of '-∅ verbs' are Patients, those of -in verbs are Locatives, and those of -ang verbs have 'Other Peripheral' semantic roles. The detailed patterns of suffixation, outlined in the following subsections, are of course more complex than this (see also 15.1).

Which roots occur with which suffix(es) is broadly predictable from the inherent semantics of a given root (see Pawley 1986 on a parallel situation in another Austronesian language, Fijian, also Wolff 1993 on Indonesian). The meaning of the resulting stem is also usually predictable. Paradigmatically the substitution of one suffix for another always implies a difference in the semantics of the verb - sometimes

---

1 The same broad 3-way opposition in suffixing possibilities and semantics was reconstructed 'for an early stage of AN' by Pawley and Reid (1979).
only this: dingeh 'hear', dingeh-ang 'listen to'. More typically, it also entails a syntactic difference, in the sense that lexemes formed from the same root morpheme with different suffixes are typically subcategorised for different Undergoer NP's, bearing different semantic roles. Compare for example tegak-in 'sit on U' (U is a Location, A is Sitter), with tegak-ang 'make U sit down' (U is a Causee/Sitter, the Location (if mentioned, is in an oblique prepositional phrase):

(4) a. nyén negak-in gedebong belus lakar=a jit-ne
   who N-tegak-in banana.stem wet will-IRR backside-3
   Whoever sits on a banana stem will get a wet backside (proverb)

b. i lutung [...] tegak-ang=a di duur kayu-né
   DET monkey sit-APP=3 LOC top wood-DEF
   The monkey was sat down by him on the wood
   [From 'Satua I Lutung Ian I Kekua', by drs Ida Kadé Suarioka, Banjar, broadcast on RRI, March 1991]

The majority of verbs bearing suffixes -in and -ang can be assumed to derive from 'elsewhere': either from -Ø transitive or intransitive verbs, or from nominal roots. To that extent, -Ø verbs could be thought to be the only inherently transitive verbs. This however requires the assumption that the function of a verbal suffix is to signal derivation from another word. I argue in 15.2 that the primary function of these suffixes is, rather, the obligatory signalling of the semantic (micro-)role of the Undergoer. Evidence for this is the fact that inherently transitive, but suffix-bearing, verbs do occur (15.2).

11.1 Monomorphemic ('-Ø suffixed') transitive stems

The vast majority of suffixless transitive verb stems ('ROOT-Ø verbs') can be considered to be inherently transitive: while they function directly as transitive verbs, to function in other word categories they must undergo further derivational processes. A small number could be products of 'zero derivation' (though there is no clear evidence for the direction of derivation):

1. A small number of -Ø verbs have a homophonous noun counterpart, including:
   gambar 'make picture of (vt), picture (n)'; bui 'prison (n), imprison (vt)'; krangkeng 'cage (n), put in cage (vt); bedil 'rifle; shoot (at)'; kapak 'axe, axe (vt)'; tumbak 'spear; spear (vt)'.

2. An equally small number have a homophonous stative intransitive verb counterpart. These include:

(5) bah 'fall over (vi); fell (vt)'
   eleg 'bent; bend (vt)'
   léb 'run free; set free'

   belah 'break (vi; vt)'
   uug 'in ruins (vi); demolish'
   nyag 'be obliterated; obliterate'
Semantically, \textsc{root-\O} verbs fall into a number of related classes. In virtually all cases 'canonical transitive verbs' (Andrews 1985, Givon 1990) are of this shape. These are verbs where the undergoer NP is a true Patient: a thing or person which is created, destroyed, profoundly changed, or totally overpowered by the action of the verb:

(6)  
\begin{tabular}{lll}
\text{uug} & 'demolish' & \text{úék} & 'tear, rip' \\
\text{lilig} & 'crush, run over' & \text{bełah} & 'smash' \\
\text{nyag} & 'melt, crush' & \text{getep} & 'cut' \\
\text{pelut} & 'peel' & \text{jait} & 'sew' \\
\end{tabular}

This class includes verbs of hitting and beating,

(7)  
\begin{tabular}{lll}
\text{daldal} & 'pound flat' & \text{tigit} & 'beat with stick' \\
\text{jagur} & 'punch' & \text{toktok} & 'rap, hammer' \\
\end{tabular}

of cooking (\text{U} is the thing cooked) and burning,

(8)  
\begin{tabular}{lll}
\text{lablab} & 'boil' & \text{goréng} & 'fry' \\
\text{kuskus} & 'steam' & \text{jakar} & 'steam (rice)' \\
\text{eñijit} & 'set alight' & \text{tuñijel} & 'burn' \\
\end{tabular}

and of eating:

(9)  
\begin{tabular}{lll}
\text{daar} & 'eat (HI)' & \text{amah} & 'eat (LOW)' \\
\text{solo} & 'swallow whole (liquid)' & \text{ajeng} & 'eat (HON)' \\
\text{uluh} & 'swallow whole (solid)' & & \\
\end{tabular}

It also includes verbs where the Undergoer NP is created by the action of the verb:

(10)  
\begin{tabular}{lll}
\text{gaé} & 'make' & \text{karya} & 'make (HI)' \\
\text{tulis} & 'write' & \text{gambar} & 'make a picture of' \\
\end{tabular}

This class also includes \text{sambat} 'utter', where the \text{U} NP is the word(s) uttered and thus 'created', and \text{baca} 'read, recite' eg \text{maca mantra} 'recite/read a mantra'.

The \textsc{root-\O} class also includes verbs with an Undergoer which is attacked, or is considered to have been 'assaulted' in some other way, either physically or verbally:

(11)  
\begin{tabular}{lll}
\text{sander} & 'swoop on (from sky)' & \text{sagrep} & 'leap on prey' \\
\text{cotot} & 'peck; bite (of snake)' & \text{cegut} & 'bite' \\
\text{temah} & 'swear at' & \text{batbat} & 'insult angrily' \\
\end{tabular}

More generally, in virtually all cases, the Undergoer of a \textsc{root-\O} verb in Balinese is viewed as a powerless, passive entity, manipulated at will. I assume that the Undergoers of all such verbs are interpreted as Patients in Balinese. (The one or two exceptions found in my corpus are discussed below).

This includes cases where the 'fabric' of the Undergoer NP is not physically changed by action of the verb, such as many verbs of lifting, holding and carrying, including:
Chapter 11

Similarly, verbs of pushing and pulling are ROOT–∅ verbs:

(13) sogok  'push'
dongsook  'elbow (vt)
kedeng   'pull'
paid     'drag'

as are certain verbs where the U is opened or closed,

(14) gagah  'open (eg package)'
tekup    'cover'
empug    'open (eg coconut)'
sengseng 'seal (eg bottle)'

and verbs where the U is of folded or woven:

(15) tampih 'fold (eg clothes)'
tabu     'weave (fibre)'
lilit     'wind'
plintir  'plait'

Also in this class are verbs where possession of the Undergoer is transferred (permanently or otherwise) via a business transaction or agreement - buying, selling, renting, borrowing and so on:

(16) adep  'sell'
séwa     'rent'
silh     'borrow'
beli     'buy'
juang    'take possession of'
idih     'get by asking for'

With '∅ transitive stems, the Actor NP is characteristically agentive: animate, usually human, and typically acting deliberately. In fact probably most ROOT–∅ verbs only occur with human Actors (or, in stories, creatures with human-like behaviour). However certain verbs do allow inanimate, typically generic, actors. These include those involving actions like crushing, or smashing by a large moving object: lilig 'crush (A is eg car, train)', gebug 'punch with downwards action; drench (A is rain)', tabrak 'run into, collide (A is eg car). I propose that all the verbs discussed so far in this section share the meaning put forward by Wierzbicka (1992:22) as 'the prototypical transitive scenario': 2

(17) A (N-[VERB]-∅ U

someone (A) did something (N-[VERB]-∅)
because of this,
something happened to all of something (U)
in the same place, at the same time
this person (A) wanted this

---

2I have added the words 'all of' to Wierzbicka's formula, to make it explicit that the Undergoer is viewed as fully affected.
The prototypical agentive Actor is thus volitional, human, and performs an aspectually punctual action, which fully affects a prototypically inanimate Undergoer/Patient. The main class of exceptions found to this proposed meaning, and thus to the generalisation that the Undergoer of a zero verb is viewed as a passive, totally manipulated entity, are certain perception verbs, verbs where the Undergoer is perceived by the senses:\(^3\)

\[(18)\]  
\[
\begin{array}{ll}
\text{dingeh} & \text{'hear'} \\
\text{not} & \text{'see'} \\
\text{adek} & \text{'smell'} \\
\text{asa, rasa} & \text{'feel, taste'}
\end{array}
\]

Here the undergoer is not affected in any way by the action of the verb, and the Actor is low in agentivity - an experiencer, often unintentionally so. Paradoxically, as in English, where these actions are specified as performed deliberately, Balinese marks the Undergoer as a non-Patient, \text{not-in} 'look at', \text{tingal-in} 'look at', \text{dingeh-ang} 'listen to', \text{asanang} 'taste', \text{adek-in} 'sniff at'. I propose that the unsuffixed verbs have the following semantic formula:

\[(19)\]  
\[
\text{A (N-\{VERB\}-\O)} \text{U} \\
\text{someone (A) did something (N-\{VERB\}-\O) because of that people can think that something happened to something (U), in the same place, at the same time}
\]

This subclass then lacks the volitional component found in (17), and the Undergoer is only 'like', rather than prototypically, a Patient.

The two other potential exceptions found in the data were \text{baang} 'give' and \text{usud} 'rub'. With \text{baang} 'give' we would expect the Theme to be the Undergoer, parallelling the behaviour of other verbs of transferred possession (see examples (16): in fact it is the Recipient/Possessor which is assigned the Undergoer role. An explanation of the exceptionality of \text{baang} is offered at footnote 32. The exceptional nature of \text{usud} is perhaps simply an artefact of the English gloss used: this is the appropriate verb to use when one person touches another in an 'invasive' way, eg one which could be viewed as sexual assault.

### 11.2 Verbs with suffix -in

The 'locative' suffix -\text{in} occurs both on syntactically transitive and on intransitive (11.2.2.2) verbs. Intransitive verbs are always of invariant shape %N-\text{ROOT}-\text{in}\%. The suffix has two main realisations: /\text{in}/ and /\text{nin}/. The realisation /\text{in}/ occurs after both consonant-final and vowel-final stems:

\[^3\text{Dixon (1979) claims that in virtually all languages the verbs 'hear' and 'see' are transitive. He does not indicate whether this is true of verbs expressing the other senses.}\]
Chapter 11

(20) kalah-in 'leave-LOC, leave U'
tegak-in 'sit-LOC, sit on'
bakti-in 'devotion-LOC, show devotion to'

However /nin/ occurs only after vowel-final roots:

(21) lali-nin 'play-LOC, pay a call on U'
beli-nin 'buy-LOC, buy from U'
margi-nin 'walk, HI-LOC, go via U (HI)'

Which form, /nin/ or /in/, occurs after vowel-final roots is lexically determined: some roots occur with either form - eg saré(n)in 'sleep on', bakti(n)in, margi(n)in, lali(n)in - others with only one: basa-*{n)in 'language-in: translate', basa-{*n)in 'spice-in: add spice to', peju-{n)in 'defecate on'.

Another sequence which could be viewed (in a classical morphemic description) as an allomorph of -in, and is not mentioned as such in other analyses, is -an (/an/ ~ /nan/). An alternative, and I believe more satisfactory, account of the relationship between -in and -an is given in 12.4.1. This suffix is found on detransitive stems:

(22) tingting-in '(A) lift U a little' me-[tingting-an] 'U be lifted a little'
tunjel-in '(A) scorch U' me-[tunjel-an] 'U be scorched'

A very small number of /al-final roots appear to undergo a 'sandi' process when affixed with -in: sengkala 'disaster', sengkalen 'waylay, trap U', beficana 'disaster' beficanén 'ambush, trap'. However this is no longer a synchronic process: the root ica 'happy.HON' is thus the historical source of icén 'grant (vt)', however icén optionally occurs with -in, icén-in, indicating that it is synchronically analysed as monomorphemic. Equally, sengkalén and beficanén are literary terms, not common in everyday speech.

Most transitive verbs of shape ROOT-in can be considered to derive either from verbs or nouns. The root morpheme can be divided into two types: free, and bound. The free morphemes occur respectively as either (i) transitive (-~) verbs:

(23) jagur 'punch' jagur-in 'punch U PL'
tulis 'write' tulis-in 'write on U'
adep 'sell' adep-in 'sell to U'

(ii) monomorphemic intransitive verbs:

(24) pules 'sleep' pules-in 'sleep on U'
demen 'happy' demen-in 'like U'
paak 'close, near' paak-in 'approach U'

or (iii) nouns:
Bound roots never occur without further affixes. Some transitive verbs bearing -in I take to be inherently transitive, with the suffix performing a purely signalling function of indicating the semantic nature of the Undergoer: only partly affected, and conceptualised as a location. These are further discussed in 15.2):

(26)  usu-usu-in 'lightly stroke (eg a child before going to sleep)'
tingal-in 'look at U; see U'
oot-in 'remove load from (or place load on) U's head'
ubet-in 'open [eg door] to (U is person)'
ulap-in 'wave to U'

Many other bound morphemes which occur with -in (eg tegak-in 'sit on U') also occur as (prefix-bearing) active intransitive verbs (eg N-tegak 'sit'). I assume that the intransitives are the source of the transitive verbs, ie that derivation is clearly directional in this case (10.4). The NP subcategorised as bearing the Actor role with the transitive verb is also Actor/subject of the intransitive verb. The intransitives carry either of the realisations of the \( \{N\} \) the Agentive Prefix, /N/- (12.2):

(27)  N-tegak 'sit'
      N-peju 'defecate'
      N-wél 'rant'
tegak-in 'sit on U'
      pejú-nin 'defecate on U'
      wél-in 'rant at, rebuke U'

or me- (12.4.3):

(28)  me-gedi 'leave (vi)'
      me-kecog 'jump'
      me-gaang 'crawl'
gedi-nin 'leave U (vt)'
      kecog-in 'jump over U'
      gaang-in 'crawl along, crawl to'

The meaning of transitives with -in is usually predictable, depending on the nature of the root morpheme, though as with derivational processes cross-linguistically, the meaning of individual derived lexemes can be idiosyncratic. In the following paragraphs I discuss in turn the meanings where verbs derive from (i) nominals, (ii) intransitive verbs and (iii) transitive verbs. (As indicated above, bound roots taking -in which are also inherently transitive verbs are discussed at 15.2)

11.2.1 Verbs with suffix -in, root is a noun

Where the root is a noun I propose that the meaning of the majority of derived \([\text{NOUN}]\)-in' verbs is (following Wierzbicka 1988):

(25)  sepatu 'shoe'
      kutilt 'skin'
      bakti 'devotion'
      raos 'speech (HI)'
  sepatu-in 'put shoes on U'
      kutilt-in 'take skin off U';
      bakti-nin 'show devotion to U'
      raos-in 'speak to U'
(29)  A [NOUN]-in U

A did something
because of this,
something happened to [NOUN] at the place U,
something happened to U because of this

The Undergoer then is conceived of as a 'location' which is affected by something
manipulated by the Actor. This is so even where U is animate. U is thus affected in the
way that locations are: for example, as a stationary surface, a source or a goal (according
to the semantics of the individual verb and its NP arguments):

(30)  udeng 'headcloth'    udeng-in    'put udeng on U'
    eificeh 'urine'       eificeh-in    'urinate on'
    kulit 'skin'         kulit-in      'put skin on/take skin from U'
    basang 'stomach'     basang-in     'take stomach from U, gut'
    atur 'speech (HON)'  atur-in       'speak to (HON)'
    omong 'speech'       omong-in      'speak to, chat up'

Another minor semantic type is found where the nominal root is (roughly) a human
'relational term'. The core meaning is, I propose:

(31)  A [NOUN]-in U

A did something (like [NOUN]s do with some people)
because of this,
people can think
something happened to U

This accounts for derivations such as:

(32)  musuh 'enemy'     musuh-in    'behave like enemy towards U'
    muani 'male, husband' muani-nin   'escort (U = woman)'
    timpal 'friend, associate' timpal-in   'accompany'

11.2.2 Verbs with suffix -in, root is an intransitive verb

I distinguish two principal derivations where the stem consists of an intransitive verb
bearing the affix -in. The first of these, described in 11.2.2.1, derives syntactically
transitive verbs with a locative undergoer; I assume these include one subclass where the
undergoer is a 'goal of emotion', conceptualised as a location. The second derivational
type, described in 11.2.2.2, produces causative verbs of emotion (the undergoer being
causated to experience the emotion). According to the individual root lexeme, the derived
verbs may be transitive (thus having both AP and UP forms), eg sakt-in 'A
(deliberately) cause pain to U' or intransitive (ngedegin 'A be annoying; A annoy people
in general').
11.2.2.1 Verbs of form INTRANS-in: undergoer is locative

The meaning of the derived -in verb can, in most cases, be explicated as follows (cf. Wierzbicka (1988)):

(33)  \[ X (N-) [INTRANS VERB]-in U \]

'\(X\) [INTRANS VERB]) happened, at [the place] \(U\) because of this, people can think that something happened to \(U\) '

The Undergoer NP then has a locative semantic role, which varies with the verb root, but includes 'stationary location', 'surface', 'goal', 'source', or 'path'. The main exceptions to this generalisation (both described below) are (a) a subset of intransitive roots which derive verbs with a causative meaning when affixed with -in, and (b) a subset of transitives which retain a true Patient Undergoer.

Where the transitive verb derives from an intransitive verb, the subject of the intransitive verb becomes the Actor of the transitive; the (obligatory) Undergoer NP of the transitive is an NP, occurring in an (optional) oblique prepositional phrase with the corresponding intransitive:

(34)  a. umah-né celep-in = maling
        house-3 enter-LOC = burglar
        His house was entered by burglars

  b. ada maling me-celep (ka umah-né)
     be burglar me-enter (to house-3)
     A burglar entered (his house)

(35)  a. sok-é pedem-in = kuluk
        basket-DEF sleep-LOC = dog
        The basket was slept on by a dog

  b. ada kuluk medem di sok-é
     be dog N-sleep LOC sok-é
     A dog slept on the basket

Although they are not obligatory, for a given intransitive verb certain locative PP's often collocate with it. The meaning of most transitive -in verbs deriving from intransitive verbs is then predictable in a straightforward way from that of the root plus collocating locative PP.4 Here are some more examples, from the three formal intransitive classes, together with prepositions they optionally collocate with (including di 'at/on/in', uli 'from', ring 'High style preposition', ke 'to', tekén 'with'):

4 An alternative analysis would thus be that -in transitives, and others, are syntactically derived, the suffix representing an incorporated preposition (Baker 1988). This analysis does not account for the great majority of verbs bearing -in and -ang, since they do not have a 'root-Ø plus PP paraphrase'. The latter include inherently transitive, but suffix-bearing verbs (15.2).
With other such verbs I could not identify such a collocating PP.\(^5\)

Some transitive verbs appear to derive from intransitive verbs of emotion in a different way: the Undergoer is the goal of the emotion, while the Actor is the experiencer of the emotion (and again subject of the intransitive):  

\[(40)\] toris-é takut-in = a tekén i kadék  
\(\text{European-DEF afraid-LOC} = 3 \text{ with DET K.}\)  
\(\text{The European is feared by Kadek} \)  
\(i \text{ kadék takut tekén toris-é}\)  
\(\text{DET K. afraid with European-DEF}\)  
\(\text{Kadek is afraid of the European}\)  

As in this example, there is a parallel with the first derivation in that in the corresponding intransitive construction the goal of the emotion typically occurs in a prepositional phrase headed by tekén'with'. Other examples include:  

\[(41)\] gedeg-in  
\(\text{'(A) feel angry towards U'}\)  
demen-in  
\(\text{'(A) feel happy-LOC; (A) like U'}\)  
lek-in  
\(\text{'(A) be shy towards U'}\)

I propose this semantic formula, which closely parallels that in (33):

\[(42)\] A (N-)[VERB OF EMOTION]-in U  
A [VERB OF EMOTION] happened,  
thinking of U,  
because of this  
people can think that something happened to U

\(^5\)The bracketed ungrammatical PPs represent ones which could be logically expected; I could elicit neither those, nor any other PPs as occurring with these verbs.

\(^6\)The derivational possibilities with intransitive verbs of emotion are complex, and vary from verb to verb. Some only form transitives with -in, others only with -ang (11.3.2.3), others occur with both. See also 11.2.2.2 on 'causative intransitives' with -in.
The similarity in semantic formula indicates that this derivation is a variant of that proposed in (33). This assumes that the component 'thinking of U' can be replaced by 'at [the place] U'; this would require that goals of emotion be conceptualised 'as though' they were locations. The element 'not because of anything else' of the proposed formula expresses the fact that the feeling felt by A may not necessarily be due to something done by the Undergoer (cf Wierzbicka (1988:239)).

As well as participating in this derivation, the root gedeg 'angry' and one or two others also derive syntactically *intransitive* verbs, discussed in 11.2.2.2, cf example (45).

Some roots (such as gedeg, dot) occur readily with both derivations, while others are apparently restricted either to the intransitive pattern (probably the majority of roots, including seneb, iseng, med 'bored', genit 'itchy'), or to the syntactically transitive derivation (lek, and perhaps demen). Further work needs to be done to determine the precise patterning of individual lexical items.

11.2.2.2 Verbs of form *intrans*-in: undergoer is a 'causee/experiencer'

In a second derivation where -in affixes to intransitive roots, semantic transitives with a *causative* meaning are produced. The intransitive root typically expresses a feeling, either physical or emotional. Its subject NP, the experiencer of the state, generally corresponds to the U of the derived transitive; the A of the derived verb is the causer. This derivation has two realisations, depending on the lexical semantics of the root verb: (i) in a smaller number of cases, syntactically transitive verbs (ie, with both AP and UP forms) are derived; (ii) in other cases, syntactically intransitive verbs of invariant shape N-[root]-in are derived.\(^7\)

1. Where the derived verbs are syntactically transitive, they have initiating, volitional actors, and are aspectually punctual; the undergoers are animate, typically people (all features associated with 'high transitivity' by Hopper & Thompson 1980):

```
(43)  sakit  'be in pain (U is body part)'  sakit-in  'torture (U is person)'
  enten  'wake from sleep'           enten-in  'remind U (eg of debt)'
  lēb    'loose, free'              lēb-in    'set U loose'
```

The meaning of this class is, I propose:

```
(44)  X (N-[VERB(Feeling)])-in U
      X did something
      because of this,
      (U [VERB(Feeling)]) happened,
      in the same place, at the same time
      Something happened to U because of this
      not because of anything else\(^8\)
```

\(^7\)Kersten's (1984:64) discussion of this latter subclass incorrectly implies that they are syntactically transitive.

\(^8\)Wierzbicka (1992) uses this component to indicate that X alone (and not for example the undergoer) initiates the activity.
Chapter 11

2. Examples of syntactic *intransitives* derived in this way include:

(45) gedeg  
    ngedegin  
    seneb  
    nyenebin  
    dot  
    ngedotin  
    iseng  
    ngisengin  
    lengeh  
    nglengehin  

'be angry'  
'cause people to feel angry, be annoying'  
'feel revulsion'  
'cause people to feel revulsion, be revolting'  
'feel desire'  
'cause people to feel desire, be desirable'  
'feel longing'  
'cause people to feel longing'  
'intoxicated'  
'cause people to be intoxicated, be intoxicating'

(46) ng-iseng-in gati kenyir-n-né  
     N-long-LOC very smile-GEN-3  

*His/her smile fills one with longing*

As these examples show, these syntactically intransitive verbs may have non-volitional, even inanimate actors; they differ too in that they are aspectually stative/durative. The undergoer is typically understood to be 'people in general'. Given the absence of other variables, it must be lexical semantic features like these which 'block' these verbs from 'fully transitive' syntactic contexts. I propose this variant semantic formula:

(47) X (N-)[VERB(FEELING)]-in

X does something  
because of this,  
(people [VERB(FEELING)]) happens,  
People can think that something happens to people because of this  
not because of anything else

An agentive component 'X does something' is nonetheless present in the semantics of these verbs: if someone is *ngedegin* 'annoying', they are in some way to blame for this. This I assume is the reason these verbs carry {N}, the agentive prefix: inanimate actors, such as intoxicating drinks (*nglengehin*), or captivating smiles (*ngisengin*, example (46)) are conceived of as initiating actors by a process of metaphorical extension, where those objects are perceived to have some special 'potency' about them. The (undifferentiated hence unexpressed) Undergoer of the derived verb is either 'people in general', or, for some verbs, a body part (see below). Verbs derived in this way are thus semantically transitive, encoding both an Actor ('someone/something') and an Undergoer in their semantic formula.

Moreover, both Actor and Undergoer are represented also in the verbal morphology, by {N} and -in respectively; however only the Actor is expressed as an NP and mapped to a grammatical function (subject/pivot).
Verbal suffixes

(48) S N-[FEEL]-in <Actor>
    Subject

An explication of the meaning, following Wierzbicka (1992), is:

(49) S N-[VERB(FEELING)]-in

S does something
because of this
people feel [VERB(FEELING)]
not because of anything else

Roots which derive intransitives in this way seem restricted semantically to those
relating to 'strong feelings': forms such as *ngiapin (kiap, 'sleepy'), *ngwaregin (wareg
'satiated (hunger)' do not occur. With roots such as these a periphrastic construction is
used instead: ngaé kiap 'make sleepy', ngaé seduk 'make hungry'.

Being syntactically intransitive, verbs formed via this derivation cannot occur with an
object NP.9

(50) *I Made ngedegin tiang/anak
    DET M N-anger/angry-LOC 1exc/people
    *Made annoys me/people

and similarly there exists no corresponding UP clause (though see footnote 9):

(51) *tiang gedeg-in-a
    1exc angry-LOC = 3
    *I was angered by him/her

As well, like other syntactic intransitives (but not transitives), these verbs can occur in
with 'adverbials of degree':

(52) I Made ngedegin gati
    DET M N-anger/angry-LOC very
    Madé is very annoying/irritating

They may also occur in the 'superlative construction', bearing the irrealis clitic, also not
found on transitives:

(53) ia ané paling ngedegin-a
    3 REL most annoying-IRR
    He is the most annoying

Some verbs such as ngedegin and nyenebin may occur with the incorporated noun
undergoer basang 'stomach [=seat of (certain) emotions]':

9While example (50) is unacceptable with the semantics there, it is grammatical, if stilted, when
glossed as Made feels anger towards me/people (cf 11.2.2.1). In elicitation verbs like (41) are
preferred in UP form; I only obtained examples like (50) with the verb in AP form in elicitation.
Chapter 11

(54) ngedeg-ngedeg-in basang(*-é) déén, jíema-né
   RED-N-ge-deg-in stomach only person-DEF
   That person does nothing but annoy (lit. 'annoy stomach')

Clearly the clause remains syntactically intransitive: the adjunct basang in (54) cannot be marked as semantically definite, and no corresponding Undergoer Pivot construction is found:

(55) *basang-é gedeg-in-a
   stomach-DEF anger-LOC = 3
   *(The/My) stomach is angered by him/her

The corresponding intransitive verb can optionally occur with basang, (or occasionally ati 'liver', keneh 'thought') as the subject NP; human subjects can occupy the same slot:

(56) gedeg gati basang-é
    angry very stomach-DEF
    'The [=my] stomach was very angry'

(57) gedeg gati i beli
    angry very DET elder.brother
    Elder brother was very angry!

These verbs can thus attain a kind of 'halfway' state intermediate between syntactic transitivity and intransitivity: both actor and undergoer may be represented in the syntax, but the undergoer (which can only be an inanimate body part, in contrast to the animate undergoers of the fully transitive verbs in (43) cannot have core NP status (bear object or subject relation).

An unusual feature of both the above derivations is their causative meaning. Their historical source may be that still found with many stative intransitive verbs, which occur with an optional pe- causative prefix, plus -in, to derive transitive verbs such as pe-gede-nin 'enlarge', pe-enggal-in 'speed up, do quickly'. These have a semantic formula identical to that in (44). The causative prefix is often dropped, indicating that it is in the process of being lost. This could explain the occurrence of '-in causatives' like sakt-in (though, if they did in the past, these apparently cannot now co-occur with pe-).

11.2.3 Verbs with suffix -in, root is a transitive verb

Where the root is a transitive verb, the meaning of the derived verb is again broadly predictable; as with nominal and intransitive verb roots, the Undergoer NP is in most cases classifiable as a Locative rather than a true Patient. For one subclass of verbs ('verbs of violence'), however, the Undergoer remains classifiable as a Patient: the affixing of -in instead adding a repetitive aspectual meaning to the verb (see below).\(^\text{10}\)

\(^{10}\)As well, some descriptions, eg Beratha 1992, ascribe the role of 'beneficiary' to the undergoer of ampak-in 'open door to U'. Pragmatically, this will be the normal context of use of this verb. The Undergoer however has the normal locative role. This can be seen from the counterpart verb ubet-in
Verbal suffixes

There are three distinct derivations. From the point of view of argument structure these can be classified, according to whether the Undergoer of the -in transitive is the same as that of the -∅ transitive root (two related derivations), or not the same (a single derivation). These are discussed in turn in the following paragraphs.

The first subclass is where the Undergoer of the -in transitive is not the same as that of the -∅ root. As with verbs derived from intransitive roots (above), the ROOT-∅ verb may collocate with an optional locative PP (example 58a), the NP of which becomes the Undergoer NP of the corresponding -in derived transitive (58b). The Undergoer's precise semantic role includes source (58b), location (59b), and other locative roles:

(58) a. i ketut meli baas sik lbu Resi
    N-beli
    DET K. AP-buy rice at.place.of l. R.
    Ketut buys rice at Ibu Resi's

b. i ketut meli-nin lbu Resi baas
    N-beli-nin
    DET K. AP-buy-LOC l. R. rice
    Ketut buys rice from Ibu Resi

(59) a. dija ke-jang gula-né?
    at.where PASS-put sugar-DEF
    Where was the sugar put (by you)?

b. kopi-né jang-in = a gula
    coffee-DEF put-LOC = 3 sugar
    The coffee had sugar put in it by him/her

Other examples include:

(60) adep 'A sell U (=theme)' adep-in 'A sell.to U (=buyer)'
tulis 'A write U (=text)' tulis-in 'A write.on U (=loc)'
duduk 'A pick up U (=theme)' duduk-in 'A levy.from U (=payer)'
jakan 'A cook U (=theme [rice])' jakan-in 'A cook.in U'

As the examples show, in verbs of this subclass, the Undergoer NP of the -∅ 'source' verb, while encoded as a Patient, is pragmatically also classifiable as a Theme (in the sense that it undergoes a change of place/owner). This NP becomes an adjunct of the -in transitive (and so, for example, cannot be marked for definiteness). Only -∅ verbs which have an Undergoer NP semantically classifiable in this way derive ROOT-in transitives of this type.

I postulate the following core meaning for these verbs:

'close the door to U', normally used of closing the door to deliberately prevent someone who wants to come in from doing so; clearly the reverse pragmatics. The derived verb ampak-ang (see 11.3.3, type 1(a) derivations) has a general benefactive sense. There is in fact no example of -in subcategorisig a beneficiary Undergoer.
Chapter 11

(61) \(X (N-)\text{[TRANSITIVE VERB]}-\text{in} \ U \ Y\)

'\((X \ N-\text{[TRANSITIVE VERB]} \ Y) \ \text{happened, at [the place]} \ U \ \text{because of this, something happened to} \ U\)'

The semantic formula (61) is almost identical to that of (most) verbs derived from intransitive roots (repeated here from (33):

(62) \(X (N-)\text{[INTRANS VERB]}-\text{in} \ U\)

'\((X \text{[INTRANS VERB]} \text{happened, at [the place]} \ U \ \text{because of this, [people can think that] something happened to} \ U)\)'

One can conclude that these are in fact variants of a single derivation:

(63) '\((X \text{[PREDICATE]} \text{happened, at [the place]} \ U \ \text{because of this, [people can think that] something happened to} \ U)\)'

The second derivational subclass, in terms of argument structure, is where the Actor and Undergoer of the derived \text{ROOT-in} transitive are the same as those associated with the -0 root. Undergoers of these \text{in} verbs are either (i) less affected than their -0 counterparts (and can be considered to have the locative semantic role of 'surface'), or else (ii) they are fully affected semantic Patients but undergo the action of -0 repeatedly. Examples of type 2(i) include:

(64) goréng 'fry' goréng-in 'lightly fry'
goréng 'fry' goréng-in 'lightly fry'
pecik 'squeeze' pecik-in 'lightly squeeze'
ulat 'weave' ulat-in 'weave more onto'
pelet 'peel' pelet-in 'peel a bit more off'
saup 'scoop up' saup-in 'scoop off surface'
jait 'make by sewing' jait-in 'repair by sewing'
adek 'sniff' adek-in 'sniff at'
gaé 'make' gaé-nin 'work on (eg field)'

A related meaning with some verbs, in some contexts, is 'do (v) a little':

(65) gagah 'open (eg package)' gagah-in 'open a little (eg to inspect), then close'
tingting 'lift up' tingting-in 'raise a bit, try to lift'
tektek 'chop up' tektek-in 'chop up a bit more'

Where the Undergoer NP consists of a group of objects, and only some of those objects are affected by the action of the verb, the group as a whole remaining intact, the -\text{in} stem is also appropriate:
Verbal suffixes

(66) pilih 'choose U' pilih-in 'choose from U'
juang '1. take (possession of) U, 
 '2. marry U' juang-in 'take (poss. of), from among U'
abut 'pull out, pluck U' abut-in 'thin out U by pulling out parts of U'
bah 'fell U' bah-in 'thin out U by felling'

(67) bes atep punya-n nyuh-é bah-in a-bedik too close tree-GEN coconut-DEF fell-in ONE-little
The coconut trees are too close together, thin them out a bit by felling

(68) mahasiswa-n-né juang-in =a
studentBl-GEN-3 take-LOC = 3
He took [=married] [one] from amongst his students

In all cases with this subclass, the Undergoer NP is only partially affected. I propose that all the verbs in sub-class 2(i) share a common meaning:

(69) X (N-) [TRANSITIVE VERB]-in U

'X did something ([TRANS VERB]), because of this, something happened to part of U, in the same place, at the same time'
X wanted this

This contrasts with derivation type 2(ii), now described. With these verbs the Actor and Undergoer of the derived ROOT-in transitive are again the same as those associated with the -Ø-root. However in this case the derived ROOT-in verbs typically continue to have Undergoers which are fully affected semantic Patients (though they too may be only partially affected). These verbs also differ in that the action of the -Ø verb is always carried out repeatedly. In this case then, the affix -in is primarily associated with an aspectual difference, rather than with a difference in the semantic role of the Undergoer NP:

(70) jagur 'punch' jagur-in 'punch repeatedly'
uug 'demolish' uug-in 'demolish (PL)'
temah 'vociferously insult' temah-in 'id. repeatedly'
tiňjak 'kick' tiňjak-in 'repeatedly kick'
sander 'swoop on' sander-in 'swoop on, PL'
lempag 'hit with long implement' lempag-in 'id. repeatedly'

As can be seen from these examples, all verbs in this last derivation are what can be broadly categorised as 'verbs of violence'. Many of these do not occur with the previous derivation.\(^{12}\)

---

\(^{11}\) uug-in umah-né! {~ house-3'} implies mass action, where one house is completely destroyed by many people; uug umah-né! on the other hand could be a command to a single person.
\(^{12}\) In elicitation I found only one or two clear examples of verbs occurring with both derivations, types 2(i) (69) and 2(ii): empug-in 'open (eg beer) bottle, then replace lid'; 'split open (coconut)', godot-in 'cut off part of (eg to shorten a length of cloth); cut up (eg cassava)'.

223
Chapter 11

Unlike verbs formed by derivation (69), verbs like jagur-in, uug-in, lempag-in cannot be used to describe a context where the action is performed only once to the (partially affected) Undergoer. Meanings like jagur-in 'hit out at, once (but not connect)', bedil-in 'shoot at, once (but not hit)' are not found: my teachers said that the ROOT-verbs (plus the appropriate qualifying statements) would still be used in these contexts. On the other hand, such verbs can occur with a partially affected Undergoer reading, as long as the action is repeated in rapid succession. The following is thus perfectly acceptable: 13

(71) jagur-in = a, kala sing kena
      punch-LOC = 3 but NEG hit
      [I] was hit at repeatedly by him, but [his blows] did not connect

In short, the iterative aspectual meaning is basic with this third ROOT-in derivation, while the affectedness of the Undergoer is not apparently marked: the undergoer may be fully affected, or not affected at all.

The question remains as to why 'iterative/repetitive aspect' should be lexicalised when -in affixes to the ROOT-verbs of violence (VV) subclass, and with no other roots, nominal or verbal. Is this aspectual reading simply an inexplicable semantic quirk? With all other '-in derivations', there is a common locative semantic element, apparently missing here.

The aspectual meaning can be seen to follow logically from both the inherent semantics of ROOT-verbs, and of the VV subclass in particular, and the pragmatics of their use. Specifically these two verb types involve the following (normally incompatible) semantic elements:

1. Verbs of violence, in both ROOT-vv- and ROOT-vv-in forms:
   A highly volitional Actor, intending to fully affect the Undergoer.

2. Other ROOT-verbs, when occurring with -in:
   An Actor intending to only partially affect the Undergoer
   (eg gorèng 'fry', gorèng-in 'lightly fry', and examples (64) - (66)).

Given then a context where an Actor wished to, say, jagur U (='punch U [affecting him completely]'), but did not connect, element (1) above in the semantics of the verb would normally lead the Actor to repeat the (intended) action, 'jagur'. The ineffective performance of an action normally describable by a ROOT-verb of violence thus involves two pragmatic elements:

(A) a partially affected Undergoer, semantically more like an affected Surface than a Patient. This is found with type 2(i) derivations, viz. 'something happened to part of U', (69) above, and

13 Verbs of cooking such as lablab-in (re)heat by boiling', gorèng-in 'lightly fry' appear also to combine both 'repetitive' and 'partially affect' meanings, since they are typically used when reheating already cooked foods. These verbs are however used when partially cooking raw ingredients - one can assume therefore that only the 'partially affect' meaning is leicalised, and that the repetitive aspect is here purely part of the pragmatics of use.
Verbal suffixes

(B) a resulting high likelihood that the action will be repeated.

I propose that an extension of this particular context of use has led to the iterative aspectual reading associated with 'ROOT-vv-in' verbs. In most instances where a violent act is likely to be repeated, I suggest, the Actor is in a state such that his wanting to affect U is not satisfied by a single performance of the action. This can be described as due to a perception of the initial action as ineffective, or a non-perception of its effectiveness.

The explanation proposed here is thus that the pragmatics of use of verbs of violence provides a natural source of the locative affix, in that the Undergoer is likely to be perceived by the Actor as only partly affected. The contextual aspectual element, 'repeated action' also derives naturally from the same context.

The affixing of -in can also be seen as registering a departure from prototypical transitivity (15.3), and from the semantics of ROOT-verbs of violence, where the initial action is not effectively carried out, or perceived to be. This marked departure from the semantics of ROOT-verbs of violence thus provides a second motivation for marking of the verb with -in. (I am thus proposing that the presence of verbal suffixes reflects not simply the semantics of the Undergoer of the verb, but also marked departures from the overall (semantically) transitive scenario.)

I propose that these various factors are sufficient to motivate the subsequent extension of the ROOT-vv-in derivation to all cases where ROOT-verbs of violence are performed repeatedly, including where the Actor acts deliberately and 'cool-headedly'.

Moreover the restriction of the 'repetitive' aspectual meaning to 'verbs of violence' only can also be seen to follow from the semantics of these verbs, and of other ROOT-verbs which cannot undergo this derivation. The action is not merely repeated, but must be performed by the same Actor (or group of Actors), on the same Undergoer, within a single event. Other ROOT-verbs, such as those of financial transaction like beli 'buy' and adep 'sell' can't be performed repeatedly in the same way: for example, once something is sold, it can only be sold again by a different Actor (the person who has just acquired it), in a distinct event. Similarly with verbs of creation like gae, 'make', tulis 'write': something can only be created once (in a given event), hence these also cannot derive ROOT-in forms with the repetitive aspectual meaning. Verbs of violence are similarly excluded from type 2(i) derivations, (69), for semantic reasons.
11.3 Verbs with suffix -ang

This is the third major transitive stem class. First the variant realizations of -ang are discussed, then the different root classes cooccurring with -ang are described, and the meanings produced in each case.

11.3.1 Realisations of -ang

The suffix -ang has three primary realisations, /ao/, /nay/, /yan/: I will use -ang as a shorthand for all of these. As well, on detransitive stems, -an (/av/ ~ /anv/) replaces the primary form.

Which primary realisation occurs with a given root is conditioned partly by phonological factors - /nao/ and /yan/ only occur after vowel-final roots (though /ao/ often occurs in identical phonological environments) - and partly by lexical and stylistic ones:

1. /ao/ is the unmarked form, occurring in all styles, after roots of all phonological shapes, both vowel-final:

<table>
<thead>
<tr>
<th>(72)</th>
<th>abā-ang</th>
<th>'carry for U'</th>
<th>beli-ang</th>
<th>'buy for U'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gedē-ang</td>
<td>'make U bigger'</td>
<td>pesu-ang</td>
<td>'make U come out'</td>
</tr>
<tr>
<td></td>
<td>kōto-ang</td>
<td>'do like that'</td>
<td>sēda-ang</td>
<td>'kill U (U is HON)'</td>
</tr>
</tbody>
</table>

and consonant-final:

<table>
<thead>
<tr>
<th>(73)</th>
<th>ajēng-ang</th>
<th>'eat H1'</th>
<th>tumbas-ang</th>
<th>'buy for (H1)'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>gedēng-ang</td>
<td>'hate U'</td>
<td>cager-ang</td>
<td>'rely on U'</td>
</tr>
<tr>
<td></td>
<td>kituk-ang</td>
<td>'move from side to side'</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>eñco-ang</td>
<td>'make U urinate'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. /nay/ occurs only after (certain) vowel-final roots, in both low and high styles:

<table>
<thead>
<tr>
<th>(74)</th>
<th>mai-nay</th>
<th>'make come here'</th>
<th>nge-liu-nay</th>
<th>'increase in numbers'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>karya-nay</td>
<td>'HI make for U'</td>
<td>kemo-nay</td>
<td>'move U to there'</td>
</tr>
<tr>
<td></td>
<td>gedē-nay</td>
<td>'enlarge (vt)'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

That the occurrence of -/nay/ is only partly phonologically conditioned, and partly lexically determined, is shown by contrasts such as peju-*(n)ay 'defecate (vt, U is excreted substance)', which informants only accepted with suffix -nang, versus pesu-*(n)ay 'make go out' where only -ang occurs. Compare also mai-*(n)ay 'make come here' and cai-*(n)ay 'address as cai'; gaē-*(n)ay 'make for U', and karya-*(n)ay 'id. HI', but abā-*(n)ay 'carry for', kēto-*(n)ay 'do that way'. Still other roots occur with either -ang or -nay: gedē-ang ~ gedē-nay 'make bigger'.

---

14The combination of prefix %N%-intransitive vb. root-ang derives various types of syntactically intransitive verb. These I take to be semantically transitive, see 11.3.2.1.
3. The form /yao/ has the most limited distribution, occurring only after certain /al/-final root morphemes, and only in formal (literary and high styles):

(75) mresidayang %pra-sida-an% 'be able to, HI'
ng-andika-yang 'AP-speak, HON'

Some roots occur with either -ang or -yang: séda-ang ~ séda-yang 'kill (U is of high status)', nyidaang (N-sida-ang, 'be able to', everyday form) nyidayang ('be able to, mid-style'). Others possibly only occur with one of these (eg N-presida-yang 'be able to, HI style').

An alternant realisation of -ang, not mentioned as such in other analyses, is as -an (/arv ~ /nan/). This is found on detransitive stems (12.4.1):

(76) tulis-ang 'k(A) write with U me-[tulis-an] 'be written with'
omong-ang 'k(A) discuss U' m-[omong-an] 'be discussed'

The distinction between -in and -ang is neutralised in the detransitive form, both being realised as -an.

11.3.2 Verbs with suffix -ang, root is an intransitive verb

Where the root verb is an intransitive verb, affixing with -ang derives semantically (and, usually, syntactically) transitive verbs with various meanings, depending on the semantic type of the root verb. I have found three main sub-classes, described in turn in the following subsections.

11.3.2.1 Verbs of form INTRANS-ang, undergoer is a causee

In combination with -ang most intransitive verbs, including verbs of physical states and bodily actions, derive semantic transitives with a meaning roughly glossable as 'A cause undergoer to (vi)'. This derivation produces both syntactically transitive, and syntactically intransitive verbs, according to verbal semantics and the degree to which the eventuality described departs from the prototypical transitive script (15.3). I discuss first syntactically transitives, and then derived intransitives. (The same phenomenon was found with intransitive roots in combination with the locative affix -in (11.2.2.2).)

Examples of derived syntactic transitives include:

(77) ada 'exist' ada-ang 'create, organise'
mati '(be) dead' mati-ang 'kill'
gedé '(be) big' gedé-nang 'make big'
me-jalan 'walk' jalan-ang 'make walk'
pules 'sleep' pules-ang 'make sleep'
mai 'come here' mai-nang 'make come here'
N-endih 'burn (vi)' endih-ang 'set alight'
Chapter 11

(78) tiang sai ling-ang-a
    1 often cry-ang = 3
    He often made me cry

As in the case of derivations with -in (11.2), bound roots taking -ang also occur as active intransitive verbs (cf me-jalan, ngeling, ngendih and meju above), the latter can be assumed to be the source of the transitive verbs, which follow the regular semantic patterns of transitives deriving from intransitives. This is also seen in the second derivational sub-class, discussed in 11.3.2.2.

I assume that in the case of the following verbs, which occur (unsuffixed) in both an intransitive and a transitive sense, the -ang form derives again from the intransitive:

(79) FORM INTRANSITIVE TRANSITIVE

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>uék</td>
<td>'torn'</td>
<td>'tear'</td>
</tr>
<tr>
<td>uug</td>
<td>'destroyed'</td>
<td>'destroy, demolish'</td>
</tr>
<tr>
<td>belah</td>
<td>'shattered'</td>
<td>'shatter'</td>
</tr>
<tr>
<td>bah</td>
<td>'fall lengthways'</td>
<td>'fell'</td>
</tr>
<tr>
<td>elung</td>
<td>'broken off'</td>
<td>'break off'</td>
</tr>
</tbody>
</table>

The derived (-ang) form of these verbs is generally associated with lower agentivity than the bare intransitive: thus uék (transitive) is only appropriate where a fully volitional Agent deliberately tears something; uék-ang is compatible with a non-intentional Actor - say, if a small child accidentally tore clothes while playing. Again, where the 'agent' is an earthquake or flood, the ang form of the verb is strongly preferred.:

(80) umah-ne uug??(-ang) = linuh/blabar
house-3 destroy earthquake/flood
His house was destroyed by an earthquake/flood

For all derived verbs like those above (in (77) and those derived from intransitives in (79), I assume the following common semantic formula:

(81) X (N-)[INTRANS]-ang U

X did something,
because of this
U [INTRANS]) came about
Something happened to U because of this

For this syntactically transitive subclass the subject NP of the intransitive verb (whether it itself is an Actor or an Undergoer) always has the Undergoer role with the derived verb. The only partial exception to this is in reflexive uses, where the reflexive pronoun (often omitted in imperatives) is Undergoer: jalan-ang (bané)! 'take yourself off!',

---

15With peju-nang and enco-ang 'make urinate', the Actor is typically an adult, who accompanies U, typically a child.
16It is possible that in Denpasar dialect uug would be acceptable in this context.
Verbal suffixes

**pules-ang (bané)!** 'get yourself to sleep!'. In all other cases, the Actor/`causer` of the derived transitive is not part of the semantics of, or an argument of, the intransitive root.

The rest of this section discusses the derivation of syntactically *intransitive* verbs. These involve essentially the same derivation as above, though with various derivational subtypes. In all cases an inchoative process is under way, which the subject of the verb is seen as having initiated, and is at the same time undergoing. I propose this formula, a variant of that in (81):

(82)  \[ S \text{ N-INTRANS-ang} \]

S is doing something  
because of this  
S is becoming [INTRANS]  
Something is happening to S because of this

Verbs undergoing this derivation are aspectually stative, either inherently so, or compatible with such an interpretation in a given context. The actor and the undergoer are moreover the same entity. Each of these features constitutes a marked departure from prototypical semantic transitivity; the combination guarantees that the clause will be syntactically intransitive - there is no syntactic expression of the undergoer.\(^\text{17}\) Verbs undergoing the intransitive derivation include:

1. **intransitive roots with adjective-like meanings, and an often inanimate subject.** The combination of \{N\} and -ang thus derives verbs with the meaning, 'become more (root)':

(83)  \begin{align*}
    \text{baat} & \quad \text{`heavy'} \\
    \text{gedé} & \quad \text{`big'} \\
    \text{mokoh} & \quad \text{`fat'} \\
    \text{waas} & \quad \text{`healed (wound)'} \\
    \text{maatang} & \quad \text{`become heavier'} \\
    \text{ngedénang} & \quad \text{`become bigger'} \\
    \text{ngemokohang} & \quad \text{`become fatter'} \\
    \text{ngwaasang} & \quad \text{`become more healed'}
\end{align*}

\text{pantes maatang} \quad \text{Bali-né}  
N-baat-ang \quad \text{B.-DEF}  
\text{fitting N-heavy-\text{ang}}  
\text{So that's why Bali's getting heavier}\(^\text{18}\)

2. **In a few cases, \{N\}...-ang occurs with a stem which is a verb of motion:**

(84)  \begin{align*}
    \text{luas} & \quad \text{`travel a long distance away'} \\
    \text{ngluasang} & \quad \text{`be in the process of travelling (etc)'} \\
    \text{mai} & \quad \text{`come here'} \\
    \text{ngemainang} & \quad \text{`be in the process of coming here'}
\end{align*}

One or two of these have a complex stem, itself containing \{N\}:

\(^{17}\)This situation differs significantly from reflexive clauses, where the undergoer is realised by a distinct reflexive pronoun; in this case, Haiman argues (ms., cited in Wierzbicka 1992), 'the grammar iconically signals the recognition of not one but two participants, and thus implies some kind of detachment from the self'. In Balinese, such clauses are syntactically transitive, see discussion of jalan-ang (bané)! above.

\(^{18}\)Said on seeing the big crowd that had turned up to watch a parade.
(85) -penek 'go up' ngemenékang %N-[N-penek-ang]%
tuun 'go down' ngenuunang %N-[N-tuun-ang]%

me-garang saling pemaalunin ngenuunang, pengaluné ajaka buah-é
me-fight RECIP go.before N-N-go.down-ang peddler-DEF with pinang-DEF
(it was as though) the peddler and the pinang fruit fought to go first as they
descended20

Since prepositional phrases can function independently as 'intransitive' predicates in Balinese (9.2):

(86) i Madé ke Gianyar
DET M. to G.
Madé has gone to Gianyar

they can therefore participate in this derivation, again deriving syntactically intransitive
verbs with the meaning: 'S be going to (nominal)'

(87) ke sisi 'to side' ngesisiang 'be going to one side'
ke tengah 'to inside' ngetengahang 'be going inside'
ke luang 'to upstream' ngeluunang 'be going upstream'
ke mbo 'go.there' ngemonang 'be going there'
k-angin 'to east' nganginang 'be going to the east'21
k-aia 'to inland' ngaianang 'be going inland'
k-lod 'to the sea' nglodang 'be going seawards'
m-ulih 'to home' ngemulihang 'be going home'22

ke jaba 'to outside palace/temple
njejabaang 'be going out of the palace/temple'

The above PP stems have as their base inherently directional/deictic nouns. As such,
these names can be said to be generic, low in referentiality. Non-deictic nouns cannot
occur with this derivation:23

(88) ke peken 'go to the market' *ngepekenang
ke uma 'go to the ricefields' *ngumaang
ke Badung 'go to Badung (district)' *ngebubung
As the glosses in example (88) indicate, PP's with non-bound nouns are ambiguous between the meanings 'to/towards (noun)' and (when used predicatively) 'go to/towards (noun)'. This includes the PP's which occur with \{N\} ... -ang such as ke sisì, ke tengah, ke luan (cf (87) and also ke puri, ke jaba. The semantic difference between a pair such as:

\begin{verbatim}
(89) a. i Madé ke puri
   DET M. to palace

b. i Madé ngepuriang
   DET M. N-to-palace-ang
\end{verbatim}

is then that (89a) is interpreted as a completed action in the unmarked case, whereas (89b) is always interpreted as ongoing/incomplete, 'on the way to the temple'. When occurring in noun phrases derived from these verbs, the 'progressive' sense is again the only acceptable one:

\begin{verbatim}
(90) di ngemulihang-né buin duduk =a bedég-é
    LOC N-m-ulih-ang-3⁴ again pick.up = 3 woven.bamboo-DEF
    On his way home (*at home), again he picked up the sheet of woven bamboo.

\end{verbatim}

To sum up the distinction between the transitive and intransitive variants of this derivation, the two differ semantically in the following respects: (i) The Actor of the transitive construction is potentially, perhaps typically, volitional, whereas that of the intransitive construction is often inanimate (perceived, I assume, as somehow initiating the event); (ii) the undergoer of the transitive construction is typically not the same as the actor; that of the intransitive construction is always the same (assuming that S in (82) can be viewed as having both semantic roles) (iii) The intransitive derivation only applies to predicates which are aspectually stative, whereas the transitive derivation applies also to aspectually punctual intransitive verbs.

These purely semantic differences indicate that underlyingly a single derivation is involved: where the event predicated is very low in features ascribed to prototypical transitive predicates cross-linguistically, (15.3), then the meaning (and associated intransitive syntactic frame) of (82) surfaces, as a variant of (81).

11.3.2.2 Verbs of form INTRANS-ang, undergoer is a 'body product'

Where the intransitive root is a 'body product' verb, (including verbs of excretion, of speech and dreaming), the understood product of the activity becomes the Undergoer of a derived transitive; the subject of the intransitive is also the Actor of the transitive:

\footnote{On this gloss, see footnote 22.}
If our ears buzz, [it means] there is someone talking about us

I propose the following semantic formula for this derivation:

\[
\text{A (N-)[INTRANS]-ang B} \\
\text{A does N-[INTRANS],} \\
\text{because of this} \\
\text{there is B} \\
\text{at the same place at the same time} \\
\text{People can think that something happens to B} \\
\text{People can think of B as part of A}
\]

Some verbs in this class also derive transitives with the causative meaning described above, 'make U (vi)'. These include \text{peju-nang} and \text{eñico-ang}. Others, like the verbs of speech do not appear to derive causatives.

11.3.2.3 Verbs of form INTRANS-ang, undergoer is 'goal of feeling'

The third subclass is where the intransitive root verb is a \text{verb of feeling}, either emotion or physical sensation. With certain intransitive roots of feeling/emotion the derived transitive means roughly 'A feel (vi) because of U'. The Undergoer of the transitive is in some way the cause (and the goal) of the feeling, but not necessarily because of anything that U has done, while the Actor of the transitive corresponds to the subject of the intransitive:

\[
\text{gedeg} \quad \text{dot} \quad \text{kéweh} \quad \text{gedeg-ang} \quad \text{dot-ang} \quad \text{kéweh-ang} \quad 'A be angry with U' \quad 'A desire U' \quad 'A feel downhearted about U'
\]

I propose the following semantic formula:

\[
\text{A (N-)[INTRANS]-ang B} \\
\text{A feels [INTRANS]} \\
\text{because of B} \\
\text{not because of anything else} \\
\text{People can think that something happens to B because of this}
\]

(Other intransitive verbs of feeling form transitives by affixing -\text{in} (11.2), and at least one forms them by both means: \text{gedeg-ang} 'be angry with U', \text{gedeg-in} 'hate U'.)
4. A closely related class is derived from intransitive verbs of bodily sensation, such as sakit 'sore, painful', genit 'itchy', dingin 'cold', where the intransitive subject is a body part (or the whole body), not the possessor of that body part. In the derived transitives the body part is Undergoer, and is semantically a Location. The possessor of the body (part) is now encoded, as Actor/Experiencer. The derived meaning is 'A feel (vi) in/at U':

(96) a. i madé, sakit basang-né
DET M. sore stomach-3
(As for) Made [topic], his stomach is aching

b. sakit-ang = a basang-né tekén i madé
sore-ang = 3 stomach-3 with DET M.
Made has a stomach ache

I propose this semantic formula:

(97) A (N-) [INTRANS]-ang B

A feels this: B [INTRANS]
B is a part of A
People can think that something happens to B because of this

A restriction is that in the AP form, as distinct from the UP (cf (96b)), the noun expressing the body part must be generic/unmarked for definiteness or possessor, and is phonologically a clitic to the verb.

(98) i madé nyakitang = basang(*-né)
N-sakit-ang
Made has a stomach ache

This incorporated body part can never be omitted: the verb is functionally intransitive. However in the UP form, where the Undergoer is the whole body it is more felicitous to omit it entirely, as in the following, said looking at horses in a field on a cold Australian winter's day:

(99) sing dingin-ang = a tekén jaran-é?
NEG cold-ang = 3 with horse-DEF
Aren't the horses cold?

Although some emotions have their 'seat' in body parts, they cannot derive transitives in the above way:
**Chapter 11**

(100) **gedeg gati basang-né**

* **gedeg-ang = a basang-né**

*His stomach was very angry

**demen ati-né**

happy liver-DEF

* **demen-ang = tiang ati-né**

Similarly neither verbs of emotion nor verbs of bodily sensation can derive causative transitives, as do apparently all other intransitives (subclass 1 above):

(101) **sakit-ang = a ia tekén musoh-né**

sore-ang = 3 3 with enemy-3

*He was hurt/tortured by his enemy

The form **sakit-in** is needed here, U = victim/location, A = human, see derivation (44).

### 11.3.3 Verbs with suffix -ang, root is a transitive verb

In combination with transitive verb roots (henceforth [TRANS]), -ang derives other transitive verbs, with variously two or three NP arguments. The meaning and syntactic behaviour of the derived verb can be broadly predicted from the semantics of the [TRANS] verb.

In all cases, the Actor of the derived verb is the same as that of [TRANS].

The Undergoer of the derived verb may be either 1) the same as that of [TRANS], or else 2) a different, 'promoted', NP. These two cases are now discussed in turn.

1. Undergoer of [TRANS] and of the derived verb are identical. In this case the argument structures and syntactic behaviour of [TRANS] and of the derived verb are the same. They differ semantically, however, either in the degree of agentivity of the Actor NP (types l(a), l(c) below), or in the degree of affectedness of the U NP (type l(b)). Derived verbs of this first type fall then into three clearly distinct, semantically-based groups.

(l(a) In most cases the Actor of the derived verb is less agentive than that of the [TRANS] root. While the Actor of [TRANS] can be considered to be the sole initiator of the event, that of the derived verb is understood to carry out the action for someone else, who is not represented in the verb's argument structure, but nevertheless represents a kind of

25 This could only mean 'he hates his stomach'.

26 This is not the case where the derived verb also carries the pe- 'CAUS' prefix; these derivations are not discussed in this thesis due to space limitations.

234
secondary initiator or causer of the event. This is a very productive derivation type. Examples include:

(102)  
\begin{itemize}
  \item \textit{cét} \quad 'paint U'
  \item \textit{getep} \quad 'cut off U'
  \item \textit{ampak} \quad 'open U'
\end{itemize}

\begin{itemize}
  \item \textit{cét-ang} \quad 'paint U [for someone]'
  \item \textit{getep-ang} \quad 'cut off U [for someone]'
  \item \textit{ampak-ang} \quad 'open U [for someone]'\textsuperscript{27}
\end{itemize}

I assume that these derived verbs have a core meaning such as (cf Wierzbicka 1988):

(103) \begin{align*}
A (N-)[\text{TRANS-∅}] &- \text{ang } U \\
\text{Someone (A) did ([TRANS-∅])} \\
\text{because that person thought another person would want this} \\
\text{Because of this,} \\
\text{something happened to something (U) in the same place, at the same time}
\end{align*}

These verbs cannot be used in a 'generic' sense, 'because A thought other people [in general] wanted this'. The identity of the beneficiary must be inferable from the immediate context. Frequently it is arrived at from pragmatic knowledge: commonly (for example in requests and orders) it is the speaker. The following 'minimal pairs' illustrate nicely the added semantic element carried by the derived verb:

(104) \begin{align*}
a. \quad \text{getep ja bok-é!} \\
\text{Cut [your] hair!}
\end{align*}

\begin{align*}
b. \quad \text{getep-ang ja bok-é!} \\
\text{Cut [my] hair!}\textsuperscript{28}
\end{align*}

(105) \begin{align*}
a. \quad \text{témbo-k-tiang-é labur = a} \\
\text{wall-1-DEF whitewash-ang = 3} \\
\text{My wall was whitewashed by him [I may or may not have wanted this]}
\end{align*}

\begin{align*}
b. \quad \text{témbo-k-tiang-é labur-ang = a} \\
\text{wall-1-DEF whitewash-ang = 3} \\
\text{My wall was whitewashed by him [and I wanted this]}
\end{align*}

Example (105) also illustrates how, where the Undergoer NP of the derived verb contains an overt possessor NP, that possessor is the understood beneficiary.

1(b). As well as type 1(a) derivations, some verbs allow a second derivation, where the Undergoer of the derived verb is conceptualised as a Theme which moves, or is restrained from movement. The undergoer of the derived verb is again identical to the

---

\textsuperscript{27}Beratha (1992) cites the sentence \begin{itemize}
  \item \textit{i belog ng-ampak-ang mémé-n-né jianan}
\end{itemize}

\begin{itemize}
  \item DET B. AP-open-APP mother-GEN-3 door
\end{itemize}

This usage of \textit{ampak-ang} with a benefactee Undergoer is rejected as ungrammatical by Peguyangan speakers.

\textsuperscript{28}The meaning contrast in examples (104a & b) derives from the use of the \textit{DEF}initeness\textsuperscript{n} morpheme to avoid both 1st- and 2nd-person pronominals, particularly where possession of body parts and other 'inalienables' is expressed. The pragmatic rules indicating how the identity of possessor is determined have not to my knowledge been described. In (104a) the understood possessor can only be 2nd person; in (78b), only 1st person. See also Appendix L.
undergoer of the root verb (which is, however, a true Patient, rather than a Theme). Such verbs include:  

(106) tegul  
ampak  
sorog  

'tie up'  
'open (eg a door)'  
'push (eg, a car)'  

lilit  
ungkab  
dongsok  

'wind'  
'lift up (a cover)'  
'shove'  

The \textsc{trans} root of verbs of this type has a 'canonical Patient' Undergoer, and so this is the verb which is used when the affectedness of (eg an animate, resisting) patient is stressed:

(107) tegul = a maling-é  
tie.up = 3 thief-DEF  

He tied up the thief

Using the derived verb is more appropriate with less affected Undergoers, including undergoers thought of primarily as changing in their position or state:

(108) sembé-né tegul-ang = a di tampul-é  
oil.lamp-DEF tie-ang = 3 LOC post-DEF  

She tied the lamp onto the post

There is no benefactive semantic element with this derivation. I propose this semantic formula:

(109) A (N-)[trans-\emptyset]-ang U  

Someone (A) did ([\textsc{trans-\emptyset}])  
Because A wanted something (U) to be in a place  
something happened to U because of this in the same place, at the same time

l(c). A third, minor, class, again where the Undergoer of the derived and root verbs remains the same, consists of certain verbs of perception, including:

(110) dingeh  
piragi  
dingo-ang  
piragi-ang  

'hear U'  
'hear.DFR'  
'listen to U'  
'listen to.DFR'

Here, in contrast to l(a) and l(b) derivations, the Actor of the derived -ang verb is more agentive than that of \textsc{trans}. With these verbs the affixing of -ang adds the meaning element 'do [\textsc{trans}] deliberately' (cf 11.1):

\footnote{I have assumed that where verbs have a homophonous intransitive and transitive form, such as \textit{uug} '1. be in ruins; 2. demolish' (see example (79) and discussion), the -ang form derives from the intransitive. The analysis of Type 1(b) would perhaps need to be changed if it could be shown that such verbs in fact derive from the transitive.}

\footnote{Alternatively, a formula closer to that deriving causative verbs from intransitive roots (see (81)) might be preferable, such as:
A N-[\textsc{trans}]-ang B  
A did something/because of this (B [\textsc{me-de}\textsc{trans}]) happened}
(111) A (N-[VERB]-ang U

someone (A) did something (N-[VERB]-Ω)
because of that, people can think that something happened to something (U),
in the same place, at the same time
A wanted this

2. The second broad derivational class is where a different, 'promoted', NP has the Undergoer role, the Undergoer of [TRANS] now having adjunct (non-core) status:

(112) beli

'buy (U = thing bought)'

beli-ang

a. 'buy (U = recipient of purchase)'
b. 'buy (U = Theme/instrument of purchase)'

The semantic role of the derived Undergoer NP is thus of two types, discussed in turn below.
2(a) The Undergoer is an Animate (generally human) Recipient:

(113) putu beli-ang salak tekén nini-n-né

DET P. buy-ang =salak. fruit with grandmother-GEN-3

Putu was bought some salaks by her grandmother

Previous descriptions describe the NP in this derivation as a Beneficiary (Fakultas Sastra 1979, Kersten 1984, Beratha 1992:177); I argue below that this is not part of the core meaning.31

2(b) The Undergoer is a Theme (transferred to an immobile Goal, syntactically an adjunct). Previous writers describe the Undergoer as an Instrument; again, I argue below that this does not represent the core meaning.

(114) puntul dapak-é, sing dadi tektek-ang bé

blunt chopper-DEF NEG CAN chop-ang meat

The chopper is blunt, [it] can't be [used to] chop meat

The meaning of derived verbs of the first type, 2(a), involves the notion of transfer of possession. Transitive verbs undergoing 2(a) derivation have, in their underived form, an Undergoer which is encoded as a Patient, but pragmatically is also a Theme:

(115) [TRANS] DERIVED VERB

<table>
<thead>
<tr>
<th>aba</th>
<th>'carry'</th>
<th>aba-ang</th>
<th>'carry.to U'</th>
</tr>
</thead>
<tbody>
<tr>
<td>bakta</td>
<td>'carry.HI'</td>
<td>bakta-ang</td>
<td>'carry.to U'</td>
</tr>
<tr>
<td>kirim</td>
<td>'send [U = theme]'</td>
<td>kirim-ang</td>
<td>'send.to U'</td>
</tr>
</tbody>
</table>

31In a similar derivation in Kimaragang (Austronesian, Sabah) 'benefactive undergoers must always be recipients in some sense' Kroeger (1991b:10).
32Kirim is borrowed from Bl, though Bl kirim-Ω and kirim-kan (corresponding to BB kirim-ang) have different syntactic frames. The semantic roles of the U NP are mirror-images:

<table>
<thead>
<tr>
<th>Bl</th>
<th>BB</th>
</tr>
</thead>
<tbody>
<tr>
<td>kirim</td>
<td>U = RECIP</td>
</tr>
<tr>
<td>kirim-kan/ang</td>
<td>U = THEME</td>
</tr>
<tr>
<td></td>
<td>U = RECIP</td>
</tr>
</tbody>
</table>
The Undergoer of [TRANS] is moreover something normally moved from one place to another, or capable of being so moved. (This is so even with the verb beli 'buy'; immovable things like land can be bought and sold, but in 99% of cases, particularly in traditional Balinese society, bought things are virtually all capable of being carried (or led) home).

This last quality distinguishes verbs potentially taking 2(a) derivation from those undergoing type 1(a), such as ampak 'open (eg door)', ungkab 'lift (covering)': these latter may have Undergoers which move, but are not usually transferred from place to place, and there is moreover no transfer of possession. Only where such a 'movable Theme' is present can verbs of Type 2(a) be derived. This also explains why many verbs form Type 1(a) 'pragmatic benefactives' but do not derive Type 2(a) verbs, cf example (105b), repeated here:

(116) témbo-k-tiáng-e la-bur-ang = a
    wall-1-DEF whitewash-ang = 3
    My wall was whitewashed by him [and I wanted this]
    *tiáng labur-ang = a témbo
    *I was whitewashed a wall by him

Similarly, an illiterate person might say (using a Type 1(a) derivation):

(117) baca-ang = a surat-tiáng-e
    read-APP = 3 letter-1-DEF
    He read my letter (for me) (lit. My letter was read by him [for me])

But the beneficiary of this action could not be 'promoted' to Undergoer, in a Type 2(a) derivation:

(118) *tiáng baca-ang = a surat
    1 read-APP = 3 letter
    *I was read a letter by him

The Undergoer of a 2(a) derived verb is then a Recipient, who at the same time acquires possession of the Theme (expressed as a clitic adjunct):

(119) i kadék cegut-ang = a poh teh-kén i nanik
    DET K. bite-ang = 3 manggo with DET N.
    Kadek had [a piece off] manggo bitten off [and given to] him by Nanik.

I propose the following semantic formula:

It seems likely too that baang 'give [U = recipient]' comes from an earlier *beri-ang, via this same derivation; synchronically it is now most simply considered to be monomorphemic.

(120) tiang baang(-in)-a pipis kéń i Madě
    1 give(-LOCJ-3 money with DET M.
    I was given money by M.

If the locative affix is added, one could appropriately ask whose money it was, since the speaker may have been given the money simply to hold; if the '-Ø' form of the verb is used, the recipient is understood to become the possessor, and the enquiry is not appropriate.
Verbal suffixes

(120) A (N-[[TRANS-0]]-ang U Y

Someone (A) did something (TRANS-0)
because of this, Y was where someone (U) was
because of this, people can think that Y became U’s Y.
people can think that something happened to U
in the same place, at the same time

Example (119) also shows how some root verbs describe actions which, when performed, create an object capable of being transferred; such verbs can thus undergo Type 2(a) derivations.

Where there is no transfer of possession, there can be no 2(a) derivation, even though there may be a transferred Theme involved. Thus verbs like gisi 'hold' and abin 'hold in one's lap' form Type 1(a) derivations:

(121) gisi-ang ja panak-tiang-é
hold-APP EMPH child-1-DEF

Please hold my child (for me).

but not Type 2(a):

(122) *tiang gisi-ang =a panak
1 hold-APP =3 child

*My child was held for me (lit. I was held for a child.)

Type 2(a) derived verbs are often described as having a Beneficiary Undergoer (eg Kersten 1984, Beratha 1992). Pragmatically, Animate Recipients are usually also beneficiaries, and this pragmatic element is thus often associated with these verbs, but this element is not a part of the inherent meaning. The primarily Recipient status of the Undergoer is suggested by example (123), where the food must be understood to be a gift brought to mémé (Undergoer), not merely something taken out of her hands, say because she was tired of carrying it:

(123) ené, cucu-né ng-aba-ang mémé ajeng-an
this grandchild-DEF AP-carry-ang mother eat.HON-NML

Here, (your) grandchildren are bringing mother (you) food
[Not *'Your grandchildren are carrying your food for you']

Similarly, the restrictions on the possible meanings of the following show that the Undergoer is a Recipient, rather than the more general 'Beneficiary':

---

33 These verbs can also undergo the more general Type 1(a) derivation. In this case, the verb has the expected benefactive sense, and the Undergoer is now a Theme:

koper-né i nanik, tuwik ané ng-aba-ang
case-3 DET N. T. REL carry

[As for Nanik’s case, Tuwik was the one who carried fit] [for her]

The understood beneficiary is, as usual, the possessor of the theme.
Chapter 11

(124) alih-ang ja tiang kutu
search.for-ang EMPH 1 lice
Go find me some lice
[not *Look for my lice for me]34

(125) getep-ang = ja tiang bok
cut-ang EMPH 1 hair
[Go] cut some hair and bring it to me
[Not *Cut my hair for me]

The non-benefactee status of the Undergoer is also indicated by the fact that the item carried to, or bought for, or brought to U could be something totally unwanted: poisoned food, a bomb, and so.

The second 'type 2' derivational class, type 2(b), is where the Undergoer of the derived verb is a moved Theme, rather than a Recipient (as in type 2(a)). It is brought towards, then into contact with an inanimate Goal.

(126) toktok 'hit U (U=patient)'
daldal 'hit (U = patient)'
slampar 'throw at (U = patient)'
jagur 'punch (U = patient)'
tektek 'chop (U = patient)'

As with type 2(a) verbs, the derived Undergoer is not part of the argument structure of the root verb. The Undergoer of the root verb, is a Patient, and pragmatically also an inanimate Goal. That participant remains a Goal in the derived verb, though as an adjunct (either in a PP or cliticised to the verb), as in the following elicited pair:

(127) ia nektekang dapak-3 di talenán-é
N-tektek-ang 3 AP-chop-ang chopper-3 LOC chopping board-DEF
He chopped his chopping knife onto the board

dapak-3 tektek-ang =a talenán
His chopping knife was chopped by him onto a board.35

The occurrence of beli-ang 'buy with U' in this derivation presumably involves a figurative extension in meaning from the above prototype, the thing purchased being treated as Goal, the money as Theme, brought towards the goal.

(128) pipisé ke-beli-ang nasi
money-DEF ke-buy-ang rice
The money was spent [by me] on rice

Previous analyses have characterised this derivation as involving an Instrument undergoer (Kersten 1984, Barber 1977). It is true that the U of the derived verb in the

34 This could be expressed by either alih-ang = ja kutu-n-tiang-é! (Type 1(a) derivation), or alih-in = ja tiang kutul, 'Look on me for lice' (11.2.3, first subclass).
35 Adjuncts which are formally marked as DEF must be in a PP. Those not so marked must be cliticised to the verb (9.5.3).
Verbal suffixes

above examples is pragmatically an instrument. Similarly, in (129a) the Undergoer corresponds to an NP marked by the preposition aji as a (peripheral) instrument when it occurs with the non-derived form of the verb, (129b):

(129) a. Batu-ne slampar-ang = a poh/ ke poh-é
rock-DEF throw-at-ang = 3 mango/LOC
The stone was thrown at a mango/at the mango

   b. la nylampar poh aji batu
      N-slampar
3 N-throw.at mango INST rock
He threw at a mango with a stone

However this is too broad a characterisation: not all such Instruments have a corresponding type 2(b) derivational counterpart:

(130) la manting kamen-né aji sabun
N-pantil')
3 N-wash cloth-3 INST soap
He washes his loincloth with soap

   *Sabun-é panting-ang = a kamen
   soap-DEF
   *The soap was used to wash a loincloth.

(131) la ngaé layang-an aji kertas
N-gaé
3 N-make kite INST paper
He made a kite with paper

   *Kertas-é gaé-nang = a layang-an

The reason that manting and ngaé cannot undergo 2(b) derivation in the above examples is because 'soap' and 'paper' are not conceptualised as 'Moved Themes', or to be more precise, not themes moved towards a goal: soap is used by moving it in contact with and across/parallel to the surface being cleaned, rather than rapidly towards it. Only transitive verbs which occur with such Moved-Theme instruments can undergo type 2(b) derivations, since that ('Moved Theme') is the semantic role of the Undergoer of the derived verb. In fact, Type 2(b) derivations can occur where the moved Theme is itself the focus of the action (the undergoer of the TRANS-Ø verb), rather than an instrument. The instrument is a stationary object, as when threshing rice:

(132) padi-né tigtig-ang-é di papan-é
paddy-DEF beat-ang-DEF LOC board
The paddy is beaten onto the board [to thresh it]

Elsewhere, there may be no 'instrument' involved at all:

(133) montor-né tomplok-ang = a ke témbok-é
car-3 collide-APP = 3 to wall-DEF
He crashed his car into a wall
Chapter 11

Examples (132) and (133) then show clearly that the Undergoer of type 2(b) derivations is not always an Instrument: this role is therefore not part of the underlying meaning of this derivation. That Undergoer is a Moved Theme, there as in all other examples, and this is therefore the semantic invariant. I propose this semantic formula:

(134) A (N-){TRANS-0}\ang U_theme Y_adjunct/goal

Someone (A) did something ({TRANS-0\}) at the place Y because of this, something (U) came to be at Y before this, U was not at Y something happened to U in the same place, at the same time

Just as pragmatic factors lead to the Undergoer of type 2(a) derivations often being described as Beneficiaries, so, since Moved Themes are pragmatically nearly always Instruments, type 2(b) derivations are often said to involve Instruments: in both cases the underlying role of the Undergoer is in fact something else.

11.4 The derivation of 'stative intransitive stems' of shape -ROOT(-an)

All previous analyses have assumed the existence of at least two distinct derivations with me-, the first using just the prefix me- and deriving stative verbs like me-jemak '(in a state of having been) taken', me-kikih '(in a state of having been) grated', and the second involving a 'confix', me-\an me-aduk-an '(in a state...) stirred', me-bah-an '(...) made to lie stretched out'.

I argue here that most verbs of shapes me-ROOT-0 and me-ROOT-an are in fact the product of a single derivational process, by which intransitive verbs with Undergoer subjects are derived from pre-existing transitives of shape [ROOT(-ang/-in)], ie the transitive stem. The structure of the derived verb I take to be me-[ROOT(-an)], the material in square brackets I refer to as the detransitive stem. There is strong evidence for directionality of derivation, and in this case at least, for the word (the transitive verb), rather than the morpheme as the basis for derivation (cf Aronoff 1976, 1994, Beard 1987, 1988).

From transitive verbs such as slampar '(A) throw-missile-at U', slampar-ang '(A) throw U', saut-in '(A) answer U', intransitive stems are derived. These detransitive stems occur mainly in combination with me-, and occasionally with ke-.

(135) slampar 'hit U with missile' > me-slampar 'be hit by missile'
slampar-ang 'throw U' > me-slampar-\an 'U be thrown'

saut-in 'answer U' > me-saut-an 'U be answered'
atur-in 'offer.HON sthg to U' > k-atur-an 'be offered sthg'
séng-in '(A) summon.HI U' > ke-séng-an 'be summoned.HI'

For simplicity of presentation, I will omit the component 'in a state of having been' from the glosses of all such derived verbs: this component should nevertheless be understood to be present.
For transitive verbs of UP shape ROOT-∅, there is of course no formal difference between that and the detransitive stem. Further, the distinction between the suffixes -ang and -in is 'neutralised' in the detransitive forms, both being realised as /nan ~ an/.

An alternative to the above analysis would be to treat -an as a realisation of two distinct derivational processes (or in traditional terms as an allomorph of both -in and -ang), which would then be unrelated to a 'third' process deriving me-ROOT verbs from ROOT-∅ transitives. The simpler solution is however that a single process of detransitive stem derivation is involved here, and that it takes pre-existing words as its input, rather than suffixes.

An analysis close to this one was first proposed by Kersten (1984:49): he argued for a subset of verbs having both a me-ROOT-∅ (ie bearing no suffix) and a me-ROOT-an form, and where both take an undergoer subject, that they derive respectively from the ROOT-∅ and ROOT-ang transitive verbs. And he argued that forms such as k-atur-an and ke-seng-an derive from ke-ROOT-in Undergoer Pivots. He describes the latter as a process in which 'ke- +-an' replaces 'ke- + in', as though one confix replaced another. I would say that in neither case is there a confix,37 the suffix being part of a complex stem, to which either me- or ke- attaches.38

The process deriving detransitive stems is in fact much more general than Kersten suggested: I give evidence below to show that it also applies to transitive verbs having only one form, either ROOT-ang, ROOT-in or ROOT-∅. I stress however that not all verbs of shape me-ROOT-an have this derivational source. The exceptions are (i) a subset of me-ROOT-∅ verbs with agentive subjects, which I take to involve inherently intransitive bases (10.4), and (ii) me-verbs deriving from noun bases (12.4.4)

The strongest evidence for this analysis comes from distributional data of verbs bearing the prefix me- and having a related transitive form or forms. Put simply, verbs of shape me-ROOT-∅ (and with an undergoer subject) have transitive counterparts of shape ROOT-∅, while those of shape me-ROOT-an have transitive counterparts of shape ROOT-ang or ROOT-in.

Thus in my sample, 97 transitive verbs of shape ROOT-∅ have an intransitive counterpart of shape me-ROOT-∅. The subject of the derived intransitive is almost invariably the undergoer of the transitive:

(136) jagur 'A hit U' me-jagur 'U be hit' pelut '(A) peel U' me-pelut 'U be peele' uug '(A) demolish U' me-uug 'U be demolished' sambat '(A) utter U' me-sambat 'U be uttered, said'

37 Beratha (1992:179) incorrectly claims that a confix analysis of ke-UP forms occurs in Clynes (1989). Neither the word, nor the notion is found anywhere in that text.
38 The process deriving verbs with me- is extremely productive: perhaps all transitive verbs have a corresponding me- form. On the other hand, verbs of shape ke-ROOT-an are less common, often restricted to High or Honorific lexis.
On the other hand, and importantly, I found no verb of shape \textit{me-ROOT-}@ and with an \textit{undergoer} subject, which had no transitive counterpart of shape \textit{ROOT-@}. (There are a number of verbs of this formal type, however they invariably have an \textit{agentive} subject, not an \textit{undergoer}: \textit{me-gedi} 'leave', \textit{me-gaang} 'crawl', and so on. I assume they constitute a distinct, non-derived, class - see 12.4.3.

As well, 128 transitive verbs in the sample of shape \textit{ROOT-ang} or \textit{ROOT-in} had an intransitive counterpart of shape \textit{me-ROOT-an}. Many, as for the following examples, have no \textit{ROOT-@(vt)} counterpart:

\begin{tabular}{llll}
   (137) & endih-ang & 'light U' & m-endih-an & 'U be lit' \\
         & gelek-ang & '(A) swallow U' & me-glek-ang & 'U be swallowed' \\
         & ulap-in & '(A) signal to U' & me-ulap-an & 'U be signalled to' \\
         & paak-in & '(A) approach U' & me-paak-an & 'U be close' \\
\end{tabular}

Crucially, I found no examples of intransitives of shape \textit{me-ROOT-an} deriving from \textit{verbal} stems which did not have a transitive counterpart of shape \textit{ROOT-ang/-in}. (I stress 'verbal' stems, since complex nominals of shape \textit{ROOT-an} do occur with \textit{me-}, eg \textit{me-[kuren-an], 'be married', kuren-an 'family-an spouse'.)\footnote{Others are ambiguous in status: \textit{eg m-atur-an} 'make offerings, pray' could derive from the nominal \textit{atur-an} 'offer-an; thing offered', or from the transitive stem \textit{atur-in} 'offer-LOC; offer to'.}

Verbs which occur with both \textit{ROOT-@} and \textit{ROOT-in/ang} forms can of course have two intransitive counterparts, one of shape \textit{me-ROOT-@}, the other \textit{me-ROOT-an} (87 in the sample were like this). As expected in such cases, the meanings of the two derived intransitives reflect their different derivational origins:

\begin{tabular}{llll}
   (138) & goreng & 'fry' & me-goreng & 'fried' \\
         & me-goreng & 'reheat by frying/cook partially by frying' & me-goreng-an & 'be reheated, be partially cooked' \\
         & tiñjak & 'kick with base of heel' & me-tiñjak & 'be kicked (etc)' \\
         & me-tiñjak & 'kick repeatedly (etc)' & me-tiñjak-an & 'be kicked repeatedly (etc)' \\
\end{tabular}

The second type of evidence for this analysis is then semantic: in the process of lexical derivation the close derivational relationship between two items is often revealed by shared meaning. This can be seen in the examples in (138), and also in the following:

\begin{tabular}{llll}
   (139) & kutang & '(A) throw (all of) U away' & me-kutang & 'be (all) thrown away' \\
         & kutang-in & 'throw (part of) away; throw away repeatedly' & me-kutang-an & 'be thrown partly away; be repeatedly thrown away' \\
         & kayeh & 'bathe (vi)' & *me-kayeh & [does not occur]
\end{tabular}

\footnote{Typically, this has the meaning 'U be summoned by a ceremony', where U is a 'spirit' \textit{(atma juita)} which has temporarily left someone's body.}
Verbal suffixes

### kayeh-ang
- 'A cause U to bathe'
- 'be in a state of having been caused to bathe'

### tegak-ang
- 'cause U to sit down'
- 'be in a state of being/having been caused to sit\(^{41}\)

### nyahnyah
- 'A dry fry U
- 'U be dry fried'

### me-nyah-nyah
- 'reheat by dry frying'

### me-nyahnyah-an
- 'be reheated by dry frying'

### slampar
- '(A) hit U by throwing object (U = target)'
- 'U be hit by thrown object (U = target)'

### me-slampar
- 'U be hit by thrown object (U = target)'
- 'U be thrown (U = missile)'

### slapmar-ang
- 'repeatedly hit (U = target) with thrown object'
- 'repeatedly hit (U = target) with thrown objects'

The base in this derivation must be a transitive verb, hence the non occurrence of *me-
[kayeh] from the intransitive base kayeh in the above example set.

A particularly transparent example where the derivational relationship is revealed by
both a shared idiosyncratic meaning, and a shared formal element (pe-
'CAUS') is the
following:

(140)
\[
\begin{align*}
\text{-gae} & \quad \text{pe-gae-nin} \quad \text{I pe-gae-nang} \\
& \quad \text{me-pe-gae-nan} \\
\end{align*}
\]

- 'work'
- '(A) marry off U\(^{92}\)
- 'U be married off'

Spontaneous paraphrases and translations from my teachers also supported the above
analysis. In eliciting the meanings of verbs affixed with me-, I often heard paraphrases
such as the following, where (141a) (with me-... ...-an verb) was paraphrased as
(141b), with an '-in verb':

(141)
\[
\begin{align*}
\text{a. } \text{suba me-eñjit-an} & \quad \text{lampa-né} \quad \text{di} \quad \text{kamar-é?} \\
& \quad \text{already me-set.alight-an} \quad \text{lampa-DEF} \quad \text{LOC} \quad \text{room-DEF} \\
& \quad \text{Have the lights in the rooms been turned on?} \\
\text{b. } \text{ba ada (anak) ng-eñjit-in?} \\
& \quad \text{already exist (person)} \quad \text{N-set.alight-in} \\
& \quad \text{Has someone turned [them] on?} \\
\end{align*}
\]

These correspondences occurred even where the item in question was said not to exist,
for example, asking about the form me-nyag-an ('me-crush-an'), I was told that 'it
doesn't exist, but if it did, it would mean nyag-in=a [crush-in=3].'

To sum up, transitive stems may be monomorphemic, when the Undergoer is a true
Patient (eg jagur 'hit', tebek 'stab') or else they can be complex, bearing either of the

\(^{41}\)The alternative, agentive, interpretation of me-tegak-an 'A be sitting down', is plausibly from the
me-[noun] derivation, me-[tegak-an], tegak-an 'seat' (12.4.4).

\(^{42}\)Both transitive forms exist: which is the source of the me- derivative (if it is not both) is unclear.
affixes -in 'LOC' or -ang 'APP'. The two affixes generally encode a distinction in Undergoer type: Locative as opposed to 'Other Peripheral' role. When the detransitive stem is derived from these this two-way distinction is neutralised, both affixes are 'replaced' by -an. There remains then a formal distinction only between verbs with Patient Undergoers (me-jagur 'be punched', me-adep 'be sold') and all other Undergoer types, which have the complex stem of shape ROOT-an.43

43 This is a slight oversimplification: transitive 'verbs of violence' retain Patient undergoers when affixed with -in, as do the resulting detransitive forms bearing -an, see examples at (139), and 11.2.
12
Verbal Prefixes

12.0 Verbal prefixes: 'zero prefix', /N/-, and me-

In this chapter I look at the three basic formal paradigmatic oppositions in verbal prefixing: (i) prefixless verbs (also referred to as 'zero prefixing', or 'bare stem' verbs) (ii) verbs carrying the agentive prefix /N/-, and (iii) verbs carrying the prefix me-. (The other common verbal prefix, ke-, found on 'type 2' Undergoer Pivots, is discussed in chapter 13.)

Balinese shows 'split intransitivity' effects similar to those found in many other languages, with inherently intransitive verbs (see 10.4.1) generally occurring with only one (very occasionally, two) of these three prefixing options. I argue below that:

(i) The subjects of intransitive verbs carrying /N/- semantically are Actors, like the subjects of transitive Actor Pivot verbs, which are also marked with /N/-.

(ii) The subjects of prefixless intransitives are semantic Undergoers, like the subjects of bare stem Undergoer Pivot transitives, which are also prefixless verbs.

(iii) Most intransitive verbs prefixed with me- are derived from other lexemes - either from transitive verbs (12.4.1) or from nominals (12.4.4). Those deriving from transitive verbs have Undergoer subjects. The small sub-class of me-verbs with Actor subjects I take to be inherently intransitive verbs, and so not derived. With these verbs the prefix can be seen as a lexically conditioned variant of /N/ (12.4.3).

In short, inherent intransitives with Undergoer subjects are prefixless, while those with Actor subjects bear a prefix (usually /N/, sometimes me-). Exactly the same contrast is found with transitive verbs in their voice alternations. As in its use of verbal suffixes, Balinese behaves as a 'head marking' language (Nichols 1986) in signalling the semantic macro-role of the subject/pivot NP directly on the verb, rather than say by case marking on the associated NP argument(s) (perhaps the more common strategy cross-linguistically (Mithun 1991)).

I therefore assume that the formal distinctions in Balinese verbal prefixing, like those found with suffixing, have a semantic rather than a purely syntactic explanation. This accords with the findings of Van Valin (1990), Mithun (1991), Dowty (1991) and other authors who have written on 'split intransitivity' effects cross-linguistically. Van Valin argues that two principal factors lead to split intransitivity: the degree of agentivity of

1 A further, minor derivation is where certain transitive verb stems are formed with the prefix pe- 'caus'; this is not analysed in this thesis (see 11.2.2.2 for a brief discussion of one function). Stems carrying pe- then may occur in either the N prefixed, bare stem, or me- forms discussed here.
Chapter 12

the subject NP and the inherent lexical aspect (Aktionsart) of the verb. The first of these factors seems to play the dominant role in the Balinese case, though lexical aspect may perhaps play a role in certain lexical classes (12.3).

12.1 'Zero prefix': bare stem intransitive verbs

Verbs which occur as bare morphemes without affixes are of two main types: 'underlyingly' intransitive verbs on the one hand:

(1)  gedé       'be big'
pules       'sleep'
demen       'be happy'
kédék       'laugh'
teka        'arrive'

and the bare Undergoer Pivot (UP) form of transitive verbs on the other. (The two can be easily distinguished, cf 10.4.1) Their common formal shape reflects a common semantic element: a subject/pivot NP which is an Undergoer.2

Semantically, the bare stem intransitive verbs are of two main types. A small minority can have agentive, volitional subjects. These include:

(2)  kayeh       'bathe'
pesu        'go out'
tuun        'come down'
mai         'come here'
kemo        'go there'
pules       'sleep'

Most members of this small agentive sub-class refer to events rather than states in their inherent lexical aspect. They occur readily in imperatives: kayeh! pesu! pules! They also occur with the modal baka~, which with a first person subject indicates the subject accidentally performed a potentially volitional action:

(3)  adi ditu baka pules!
     PART there baka sleep
     How did [I] manage to [accidentally] sleep there!

However the great majority of bare intransitives have non-volitional/non-controlling subjects. As such, they cannot occur in imperatives, or with the 'accidental' usage of baka:

2As well, many 'complement taking verbs' are bare stem verbs. These are verbs which occur with a clausal complement, and have stative aspecautal or modal meanings. They include suba 'already (COMP)', suud 'finished (COMP)', dot 'want very much to (COMP)', takaar 'will (COMP)', oka 'intend to (COMP)', bisa 'be able to (COMP)', dadi 'be allowed to (COMP)'.

248
While some of these verbs are semantically eventive, most have the stative meanings encoded by an 'adjective' word class in many languages. (No distinct adjective class is assumed in this analysis, since the semantically appropriate morphemes do not show formal/syntactic properties setting them apart from other verbs.)

All bare intransitives can be readily used as nouns without further derivational morphology:

(5)

a. tan carita-yang unduk gedeg i lutung-é
    NEG.HI relate-ang matter angry DET monkey-DEF

The matter of the monkey's anger will not be related.

b. aéng demen-né naar rujak
    strange happy/like-3 eat rujak

It's amazing, his liking for eating rujak [a spicy fruit dish]

I nonetheless consider these to be examples of 'conversion' or zero-derivation: true 'underlying' nouns freely occur with me- to form intransitives (see 12.4.4), bare intransitives never do: *me-gedeg, *me-demen.

When transitive verbs bearing the 'applicative' morpheme -ang have an intransitive counterpart, the S of the intransitive corresponds to the Undergoer NP of the transitive.

The meaning is (A) cause U to (vi):

(6)

teka 'come'
    teka-ang 'make U come'
kayeh 'bathe (vi)'
    kayeh-ang 'bathe U (vt)'
ulung 'fall'
    ulung-ang 'make U fall, drop U'
mati 'dead'
    mati-ang 'make U dead, kill U'

Kana (1986) argues from parallel facts in Indonesian that bare intransitives in that language are underlyingly 'unaccusative', bearing an underlying Object grammatical relation. But Kana's argument overlooks other evidence: in combination with the 'locative' affix -i (or in the case of Balinese, its counterpart -in), the same roots derive transitives with the approximate meaning 'A do (vi) to/at/in/on U', where the S of the
intransitive root now corresponds to the $A$ (Actor/"transitive subject") NP of the transitive verb, not to the $U$ NP. Thus, in Balinese:

\[
\begin{array}{llll}
\text{teka} & \text{'come'} & \text{teka-in} & \text{'A come to U'} \\
kayeh & \text{'bathe (vi)'} & \text{kayeh-in} & \text{'A bathe in U'} \\
ulung & \text{'fall'} & \text{ulung-in} & \text{'A fall on U'} \\
\text{negak} & \text{'sit'} & \text{tegak-in} & \text{'A sit on U'}
\end{array}
\]

Kana's analysis would require that the forms with -ang be somehow more basic than those with -in. I know of no reason to assume this.

Moreover, the fact that both agentive (eg negak (%N-\text{tagak} %)) and non-agentive (eg mati) intransitives behave in the same way when affixed indicates that the derivational relationship between the NP arguments of the intransitive and transitive verbs ($S = A$ of -ang verbs; $S = U$ of -in verbs) is a product of the respective derivational processes themselves, rather than being due to the putative underlying grammatical relation (or semantic role) born by the $S$ NP.

To sum up so far, except for a handful of exceptional cases where the subject NP may be considered an Actor (see also 12.3), bare stem intransitive verbs regularly have Undergoer subjects. (Evidence from derivational processes, such as that adduced by Kana for Indonesian, is, I believe, irrelevant for determining this.)

The same situation is found, this time without exception, with semantically transitive verbs. On these verbs, the prefixless bare stem form is that where the Undergoer NP (rather than the Actor NP) is functioning as syntactic pivot (9.4.2).

\[
\begin{array}{llll}
\text{a. } & \text{batis-né} & \text{kepak =a} & \text{'kén konyong-é} \\
\text{foot-3} & \text{chew.UP-3} & \text{with puppy-DEF} & \text{his foot was being chewed by the puppy} \\
\text{b. } & \text{pedanda, apa déén uning-in = a} & \text{Priests, everything is understood by them}
\end{array}
\]

This 'Undergoer-is-Pivot' interpretation is consistent then with the semantics of prefixless verbs generally, be they intransitive or transitive.

### 12.2 The agentive prefix {N}

Just as '0-' signals that the subject/pivot verb is an Undergoer, irrespective of the syntactic transitivity of the verb, so the {N} prefix signals 'Pivot-is-Actor', on both transitive and intransitive verbs. Evidence for this is presented in the following subsections, including 12.3, where problematic data is discussed. This analysis also assumes that the prefix me-, where it occurs on agentive intransitives (12.4.3), is a lexically conditioned realisation of the same agentive derivation as is realised by {N}.

250
I will refer to the usual realisations of \{N\} as '/Nf/'. They are described in 6.2. In very formal archaising literary works \{N\} often occurs as <\text{maN}>: the 'N segment' is realised in the regular way, but it is preceded by an extra syllable, usually pronounced /ma/:

(9) maka pengawit titiang mangaturang puja panganjali
    as beginning.HI 1.HON N-offer.HON-ŋ prayer praise

To begin I offer a prayer of praise\(^5\)

(10) kénak palungguh batara mang-ampura kaiwangan titiang
    be.happy.HON lord god forgive mistake 1.DEF

See fit, oh Lord, to forgive my mistakes\(^6\)

This 'maN' literary form retains the older form of the morpheme (Goris 1954), the initial two segments having been largely lost in the spoken language (though see the next paragraph). The optional retention of the extra syllable is perhaps due to the flexibility it gives when writing in poetic metre, though it is found in prose texts too, as the above examples show.

The one exception to the loss of the initial /ma/ or /ma/ in everyday speech is before roots beginning with a labial stop. For a subset of these, '/məŋ/'-forms are still common, either as optional variants:\(^7\)

(11) maunci, memaunci %N-paːciŋ % 'N-fishing pole: angle'
    meteng, memeteng %N-petəŋ % 'N-dark; be in darkness (eg a house)'
    maak, memaak %N-baak % 'N-robber; rob'
    mula, memula %N-pula % 'N-plant (v)'

or as the only permitted realisation:

(12) memules %N-pulas % 'N-sleep; do nothing but sleep'
    memunyah %N-puːnaŋ % 'N-drunk; habitually get drunk'

The derived verbs carrying '/məŋ/' are often intransitive, but derived transitives also occur. Examples like this rule out another possible analysis, that these are CV-reduplications, expressing say, durative aspect:

(13) anak ten tiang memaling bé-n kaki-né
    EMPH NEG.HI 1.EXC.HI N-steal fish-GEN grandfather-DEF

It wasn't me (who) stole grandfather's fish\(^8\)

---

\(^5\)From I Madé Bija Arya Bang Pinatih (c.1989) *Kidung Panca Yadnya* Yayasan Kawi Sastra Mandala, Singaraja.

\(^6\)From 'Prabu Pandu', Ketut Mandi Pinatih, ms., written in 1960's.

\(^7\)A surprising analysis common to all previous descriptions is that forms such as memula carry the prefix me- (see 12.4). They neglect to account for the nasalisation of the root-initial consonant.

\(^8\)Kirtya No. 1976 'Idjung Lantang', from Delodpeken, Singaraja.
Whether \{N\} is realised as monosegmental '/N/' or /maN/ with labial-initial roots is lexically determined; where both are allowed I could elicit no meaning difference. Most commonly, only the mono-segmental realisation, where /m/ 'replaces' the initial labial obstruent, is found.

12.2.1 {N} on monomorphemic intransitive verbs

Like other intransitive verbs, those with \{N\} are invariant: whether they be used in imperatives, questions or any other context, they retain the initial nasal segment. This contrasts with transitive verbs, where the presence or absence of \{N\} is influenced by a combination of factors such as voice, mood and the referential status of the undergoer NP (see for example 9.5.2, 15.3).

Typically intransitive verbs carrying \{N\} have animate (usually human) subjects and are eventive rather than stative in lexical aspect. The meanings of these verbs are broadly predictable from those of their roots, which include both verbs and nouns.

1. Where the root is a verb (or can be assumed to be so)\(^9\) it often relates to body movements or functions. The intransitive has meanings including:

1(i) 'do bodily action':

(14) negak %N-tegak % 'sit down'
    meju %N-peju % 'defecate'
    ngadeg %N-adeg % 'stand (HON)'
    ngigel %N-igal % 'dance'
    ngidem %N-kidem % 'close one's eyes'
    nyengangal %N-seangal % 'gape, hold mouth open (of eg dog)'
    nylémpoh %N-jémpoh % 'collapse onto one's knees'

This semantic category includes verbs expressing the production of sounds:

(15) ngelur %N-gelur % 'yell, howl'
    ngremek %N-gremek % 'mutter, mumble'
    nyerit %N-jerit % 'scream'
    ngwel %N-wel % 'complain loudly and angrily'
    ngongjong %N-kjong % 'bark (of a dog)'

1(ii) a second important category, 'move (with respect to a place)':

(16) neked %N-taked % 'reach a place'
    menék %N-penék % 'climb'
    nglipet %N-lipet % 'return'
    ngoyong %N-ojong % 'remain still; reside'
    nglik %N-glik % 'roll'
    ngeléb %N-leb % 'run away'

\(^9\)Where bound root morphemes are bound and cannot function as nominals without further affixing, as in all of the examples in (14), I assume they are verbs underlyingly: see sections 10.4, 15.2.
Verbal Prefixes

1(iii) certain verbs describe the change in state as an ancestor reincarnates in a subsequent generation, or as gods take on earthly manifestation:

(17) nitis %N-titis% 'take on earthly form'
numitis %N-tumitis% 'reincarnate'
numadi %N-dumadi%10 'reincarnate'

In the same semantic class is nadi %N-dadi% ('N-become') with meanings including 'manifest oneself (of a god); (deliberately) become possessed by a spirit (of a medium)':

(18) lantas ida sang hyang gana nadi dadi Ganamurti
then 3HON. S. H. G. N-become become G.
Then Ida Sang Hyang Gana (Ganesa) manifested himself, becoming Ganamurti

The minimal pair dadi 'become', and nadi %N-dadi% 'deliberately become(...)’ clearly indicates the agentive semantic component associated with the /N/ prefix.

1(iv) Other such verbs (ie formed from a verbal root) have an implied undergoer which is not the subject's body or body part:

(19) ngijeng %N-ijao% 'guard a place (eg one's home)'

ngijeng is a true intransitive: its transitive counterpart is (ng-)ijeng-in 'guard (vt)’. On the other hand some apparently intransitive verbs have a homophonous Actor Pivot transitive counterpart (and so also a corresponding Undergoer Pivot form):

(20) numbeg %N-tumbag% 'hoe (vi, vt)'
nunun %N-tunun% 'weave (vi, vt)'
nyakan %N-jakan% 'cook (cereals such as rice) (vi, vt)'
ngabén %N-aben% 'cremate (vi, vt)'
ngajeng %N-ajan% 'eat (HI ~ HON; vi,vt)'
ngaduk %N-aduk% 'stir (up trouble) (vi); stir (vt)'
nolih %N-tolih% 'look around; look around at'
ingeh %N-dinah% 'hear (vi, vt)'
ngenot %N-not% 'see (vi, vt)'

Whether verbs like those in (20) should be considered 'true' intransitives is unclear: I assume here that all are best analysed as underlyingly transitive, with ellipsis of an undergoer NP which is either generic or inferable from context.

Most or all of the verbs in examples (14) to (20) can potentially have a volitional, controlling subject. Still, some verbs of motion taking {N} refer to involuntary actions. They have either an inanimate subject, or else a noncontrolling human subject:

10Historically tumitis and dumadi derive by infixation of *um from titis 'be manifested' and dadi 'become'. um-infixation is no longer a synchronic morphological process.
I propose that all of the above 'N-INTRANS' verbs share the following semantic components, either:

(22) someone did something

or else, where the notion Actor is extended to include inanimates:

(23) something did something

2. Where the root is a noun, the meanings of the derived forms include, roughly:

2(i) 'do what one typically does with (inanimate noun)/ use (noun)'

(24) nyampat %N-sampat% 'N-broom; sweep'
    ngapang (me)mancing %N-papcil% %N-papcin% 'N-fishing rod; angle'
    ngubu %N-kubu% 'N-hut; live in a hut'
    timba %N-timba% 'N-bucket; draw water with a bucket'
    ngroko %N-roko% 'N-cigarette; smoke'
    nongos %N-tongos% 'N-place; remain in one place, reside'
    memek %N-pakan% 'N-market; go to the market'
    ngopi %N-kopi% 'N-coffee; drink coffee'
    nyekah %N-sekaha% 'N-representation of dead person; perform rite using this'
    ngluah %N-luah% 'bring up milk (baby); erupt (mountain)'
    nyekala %N-sekala% 'N-earthly world; take earthly form (god)'
    nyeneng %N-janal% %N-janal% 'N-form, shape.HON, be alive.HON'

2(ii) 'make (inanimate noun)'

(25) nembok %N-tembok% 'N-wall; make a wall'
    numpeng %N-tumpa% %N-tumpen% 'N-rice cone; make rice cone'
    nyongkét %N-songket% %N-songket% 'N-songket cloth; make songket'
    mubuh %N-bubuh% 'N-rice porridge; make rice porridge'
    napé %N-tape% 'N-fermented rice; make tape'
    nodol %N-dodol% 'N-dodol cake; make dodol cakes'
    nyaja %N-jaja% 'N-rice cake; make jaja'

In most of these cases, the activity is part of customary behaviour (for example numpeng, napé, nodol and nyaja are only used to describe these actions when done in...
Verbal Prefixes

preparation for a religious holiday), or else are virtually the only activity regularly associated with the root noun (nembok). As such, these can all be assumed to be pragmatically determined realisations of meaning 2(i), 'do what one typically does with (inanimate noun)'.

Where the root is a body product noun, the meaning is 'produce (body product)', though again, this semantic type can be seen as being subsumed under, 'do what one typically does with (inanimate noun)'.

\[(26)\]

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>ngomong</td>
<td>%N-omoŋ%</td>
</tr>
<tr>
<td>ngeling</td>
<td>%N-liŋ%</td>
</tr>
<tr>
<td>ngentut</td>
<td>%N-entut%</td>
</tr>
<tr>
<td>ngenceh</td>
<td>%N-ŋčah%</td>
</tr>
<tr>
<td>ngutah</td>
<td>%N-utah%</td>
</tr>
</tbody>
</table>

2(iii) where the root is an animate noun, the meaning is 'act like (noun)':

\[(27)\]

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meanings</th>
</tr>
</thead>
<tbody>
<tr>
<td>noris</td>
<td>%N-toris%</td>
</tr>
<tr>
<td>nyupir</td>
<td>%N-supir%</td>
</tr>
<tr>
<td>ngarjuna</td>
<td>%N-arjuna%</td>
</tr>
<tr>
<td>nyangut</td>
<td>%N-saŋut%</td>
</tr>
<tr>
<td>nyelam</td>
<td>%N-selam%</td>
</tr>
<tr>
<td>nglegu</td>
<td>%N-ŋlegu%</td>
</tr>
</tbody>
</table>

These various meanings are reducible to the following common semantic formula, where 'Y' is the set of conventionally accepted behaviours associated with N-[NOUN]:

\[(28)\]

\[
\text{S did something (Y)} \\
\text{Thinking of [NOUN], people think it is normal to do Y}
\]

Very similar agentive meanings to those proposed above are found on a subset of verbs of structure me-[NOUN]: eg me-suah 'me-comb: use a comb', me-kutu 'me-louse: look for lice'. One might consider those verbs and verbs of structure /N/-[NOUN] to be lexically and morphologically conditioned variants of a single derivation process. In 12.4.4 I examine the evidence for this and conclude in the negative.

### 12.3 The semantics of 'N-' and '∅-' verbs: Balinese as a 'split-S' language

I have argued that the subjects/pivots of 'zero prefix' verbs are generally semantic Undergoers, while (N)-affixed verbs generally have Actor subjects: a classic 'split-S' pattern (Dixon 1979). Here I look at this question in closer detail. First extra evidence in favour of this hypothesis is presented. Then I attempt to account for some otherwise problematic data. Most 'exceptions' to the above hypothesis can be seen as predictable

---

\(^{11}\)Arjuna is a character from the Mahabhrata stories; Sangut is from wayang (shadow theatre).
departures from a prototype. This assumes that semantic roles such as Agent/Actor and Patient/Undergoer are complex notions each involving several properties, not all of which need be associated with a given NP argument of a given verb (cf Mithun (1991), Dowty (1991), Wierzbicka (1992)).

The clearest case of the contrasting semantic value of \( {N} \)- and \( \emptyset \)- is of course when they alternate on transitive verbs, the presence of \( /N/ \) indicating that the Actor NP is functioning as syntactic pivot/subject (29a), and \( \emptyset \) that the Undergoer NP is subject (29b):

\[
\begin{align*}
\text{(29) a. } & \quad i \quad \text{Madé} \quad \text{meli} \quad \text{bunga} \\
& \quad \text{N-beli} \\
& \quad \text{DET M. AP-buy flowers} \\
& \quad \text{Made is buying flowers} \\
\text{b. } & \quad \text{Bunga} \quad \text{beli-na} \quad \text{tekén} \quad \text{i} \quad \text{Madé} \\
& \quad \text{flower buy(UP)-3 with} \\
& \quad \text{she bought flowers (lit. Flowers were bought by her)}
\end{align*}
\]

Intransitive verbs do not generally occur in both a bare stem and an \( {N} \)-prefixed form, however where they do there is a similar contrast in semantics, as in these examples:

\[
\begin{align*}
\text{(30) } & \quad \text{bah} \\
& \quad \text{'fall lengthways (eg of a tree)'} \\
& \quad \text{ngebah} \\
& \quad \text{'id., deliberately; lie down to rest'} \\
& \quad \text{pules} \\
& \quad \text{'sleep'} \\
& \quad \text{memules (\%N-pules\%)} \\
& \quad \text{'do nothing but sleep, deliberately'} \\
& \quad \text{joh} \\
& \quad \text{'be far off'} \\
& \quad \text{ngejoh} \\
& \quad \text{'deliberately distance oneself'}
\end{align*}
\]

Semantically the 'minimal pairs' in example (30) contrast for two factors generally recognised as affecting degree of agentivity: animacy and volition (see eg Mithun 1991, Dowty 1991). Where in (30) the 'zero-prefixed' verb's meaning is compatible with either an animate or inanimate subject, (bah and joh), their \( {N} \)-affixed counterparts ngebah and ngejoh must have human/animate subjects. In some cases, only the \( {N} \)-form allows the animate subject (body parts being treated as inanimate with this derivation):

\[
\begin{align*}
\text{(31) } & \quad \text{kebus} \\
& \quad \text{'be hot (of an object (including the body/a body part))'} \\
& \quad \text{ngebus} \\
& \quad \text{'have a fever (of a person)'} \\
& \quad \text{sakit} \\
& \quad \text{'hurt (S = body part)'} \\
& \quad \text{nyakit} \\
& \quad \text{'experience pains of childbirth (S = person)'} \\
& \quad \text{nyakitang[ = body part] } \quad \text{'experience pain in [body part] (S = person)'}
\end{align*}
\]
(32) sakit kuping-ne
sore ear-3

His/her ear is sore

nyakitang = kuping i Madé
N-sakit-ang
N-sore-ang = ear DET M.

Madé has an earache

And in all three cases in (30), where verbs carry {N}, the subjects exert volition/control over their actions, whereas no such control is implied by the bare stem forms. Moreover, the unprefixed form can be incompatible where volition is implied on the part of the subject:

(33) uli telu-ng dina sing meju-meju (??peju-peju)
from three-LNK day NEG RED-N-defecate

(he) hasn't (been able to) defecate for three days

The unaffixed form would be inappropriate in (33), since it implies a lack of control incompatible with the 'desiderative/volitional' import of the sentence. A person with diarrhoea on the other hand might say:

(34) tiang jani kéto, peju-peju dogén
1exc now like.that RED-defecate just

I'm like that now, all I do is defecate (uncontrollably)

Another basic contrast bought out by the pair joh, ngejoh in (30) is the distinct lexical aspects prototypical of each type: joh 'be a long way away', like verbs taking an undergoer cross-linguistically, is semantically stative, whereas ngejoh 'move away' is eventive. Van Valin (1990) and Mithun (1991) both show how this aspectual contrast underlies split intransitivity effects in a variety of languages.

In some cases, {N}-affixed forms are used of inanimate subjects, and so are problematic for the analysis above. It seems that in these cases, the subject NP is viewed as, while not volitional, in some way initiating the action/state described by the verb. The subject NP is viewed as having the agentive property of 'causing an event or change of state' (cf Dowty 1991:572), and this is sufficient in Balinese for the verb to carry {N} marking. This includes the N-(ROOT)-in intransitives (11.2.2.2), the subjects of which are not necessarily volitional, or even animate: ng-lengeh-in 'be intoxicating', ngenitin (%N-genit-in %) 'cause itchiness'.

In other cases {N}-intransitives occur with inanimate subjects apparently via a process of metaphorical extension. Thus while memeteng with human subjects means 'do (complement clause) in the dark, be (deliberately) in the dark' (35), it is often used to refer to places, particularly those normally illuminated (36):
Chapter 12

(35) icang suba sai-sai memeteng ngamah
     1 already RED-often N-dark eat
     I have often eaten/often eat in the dark\(^{(35)}\)

(36) memeteng umah-é
     N-peteng house-DEF
     (The/my) house is in darkness

It was however stressed to me that in example (36) 'it is as though the house is deliberately in darkness', whereas if peteng 'dark' were used in the same context no such connotation would be there.

Similarly nadi \(\%N\text{-dadi}\) ('N-become') has volitional meanings including 'manifest oneself (of a god)' (see example (18), as well as '(deliberately) become possessed by a spirit (of a medium)'). Metaphorical extension is clearly the source of a third meaning 'rise (of dough)': again, it is as though the dough (like the house in (36)) is capable of deliberate, internally initiated action. This presumably accounts for other pairs such as teked, neked 'reach destination (of person or, eg, a letter, a bus).

Metaphorical extension presumably also gives rise to the use of \{N\} even where the subject NP is always or is typically inanimate, as in the doublet tasak, nasak 'ripe, become ripe'. Again, as with nadi and the N-(ROOT)-in 'causative intransitives' discussed above, I assume that the inanimate subject is seen as independently initiating the change of state. In the same way many of the \{N\}...-ang inchoatives (11.3.2.1), such as ngedénang \(\%N\text{-gade-nal}\) 'get bigger', nyelemang \(\%N\text{-salam-an}\) 'get blacker/darker', are freely used with inanimate subjects.

A related 'problematic' group of \{N\}-intransitives contains verbs involving spontaneous, often sudden, movement even though the subject is inanimate, or in a non-controlling state:

(37) ngetel \(\%N\text{-katel}\) 'drip (of a liquid)'
    nquad \(\%N\text{-uad}\) 'stretch (of a material)'
    ngejer \(\%N\text{-gejor}\) 'tremble (eg body part or the earth)'
    nylekak \(\%N\text{-slakak}\) 'suddenly stop moving, become stuck (eg of a bone in the throat)'
    ngundap \(\%N\text{-undap}\) 'nod (from drowsiness)'
    ngamuk \(\%N\text{-amuk}\) 'run amuck'
    nungkak \(\%N\text{-tul\text{'}}kak\) 'be brought to a halt before completion (eg work)'

Dowty (1991) argues that moving subjects are associated with prototypical agentivity. In related languages there is a similar class of verbs, also carrying the (counterpart to the) \{N\}-affix, cf Indonesian/Malay menétés \(\%m\text{-N-tetes}\) 'drip', meng-antuk 'be drowsy', meng-amuk 'run amuck'. I conclude that that verbs like those in (37) are also conceptualised as involving a subject NP which, if not controlling the action, initiates it,

\(^{12}\text{Kirtya, Satua I Taniyeng.}\)
Verbal Prefixes

and that this is sufficient in Balinese to for that NP to be assigned Actor status (cf discussion of N-(ROOT)-in intransitives above and (11.2.2.2)).

In my data there remains a small group of {N}-affixed intransitives which are less readily accounted for: these include ng-enah (%N-nah%) 'be visible', ng-ills 'be clearly visible' and the complement taking verb ng-aat 'lust to'; the honorific (deprecating) tambet 'stupid' occurs also as nambet, perhaps better glossed as 'do things stupidly'.

To sum up, and assuming that this last small group can either be discounted as exceptional, or will yield to further semantic analysis, I see no need to posit two distinct but homophonous morphemes of shape /N/, the one associated with intransitive stems, the other associated with transitive stems; the latter analysis is proposed for Balinese by Kersten (1984:33-34). Instead, I assume a single word formation process, with all verbs carrying {N} having as their subject NP an Actor, which, if not animate and/or volitional, is always viewed as initiating the predicated action.

Potentially problematic 'zero-prefix' verbs are equally few, limited perhaps to the following, which could be expected to have agentive subjects, and hence take {N}:

(38)  
| a.    | tuun   | 'come down' | pesu   | 'go out'  |
|      | kemo   | 'go there'  | mai    | 'come here' |
|      | teka   | 'arrive'    | laku   | 'go'      |
|      | lunga  | 'go.HON'    |        |           |
| b.    | kayeh  | 'bathe'     |        |           |

The verbs listed under (38a) are all verbs of motion. They have counterparts with parallel behaviour in languages such as Italian (cf Van Valin (1990:231-237), analysing the latter in terms of lexical aspect). The other potential exception is (38b), kayeh, which can be explained in the same way, though via a historical explanation, assuming that it derives from a reanalysis of ke yēh '[go] to water' (ke was formerly realised as /ka/). Mithun (1990) argues that idiosyncratic historical developments like these often obscure the underlying semantic basis of split intransitivity.

All in all in Balinese the split remains clear, and is a reflection of synchronic processes.

12.4 The prefix me-: introduction

The variant realisations of the verbal prefix me- are described in 7.6. This prefix presents a variety of analytical challenges. The first is to determine how many derivational processes use me-. (I assume that me- does not in itself bear meaning,

---

13 The 'homophonous' morpheme analysis is sometimes proposed for Malay, eg by Prentice (1987:922, 933). The 'single morpheme' analysis proposed here applies also in that language, I believe.

14 It still is in certain lexicalised expressions: ka uma [kaumal] 'to the ricefields. The sequences /ye/ and /aCe/ are highly disfavoured morpheme-internally (5.4.2 and 6.3.1 respectively), hence perhaps the change *kayeh/ > kayeh/.

259
contra Fakultas Sastra (1977) who claim that there are 6 different homophonous me-
affixes, 5 of them realising distinct synchronic derivations.\textsuperscript{15} There are different meanings associated with verbs derived with me-, however these meanings are largely predictable, according to whether the stem is an inherently agentive intransitive (12.4.3), a detransitive verb (12.4.1), or a noun (12.4.4).

A second, related, issue is whether there is a 'me-...-an' confix. I argue that almost all verbs of shape me-ROOT-an involve either me- prefixed to a complex detransitive stem of form ROOT-an (12.4.1), or else to a complex nominal (12.4.4), rather than a confix plus ROOT (pace previous analyses, eg Beratha 1992). There are candidates for 'true' me-
an confix status, though these are very few (12.5).

To sum up, verbs carrying the affix me- are always intransitive.\textsuperscript{16} me- attaches to (i) stems deriving from transitive verbs, (ii) bound intransitive roots (in a few cases) and (iii) nouns. These three types are discussed in turn in the following subsections.

\section*{12.4.1 me- on detransitive stems}

Prefix me- occurs with detransitive stems (11.4) to derive stative verbs with Undergoer subjects, with the meaning 'U be in a state of having been [VERBed}'. Transitive verbs of shape ROOT-\(\ominus\) have detransitive stems identical to the transitive root ((39)a), while those of shape \(\text{ROOT-an}\) and \(\text{ROOT-in}\) both have stems of shape \(\text{ROOT-an}\) ((39)b).

(39)
\begin{itemize}
  \item[(a)]
    \begin{align*}
      \text{tulis} & \quad \text{ '(A) write U'} \quad \text{me-tulis} \quad \text{'U be in a state of having been written'} \\
      \text{duduk} & \quad \text{ '(A) pick up U'} \quad \text{me-duduk} \quad \text{'U be in a state of (...) picked up'} \\
      \text{timpug} & \quad \text{ '(A) throw at U'} \quad \text{me-timpug} \quad \text{'U be in a state of (...) thrown at'}
    \end{align*}
  \\
  \begin{align*}
    \text{b.} \\
    \text{duduk-in} & \quad \text{ '(A) pick up (PL) U'} \quad \text{me-[duduk-an]} \quad \text{'U be (...) picked up (PL)} \\
    \text{timpug-an} & \quad \text{ '(A) throw U (at sthg)'} \quad \text{me-[timpug-an]} \quad \text{'U be (...) thrown'} \\
    \text{tulis-in} & \quad \text{ '(A) write on U'} \quad \text{me-[tulis-an]} \quad \text{'U be (...) written on'} \\
    \text{tulis-ang} & \quad \text{ '(A) write with U'} \quad \text{me-[tulis-an]} \quad \text{'U be (...) written with'}
  \end{align*}
\end{itemize}

No such verbs were found where the subject was an Undergoer, and there was not a corresponding transitive verb as a derivational source. Given the existence of the detransitive stem, the analysis that such me-\{[...]\}-an verbs contain a confix is unnecessary.

This is a very productive derivation, with virtually all transitive verbs having a 'me-derived' counterpart. For one or two transitive verbs dictionaries list only agentive intransitive counterparts: me-daar 'eat (vi)', me-gaé 'work (vi)'. For both of these however I did observe 'S = undergoer' usages in spontaneous conversation: me-daar 'be

\textsuperscript{15}As with perhaps all previous writers, the /men/ realisation of \{N\} is incorrectly analysed by Fakultas Sastra 1977 as a variant realisation of \{me\}. Transitive verbs of underlying structure \%N-[pa-(...)]\% (in Singaraja at least) are also often confused for verbs with \{me\}, eg mebalih 'look at, watch' from \%N-[pa-balih]\%, me-liwat 'pass by' from \%N-[pa-liwat]\%.

\textsuperscript{16}One possibly transitive exception to this is me-karya 'make' discussed in footnote 25.
partly eaten' was said of a durian fruit, and me-gaé 'be made', of temple offerings. Some speakers when questioned accepted these as perfectly grammatical, others rejected them. The agentive forms could derive from a variety of sources: ellipsis of incorporated undergoer nominals (12.4.2), or from a homophonous intransitive root (12.4.3) or from a nominal stem (12.4.4).

The verbs affixed with me- in (39) all have undergoer subjects, and cannot cooccur with an Actor NP, even in an oblique PP.\(^\text{17}\)

\[
\begin{align}
\text{(40) } & \text{pipis-é suba me-duduk (*tekén i Madé)} \\
& \text{money-DEF already me-pick.up (by DET M.)} \\
& \text{The money has been picked up (*by Madé)}
\end{align}
\]

Equally, they cannot be used to merely imply, without stating explicitly, that an actor was involved in the event leading up to the predicated state. It could be hypothesised that this impossibility of reference to an Actor follows from the stative aspect of these verbs, rather than from a lack of reference to an Actor in their semantic structure (Dowty 1991:558). I prefer rather to assume the latter analysis: Bare-Stem- and ke-Undergoer Pivots are used extensively to imply the presence of, but avoid direct reference to, an Actor (chapter 13); one would expect that me-dettransitives would be used in the same way, if they did contain reference to such a participant in their semantic structure; they are never used in that way. I therefore assume that verbs of structure me-[DETRANS] are true semantic (and therefore also syntactic) intransitives (15.3). Of Wierzbicka's prototypic transitive scenario (15.3),

\[
\begin{align}
& \text{(41) } A \text{ VERB U} \\
& \text{someone (=A) did something} \\
& \text{because of this,} \\
& \text{something happened to something (=U)} \\
& \text{in the same place, at the same time} \\
& \text{this person wanted this}
\end{align}
\]

they retain only the element:

\[
\begin{align}
& \text{(42) } U \text{ me-[VERB]} \\
& \text{something ([VERB]) happened to U}
\end{align}
\]

This formula is better expressed as follows, to capture the stative or aspectually durative element associated with these verbs:

\[
\begin{align}
& \text{(43) } U \text{ me-[VERB]} \\
& \text{something ([VERB]) has happened to U}
\end{align}
\]

\(^{17}\text{Instrumental adjuncts can cooccur, but apparently only in PP's: me-godot *(aji) tiuk 'me-cut *(with) knife', me-cuci *(aji) sabun me-wash *(with) soap', *me-bayah pipis 'me-pay money'). The following was nevertheless heard in a performance by JDP, at Lovina Beach, November 1991:}

\[
\begin{align}
tai-n & \quad \text{kebo me-kaput plastik} \\
dung-GEN & \quad \text{buffalo me-wrap plastic.bag}
\end{align}
\]

\text{Buffalo dung wrapped in a plastic bag.}

\text{Other speakers require me-kaput aji/baan plastik.}
12.4.2 me- on detransitive stems, with an incorporated undergoer NP

Occasionally me- attaches to a detransitive stem which includes an incorporated undergoer NP (corresponding to the Undergoer NP of the transitive verb). In such cases, the intransitive subject NP is semantically an Actor ('SA' in glosses), not an Undergoer, as is the case when there is no incorporated NP (12.4.1).

(44) me-kancing  
    me-kancing jlanan  
    'SU be locked'  
    'SA (be in a state of) keeping the doors locked'

me-pecik  
me-pecik manggis  
'SU be squeezed'  
'SA "squeeze manggis": pay homage to a superior'  

me-timpug-an  
me-timpug-an batu  
'SU be thrown'  
'SA be throwing stones (durative)'

me-téngténg-an  
me-téngténg-an tas  
'SU be carried in the hand'  
'SA be carrying a handbag'

(45) me-salin  
    me-salin rupa  
    'SU be changed/exchanged'  
    'SA have changed one's appearance'

kekua-né  
me-salin = rupa  
= dadi  
= lutung  
tortoise  
me-change = appearance  
become  
monkey

The detransitive form again has a stative sense not found with the AP form:

(46) ia  
    me-kancing  
    jlanan  
3  
    me-lock  
    door

he is 'with' [=behind] locked doors

ia ngancing jlanan  
he locks/locked/has locked/etc the doors

ia  
    pules  
    me-kancing  
    jelan  
    sleep

he slept with the doors locked

?ia pules ngancing jlanan  
?he slept and locked the doors

The incorporated nominal must be semantically non-referential, and so cannot be formally marked for definiteness. It clearly does not function as a distinct NP argument of the verb (pace Beratha 1992:158). Contrasts such as me-timpug-an 'SU be thrown' versus me-timpug-an batu 'S throw stones' suggest strongly that while the non-incorporating form derives from the UP form of the transitive, the incorporating form derives from the AP (with a U NP incorporated prior to me- derivation).

18 From the habit of clasping the clenched right hand in the left (the way one squeezes a manggis fruit to open it) when speaking to a feudal superior or a high priest.

19 Satua Lutung lan I Kekua', by drs Ida Kadé Suarioka, Banjar village, Buléleng, broadcast on RRI, March 1990.
This type of incorporation seems limited in productivity; not all generic activity can be described using it. In most cases, the AP form with incorporated Undergoer NP is used instead:

(47)       NON-OCCURRING                      OCCURRING
*me-gandong panak          ngandong panak
            carry.on.hip child       
*me-baca buku            maca buku
                read book        
*me-tigitig padi         nigtig padi
                    thresh rice  

12.4.3 me- on intransitive stems

A number of intransitive verbs of shape me-ROOT occur only with Actor subjects. Examples include:

(48)      me-rérén    'stop'
          me-kaad    'clear out, leave'
          me-kecuh  'spit'
          me-tari    'propose'
          me-tari    'propose'
          me-solah   'dance.HI'
          me-solah   'dance.HI'
          me-tangi   'wake up'
          me-tangi   'wake up'

They have neither an unaffixed nominal or intransitive alternant, nor (unlike me-ROOT verbs where the subject is an Undergoer) a ROOT-Ø transitive verb counterpart. I take these verbs to be formed from inherently intransitive, but bound, roots (see also 10.4, 15.2 for arguments for this analysis). The prefix me- here can be considered to have the same signalling function as prefix /N/, indicating the presence of an Actor in the verb's semantic structure.

The evidence for this is from semantics and from distribution patterns. To avoid confusion, I will use '/N/' to refer to the 'usual' realisations of {N}: 'replacement' of the initial obstruent consonant of the stem by the homorganic nasal consonant, elsewhere, prefixing of /N(a)/ (see 7.2).

1. Agentive intransitive verbs occur either with /N/:

(49)      ng-oyong        'remain, settle'
          ng-igel          'dance.LO'
          nulis (%N-tulis%) 'write'
          ngongkong (%N-kon% %) 'bark'

or with me-:

(50)      me-solah        'dance.HI'
          me-rérén        'stop'
          me-gedi         'leave'
          me-lali         'go on a jaunt'

but probably never with both.20 This is the first evidence that /N/ and me- are in lexically determined complementary distribution, both realising a single agentive word

20 The verbs me-daar 'eat (vi)' and naa (eat (vt, AP)' both occur, but naa is never used intransitively. The only possible exception found so far is ngalah (%N-kalah%) 'give in', (from kalah 'lose') and me-kalah 'leave', though since the meanings are so distinct these surely derive from distinct, homophonous, bases.
formation process, with some bound intransitives taking one affix shape, and some the other.

2. Doublets realising the same basic lexical meaning (involving an agentive subject), but used in different styles, sometimes 'disagree' in prefix:

(51)  
| a. me-solah   | 'dance.HI' | ng-igel | 'dance.LO' |
| b. me-daar    | 'eat.LO/MID' | ng-ajeng | 'eat.HI' |
| c. me-jujuk   | 'stand.LO' | ng-adeg | 'stand.HON' |

3. The meanings of verbs formed when me- attaches to an intransitive verbal root closely parallel those produced when N- attaches to the same type of root (see 12.2.1, for parallel examples with /N/-):

(i) 'do bodily action':

(52)  
| me-daar | 'eat' | me-jujuk | 'stand' |
| me-kecuh | 'spit' | me-tangi | 'wake up' |
| me-solah | 'dance.HI' |

(ii) produce speech/sounds:

(53)  
| me-suryak | 'cheer' | me-gending | 'sing' |
| me-saut  | 'answer' | me-tengkéém | 'clear throat' |
| me-tari  | 'make a proposition' | me-kisi-kisi | 'whisper' |

(iii) move (with respect to a place):

(54)  
| me-gedi | 'leave' | me-kaad | 'clear out' |
| me-kecog | 'jump' | me-kirig | 'retreat' |
| m-ulih | 'go home' | me-keber | 'fly' |

(iv) verbs involving spontaneous, often sudden, movement, where the subject is inanimate or in a non-controlling state (12.3):

(55)  
| me-bangkes | 'sneeze' | me-cuab | 'spurt out (eg blood)' |
| me-kebyos | 'spill over (of liquid)' |

There are then three types of evidence suggesting that the prefixes of shape me- and /N/- here perform the same function. They can be seen as lexically determined variants, having the same purely iconic function of signalling the presence of an agentive component inherent in the semantics of the root morpheme (15.2). More problematic is a small number of verbs which appear to allow either actor or undergoer subjects, and also have no monomorphemic transitive counterparts which could function as derivational source for the undergoer-subject reading:

(56)  
| me-jujuk | 'stand; be stood up against sthg' |
| me-sedédéég | 'lean against; be leant against (S = U)' |
| me-bading | 'turn around; be turned around (S = U)' |
| me-celep | 'go in; be put in' |
I assume that these are again inherently agentive, bound intransitives, and that they have developed the Undergoer subject readings by analogy with the many me-ROOT verbs of that type. (All the 'problematic' me- verbs - those not analysable in terms of the detransitive or nominal stem derivations - have either just an Actor subject, or both Actor and Undergoer subjects. I have found no examples of the remaining type: me-ROOT with only an Undergoer subject, and no monomorphemic transitive counterparts. This supports the view that the 'Undergoer subject' readings of the above verbs are an extension from a more basic Actor subject orientation.)

I found only one doublet clearly showing the reverse direction of derivation: me-jemuh 'be dried; sunbake'. This has a transitive source jemuh 'dry (vt)'. The agentive sense has apparently been added only recently, since mass tourism came to Bali, and is probably calqued on Indonesian ber-jemur 'be dried; sunbake'.

The verbs me-daar 'eat (vi)' and me-gae 'work (vi)' are also apparently exceptional in having both agentive subjects and transitive counterparts (-daar, -gae). Me-gae plausibly derives regularly either from a nominal base (12.4.4) or from a homophonous agentive intransitive, while me-daar may simply involve ellipsis of an incorporated undergoer nominal (12.4.2). In its unmarked sense it means 'eat rice', and coexists with the incorporating form:

(57) suba me-daar nasi?
already eat rice
'Have [you] eaten/had a meal?'

I learnt about this ellipsis one day while eating a piece of fruit at a stall: a friend approached and asked the disconcerting question sing me-daar? ('aren't [you] eating [rice]?'). I take it that the very high frequency of use of this verb, with the meaning in (57), has lead to the ellipsis.

There appear at first to be good reasons to extend this analysis to cases where the root morpheme is a noun. For example, as in the case of intransitive verb roots, meanings typical of %N-[NOUN]% combinations (12.2.1) also occur with many verbs of structure me-[NOUN]:

(i) 'do what people typically do with (inanimate noun)':

(58) me-kutu 'look for kutu (lice) on another person's head'

(ii) 'make (inanimate noun)':

(59) me-je-jait-an 'make je-jait-an (palm-leaf offerings)'

(iii) 'produce (body product/utterance)':

(60) me-munyi 'produce munyi (sound); speak'
    me-warih 'urinate (warih 'urine.HI')'
Chapter 12

(iv) 'act like (animate noun)'

(61)  
me-musuh  'act as a musuh (enemy)'
me-dagang  'earn a living as a dagang (shopkeeper/peddler)'

In 12.4.4 I argue that these verbs in fact are better analysed as products of a different, Locative, derivation, with the agentive meanings the product of secondary, pragmatic overlay.

12.4.4 me- on nominal stems

The prefix me- combines with nominal stems to derive intransitive verbs. The stems can be either monomorphemic (me-kamen 'me-loincloth: wear a loincloth') or complex me-[ke-kedék-an]'me [RED-laugh- NOM], melaughter: laugh, banter (vi)'). The meaning of the derived verb depends on the meaning of the nominal stem ([NOUN]), and on the pragmatic relationship holding between the S NP and [NOUN]. The semantic role of the pivot/subject NP is in all cases, I will argue, Location. Often, though, it appears to be better described as Actor, however this is the result of secondary pragmatic factors.

1. In the majority of cases, the predication me-[NOUN] of an S is possible only where the stem nominal is thought of as 'naturally occurring with/at' S (hence the locative semantic role associated with S). The notion 'naturally occurring with' subsumes a variety of non-linguistic relationships, including:

(i) where the referent of the nominal is thought of as an inalienable part of S. Thus [NOUN] can include body parts, thoughts, one's breath:

(62)  
bo  'odour'
asa  'flavour, feel'
bulu  'feather/fur'
keneh  'thought'
adan  'name'

as well as some family relations:

(63)  
kuren  'spouse'
kuren  'be married'
nyama  'sibling'
nyama  'be related as sibling'
panak  'offspring'
panak  'to have offspring'  
matua  'parent-in-law'
matua  'have as matua'

(64)  
tiang me-matua ajak ia
1 me-parent.in.law with 3
I have him/her as parent-in-law

also parts of inanimate objects, flowers on a tree, and so on:

(65)  
raab  'roof'
raab  'have a roof'
bunga  'flower/interest'
bunga  'flower (vi)/accrue interest'

21For pragmatic reasons, this is limited to animal subjects only, although panak also means 'child'.
One can thus say

(66) \textit{batis-né me-bulu}  
feet-3 body.hair  
\textit{His feet are hairy}

but not:

(67) *\textit{mata-n-né me-bulu}  
for 'He has a lash in his eye'

As pointed out by Kersten (1984), the relationship of everyday (alienable) possession must be expressed using a different construction with the transitive verb -\textit{gelah} 'possess':

(68) \textit{ia ngelah uma}  
3 possess ricefield  
\textit{He owns (a) ricefield}

*\textit{ia m-uma}

More generally, the S NP and the nominal stem ([NOUN]), while not in a relationship of inalienable possession, occur naturally and appropriately together. Such a relationship occurs for example, if the [NOUN] is a manufactured object, and S is using NOUN for the purpose for which [NOUN] was made: eg if the nominal stem is an item of clothing, the S NP must be a 'wearer':

(69) \textit{sepatu 'shoe' me-sepatu} 'be wearing shoes'  
\textit{kamen 'sarong' me-kamen} 'be wearing a loincloth'  
\textit{pupur 'face powder' me-pupur} 'be wearing powder'

It is not simply enough for the item indicated by the incorporated nominal to be accidentally located on the subject NP:

(70) *\textit{lima-/kamen-n-né me-pupur}  
hand-/cloth-GEN-3 me-powder  
*His hand/loincloth had powder on it.

However nor is it a requirement that the S NP be a volitional Actor; such verbs are appropriately used where the subject is, say, a baby who as just been bathed:

(71) \textit{suba me-pupur, i komang?}  
already me-powder DET K.  
\textit{Is Komang wearing powder?}

or even a statue, or an animal:
Chapter 12

(72) **tohog-e me-kamen poléng**
    statue-DEF me-loincloth chequered
    *The statues have chequered loincloths on them*\(^{22}\)

    **kucit-e me-planka**
    piglet-DEF me-collar
    *The piglet has a wooden collar around its neck* (to stop it wandering)

    **botol-e me-sengseng**
    bottle-DEF me-stopper
    *The bottle has a stopper*

The meaning of me-[NOUN] verbs formed in all the ways described above can be summarised as:

(73) S me-[NOUN]:

    a. [NOUN] is at S
    b. People can think of [NOUN] as normally being with S

There are thus two important components: the S NP bears a locative semantic role, it is the place where the NOUN is located; and NOUN 'belongs with' S in some way. With these verbs then the degree of agentivity of the S NP is not important. An agentive semantic element may be apparently involved, but this, if present, is secondary 'pragmatic overlay', rather than part of the inherent verb semantics.

And though the above examples all involve durative states, aspectually 'eventive' verbs can also be formed with this construction: thus me-suah 'use a comb/have one's hair combed (eg, of a baby)'.

When a NOUN 'belongs with' another is clearly a highly culture-specific perception, cf (72), also siap-e me-taji 'the [fighting] cock is wearing spurs'. In other cases the reasoning is perhaps universal. For example, mud occurs naturally on unsealed paths, so rurung-e me-endut 'the path is muddy' is acceptable. On the other hand, while mud does occur naturally on feet, it does not belong there, hence the unacceptibility of *batis-né me-endut* 'His feet are muddy'. Similarly the relationship between many other objects which commonly associate, but in an undesirable way, such as dandruff or lice with hair, cannot be expressed with this construction:

(74) **bok-né me-kutu/me-kepur**\(^{23}\)
    hair-3 me-louse/me-dandruff
    *His hair had lice/dandruff in it.*

Contrast the following two elicited examples, only the first being acceptable:

---

\(^{22}\)On some religious holidays statues are dressed in loincloths.

\(^{23}\)me-kutu does occur, but with a different meaning, see below.
Verbal Prefixes

(75) sing me-getih, mua-n-né
NEG me-blood face-GEN-3
'His face was bloodless [pale from shock, eg on being told bad news]

*lima-n-né me-getih
hand-GEN-3 me-blood
*His hand had blood on it [eg after slaughtering a pig]

The S me-[NOUN] construction is also not used where S is a food and NOUN is one of the ingredients used to make it: I presume because the ingredients do not co-exist with S, but rather are destroyed to become S.

In those cases where a noun 'X' does not meet the criterion of 'belonging naturally to' the S NP, a variant of the above construction is used: S m(e)-isi X, where isi is a noun stem variously glossable as 'contents [of S]', surface matter on [S]; thing in contact with [S]', and X is an incorporated nominal. The grammatical counterparts of the ungrammatical sentences above are thus:

(76) lima-n-né m-isi pupur
His hand/loincloth had powder on it.

batis-né m-isi endut
His feet had mud on them.

bok-ne m-isi kutu!
His hair had lice.

lima-n-né m-isi getih
His hand had blood on it.

2. With some me-[NOUN] verbs the S NP appears to be a true Actor rather than a Locative. Here are some examples:

(77) kutu 'louse' me-kutu 'look for lice (on s.o. else)'
ceki 'card game' me-ceki 'play ceki'
se-sangi 'vow to gods' me-se-sangi 'make a vow to the gods'
de-demen-an 'lover' me-dedemen-an 'court (vi)'
atur-an 'offering' m-atur-an 'make offerings'

Pragmatic factors again determine the precise meaning of each individual verb. I originally analysed verbs like this as being produced by a lexically variant of the {N} agentive derivation, cf (28), section 12.2.1:

(78) S me-[NOUN]:
S did something (Y)
Thinking of [NOUN], people think it is normal to do Y

If this were the case, these verbs would pattern with the subset of agentive intransitives taking prefix me- (12.4.3). On reflection, the more economical analysis is that these me-[NOUN] verbs also derive straightforwardly from (73). An agentive pragmatic
interpretation, such as that proposed in (78), is no doubt added when these verbs are predicated of potentially initiating actors. A quick survey of the verbs in examples (62) to (76) will show that most of them naturally collocate with inanimate subjects, or with 'non acting' animate subjects. Those in (77) on the other hand naturally collocate with deliberately acting Subjects.

Verbs which have both agentive and 'passive' subjects, according to context, such as me-suah 'comb one's own hair/have one's hair combed', me-pupur 'put on powder/have powder put on (by another)', and me-kamen 'put on a loincloth/have a loincloth draped around one (by another)', give the best clue to the underlying unity of all verbs of structure me-[NOUN].

More problematical is a verb like me-kutu, 'look for lice', where the subject of the verb can only be the searcher, never the person being inspected for lice. This does not at first glance appear to satisfy the '[NOUN(kutu)] is at S' component of (73). This again is a question of pragmatics: lice in Balinese are appropriately found 'at' (ie 'with', when caught by) someone looking for them in another's head.

What at first appears to be another class of stem nominals occurring with Actor subjects is that of terms of address (including pronouns and certain kin terms). Where these form the stem, the resulting meaning of S me-[NOUN] tekén (with) X is 'S call X by [NOUN(address term)]':

(79)  m-embok  'me-elder.sister: address (someone) as embok\[24\]
     me-cai    'me-2: call someone cai ('you')'

These can also be seen as deriving from (73), with the added pragmatic element of (78): the set of conventionally accepted behaviours associated with terms of address is precisely to use those terms of address towards particular people.

Semantically then, while me-[NOUN] verbs seem at first to be divisible into two types: those with Undergoer (Locative) subjects, and those with Actor subjects, they are in fact better analysed as products of a single Locative derivation, with further pragmatic overlay, depending on the nature of the nominal occupying the Subject slot.

Formally, the nominal stem can be either monomorphemic, as with nearly all the examples discussed so far, or - in contrast to /N/-[NOUN] verbs (12.2.1) - morphologically complex:

---

\[24\]There are thus two homophonous verbs m-embok (tekén X 'with X'), the first with a Locative subject, meaning 'have (X) as elder sister', the second (here) with an Actor subject. These derive from the different meanings of embok: 1. a kin relationship; 2. a term of address (used towards someone in that kin relationship).
Verbal Prefixes

(80) angkih- (bound root) -sangi (bound root)
  angkih-an 'breath' se-sangi 'oath'
  m-angkih-an 'breathe' me-se-sangi 'make oath'

kapal 'boat'
kapal-kapal-an 'toy boat'
me-kapalkapalan 'play with toy boat'

-sangi se-sangi me-se-sangi

Verbal Prefixes

(81) me-[batis lantang]
  me-[bok dadua]
  'me-feet long: have big feet'
  'me-hair-two: be going grey'

or by other nouns:

(82) me-bo-enceh
  me-mua-bojog
  'me-odour-urine: smell of urine'
  'me-face-monkey: have a monkey face'
  me-[panak bikul]
  me-bo-andus
  'me-offspring-mouse: be spoilt, be a crybaby'
  'me-odour-smoke: smell of smoke'

or by both:

(83) me-kurenannan cina sugih tua bongol-bongolan
  me-spouse china rich old RED-deaf-an
  'have a rich old half-deaf chinese spouse'
  [from 'Mirah', by Putu Sedana, in Bagus and Ginarsa 1978]

However neither the nominal base nor the incorporated nominal may be definite:

(84) *me-sepatu-ne 'me-shoe-DEF'
    *me-kurenannan cina-ne 'me-spouse-chinese-DEF'

Names qua names behave as non-definite in this respect. Contrast the perfectly acceptable:

(85) ia m-adan Luh Susilawati
    3 m-name L. S.
    'She is named Luh Susilawati'.

with the ungrammatical:

(86) *ia me-kurenannan Luh Susilawati
    3 me-spouse L. S
    (For 'he is married to Luh Susilawati')

Nor can the nominal root occur with {-n}, the 'genitive' linker:

25See discussion of m-ulih, 11.3.2.1, on why m-ulih-6 ['come] to the [=my] house' is not an exception to this statement. The one potential exception noted was: me-karya 'work [vi]; make.HI (vt)'; sira me-karya banten-6 'who made the offerings?'. This could be accounted for if me-karya, also exceptionally, derives from a complex nominal base pe-karya 'work (n)', plus prefix {N}.
This is despite the fact that in most non-derived noun-noun compounds, occurrence of {-n} is obligatory or at least permissible after a vowel-final head: bo-*{n} enceh 'odour of urine', mua-*{n} bojog-e 'face-GEN-monkey-DEF: the monkey's face'.

12.5 The confix me-[NOUN]-an

I have argued that nearly all verbs elsewhere considered to have a 'me-...-an' confix are better analysed as involving me- prefixed to a complex stem of shape ROOT-an: either the detransitive stem (12.4.1), or else to a complex nominal (12.4.4). There remains a small residue of cases which cannot be accounted for in this way. These are all of structure me-NOUN-an, and have a meaning broadly 'do what one habitually does with NOUN' (cf the derivation type (78)). Most of the examples I found involved travel:

(89) me-montor-an 'me-car-an; travel by car'
me-jaran-an 'me-horse-an; travel by horse'
me-dokar-an 'me-buggy-an; travel by buggy'

but not all:27

(90) cara anak me-drama-an
like person act.in drama.
[you look like] someone acting in a drama ['folk opera']

Complex nominal stems of shape *montor-an, *jaran-an, *drama-an and so on do not occur. Equally, there exist no transitive verbs of shape *montor-ang/-in, *jaran-ang/-in, *drama-ang/-in, which could be the source of a detransitive stem. I therefore assume that in this small subclass only, the derivation does involve the simultaneous affixation of both me- and -an.

Based on the nominal roots occurring with it, this small subclass looks as though it is a recent innovation (Malcolm Ross, p.c.). It may have arisen on analogy with the many derived nominal bases which carry the suffix of shape -an (12.4.4). Cohn (1989) argues for the existence of confixes in the related language, Indonesian: it is not unlikely then that the same morphological type could arise in a related language. There is strong evidence, however, that such an analysis is not motivated for the vast majority of lexemes of shape me-...-an (11.4, 12.4.1).

26 {-n} is not realised after a consonant, and in some lexicalised compounds is either optional or does not occur (bal6{-n} bafljar 'village hall').
27 Kersten (1984) lists reduplicated forms like me-ubad-ubad-an 'me-RED-medecine-an: work as a folk doctor' - this could contain a reduplicated detransitive stem, from ubad-in 'give medecine to'.
13

Pragmatics and morphosyntax:
the meanings of actorless Undergoer Pivot constructions in
Singaraja Balinese

13.0 Introduction

The discussion of the principal verb prefixes in chapter 12 is incomplete, in that the prefix ke- is not discussed. This prefix occurs on Undergoer Pivot verbs, and is of extremely high frequency in everyday use. Previous descriptions of this prefix, and of the very similar functions of both type ke-UP and Bare Stem ('zero-prefix') Undergoer Pivot verbs, are inaccurate (including Clynes 1989). This is because, as it turns out, the meaning or 'discourse function' of both Undergoer Pivot types (9.4.2, 9.4.3) varies according to discourse genre - a phenomenon predicted to occur by Foley and Van Valin (1984):

It is entirely possible and even probable that the function of a form or construction could vary across different discourse genres (1984:2)

Here I show that to correctly interpret the implied actor of both Bare Stem and ke-Undergoer Pivot verbs - where they occur without an overt actor NP - information is required not only about discourse genre (narrative, conversational; high style or low, and so on) but also information about the syntactic context, the semantic class of the verb, the aspect and modality, and so on.

The patterns described in this chapter show, once again, how meaning assignment in Balinese cannot be described in simple 'agglutinating' terms (15.1, 15.2); in the cases described here meaning is assigned not at the level of the morpheme, nor even of the word/derivation, but only as a function of the detailed grammatical and discourse context. Foley & Van Valin's further assumption is correct: adequate linguistic description, be it of verbal morphology, syntax, or any other area, can only be done given information from a variety of pragmatic contexts.

The chapter offers, I hope, an improved analysis not just of the ke- morpheme, but of both Bare Stem- and ke-Undergoer Pivot constructions, in their various contexts of use. The analysis in Clynes (1989), which assumed two homophonous prefixes of shape ke-, is not supported; nor are other analyses proposed in the past (13.2.1).
The chapter also brings out differences in the syntax and semantics of Undergoer Pivots in main clauses as opposed to subordinate clauses in Singaraja Balinese. These differences are unnoted in previous descriptions of Balinese (Beratha 1992, Kersten 1984, Oka Granoka et al. 1985, Barber 1977, Antara 1989, Artawa 1992).

Motivating the phenomena I will describe are conventions pervading and informing language use in Bali. These require that status distinctions between speaker, addressee, and third parties, (as well as the perceived formality of context) be expressed through a web of lexical choices: this in turn produces distinct High, Mid and Low speech ‘levels’ or ‘styles’ (Clynes 1989). To give a simple, if hackneyed example, the core meaning ‘s/he has eaten’ can be rendered in a variety of ways, according to the relative status of speaker/addressee and eater, and the nature of the relationship between between speaker and listener. Just some of these ways include:

\[
\begin{array}{ccc}
\text{ia suba medaar} & \text{neutral} & \text{non-formal} \\
3 & \text{already eat} & \\
\text{ipun sampun nunas} & \text{humble} & \text{formal} \\
3\text{DFR already.HI eat.DFR} & \\
\text{ida sampun ngrayunang} & \text{high caste} & \text{formal} \\
3\text{HON eat.HON} & \\
\text{ida suba ngrayunang} & \text{high caste} & \text{non-formal} \\
\text{ia) sampun ngajeng}^1 & \text{respected} & \text{formal} \\
\text{eat.HI} & \text{non-high caste} & \\
\end{array}
\]

This chapter focusses largely on usage in the everyday, 'Low' style, exemplified by the first sentence in (1): this is the language of intimacy, a kind of 'refuge' from the status-ridden Mid and High styles (cf all other sentences in (1)). Yet even in this style, the Balinese are often at pains to avoid the many 1st- and 2nd-person pronouns, which are all, in most contexts, pragmatically marked as 'too crude', 'too distant', 'too deferential', and so on.\(^2\) A variety of strategies are used to avoid these pronominals, in Singaraja and, I believe, generally, in lowland dialects at least. The use of -né the 'definite' morpheme on nominals is one of these, also to my knowledge not previously described (see Appendix L).

One of the principal strategies is the subject of this chapter: the use of the two actorless Undergoer Pivot types. The first of these is the bare stem Undergoer Pivot (9.4.2):

---

1. In this context the unmarked 3rd person pronominal would be avoided, hence the brackets - titles/names would be used instead. The choice of ng-ajeng is appropriate where the eater is respected but not necessarily of higher caste.

2. In Peguyangan, regularly used 1st person pronominals include: (aw)aké, iraga, (t)iang, titiang, i dêwék and (in folk literature) icang; each has a variety of pragmatic functions.

274
Do you know that? (lit. Is that known by you)
I know! (lit. It is known by me)

and the second the ke-Undergoer Pivot (9.4.3):

What did you take? (lit. What was taken? understood: by you(*me))
I took this (lit. This was taken by me(*you))

As the examples suggest, the participation of an Actor is always implied when these are used in matrix clauses; and in most contexts that implied Actor can only be one of either Speaker or Addressee, not both, and not a 3rd party. Rules of pragmatic interpretation supply the identity of the actor, according to whether the Undergoer Pivot stem carries a prefix or not. To accurately characterise each Undergoer Pivot type, then, one must consider them together, looking at the contrastive ways the two are used when occurring in the same context.

The chapter has the following structure. Additional background facts about both Undergoer Pivot constructions are provided in 13.1. After considering previous descriptions (13.2.1), contexts where the two Undergoer Pivot verbs can occur, with and without a difference in interpretation, are considered in 13.2. Complications due to factors such as Style, Verb Class, and pragmatic factors are described in 13.3. Finally, I draw some more general conclusions in 13.4.

13.1 Two Undergoer Pivot types in Balinese: a sketch description

The two Undergoer Pivot verb types were briefly described in 9.4.2 and 9.4.3. Here I review and add to those descriptions, focussing in particular on the use of these forms in Actorless clauses. (The third Undergoer Pivot-like derivation, the me-detransitive (12.4.1), derives stative verbs which are semantically actorless; the use of me-detransitives is therefore inappropriate where the speaker wishes to imply the participation of an Actor. This type is not considered further in this chapter.)

In glosses and translations the active sense is used, for brevity and naturalness. There are no inanimate pronouns. These are always represented by a 'null pronoun'.

275
Chapter 13

13.1.1 Bare stem-Undergoer Pivot verbs

As the name indicates, bare stem Undergoer Pivot verbs consist simply of a bare verb stem, without a specific 'Undergoer Pivot' marker. They may occur with a cliticised actor pronominal (see also 9.4.2). A typical simple bare stem Undergoer Pivot clause has the structure:

(4) [Undergoer]NP Verb = [Actor] ([Actor]PP)

as in these examples:

(5)  
a. I Made jagur = awaké  'Madé was hit by me.'
    DET M. hit = 1

b. I Made jagur = cai  'Madé was hit by you.'
    = 2

c. I Made jagur = a  'Madé was hit by him.'
    = 3

d. I Made jagur = a tekén i Putu  'Madé was hit by Putu.'
    DET P.

While examples (5c & d) are perfectly natural, there is something artificial or awkward about sentences like (5a & b), in most contexts - except in oral literature and formal high style, where they are frequent. This is because they use 1st- and 2nd-person pronouns. Attempts to elicit or use sentences like these are usually rejected as 'not Buleleng speech'. In their place, actorless Undergoer Pivot verbs, either bare-stem or ke-Undergoer Pivots, are typically used.

13.1.2 'Actorless' bare stem-Undergoer Pivots

Actorless bare stem Undergoer Pivots do not carry a cliticised pronominal. Even though they are formally 'actorless', in main clauses they can only be interpreted as involving a 1st- or 2nd-person Actor in the predicated action:

(6) dija lakar jang?

where will put

Where will I/you(/*he) put [it]?

Where the Actor of a bare-stem Undergoer Pivot is 3rd person, the pronominal clitic = a cannot normally be omitted, even where there is no specific referent: 4

(7) nangka ané nu nguda ketéwel-ang = a

nangka REL still young ketéwel-APP = 3

Nangka fruit which is still young is called 'ketéwel' they call young nangka fruit 'ketéwel'

---

4 The 3rd person pronominal is omitted where a generic actor is incorporated directly into the Undergoer Pivot verb:  i komang cegut = legu
DET K. bite = mosquito
Komang was bitten by a mosquito.
Such forms are best glossed with generic 'they' [= 'people in general'] to render the effect of the clitic pronominal. The only case I have found in which forms like Bare Stem Undergoer Pivots (and occurring without pronomial clitics) have truly 'actorless' semantics is in certain subordinate clauses (13.3.4.2).

From the facts outlined in the preceding paragraphs, it could be concluded that 'actorless' bare stem Undergoer Pivots in main clauses are simply those where a 1st- or 2nd-person pronominal has been optionally ellipsed. Kersten (1984) assumes this; however other hypotheses are possible, for example:

- Actorless bare Undergoer Pivots are truly so underlyingly, without ellipsis of a pronominal.
- Bare Undergoer Pivots derive from ellipsis of the ke- prefix (at least in those contexts where the two are mutually substitutable, cf below).

Going against the second of these hypotheses is the evidence that the Bare Stem form is the 'unmarked', and therefore, more basic, form (13.3.2).

13.1.3 ke-Undergoer Pivot

The second Undergoer Pivot type is the ke-Undergoer Pivot, that where the verb stem carries the prefix ke- (9.4.3). In the formal styles - either the High spoken language (9), or Low Literary style ((10) - the syntactic frame is generally:

(8) [Undergoer]NP ke-Verbstem ([Actor]PP)

(9) \textit{ida betara Rama ke-iring olih ida sang Laksamana}  
\textit{Lord Rama was followed by Lord Laksamana}

(10) \textit{Lentir [...] ka-tutug-in baan i buyung}  
\textit{Lentir was followed by Fly}\textsuperscript{5}

A second possibility is where a 1st/2nd-person Actor pronominal occurs as a postclitic:

(12) \textit{[i seruni] sampun ja ka-enten-in = tita}\textsuperscript{6}
\textit{I've already warned Seruni.}

The syntactic frame is quite different in the everyday Low Style: the ke- Undergoer Pivot must be formally actorless in main clauses\textsuperscript{7}. The actor cannot occur in a separate PP:

\textsuperscript{5}From 'Satua I Lengkir', told by Madé Narsin, Peguyangan, Singaraja, in Bawa & Jendra 1981:63. It should be noted that the transcription in Bawa & Jendra has many errors of transcription, and cannot be relied on as an accurate reflection of Peguyangan usage.

\textsuperscript{6}From the play Nang Kepod, by Ketut Aryana of Banyuning, in Bagus & Ginarsa 1978.

\textsuperscript{7}Relative clauses allow different possibilities, see section 13.3.4.1.
Chapter 13

(13) I Made ke-jagur (*tekén awaké)
DET M. ke-hit with/by 1

[I] hit Made.

or as an enclitic:

(14) I Made ke-jagur(* = awaké)

There is a second striking distinction between the use of ke-Undergoer Pivots in everyday speech as opposed to formal styles, this time pragmatic. In everyday Low style, the understood Actor of an actorless ke- Undergoer Pivot, like that of a bare Undergoer Pivot, must be 1st or 2nd person (as in examples (16a & b):

(16) a. I Ketut k-ajak?
DET K. ke-take
Was Ketut taken [by you/*him]? 

b. aa, k-ajak
yes
Ketut was taken [by me/*him])

In contrast, in formal styles, actorless ke-Undergoer Pivots are most usually understood to be 3rd person, as in this stage direction (only the words in bold are glossed):8

(17) Pan Langkir (Kagelekang poosé. Masaut munyiné baat): [...] 

ka-gelek-ang poos-é
ke-swallow-APP saliva-DEF
[He] swallowed [his] saliva

Note that the vocabulary is from the low style, but is used in the formal context of a literary text. The same phrase, if uttered in conversation (or indeed by one of the characters in the play), could only mean '[I] swallowed [my] saliva', never the 3rd-person Actor interpretation. (Bpk Gede Dharna, the author of the text, kindly confirmed this fact for me.)

13.2 Are the two Actorless Undergoer Pivots freely interchangeable?

In Low Balinese then there are 2 kinds of 'actorless' Undergoer Pivots, both understood to involve either 1st- or 2nd-person Actors, and not third-person Actors. Both are of high frequency in everyday discourse. Can one simply use either form, interchangeably? Are both 1st-person ('1ACT') and 2nd-person Actor ('2ACT') readings possible, with both forms? The answers to these two questions turn out to be 'sometimes yes, sometimes no'.

8From Kobaran Apine, by Bpk Gede Dharna, of Sukasada, in Bagus & Ginarsa 1978.
13.2.1 Previous descriptions of Actorless Undergoer Pivots

Previous descriptions that I am aware of hardly address the questions posed above. On bare Undergoer Pivots, Kersten (1984:94) says:

[jika pelaku orang pertama dan kedua] orang kadang-kadang melalaikan menyebut pelaku, jika cukup terang siapa yang dimaksud.

[j when it is 1st- or 2nd-person] people sometimes omit to state the actor, if it is clear enough who is intended.9

The factors which determine when it is 'clear enough who is intended' are not stated. Similarly, on ke-Undergoer Pivots, Clynes (1989:168-9) incorrectly suggests that 1 ACT and 2 ACT readings are both possible in all contexts in Low style. The same is suggested by Beratha (1992:163, examples 5-43a & b). Kamus Bali (1990:x) implies that the prefix ke- marks either actorless/3rd-person Undergoer Pivot verbs (BI: di+verb) or Undergoer Pivots with 1st person singular Actor (BI: ku+verb), however it gives no indication of the contexts in which each meaning occurs. It does not mention the 2 ACT interpretation.

13.2.2 Some first answers: Can the two Undergoer Pivot types be freely interchanged?

The distribution and interpretation of the two Undergoer Pivot types depends in fact on a variety of factors, including the following:

• style
• syntactic frame
• verb class
• aspect/modality
• other discourse factors

In the following sub-sections I briefly illustrate some of these factors.

13.2.3 Cases where the two Undergoer Pivots are interchangeable, with no meaning difference

In some contexts one can use either bare Undergoer Pivot or ke-Undergoer Pivots interchangeably without giving rise to a meaning difference. This is generally the case in clauses where aspect/mood is marked lexically (cf words underlined in the following examples):.

9See Barber 1977:218 for a similar comment: Barber's work is highly derivative of Kersten's.
10These examples all involve 1 ACT readings, but from elicitation it appears that 2 ACT readings are possible in the same contexts.
Chapter 13

(18) a. tendas-né lakar ka-sepeg-in [...] 
   head-3 will ke-split

   b. lakar sepeg-in tendas-né 
   [I] will split his head open.\textsuperscript{11}

(19) a. lesung ja malu! nyan masi tulung-in nyidi 
   pound PART first later too help-LOC sieve 
   'You] pound it first, later [I'll] help sieve'\textsuperscript{12}

   b. nyan ke-takon-in Putu 
   later ke-ask-LOC P. 
   Later [Till] ask Putu

(20) suba (ke-)beli-ang jaja, jani sprite tagih=a 
   already (ke-)buy-sc cake now S. ask.for= 3 
   [I've already bought you cake, now you ask for Sprite!]

13.2.4 Interchangeable: but meaning difference

In everyday speech, simple clauses unmarked lexically for aspect or mood occur with high frequency. (I refer to these as 'unmarked clauses' in the rest of this chapter.) Both Undergoer Pivot types can generally occur in them, but they are not interchangeable: use of one form will imply 1ACT, and of the other 2ACT.\textsuperscript{13}

(21) a. ke-juang? 
   take 
   2 *[you/took it]? *1/*3 
   *I/he took it.

   b. ke-juang. 
   1 *[I/took it] *2/*3 
   *you/he took it.

(22) a. juang? 
   1 *[I'll take it]? *2/*3 
   *you/he take it?

   b. juang! 
   2(IMP) *[you/took it]! *1/*3 
   *I/he'll take it

That is, in simple questions ke-Undergoer Pivots almost always imply a second person actor (example (21a)), while in simple declarative clauses, they are understood to involve a 1st person actor, not 2nd or 3rd person (example 21b).\textsuperscript{14} Bare stem forms in the same intonational contexts require the reverse interpretations (examples (22a&b)).\textsuperscript{15}

\textsuperscript{11}Examples (18a & b) are both from Satua Pengalu Mati Tepen Buah, Kirtya Vlb 1317/4 (Br Dangin Peken [=Delod Peken], Singaraja). 2ACT readings for both clauses would also be possible.

\textsuperscript{12}Satua Bawang Kesuna, told by Ibu Ketut Supani, Peguyangan, March 1991.

\textsuperscript{13}See sections 13.3.2.1, 13.3.2.2 and 13.3.5.2 for some qualifications.

\textsuperscript{14}To make a simple declaration that so-and so was done by "you" (2nd person), the only acceptable avoidance strategy is to use 3rd person pronouns; ke- and @Undergoer Pivots cannot be used here, see example (20) and Appendix L.

\textsuperscript{15}Strictly, example 22b) is an imperative, rather than a Undergoer Pivot: a 2nd-person pronoun cannot be present, as it can optionally in non-imperatives.
Another context in which substitution of one form for the other produces a meaning difference is in 'interrogative' clauses preceded by the particle (d)adi, which adds a meaning roughly like:

(23)  
\text{dadi} [X]! (X = \text{clause})

I did not think that [X] would happen,  
I feel something because of this  
I ask 'why [X]'?

Note that the meanings assigned (at least in elicitation) are the reverse of those assigned in simple unmarked interrogatives:

(24)  
\text{adi} \text{ jagur adi-n tiang-é}  
\text{PART hit yngr.sibl.-GEN 1-DEF}

'Why did you [*I] hit my adik?'

(25)  
\text{adi} \text{ ke-jagur adi-n tiang-é}  
\text{PART hit yngr.sibl.-GEN 1-DEF}

'Why did I [*you] hit my adik?'

One is tempted to begin constructing paradigms of verbal 'inflection', as in the following table:

<table>
<thead>
<tr>
<th>Actor</th>
<th>Declarative</th>
<th>Interrogative</th>
<th>'Adi Interrogative'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ke-</td>
<td>Ø</td>
<td>ke-</td>
</tr>
<tr>
<td>2</td>
<td>a</td>
<td>ke-</td>
<td>Ø</td>
</tr>
<tr>
<td>3</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

Table 13.1: Person of implied Actor of actorless Undergoer Pivots, and verb affix shape, in simple clauses unmarked for aspect: a first approximation.

However this table, while accounting for much of the data, is clearly insufficient. For example, it incorrectly suggests that actorless bare-stem Undergoer Pivots do not occur in unmarked declaratives (see 13.3.2.1), and oversimplifies the meanings attributable to verbs occurring in these contexts (see 13.3.3).

---

16 Both forms also occur in subordinate clauses introduced by yén, lén, amon, 'if', again with different meanings. The bare-stem form is unusual in being truly semantically 'actorless' (see 13.3.4.2).

17 See footnote 14.
13.2.5 Not interchangeable: only one form grammatical

The third, minor, case is where one of the two Undergoer Pivot types simply cannot occur. Thus bare stem verbs in most simple clauses (unmarked lexically for aspect/mood) must be interpreted functionally as imperatives, not as declarative Undergoer Pivots: see example (22b) above. This also accounts for the unacceptibility of (26), when uttered as a simple declaration:

(26) \(i \text{ ketut takonang} \)
\(\text{DET K.} \)
\(*Ketut was asked-about [by me/you/him].\)

This sentence can only have the imperative reading 'Ask Ketut!'. To get the declarative reading, the ke-Undergoer Pivot must be used:

(27) \(i \text{ Ketut ke-takon-ang} \)
\(\text{DET K. ke-ask-APP} \)
\([I] asked about Ketut\)

In certain other cases 'bare-Undergoer Pivots' occur, but not the ke-Undergoer Pivot alternants. These are perhaps infinitival constructions - see 13.3.4.2.

13.3 Other Complicating Factors

The following sub-sections briefly summarise some other factors influencing occurrence and interpretation of the two actorless Undergoer Pivots.

13.3.1 Style

The striking difference in the interpretation of ke-Undergoer Pivot verbs in 'formal' versus Low conversational styles was discussed in section 13.1.3 (see examples (16) and (17)).

In elicitation, my consultants preferred to reserve ke-Undergoer Pivots for 3rd person actors in High style; for 1st- and 2nd-person Actors in High style, bare stem Undergoer Pivots without ellipsis of the pronominal were preferred. This is presumably because the pragmatics of explicit (deferential) 1st- and (honorific) 2nd-person pronominal use accords with the overall pragmatics of high style use.

13.3.2 Exceptional Verb classes

At least two verb classes, verbs of perception and of commanding, are exceptions to the patterns outlined in table 13.1, in that actorless bare-stem Undergoer Pivots of verbs of

---

18 Verbs of certain classes are exceptions to this: see section 13.3.2.1.
these classes can occur in simple declaratives. The two classes are discussed in turn. I
argue that the exceptionality of these classes has a pragmatic explanation. From the
phenomena described here it can be concluded that the bare stem UP is indeed the
unmarked UP form (as its lack of affixation suggests).

13.3.2.1 Verbs of perception

I was once asked the following (perfectly grammatical) question, which uses first a bare
Undergoer Pivot, then a ke- Undergoer Pivot:

(28) ingsap-in apa ke-dep-in?
    forget-LOC or ke-leave-LOC
   [Did you] forget [it], or leave [it] there [on purpose]?

Table 13.1 predicts that example (28) should be ungrammatical, that both verbs will
carry ke-, as in (29), which is equally acceptable:

(29) k-ingsap-in apa ke-dep-in?
    forget-LOC or ke-leave-LOC

Compare (30), where the sentence becomes inacceptable if ke- is not prefixed to dep-in:

(30) *ingsap-in apa dep-in?

The reply to such a question could also use either the bare stem or ke-Undergoer Pivot,
even though elsewhere only the ke-Undergoer Pivot would be acceptable (see section
13.2.5):

(31) (k-)ingsapin
    [I] forgot

ingsap-in then can occur with or without ke-, but not dep-in. In fact ingsap-in is just
one of several verbs, all to do with mental perceptions, for which ke- is optional and
usually omitted. Other verbs patterning in this way include: 19

(32) dingeh 'hear' tepuk-in 'see'
kadén 'believe' tingal-in 'see'
tawang 'know' uning 'know.HI'
takeh 'guess' adek 'smell (vt)'

The exceptionality of this class is straightforward and derives from their pragmatics.
For these 'primary' perception verbs, it simply does not make sense to ask another person
naive information-seeking questions like

(33) 'do I hear/see/forget/know [X]?'

It equally does not normally make sense, in simple (non-ironic) declaratives to inform
people:

19 Transitive verbs of emotion, such as demen-in 'like', gedeg-ang 'hate', do not pattern in this way.
(34) 'you (can) hear/see/forget/know X'

With actorless Undergoer Pivots involving perception verbs, then, when a 3ACT reading is ruled out, there is no further potential ambiguity: I tell you about my perceptions, I ask you about yours, not the reverse. So declarations about perceptions using Actorless Undergoer Pivots must imply 1st-person actors, and questions must imply 2nd-person actors. A single Undergoer Pivot form can therefore be used unambiguously for both tasks: the bare root Undergoer Pivot, elsewhere reserved for 1st person usage in unmarked questions, is used, though the ke-prefix can be optionally used in the same way as it is with other verbs (Table 13.1).

This use of the bare stem UP in contexts where ambiguity is not possible suggests strongly that the bare stem UP is the unmarked UP form (as its lack of affixation indicates).

13.3.2.2 Verbs of commanding

Another minor class of verbs which regularly occur without the ke-prefix expected in declaratives are verbs of commanding, used with a perfective sense:

(35) orah-in sing!
    tell-LOC no
    '(I) told (you) 'no'!

(37) tunden nyemak pipis
    order take money
    (I) told (you) to get the money

The reason for the optionality of the ke-Undergoer Pivot with 1st-person Actors perhaps derives from the performative function of these verbs: performatives, as in other languages, require a 1st person Actor. The high frequency of 1ACT readings with these verbs may have made this the 'unmarked' reading. As in the case of perception verbs, once there is little or no question of ambiguity between 1ACT and 2ACT readings, the bare Undergoer Pivot becomes available for use in simple declaratives.

These verbs also have the peculiarity that in declaratives, bare Undergoer Pivots are normally understood to involve a 1st-person Actor plus 2nd-person Addressee/Undergoer, while ke-Undergoer Pivots are usually understood to have a 3rd person Addressee:

(38) ke-tunden nyemak pipis
    order take money
    (I) told (him) to get the money

These alternate 'portmanteau' readings are apparently merely a strong tendency rather than an absolute requirement, at least with the ke-forms.
Again, the use of the bare stem UP in contexts where ambiguity is not possible supports the conclusion that the bare stem UP is the unmarked UP form.

### 13.3.3 Interplay of Aspect & Modality

As mentioned above, Table 13.1 oversimplifies the semantics of actorless Undergoer Pivots used in unmarked clauses. ke-Undergoer Pivots in unmarked declaratives imply not simply 1st-person actor, but also perfective aspect:

\[(42) \]
\[\begin{align}
   a. & \quad \text{ke-juang?} \\
      & \quad \text{take} \\
      & \quad [\text{you} \text{ took } \text{it}]? \\
      & \quad *\text{are you taking/will you take it?}
   \\
   b. & \quad \text{ke-juang} \\
      & \quad [I \text{ took } \text{it}] \\
      & \quad *\text{I will take/am taking it}
\end{align}\]

Similarly unmarked bare-Undergoer Pivots mark not just 1st-person Actor, but also an added modal element. With question intonation, they normally convey: 'do you want me to [verb]' (though 'you' may be replaced by a third person, in marked uses), while with declarative intonation they function as imperatives ('I want you to...'), cf Goddard (1994):

\[(43) \]
\[\begin{align}
   a. & \quad \text{juang?} \\
      & \quad \text{take} \\
      & \quad \text{Should/Will I take it? [Not *'Did I take it? *Am I taking it?'(etc)]}
   \\
   b. & \quad \text{juang!} \\
      & \quad \text{take [it!}]
\end{align}\]

Clearly then, aspectual and modal components are part of the interpretation of simple unmarked clauses containing actorless Undergoer Pivots. However cooccurrence patterns suggest that these elements are not part of the inherent semantics of these morphemes/constructions. The ke-Undergoer Pivots, for example, occur freely with lexical (non-perfective) aspect markers:

\[(44) \]
\[\begin{align}
   & \quad \text{montor-é nu ke-sutsut} \\
   & \quad \text{car-DEF still ke-wipe} \\
   & \quad [I'm] \text{ still polishing the car}
\end{align}\]

See also examples in 13.2.3. That the ke-Undergoer Pivot is not inherently perfective is also shown by its use in threats, otherwise unmarked for aspect/mood:

\[(45) \]
\[\begin{align}
   & \quad \text{ke-cakcak tendas-né!} \\
   & \quad \text{ke-chop.up head-3} \\
   & \quad [I'll] \text{ chop up his/your head}
   \\
   & \quad \text{ke-tig-tig polo-n-né!} \\
   & \quad \text{ke-beat brain-GEN-3} \\
   & \quad [I'll] \text{ beat his/your brains!}
\end{align}\]
The cooccurrence of both ke-Undergoer Pivots and bare Undergoer Pivots (13.2.3) with a variety of aspect/mood markers, indicates that the aspectual readings associated with unmarked clauses derive from a secondary semantic overlay. If this were not the case we could expect the inherent aspectual content to block or restrict the cooccurrence of bare Undergoer Pivots with incompatible lexical aspect markers.

13.3.4 Clause type: main versus subordinate clauses

The patterns described in the previous sections applied to main clauses. In some subordinate clauses the syntax of Undergoer Pivot verbs differs from that of main clauses.

13.3.4.1 Relative clauses

In relative clauses, Undergoer Pivot verbs occur in structures not found in matrix clauses: (i) the actor pronominal can co-occur with ke-Undergoer Pivots in everyday Low style, either as a post-clitic (as in main clauses in High style):

(46) apa ka-alih cai? what ke-look.for 2 What are you looking for?20

or even pre-verbally:

(47) anak-é ané Tut ke-rekam
     person-DEF REL T. ke-record
     The person Tut [=you] recorded21

(ii) Actorless bare stem Undergoer Pivots may occur in relative clauses, implying 1st-person actor in declaratives:

(48) (dedaaran) ané sai-sai daar di Bali
     food REL daily eat LOC
     (...food) which [we] usually eat in Bali

and 2nd-person Actor in questions:

(49) nyén ané takonang ituni?
     who REL ask.about earlier?
     Who was the one [you] asked about earlier?

These interpretations contrast with those of bare-stem Undergoer Pivots in unmarked main clauses (Table 13.1).

20 Satua I Lengkir, told by Madé Narsin, Peguyangan, Singaraja (Bawa & Jendra 1981). Although this example occurs in a literary text, the usage is also found in everyday Buléleng speech.
21 Tut is a 'pro-pronoun': names and titles often replace true pronouns - another avoidance strategy.
13.3.4.2 Infinitival use of bare Undergoer Pivots?

Other cases involving what appear to be 'bare-Undergoer Pivots' in subordinate clauses appear analysable as types of infinitival constructions. They are distinct from all other cases in being truly actorless semantically. *ke-Undergoer Pivots cannot be used in the same context:

(50) sing dadi (*ke-)orahin, i Putu
    NEG can (*ke-)tell DET P.
    I Putu can't be told what to do! [P. won't listen if rebuked]

(51) tiuk puntul patut (*ke-)sangih
    knife blunt should sharpen
    Blunt knives should be sharpened

The subordinate clause may occur in a higher VP, as in the preceding examples, or in an NP:

(52) 'kar meli kamen (*ke-)anggo ke pura
    FUT buy cloth wear to temple
    [I'll] buy a cloth to wear to the temple.'

(53) sing ngelah nasi (*ke-)baang nyai
    NEG have rice give 2FEM
    [I] don't have any rice to give to you.

In subordinate clauses introduced by *yén, lén, amon, 'if', both bare form and *ke-Undergoer Pivots may occur, but with different semantics, again suggesting that the bare-form occurrences are infinitives:

(54) pedalem nasi-né lén sing daar
    pity rice-DEF if NEG eat
    It would be a pity if the rice were not eaten

(55) pedalem nasi-né lén sing ke-daar
    pity rice-DEF if NEG ke-eat
    It would be a pity if the rice were not eaten [by me/*you/*other people]

Sentences like the following I take to be examples of ellipsed *yén/lén clauses

(56) seket laci-né kedeng
    stiff drawer-DEF pull
    The drawer is stiff [iff pulled] (Kersten 1984:522)

[ = seket laci-né *yén kedeng]

(57) katos nasi-né daar
    hard rice-DEF eat
    The rice tastes hard

With *ke-Undergoer Pivots they acquire the usual 1ACT reading:
Another minor clause type allowing bare Undergoer Pivots, again with true actorless semantics, are those introduced by *buka/cara* 'like':

(59) *buka* siap-é sambeh-in ifijin
like chicken-DEF scatter-LOC glutinous.rice

*Like the chicken fed glutinous rice* [puzzled, confused]

(60) kendel pesan buka oton-in keneh-né Mén Paluk
happy very like birthday-LOC thought-3 M. P.

*Men Paluk’s thoughts were very happy, like being given a birthday treat*

Satua Pan Paluk, Kirtya 1316/4, Banjar Dangin Peken, Singaraja

### 13.3.5 Other discourse factors affecting the interpretation of ke-Undergoer Pivots

#### 13.3.5.1 Interaction with basic principles permitting ellipsis

As Table 13.1 and examples like (21a&b) indicate, ke-Undergoer Pivots in unmarked declaratives imply 1ACT, while in interrogatives they imply 2ACT. Should two distinct 'portmanteau' morphemes be assumed? Here I argue (contra Clynes 1989) that the 1ACT or 2ACT interpretations derive purely from 'pragmatic' overlay, and are not part of the inherent semantics of the ke-affix.

Consider the unmarked interpretation of the following (non-Undergoer Pivot voice) sentences:

(61) a. *suba me-daar?*  
already eat

"Have you already eaten?"

b. *suba me-daar*  

"I’ve already eaten"

The same interpretations apply in English:

(62) a. Want some?

b. Been there, done that.

It seems then that, due to general principles of information processing, *and independently of the voice used*, in perfective statements the ellipsed Actor will generally be interpreted to be 1st-person, and in questions to be 2nd-person. Put more formally, I assume a principle such as the following is operating:
Pragmatics and morphosyntax

In a context where involvement of third-persons is ruled out:

a. Interpret perfective statements to be about the speaker's actions
b. Interpret perfective questions to be about the addressee's actions.

This seems to apply both in Balinese and in English, and perhaps universally, when NP's are ellipsed. Principle (63) in turn follows from the following probably universal pragmatic assumptions. Firstly, about the Speaker's and Addressee's relative knowledge of their own, and the other's, past actions:

a) I know more about what I have done than about what you have done
b) You know more about what you have done than about what I have done

Secondly, and following directly from (64a), about the normal pattern of information exchange between speaker and addressee:

a) I tell you about my actions
b) I ask you about yours.

Given these assumptions, the interpretation of the sentences in both (61) and (62) is straightforward. The interpretation of examples (21a&b), and similar cases follows then from (i) the understood perfective meaning of ke-Undergoer Pivots in unmarked main clauses, in combination with (ii) principle (63).

The ke- prefix does not then in itself mark 1ACT (pace Kamus Bali 1990), 2ACT, or 3ACT, and there is no need to posit 'portmanteau' morphemes. Similarly, there is no need to assume (as did Clynes 1989) the existence of two distinct, homophous, ke-morphemes, the one found in high style and a general Undergoer Pivot marker, the second found in Low style and incorporating a '1ACT/2ACT' semantic element.

13.3.5.2 Other contextual factors override 'unmarked' interpretations

In support of the analysis in the previous sub-section is the fact that that the interpretations of examples (21)a&b, repeated here as example (66) can be overridden in certain contexts:

a. ke-juang? take *1/*3
   [you] took [it]?
   *I/*he took it
b. ke-juang *2/*3
   [I] took [it] *you/*he took it
   *you/*he took it

Informants told me that a sleepwalker for example could plausibly ask (66a) with first-person Actor reading, if he were asking about actions he had carried out while...
sleepwalking, and 2ACT readings were clearly ruled out. Similarly, while the following question will normally imply 2nd-person Actor,

(67) **suba ke-jemak-ang nasi?**  
    already ke-take-APP rice  
    [Have you] already taken some rice for [X]

one of my consultants pointed out that a 1ACT reading could be possible, in a context where the Speaker is the person who normally fetches the rice, and the Addressee is the person the rice is fetched for (say, an elderly person). In this context, the 2ACT interpretation can be 'displaced' by the 1ACT reading. If however the Addressee usually gets the rice for, say, the grandchild he is holding, then 2ACT reading would again get priority.

The rule (at least for everyday/informal styles) seems to be something like:

(68) Interpret the Actor of questions containing ke-Undergoer Pivots as 2ACT wherever possible, except where such an interpretation is impossible, in which case interpret it as 1ACT

Similar rules no doubt apply in the other cases. Nowhere is a 3ACT reading possible with ke-Undergoer Pivots in everyday speech.

### 13.4 Conclusions

It was argued above that while in unmarked clauses ke-Undergoer Pivots often are associated with (i) perfective aspect, and with either (ii) 1ACT or 2ACT interpretations, these interpretations derive from secondary pragmatic 'overlay' (see 13.3.3 and 13.3.5.1). Can we then assign a meaning to the ke-morpheme itself? Since bare-stem verbs already function as Undergoer Pivots, it is difficult to argue that the addition of the ke-affix to that stem adds a semantic element 'Undergoer Pivot', particularly where both forms can be used interchangeably:

(69) **éné lakar (ke-)juang**  
    this will (ke-)take  
    This one will be taken [by me/you].

Assigning a distinct meaning to the ke-morpheme seems as problematic as attempting to attribute a single meaning to the bare verb stem: clearly the latter has multiple functions: as Undergoer Pivot verb, as imperative, as an infinitival form. I prefer to assume the correctness of the approach of Aronoff (1976) or Beard (1988): the minimal level at which meaning can be consistently assigned in Balinese is the word, not the morpheme - and even words, as this study shows, need to be embedded in a grammatical and discourse context to be more than potentially meaningful. Compare Wierzbicka, (1988:8):
Meaning is conveyed by grammatical constructions as much as it is by words, and it is conveyed jointly by all levels of linguistic structure. There is no such thing as 'grammatical meaning' or 'lexical meaning'. There are only lexical and grammatical means of conveying meaning [ ... ] In fact the notion of the meaning of a word in isolation is [ ... ] a fiction'.

As argued in 15.1, a 'word & paradigm' approach to morphology (where the relationship between meaning and form is indirect, and is assigned at word level minimally), though developed to describe highly inflectional languages, again appears equally applicable to an agglutinating language like Balinese (cf Koch 1990). It is also well suited to describing the 'overlay' meanings acquired by actorless Undergoer Pivots in specific stylistic (and other) contexts.

This chapter records then a variety of previously unremarked interactions of pragmatic meaning and morpho-syntax in Balinese. It shows that, contrary to claims made in previous studies (including Clynes 1989), the two actorless Undergoer Pivots in Balinese cannot be freely interchanged. Rather, a complex of syntactic and discourse factors determines which is used, and with what (implied) meaning. The chapter also brings out differences in the syntax and semantics of Undergoer Pivots in main clauses as opposed to subordinate clauses in Singaraja Balinese. These differences are previously unnoted in descriptions of Balinese.

In response to the imperative to express social relationships holding between Addressee and Speaker, Balinese has developed complex High and Honorific vocabularies, and a distinct High style (Bagus 1979, Clynes 1989). The study shows, I believe for the first time, that even everyday 'Low' Balinese has been profoundly affected by the same imperative, though it has lead to the elaboration of strategies for avoiding, rather than expressing, such relationships.

At the same time, the study indicates that a previously unrecorded distinction between Formal (High and Low literary registers) and Non-formal styles is needed to accurately describe meaning assignation (13.1.3): this cuts across the traditional 'High' versus 'Low' (alus/kasar) view of speech register oppositions in Balinese.

This chapter is preliminary in some respects. The use of the (again, previously unrecorded) alternative avoidance strategies outlined in Appendix L has yet to be described adequately, as has the use of avoidance strategies in High Style. Further studies could look at dialectal and sociolinguistic variation in the use of avoidance strategies; as well, a plausible diachronic account of the emergence of the two Undergoer Pivots is needed (see Clynes 1989, and Beratha 1992 for some speculations).

Another conclusion of this chapter concerns the methodology of descriptive linguistics as it applies to Balinese, and no doubt to all languages. That basic alternations such as those illustrated in examples (21) and (22) had not been reported before is apparently due to an over-reliance on written texts by previous writers, even native speakers. As Foley & Van Valin (1984) point out, data from a wide variety of discourse genres,
including the everyday spoken language, is clearly needed if one is aiming for true 'descriptive adequacy'.
14

Undergoer primacy and the morphosyntax of Balinese

14.0 Introduction

Here I give a variety of evidence from morphology, syntax, and discourse patterns to indicate that the Undergoer is the 'unmarked' macrorole in Balinese, and the (Bare Stem) Undergoer Pivot the basic voice diathesis. In 14.2 I go on to argue that the Undergoer Pivot is not a syntactically transitive construction, and that claims that Balinese is syntactically ergative which depend on that analysis are therefore unfounded.

The conclusions made here, that Undergoer Pivot verbs are (often) non-derived, are more basic than their Actor Pivot counterparts, and are intransitive, are important to the arguments made later in chapter 15. There I describe what I see as the principal function of affixation, and its relation to the notion of transitivity in Balinese.

14.1 Undergoer Pivot as the basic voice in Balinese

The facts of verb morphology suggest a basic 'Undergoer primacy' in Balinese:

1. On both transitive and intransitive verbs, where the Subject NP is an Undergoer, the verb is generally unmarked. Where the Subject NP is an Actor, the verb carries a prefix, %N% (9.3.2, 12.3).

2. On transitive verbs the presence of an Actor is only signalled by the morphology when that Actor is realised by the Subject NP. The semantic role of the Undergoer participant on the other hand is always marked on the transitive stem (by the suffixes), irrespective of whether that participant is realised by an NP in a Subject, Object or oblique relation; indeed, it continues to be marked even when there is no Undergoer NP at all in the surface syntax - see 11.2.2.2 and 11.3.2.

In this section I illustrate some ways in which this Undergoer primacy also accounts for unusual patterns in the syntax. Specifically, the Undergoer Pivot and Actor Pivot constructions in Balinese are atypical, when compared with their passive and active counterparts cross-linguistically (cf Keenan 1985). I conclude that these differences can be explained if the Undergoer Pivot is taken to be the more basic of the two constructions.
1. Unlike passives cross-linguistically, the (Bare Stem) Undergoer Pivot form is morphologically simpler than the Actor Pivot, which carries the {N} prefix.  

If one can talk of a direction of derivation in this case, it is more plausible therefore that the Actor Pivot derives from Undergoer Pivot, rather than the reverse.

<table>
<thead>
<tr>
<th>Undergoer Pivot</th>
<th>Actor Pivot</th>
</tr>
</thead>
<tbody>
<tr>
<td>'eat'</td>
<td>daa</td>
</tr>
<tr>
<td>'kill'</td>
<td>mati-ang</td>
</tr>
<tr>
<td>'chop'</td>
<td>tektek</td>
</tr>
<tr>
<td>'hit'</td>
<td>jagur</td>
</tr>
<tr>
<td>naa</td>
<td>%N-daar%</td>
</tr>
<tr>
<td>nge-mati-ang</td>
<td>%N-tektek%</td>
</tr>
<tr>
<td>%N-jagur%</td>
<td></td>
</tr>
</tbody>
</table>

2. As in many other Austronesian languages, Undergoer Pivot forms occur with higher frequency in narrative and conversation than Actor Pivot forms. Blake (1993:30) claims this, though without giving any supporting evidence. Beratha (1992:286-288) gives the following figures for occurrences of AP and UP forms in the text Dang Hyang Nirartha (HKS 4105, University of Sydney):  

<table>
<thead>
<tr>
<th></th>
<th>Total UP</th>
<th>Total AP</th>
<th>Total transitives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>134</td>
<td>105</td>
<td>239</td>
</tr>
<tr>
<td>Total UP</td>
<td>56%</td>
<td>44%</td>
<td></td>
</tr>
</tbody>
</table>

3. Depending on the mood of the clause, if the Undergoer NP is definite, Undergoer Pivot structures are either highly preferred, or obligatory. In simple declaratives, if the Undergoer NP is definite, the UP structure is strongly preferred. I do not have statistics to support this, though spontaneous comments to this effect were made several times by my teachers, as they corrected my speech.

In simple (non-clefted) questions, the situation is more categorical: if the Undergoer is definite, the UP construction is apparently obligatory:

(2) k-ade\textsubscript{p} kesela-né?
    UP-sell s.potato-DEF
    'Is the sweet potato sold (by you),'

hence the ungrammaticality of ((3):

(3) *ng-ade\textsubscript{p} kesela-né?
    AP-sell sweet.potato-DEF
    'Do (you) sell *the sweet potato?'

The Actor Pivot form must be used in such questions if the Undergoer is indefinite:

(4) ng-ade\textsubscript{p} kesela?
    AP-sell sweet.potato(*-DEF)
    'Do (you) sell (*the) sweet potato?'

---

1 Cf Keenan (1985:250): 'a passive will consist of a strict morphological modification of a transitive verb'.

2 These figures are extracted from three different tables; those tables count the relative order of Actor and Undergoer with respect to the verb in both transitive and intransitive clauses.
A parallel situation is found with imperatives: where the Undergoer is formally definite, a bare transitive stem, formally identical to the Bare Stem UP must be used.

(6)  
\text{k-adep} \text{kesela?}  
\text{UP-sell} \text{s.potato}  
'Is sweet potato sold (by you)'

Where the Undergoer is indefinite, the Actor Pivot form must be used:

(7)  
\text{ng-adep} \text{kesela=ja!}  
\text{UP-sell} \text{s.potato=PART}  
'Sell sweet potato!'

Balinese can thus be analysed as not having a formally distinct imperative construction. Rather it uses either AP or Bare Stem UP constructions, according to the referential status of the Undergoer NP.

One can thus capture the patterning of both Interrgoatives and Imperatives with a single generalisation: the presence of a definite Undergoer requires the use of UP constructions in the 'irrealis' moods (see also 9.5.2).

4. As well, Balinese prefers UP in many other contexts where English prefers to use active voice, for example with 1st- and 2nd-person Actors:

(8)  
\text{ane } \text{eñcén demen-in=cai?}  
\text{REL which happy-LOC = 2}  
'which one is liked by you'

For this reason, UP has a high frequency in conversation. Even where understood 1st- and 2nd-person Actors are ellipsed or deleted, Balinese prefers to use Undergoer Pivot forms (see also chapter 13 on pronominal avoidance strategies):

(9)  
\text{a. men, apa ke-demen-in?}  
\text{PART what ke-happy-LOC}  
'well, which one is liked [by you]'  
[from 'Mirah' by Putu Sedana, in Bagus & Ginarsa 1978:8]

\text{b. ané ené ke-demen-in}  
\text{REL this ke-happy-LOC}  
'this one is liked [by me]'

5. UP verbs are preferred also in third-person narrative, where there is a rapid sequence of aspectually punctual events, with a continuity of one or more participant(s). There may be an AP or simple intransitive verb in the first clause, but UP forms are preferred in subsequent clauses, with topical Undergoers becoming the (typically ellipsed) subject of the UP (10), (11), and topical Actors becoming clitics to the UP verb (10), (12):
In English and other European languages such sequences are unexpected. Beratha (1992:131), following Verhaar’s (1988) analysis of similar phenomena in Indonesian, describes patterns such as (10) and (11) as ‘ergative’: the ellipsed subject in the second clause being coreferential with the object NP of the first clause. Nonetheless, as Beratha herself recognises, it is easy to find examples such as the following, which clearly reflect neither an ergative nor a nominative pattern of coreference. In (13) the ellipsed subject of the second clause has an adjunct/non-core NP [bubuh] as antecedent:

Evidence of the non-core status of bubuh ‘porridge’ in the first clause of (13) comes from its inability to either (i) occur with the definiteness marker, -é:

or (ii) to fill the subject position with the corresponding Undergoer Pivot verb:

"Porridge is what was made for His Majesty."
Kroeger (1991a), writing on similar anaphora patterns in Tagalog, suggests that they are 'a fact of discourse grammar (45)', rather than being determined by syntactic constraints, such as an ergative syntactic orientation. The co-reference patterns in the above examples (and other data problematic for a purely syntactic account, such as in example (20), section 9.2.1) appear to be explicable in terms of 'Undergoer prominence': the preference for placing topical Undergoers into the Subject position where possible, rather than topical Actors.

To sum up, the simpler morphology of (Bare stem) Undergoer Pivot verbs suggests that this morphsyntactic type is more 'basic' than the Actor Pivot in Balinese. The obligatory or strongly-preferred occurrence of UP forms (of both types) in certain syntactic frames, together with their higher discourse frequency, suggest also that Undergoer Pivot structures are more basic at the syntax and discourse levels. To this extent they differ from passive verbs in many languages, while resembling them in having Undergoer subjects.

14.2 Syntactic orientation of Balinese: ergative, nominative or none of the above?

A number of writers have either assumed (Hunter 1988) or argued (Beratha 1992, Blake 1993) that Balinese shows an ergative orientation in its syntax. One of Beratha's claims has been examined (and rejected) in 14.1. Here I discuss a second argument made by both Beratha (1992:130) and Blake, which is based on perceived parallels in the treatment of NP arguments in intransitive and 'transitive' clauses. I will refer only to Blake's discussion, which is more elaborated than Beratha's, though both make the same, and I will argue, equally erroneous, claims.

Blake (1993, 1994) argues that Balinese shows ergative syntax because 'the unmarked transitive construction [...] shows an alignment of P ['the patient argument of a transitive verb' (19)] with S ['the sole argument of an intransitive verb (19)'] (1993:30)'. In nominative-accusative systems, he says, the A ['the agent argument of a transitive verb'] is treated in the same way as S (1993:19).

In intransitive clauses, the S occurs preverbally in unmarked word orders:

(16) **Nyoman** ulung
    N. fall
    Nyoman fell. [ = Blake's example number (29)]

In the Bare Stem Undergoer Pivot construction, which Blake calls 'the unmarked transitive construction', the preverbal argument is the Undergoer, which Blake assumes is a P, 'the patient argument of a transitive verb (19)'.

297
Blake's claim then depends on the correctness of the identification of the Bare Stem Undergoer Pivot construction (9.4) as 'the unmarked transitive clause'. Blake does not however demonstrate that this construction is transitive, nor even define the notions 'transitive verb/clause'. I class all Bare Stem and ke-Undergoer Pivot verbs as semantically transitive 15.3; it does not however follow from this that all, or any, of the clauses in which they occur will necessarily be syntactically transitive (see 15.3.3). 3

I assume that a transitive clause is one containing two core syntactic arguments, as defined in 9.5.3. By this criterion, the Actor pronominal in clauses like (17) must have core status for Blake's claim to be correct. The evidence in fact indicates that all Actor participants in Bare Stem Undergoer Pivot clauses behave like non-core arguments (9.5.3):

1) If they are non-definite nouns, they cannot occur in a separate syntactic constituent (either an NP or a PP), but must cliticise to the verb, as though incorporated into the verb (cf Myhill 1988): 4

(18) batis-é cegut = legu
leg-DEF bite = mosquito
*My leg was bitten by a mosquito

The verb and the clitic Actor cannot be separated by other constituents:

(19) batis-é cegut = legu ibi
leg-DEF bite = mosquito yesterday
*My leg was bitten by a mosquito yesterday

(20) *[batis-é cegut ibi legu
*!cegut batis-é legu ibi

Non-definite oblique NPs in other contexts behave in a similar way, forming part of a non-interruptable VP constituent (cf discussion of example (72), section 9.5.3).

Pronominal Actors of Bare Stem Undergoer Pivot verbs must also cliticise to the verb (see also discussion of example (39), section 9.4.2), though the Balinese orthography is misleading on this point. The verb-Actor sequence cannot be interrupted:
Blake's example sentence ((17) above) is thus more accurately transcribed as follows:

(22)  \textbf{Nyoman lempag = tiang}

\begin{tabular}{l}
N. \hfill hit = 1exc \\
Nyoman was hit by me.
\end{tabular}

2) Again like non-core NPs (cf example (68), section 9.5.3), Actor NP's marked for definiteness must be relegated to a separate PP, though with an obligatory 3rd-person actor clitic on the verb:

(23)  \textbf{Nyoman lempag = a ajak kuren-an-né}

\begin{tabular}{l}
N. \hfill hit = 3 \hfill with \hfill spouse \\
Nyoman was hit by his spouse.
\end{tabular}

This again contrasts with the behaviour of core NPs, which can never occur in a PP (9.5.3).

The only possible evidence I have found that the Actor in such clauses is a core argument comes from the fact that it is obligatory, in Singaraja Balinese at least. However, Andrews (1985) in his cross-linguistic survey shows that obligatoriness is not exclusively the property of core arguments: oblique NPs are often also obligatory.

All the evidence is therefore that the Actor of the Undergoer Pivot clause is syntactically an oblique. The same conclusion is reached by Artawa: 'this NP argument behaves like an oblique' (1992:28). In that case, Bare Stem Undergoer Pivot clauses are 	extit{not} transitive, by the definition above. Claims about Balinese syntax which rely crucially on a transitive analysis of such clauses, such as those of Blake (1992, 1994), and even, surprisingly, Artawa (1992), are therefore unconvincing.

The only candidate for an 'unmarked transitive construction' is then the Actor Pivot construction (9.4). If that identification is correct, Balinese then appears to show nominative-accusative orientation, at least in terms of Blake and Beratha's word order test:

(24)  \textbf{Nyoman ulung}

\begin{tabular}{l}
N. \hfill fall \\
Nyoman fell.
\end{tabular}

(25)  \textbf{Nyoman ng-lempag i Nengah}

\begin{tabular}{l}
N. \hfill AP-hit \hfill DET \hfill N. \\
Nyoman hit Nengah.
\end{tabular}

---

5 Note that this analysis differs from that proposed by Kroeger (1991a) for Tagalog: he gives evidence that Actor NPs remain core arguments/"terms" in Tagalog.

6 I have not been able to consult Artawa (1994).
Chapter 14

It would however be premature to draw conclusions at this stage: the claim that Actor Pivot constructions are syntactically transitive relies at the moment on a single test for Object status (9.5.2)), one which could ultimately turn out to be a fact about Undergoer participants, rather than the syntactic entity Object. Further exploration of this area is beyond the scope of this thesis. To my knowledge, no convincing arguments have yet been put which would enable Balinese to be classified as syntactically ergative, nominative, a mixed system (as Artawa 1992 concludes), or none of the above (cf Durie's analysis of Acehnese (1985, 1987)).
15

Verbal Morphology in Balinese - some conclusions

15.0 What does affixing do, and not do, in Balinese?

In this chapter I draw some broad conclusions about how affixation 'works' in Balinese, based on the detailed analyses presented in the previous chapters, plus other evidence. Amongst other things, I conclude that: 1) even though Balinese is typologically an agglutinating language, a classical morphemic analysis does not 'work' for Balinese; 2) the presence of affixes does not in itself signal that derivational (or inflectional) processes have occurred in Balinese; 3) the often-made claim that affixes are associated with syntactic transitivity in Balinese and other Western Indonesian languages is incorrect.

Instead I argue that meaning is assigned via derivational processes, which may or may not involve affixing (15.1, 15.2); and that affixes instead signal information about the semantic structure of both derived and nonderived lexemes (15.2), (15.3).

Syntactic transitivity is then irrelevant to the functioning of affixation in Balinese. There are, for example, no 'object-creating' morphological operations in Balinese (and no doubt, pace Chung (1975)'s analysis of parallel facts in that language, in Indonesian, see 15.3.4).

15.1 Meaning as a function of lexemes and derivations, rather than of morphemes

According to Nida, who I will take to represent the classical approach to morphological analysis 'morphemes are the minimal meaningful units which may constitute words or parts of words (1949:1)'. Meaning is thus a key element in Nida's definition: morphemes are identifiable in part by their 'common semantic distinctiveness' (see his Principles 1 and 2, pp7 & 14).

Given the evidence presented in chapters 11, 12 and 13, it would be difficult to characterise verbal affixes in Balinese as morphemes by Nida's definition, since most affixes cannot be assigned a constantly recurring single meaning. While this can be done for lexical morphemes (see 10.4), of the affixes perhaps only the 'agentive' prefix (12.2) can be said to have a specific meaning.

It is true, for example, that the three-way paradigmatic contrast in transitive verb suffixing ('-∅, -in, -ang) does generally correspond generally to a distinction between
Patient Undergoer, Locative Undergoer, and 'other peripheral semantic roles' Undergoer. Nonetheless, affixes like -in (11.2) and -ang (11.3) are associated with a wide variety of meaningful functions. The Undergoers of verbs bearing -in include 1) 'affected locations' (11.2.1):

1) eñče-in 'urinate on U'
kulit-in 'put skin on/take skin from U'
atur-in 'speak to U'

2) animate experiencers (11.2.2.2):

2) sakit-in 'torture U'
enten-in 'remind U (eg of debt)'

and 3) fully affected Patients (11.2.3):

3) jagur-in 'punch U repeatedly'
temah-in 'vociferously insult U, repeatedly'
tïfjak-in 'repeatedly kick'

Analogous comments can be made in the case of affixes like -ang (11.3), me- (12.4) and ke- (chapter 13). In the vast majority of cases, then, it is not the case that the meaning of a lexeme is simply the sum of meanings attributable to its root morpheme and affixes. On the other hand, the evidence of chapters 11 and 12 shows that it is possible to predict the meaning of a lexical item produced when a base combines with a verbal affix, once the broad syntactico-semantic type of the base is known.

I conclude therefore that in derived lexemes, meaning is assigned to a large extent by derivational processes, rather than by affixes. Derivational processes are complex processes which refer to the semantic class of the base, to derive other lexemes. Affixation may, or may not, be involved in such processes (cf 15.2).

The description of word-formation processes in Balinese, as in other agglutinating languages (Koch 1990), requires then an approach which allows for an indirect relationship between meaning and form, such as those in the 'word and paradigm'-tradition of analysis (for example Matthews 1991, Aronoff 1994, Beard 1988). Such an approach is needed in any case, at all levels of description, to account for the way context, linguistic and otherwise, influences meaning (see particularly chapter 13), and, often even the phonological shape (chapter 3) of a given linguistic unit.

15.2 The primary function of affixes in Balinese: signalling, not deriving

I will give evidence here that the primary function of verbal affixes in Balinese is to obligatorily signal the presence of semantic elements in the lexeme (or in some cases, word, if there are inflectional uses of morphology in Balinese, 10.3). Affixes do not function to signal that derivational or inflectional processes have occurred, as is the
close-to-universal assumption in the analysis of the morphologies of Western Austronesian languages.

I nevertheless follow those who assume that derivational processes do derive one lexeme from another. There is good evidence that this is true for Balinese, where the base occurs independently, as a 'free lexeme'. With intransitive verbs bearing the -an suffix, for example, there is very strong formal and semantic evidence for directionality of derivation, from transitive verbs (see 11.4). Similarly, monomorphemic intransitive verbs virtually all have a transitive counterpart, bearing one or more of the affixes -ang 'APP', -in 'LOC', and pe- 'CAUS':

(4)  

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pules</td>
<td>'sleep'</td>
</tr>
<tr>
<td>pules-in</td>
<td>'sleep on U'</td>
</tr>
<tr>
<td>pules-ang</td>
<td>'cause U to sleep'</td>
</tr>
<tr>
<td>gedé</td>
<td>'big'</td>
</tr>
<tr>
<td>pe-gedé-nin</td>
<td>'cause U to be bigger'</td>
</tr>
</tbody>
</table>

These transitive verbs can uncontroversially be considered to derive from the intransitive, since they bear more morphological material, and since the transitive verb is definable in terms of the intransitive, but not it seems vice versa.

It is wrong, nonetheless to equate the process of affixation with such derivational processes. Affixation and derivation are in fact two distinct processes, as Beard has argued (eg 1988). Affixation can occur without derivation, and derivation without affixation. Evidence for the occurrence of derivation without affixation, or 'zero derivation'/'conversion' in Balinese is given in 12.1. There it is argued that all bare stem stative verbs can be used as nominals, via that mechanism: demen 'happy/liking for', gedeg 'angry/anger', and so on.

The second logical possibility, the occurrence of affixation independently of derivation, also occurs in Balinese. One kind of evidence adduced by Beard (1987:40) for this is the occurrence of 'morphemes without functions' (my 'complex formants', see chapter 4 for detailed arguments for this). Affixation without derivation of a different sort is also found with bound lexical morphemes if, as I argue in 10.4, these are not analysable as 'precategorial'. If bound lexical morphemes in Balinese have full lexical specifications, like other lexical morphemes, then the affixes which occur with them are either functionless, or else they have the obligatory signalling function proposed in the first paragraph of this subsection. I will now give evidence for the latter analysis.

The evidence from prefixing is that Balinese is a head-marking, split-S language (see 12.3). As such, it obligatorily signals the semantic macrorole (Actor or Undergoer) of the subject NP of inherently intransitive verbs. This is done through obligatorily marking verbs with Actor subjects with either of the realisations of the {N} derivation, %N% or me-. Agentive intransitive verbs such as ng-igel (%N-igel %) 'dance', me-jujuk 'stand up' thus consist of a bound root morpheme, plus an affix, while non-agentive intransitives occur as bare morphemes, without affixation: ulung 'fall', kéweh '(be) difficult' (12.3).
Chapter 15

The agentive intransitives, while never having a free, unaffixed alternant, seem otherwise to be no less 'inherently verbal' than the non-agentives. Semantically, they correspond to intransitive verbs in many other languages. In combination with the affixes -ang and -in their behaviour parallels that of free monomorphemic intransitive verbs (10.4). Moreover, the corresponding nominal forms are complex, or morphologically unrelated. This is true whether the prefix be {N} or me-

(5) VERB NOUN
negak (N-tekak) 'sit' tegak-an 'seat'
ng-igel 'dance (vi)' igel-igel-an 'dance (n)'
nyerit (N-jerit) 'scream' jerit-an 'scream (n)'
meju (N-paju) 'defecate' tai 'faeces'

(6) VERB NOUN
me-laib 'run' laib-laib-an 'way of running'
me-bangkes 'sneeze' bangkes-an 'sneeze (n)'
me-solah 'dance.HI' se-solah-an 'dance (n.)'
me-kecuh 'spit' poos 'saliva'

I conclude that the root morphemes here are 'already' (inherently) intransitive verbs, with the prefixes {N} and me- functioning purely to obligatorily signal the presence of an 'agentive' element in their semantic structure. By default the absence of such a prefix on an inherently intransitive verb signals the 'unmarked' state, the absence of this agentive semantic element (12.0, 12.3). This then accounts for the skew in markedness, inherently patientive intransitive verbs being free morphemes, and inherently agentive ones bound.

Similarly, I give evidence in chapter 11 (passim) that the semantic role of the Undergoer NP of transitive verbs is obligatorily signalled by suffixing in Singaraja Balinese.¹ Only if U is a true Patient is there no suffix. Inherently transitive verbs must therefore carry a suffix, where the Undergoer is conceptualised as partially affected, or not at all affected. Examples of such verbs are anti-ang 'wait for U', ateh-ang 'escort CT/accompany U'², orah-in 'tell U', which have no suffixless counterpart, and no simple intransitive or nominal form:

(7) INTRANSITIVE NOMINAL
anti-ang ? ?
ateh-ang ? ateh-ateh-an 'way of accompanying'
orah-in me-pi-orah 'inform' pi-orah 'informing (n, literary)'

There are a good many such verbs, all plausibly involving 'unaffected' or partially affected Undergoers, and lacking monomorphemic verbal or nominal alternants; other examples include:

---
¹ This is not the case in dialects in the Badung/Denpasar region, where final suffixes are consistently elided or dropped completely (5.8).
² Warna et al (1990) lists unsuffixed ateh (also anti and ubet, cf below) as a transitive verb. These are rejected by Singaraja informants as 'Badung' (Denpasar) speech; the -ang/-in distinction is apparently being lost in the (Low) Denpasar dialect.
Many verbs like this, including anti-ang, ateh-ang, oot-in, ubet-ang, usu-usu-in, and entung-ang, appear to have no nominal counterpart at all (or not one readily elicited). I conclude then that these are inherently specified as verbs.

The function of suffixes on verbs like these, and logically therefore on all verbs, cannot be qualified as derivational, in the traditional sense of the word. Rather, suffixes must function to signal the presence in the verb's semantic structure of an Undergoer, viewed as only partially affected 15.3). The same conclusion applies, mutatis mutandis, where inherently intransitive verbs carry the Agentive prefixes {N} and me-. There is no derivational process involved in either case. Nor can this signalling be described as inflection, since it is primarily information about the internal semantic structure of the verb which is being signalled.

The obligatory nature of the signalling function, expressed through the presence of either prefixes or suffixes, or both, necessarily precludes some bases from 'surfacing' as free morphemes - these bear affixes, and yet are not derived verbs. The apparent derivational function of affixing is then a secondary effect, a by-product of the signalling function.

Descriptions which claim that the function of verbal affixes is to 'repackage the role structure', 'increase the valency', and so on, all assume a process model of affixation, with one lexeme as the input, and a new, different one as output. Such models clearly cannot account for the presence of affixing on non-derived verbs.

15.3 Transitivity and Affixation

In this section I give evidence that verbal affixation in Balinese does not refer to syntactic transitivity. I first define two types of transitivity, semantic and syntactic (15.3.1), and discuss possible types of interaction between the two. I then give evidence in 15.3.2 and 15.3.3 that affixation refers to participants in semantic structure, rather than to NP arguments realised in the syntax. Finally, in 15.3.4 I compare this analysis with other analyses of Western Indonesian languages which claim that verbal affixes mark syntactic information.
15.3.1 Semantic transitivity versus syntactic transitivity

Here I argue that one can distinguish between a semantically transitive event, and its syntactic realisation, which can be either syntactically transitive, or intransitive. I will use the term semantically transitive to refer to a predicate which refers to two salient entities in its semantic structure, both potentially encodable in the syntax as core NP arguments (9.5.3). The more agentive of these two entities bears the Actor role, and the less agentive, the Undergoer role. Semantically intransitive verbs are those with just one entity in the semantic structure potentially encodable in the syntax as a core NP argument (9.5.3).

I use the term syntactically transitive to refer to a clause (or a verb in such a clause) where both of the the grammatical relations, subject (9.5.1) and object (9.5.2) are assigned to NP arguments of the verb.

Mismatches between the two levels constitute evidence for the view expressed by Wierzbicka (1981, 1988, 1992), Hopper and Thompson (1980), Pawley (1986), and Givon (1990), that the terms 'transitive' and 'intransitive' are fundamentally semantic, rather than syntactic, notions. I first review some of their claims, before considering the evidence from Balinese.

Another important claim by Givon (1990) and Wierzbicka (1992) is that transitivity is a prototypical notion. Givon attributes the following 'three semantic dimensions' to the 'prototypical transitive event' (1990:565):

(9) a) Agent: The prototypical transitive clause involves a volitional, controlling, initiating, active agent, one that is responsible for the event, i.e. its salient cause.

b) Patient: the prototypical transitive event involves an inactive, non-volitional, non-controlling patient, one that registers the changes-of-state associated with the event, i.e. its salient effect.

c) Verb: The prototypical transitive clause involves a compact (non-durative), bounded (non-lingering), realis (non-hypothetical) verb and tense-aspect-modality. It thus represents an event that is fast-moving, completed, and real. i.e. perceptually and cognitively salient.

Wierzbicka (1992:22) formulates 'the prototypical transitive scenario', as follows:

(10) someone did something because of this, something happened to something in the same place, at the same time this person wanted this.

It is not the case however that predicated events must have all of the characteristics identified by Givon or Wierzbicka, to be encoded in syntactically transitive clauses. If for example the prototypical actor is a Givonian Agent (9a), Balinese nonetheless extends the use of {N}, the agentive prefix to include experiencer subjects, and even
occasionally inanimate subjects, which in certain cases are encoded as though they were initiators (12.3). On a language-specific basis, though, extreme departures from the prototypical semantic scenario lead to departures from the syntactic encoding of transitivity, of the types catalogued by Hopper and Thompson (1980).^3^ The nature of the syntactic coding of a semantically transitive scenario - whether both entities receive overt expression in the syntax, and whether they are assigned to core NPs (bearing subject or object grammatical roles) or to non-core NPs, or are incorporated into the verb phrase - depends on a complex of factors including the semantics of the predicated event/state, and whether or not the entities are referential, individuated, animate, initiating, and so on: in short, the degree to which they conform to, or depart from, the prototype.

In Balinese, when entities are not uniquely identifiable, for example, they often behave like syntactically incorporated nominals: in AP verbs, generic undergoers are incorporated (9.5.2, and 15.3.3), in UP, generic actors (14.2, footnote (?)). I have described how departures from the prototypical notions of Actor and Undergoer are encoded in the morphology (aprototypical Actors in 12.3, and Undergoers in chapter 11, especially 11.2 and 11.3).

The analysis of the Actor and Undergoer as 'macroroles' (Foley & Van Valin 1985, 9.3.1) does not account for their variable realisation; this is predicted by the prototype model, and I therefore adopt that view: the notions Actor and Undergoer are best considered to be prototypical notions in Balinese.

In the syntactic realisation of a semantically transitive scenario, there are therefore a variety of logical possibilities. Either only the Actor, or only the Undergoer, or both of these semantic entities, can be represented in the syntax (ellipsis as a surface phenomenon can of course subsequently apply). Assuming that the grammatical relations 'Subject' and 'Object' are (or may be, in the case of Object) assigned to given NPs, there can be a variety of mismatches (cf Bresnan 1994). Possible mappings, where verbs refer to both Actor and Undergoer in their semantic structure, include the following, all found in Balinese:

<table>
<thead>
<tr>
<th>(11)</th>
<th>Mapped to syntax</th>
<th>GR assigned</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Only Actor</td>
<td>S</td>
<td>'Verbs of caused emotion' 11.2.2.2, inchoatives 11.3.2.1</td>
<td></td>
</tr>
<tr>
<td>b. Only Undergoer</td>
<td>S</td>
<td>(i) Formally agentless ke- Undergoer Pivots, (ii) UP, incorporated Actor AP, definite Undergoer</td>
<td></td>
</tr>
<tr>
<td>c. Both Actor &amp; Undergoer</td>
<td>S, O</td>
<td>AP, incorporated Undergoer (see discussion below)</td>
<td></td>
</tr>
<tr>
<td>d. Both Actor &amp; Undergoer</td>
<td>S only</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^3^I prefer not to see syntactic transitivity as a scalar notion: cf the definition below.
In terms of the two types of transitivity, semantic and syntactic, the following patterns of interaction are found:

<table>
<thead>
<tr>
<th></th>
<th>realised as syntactically intransitive clause</th>
<th>realised as syntactically transitive clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>semantically intransitive</td>
<td>✓</td>
<td>doesn't occur</td>
</tr>
<tr>
<td>semantically transitive</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 15.1 Interactions of semantic- and syntactic transitivity in Balinese ('✓' = the combination does occur)

As well as uncontroversial cases where both semantic and syntactic transitivity values match, Table 15.1 allows for two kinds of mismatches. I take it that the first, the realisation of a semantically intransitive event (involving only a single entity potentially encodable in the syntax) as a transitive verb/clause, is a logical impossibility and does not occur. On the other hand, the second predicted mismatch, the encoding of a semantically transitive event by a syntactically intransitive clause, is common, and occurs in a variety of ways. In the following subsections I give evidence, both morphological and syntactic evidence for this view.

15.3.2 Evidence for transitivity as a semantic notion - 1

Formal features of Balinese verb morphology support the view that transitivity is fundamentally a semantic notion. As outlined in 9.4, Balinese has three passive-like verb types:

(12) a. 'bare Undergoer Pivots':

jagur-in  'be hit (PL)'

b. 'ke-Undergoer Pivots':

ke-jagur-in 'be hit (PL)'

c. 'me-dettransitives':

me-jagur-an  'be hit (stative)'

While passive verbs are often assumed to be intransitive, cross-linguistically, I refer to both Bare Stem Undergoer Pivots and ke-Undergoer Pivots as transitive. Only the third type, those prefixed with me-, do I take to be true (semantic) intransitives. This is for both formal and semantic reasons. Firstly, both Bare Stem- and ke-Undergoer Pivots
share a common stem with the Active/AP diathesis, often distinct from that of the intransitive stems:

(13) **Intransitive form**

<table>
<thead>
<tr>
<th>Intransitive form</th>
<th>Transitive forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>selem 'black'</td>
<td>1 selem-ang 'blacken'</td>
</tr>
<tr>
<td></td>
<td>2 ke-selem-ang 'id.'</td>
</tr>
<tr>
<td></td>
<td>3 N-selem-ang 'id.'</td>
</tr>
<tr>
<td>demen 'be happy'</td>
<td>1 demen-in 'like (vt)'</td>
</tr>
<tr>
<td></td>
<td>2 ke-demen-in 'id.'</td>
</tr>
<tr>
<td></td>
<td>3 N-demen-in 'id.'</td>
</tr>
</tbody>
</table>

The AP and UP forms are thus based on the same stem (eg selem-ang), and are different from the non-derived intransitive form (eg selem). The same pattern is found with intransitives derived from the transitive stem (11.4):

(14) a. transitive stem

- nyagurin N-[jagur-in] 'A hit (repeatedly) U'
- jagur-in 'U be hit repeatedly (by A)'
- ke-[jagur-in] 'U be hit repeatedly (by A)'

b. detransitive stem

- me-jagur-an 'U be in a state of having been hit repeatedly (**by A)**

There is thus a clear formal distinction between the derived intransitive verb stem, and that of the Undergoer Pivot verbs. If the latter were also derived intransitives, we could expect them to share a common stem with the detransitive, but this is not the case.

Moreover, as the glosses in (14) and the 'someone did something' component of (10) indicate, the Bare Stem- and ke-Undergoer Pivot types always involve an actor as well as an undergoer in their semantic structure, even where only one of the two is assigned a grammatical role. Strong evidence for this view is given in chapter 12, where it is shown that formally agentless Bare Stem and ke- Undergoer Pivots are always assumed to involve an agent in the predicated action. The bracketing in a gloss like 'U be hit repeatedly (by A)' indicates merely that that 'A/someone' may be unrealised in the syntax.

Other evidence that the Bare Stem Undergoer Pivot always involves an agent in the predicated action is that even where there is no specific actor referent the verb obligatorily carries a 3rd-person clitic pronoun =a:

(15) anak deriki ten nyak gusti-ang-a

people here.HI NEG.HI want [title]-APP =3

People here don't want to be addressed as 'gusti' [by people in general]/

The flavour of this generic use can often be rendered in English by the indefinite pronouns 'they' or 'one':

309
Chapter 15

(16) nangka ané nu nguda ketéwel-ang=a
nangka REL still young ketéwel-APP = 3

Nangka fruit which is still young is called 'ketéwel' [by people].
They call / one calls young nangka fruit 'ketéwel'.

Bare stem-and ke-Undergoer Pivot verbs therefore (i) share a common stem with Actor Pivot verbs, a stem distinct from that of intransitives (ii) are always understood to involve the participation of an Actor, which may be represented in the clause (by a non-core nominal/NP, cf 14.2).

The third passive-like type, the 'me-detransitive', on the other hand, is both formally and semantically distinct. It is based on a distinctly different stem (11.4), and has a stative meaning not found with the two Undergoer Pivot and the Actor Pivot forms:

(17) me-jagur-an 'U be in a state of having been hit'

Crucially, the me-detransitives never encode an actor, either semantically or syntactically: an actor may not occur, even in an optional PP:

(18) i Made me-jaguran *(teken i Wayan)
DET M. with DET W.

Made had been hit *(by Wayan)

It can be concluded that, of the scenario in (10), me-detransitives encode only the component:

(19) U me-[VERB] something ([VERB]) happened to U

The me-detransitives then are 'true intransitives', assuming that such verbs are those which have only one entity in their semantic structure available for encoding as a core NP in the syntax: pules, 'sleep' negak (%N-təgak%) 'sit', me-daar '1. eat; 2. be (partly) eaten', me-jagur 'be hit'.

On the other hand, the formal and semantic features shared by both Bare stem- and ke-Undergoer Pivots with Active transitive verbs indicate that they share a common semantic frame, compatible with that in (10): the action they conceptualize involves both an undergoer and an actor - even if only the former is necessarily always encoded in the syntax.

15.3.3 Evidence for transitivity as a semantic notion - 2

Here I discuss two further cases of mismatches between semantic and syntactic transitivity, both involving 'formally transitive' Actor Pivot verbs which behave as syntactic intransitives (i) where the undergoer nominal is present, but incorporated, and (ii) where the undergoer is not represented in the syntax at all.
1. Even when a verb is in the (cross-linguistically) 'canonically transitive' Actor Pivot form, it behaves as though it is syntactically intransitive, if the Undergoer NP is non-referential or generic in reference. The evidence for this was given in 9.5.2. Here I briefly review that evidence.

'True' intransitive verbs (those only ever occurring with a single NP argument) which carry the AP prefix \{N\} retain this prefix in the imperative:

\[(20)\] ngoyong! \%N-oyon\% 'stay!'

\[\text{negak!} \%N-tagak\% 'sit!'\]

Transitive imperatives on the other hand have two forms: where the undergoer argument is definite/referential, the bare transitive stem is used:

\[(21)\] baca ja buku-ne! 
\text{read} \text{PART} \text{book-DEF}

Read the book!

However where the undergoer argument is indefinite/generic, the verb stem carries the \{N\} affix, just like the true intransitives do when occurring in imperatives:

\[(22)\] maca buku ja!
\text{read} \text{book} \text{PART}

Read books/do book-reading

One can conclude then that indefinite Undergoer nominals are incorporated into the AP verb, which therefore behaves like a syntactically intransitive predicate.

The position of the 'emphatic particle', ja, in the above two examples is another indication that the indefinite undergoer is incorporated into the verb: ja normally comes immediately after the first constituent in the clause. This is the verb, where the Undergoer is definite, as in (21); however it must come after the Verb-Undergoer NP unit where the Undergoer is indefinite. Final evidence for incorporation is phonological: whereas baca in (21) carries stress \[\text{ba'c}\], in (22) maca buku behaves as a single unit in only allowing stress on the final syllable: \[\text{macabu'ku}\].

If, on the other hand the Undergoer is definite, the form of the verb remains identical to that where the Undergoer is incorporated, but it is now syntactically as well as semantically transitive. I argue in (9.5.2) that Undergoer NPs formally marked for definiteness carry the Object grammatical relation:

\[(23)\] tiang suba maca buku-ne ento
\text{I} \text{already read} \text{book-DEF} \text{that}

I have already read that book

\[\text{4Even if it could be shown that this is not the case, the basic argument being made here, that transitivity is primarily a feature of semantic structure, is not affected. In fact it would then be a feature only of semantic structure.}\]
Chapter 15

The mapping of semantics to syntax in (23) can be represented informally as follows (cf Bresnan 1994, Jackendoff 1990):

(24) $[\text{tiang}_\text{actor} \text{ suba maca } [\text{buku-}\text{né ento}_\text{undergoer}]

semantic structure: $<\text{actor, undergoer}(...)>$

GR assignment $\text{Subject} \text{ Object}$

The equivalent simple declarative clause, but with an unmarked undergoer, will have a necessarily generic meaning:

(25) $\text{tiang} \text{ suba maca \,= \, buku}$

I already read = book
$I \, have \, already \, 'done \, book \, reading'$

and, given evidence such as (22), a different mapping of grammatical roles, the undergoer being represented in the surface syntax (cf Bresnan's c-structure), but not being assigned a core grammatical role:

(26) semantic structure: $<\text{actor, undergoer}(...)>$

GR assignment $\text{Subject}$

2. Semantically transitive verbs also surface as syntactic intransitives in a second context. This regularly happens, with derivations producing verbs of shape $N-[\text{VERB(FEELING)}]-\text{in}$ and $N-[\text{INTRANS}]-\text{ang}$. (i) Intransitives of shape $N$-ROOT-in I take to involve an undifferentiated, generic 'causee' undergoer (11.2.2.2):

(27) $\text{ng-lenghe-in} \, \text{ngedegin}$

'$\text{AP-intoxicated-LOC: be intoxicating}'$

$\%N$-gadag-in$\%$ $'\text{AP-angry-LOC: be annoying}'$

Those of shape $N$-ROOT-ang have a non-surfacing, reflexive undergoer, (11.3.2.1):

(28) $\text{ngedé-nang} \, \text{ng-waas-ang}$

'$\text{AP-big-APP: become bigger}'$

$\%N$-gade-naŋ$%$ $'\text{AP-healed-LOC: become healed}'$

It is important to remember that the $N$-ROOT-ang derivation is a subvariant of one which 'produces both syntactically transitive, and syntactically intransitive verbs, according to verbal semantics and the degree to which the eventuality described departs from the prototypical transitive script' (11.3.2.1); such departures could include non-volitional, sometimes inanimate Actors, and/or reflexive, and therefore animate, Undergoers, as well as other features such as non-bounded verbal aspect.

Some verbs, such as $\text{uék-ang}$ 'torn-APP; tear' do not have an intransitive form, because their inherent semantic content is 'too close' to the transitive prototype: inanimate Undergoer, bounded aspect, and so on.

312
However, in other cases, a given base occurs with both intransitive and transitive uses. The intransitive use has the reflexive undergoer, the transitive use a non-reflexive, often inanimate, undergoer:

(29) INTRANSITIVE
    ngedé-nang 'become bigger'
    ngemainang 'be coming here'
    ngemenékang 'be climbing up'

TRANSITIVE
    gedé-nang ' (A) enlarge U'
    mai-nang ' (A) bring U here'
    penék-ang ' (A) raise U'

The fact that with these verbs the syntactic realisation, now intransitive, now transitive, depends purely on semantic factors, is very clear evidence for the primarily semantic basis of the notion transitivity in Balinese. Where such verbs are used intransitively, I assume the following mapping from semantics to grammatical relations:

(30) semantic structure: <actor, undergoer (...)>
    GR assignment Subject

In each of these verb classes, N-ROOT-in and N-ROOT-ang, the verb retains the form of a transitive verb, with the semantic role of an Undergoer participant signalled by the suffixes - however in neither case is that Undergoer represented in the syntax.

To sum up the evidence of this subsection, the intransitive syntactic behaviour of (i) AP verbs with indefinite Undergoers (eg, maca buku) (ii) some AP verbs of shape N-ROOT-in/ang supports the conclusion that 'transitive' verbs are so primarily at the level of semantics. The shape of the transitive verb encodes the involvement of two participants, an Actor and an Undergoer, at at the level of semantic structure: departures from the 'prototypical transitive scenario' then give rise to the intransitive syntactic behaviour of the verb.

15.3.4 Conclusions, and comparison with other analyses

Two conclusions then fall out from the evidence presented in 15.3.2 and 15.3.3:

(i) Transitivity is primarily a semantic notion in Balinese, and

(ii) Verbal affixation in Balinese does not refer to syntactic transitivity

This second conclusion follows from the demonstration that (i) the suffixing on transitive verbs refers to the semantic role of the Undergoer participant, irrespective of whether that participant is realised in the syntax; and (ii) as was demonstrated in section 12.3, the prefixing contrasts in Balinese refer only to the semantic role of the Subject NP, irrespective of whether the predicate is transitive or intransitive.

Analyses of Balinese affixes which treat either prefixes or suffixes as signalling facts about syntactic transitivity are therefore erroneous. This view for example leads Kersten (1984:33-34) to posit two distinct (but necessarily homophonous) nasal prefixes, the one associated with intransitive verbs, the other with transitive verbs. Similarly, in her
treatment of verbal suffixes, Beratha (1992: 165) erroneously claims that 'the main function of -ang is valency increasing', and that -in functions 'to rearrange and increase the valency of the verb' (167). Beratha does not discuss the intransitive verb classes which carry those affixes (15.3.3).

Descriptions of the verbal morphology of other Western Indonesian languages, such as Indonesian/Malay, are also generally 'syntactically driven', in that they assume that the function of affixes is to signal the assignation of grammatical roles to particular NP arguments. This is the case, for example, in Chung's (1976) analysis of Indonesian Malay, and Alsagoff (1992) and Nik Safiah Karim (1993) on Malay.

All such descriptions encounter major problems in that the 'transitive' affixes, be they prefixes (like %N%, the agentive prefix, but often described as 'active voice') or suffixes (like -in, ang), are also regularly found on many syntactically intransitive verbs.

Chung (1976) and Alsagoff (1992), for example, both characterise the 'meng-' affix of Indonesian (cognate with the Agentive prefix, %N%, in Balinese) as a transitive prefix, despite the many intransitives which carry meng-.

(31)  

<table>
<thead>
<tr>
<th>ROOT</th>
<th>MENG-INTRANSITIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>tangis</td>
<td>menangis</td>
</tr>
<tr>
<td>raung</td>
<td>me-raung</td>
</tr>
<tr>
<td>turun</td>
<td>menurun</td>
</tr>
<tr>
<td>kuning</td>
<td>menguning</td>
</tr>
<tr>
<td>putih</td>
<td>memuthi</td>
</tr>
<tr>
<td>didih</td>
<td>mendidih</td>
</tr>
</tbody>
</table>

'the sound of weeping' 'weep' 'roar (vi)' 'descend (inchoative)' 'turn yellow' 'turn white' 'boil'

These apparently exceptional intransitives are not discussed at all by Chung; they are (highly implausibly) explained away as not formed by synchronic processes by Alsagoff (1992:12). Alsagoff claims that meng- can only occur on transitive verbs. She mentions

'a few exceptions, where intransitive verbs apparently occur with meng-: eg menari 'dance', menangis 'cry', membatuk 'cough'. However these forms are in fact monomorphemic, since they never appear without the so-called prefix meng-, i.e. never in the 'root' form, viz *tari, *tangis, *batuk. (1992:12, footnote 4)

This claim is an extraordinary one. Each of these roots does in fact regularly occur without the prefix, as (probably zero-derived) nouns: tari piring 'dance-plate: the plate dance'; suara tangis 'sound-weep: sound of weeping', obat batuk 'medicine-cough: cough medicine'.

Prentice (1987:922, 933) recognises that both intransitives and transitives occur with the meng- form, but proposes essentially the same 'homophonous morpheme' analysis as is suggested by Kersten for Balinese %N%. Prentice bases his argument on the

---

5 On page 13, Alsagoff cites menurun (meng+turun) as a non-occurring intransitive verb. Yet this same verb is given as an example of an intransitive verb carrying meng- by Nik Safiah Karim et al (1993:153).
exceptional allomorphy of meng- when it occurs with (i) transitive verbs bearing the
prefix per- (ii) loanwords used to form transitive verbs; I would say that such
morphologically/lexically conditioned realisations are perfectly consistent with a single
prefix analysis.\(^6\)

On verbal suffixes, Nik Safiah Karim et al (1993) claim that both -kan (p177) and -i
(p178), whose functions closely parallel those of Balinese -ang and -in respectively, are
used to form transitive verbs. Similarly, Chung in an influential work claims that
Indonesian -kan, and sometimes -i, manifest a rule of Dative (1976). This, she claims, is
an 'object creating rule', turning an indirect object/benefactive into an object. Again, this
analysis at best accounts for only part of the data, either in Indonesian or, as I have
shown, if it were to be applied to Balinese. Crucially, it does not account for Indonesian
verbs which remain syntactically intransitive, despite the affixation of -kan. There are
many of these in Indonesian, all of which are semantically transitive, like their Balinese
counterparts (11.3.2.1, 15.3.3). The Indonesian verbs, again like their counterparts in
Balinese, are formed by derivation, from a stative verb:

\[(32)\]  
**SOURCE VERB:**  
| bosan | 'bored' | mem-bosan-kan 'make people feel bored, be boring' |
| gembira | 'delighted' | meng-gembira-kan 'make people feel delighted, be delightful' |
| ngeri | 'terror' | me-ngeri-kan 'make people feel terrified, be terrifying' |
| terkejut | 'startled' | mengejutkan 'make people feel startled, be startling' |
| malu | 'shy, ashamed' | me-malu-kan 'make people feel shame, be shameful' |
| kesan | 'impression' | mengesankan 'make people feel impressed, be impressive' |
| jengkel | 'irritated' | menjengkelkan 'make people feel irritated, be irritating' |
| kuat | 'strong' | mengkuatkan 'make people feel strong, be strengthening' |
| senang | 'happy' | menyenangkan 'make people feel pleased, pleasing' |

All the above examples are taken from Kridalaksana (1990:61), who classifies them as
deverbal adjectives. I have glossed them with the meaning 'make people feel...'; it
should be remembered that this 'generic undergoer' is part of the semantic structure only;
as the term 'deverbal adjectives' indicates, these verbs do not occur with an
object/Undergoer NP.

Moreover, Chung's analysis does not account for the following two facts:

- In Balinese, and probably also in Indonesian, the Actor Pivot/active verb is
  formed from the Undergoer Pivot form (14.1).

- The Bare Stem Undergoer Pivot is syntactically intransitive in Balinese (14.2), as
  is the equivalent construction in Indonesian (her 'object preposing' construction).

\(^6\)To make a case, Prentice would need to demonstrate for example that loanwords used to form
intransitive verbs have a distinct allophony pattern from that found with similar derived transitives.
Chapter 15

The UP verb form therefore carries the so-called 'object creating' suffixes, without ever having had the Object relation assigned to a subcategorised NP. A different account is clearly in order, both for Balinese, and for Indonesian.7

The present analysis therefore has greater descriptive adequacy than 'syntax-prior' accounts, in that it explains the occurrence of verbal affixes in both syntactic contexts of use, in intransitive and transitive clauses. This analysis, which follows from the fact that transitivity is primarily a semantic notion (15.3), is consistent with the other main conclusion of this chapter, that the primary function of verbal affixes in Balinese is to obligatorily signal the presence of elements in the semantic structure of lexemes, independently of whether these elements are realised in the syntax of the clause (15.2).

7While Chung (1976) has been very influential, its conclusions are to me entirely unconvincing. Repeatedly, ungrammatical elicited sentences are offered by Chung as grammatical: in the first few pages, for example, sentences (13), (35b), (36a&b) and (58). (58) is predictably ungrammatical by a rule that Chung herself proposes at (39). Chung claims that one feature that distinguishes direct objects from other NPs is that they are prepositionless (p220) yet, on the same page, example (2b) contains a non-object NP which is prepositionless. A detailed discussion would show similar flaws in the analysis on every page. I agree fully with the Indonesian linguist, Kaswanti Purwo [1994:4, my translation]: 'This work [Chung (1976)] is very well known and often quoted by linguists of a variety of theoretical schools. [...] For linguists who do not know Indonesian this an extremely readable, interesting work. But for linguists who know the language, it gives rise to feelings of irritation and anger. Many of the example sentences filling that work do not 'feel' Indonesian. Moreover, a variety of important constructions in Indonesian are passed over. [...] As a result, the complex behaviour of the suffixes -kan and -i is not adequately brought out.'
Epilogue

This thesis can be seen as an exploration of ways in which meaning receives form in Balinese. Here I summarise some of the main points which emerge.

In Part I features of the phonology are described.

1. I argue that expressive meaning is signalled regularly and productively by phonological structures, in a way quite different from the conventional linking of meaning with form in language (Chapter 3). I have used statistical analyses to support my arguments, and presented my data in full in the Appendices. I hope that my findings will be confirmed by other studies, using other statistical tests.

2. I have argued for parallel hierarchical structures in the phonology: a hierarchy of non-prosodic formants, from feature to morph (chapters 2, 4, 5, 6), and a prosodic hierarchy (chapter 5, 8). Meaning-assignation processes can target any level of these hierarchies: meaningful allophonic processes can manipulate individual feature values (chapter 2), whereas the minimal lexeme is usually a disyllabic morph (chapters 4, 5 and 6).

In Part II features of verbal morphology and clause syntax are described.

3. The analysis of affixation processes supports a model of linguistic structure in which derivation and affixation are separate processes (Beard 1988). While I argue that affixing in Balinese conveys information primarily about semantic or conceptual structure, I have not speculated about the nature of that structure, in particular, the principles determining how participants are mapped to the syntax. Clearly, there is evidence for a structured semantic component (cf Jackendoff 1990), with Actor and Undergoer participants being given special status in 'mappings' to other components of linguistic structure.

4. The semantic motivation for morphosyntactic phenomena argued for by Wierzbicka (1988) is supported both by the analysis of transitivity effects proposed here (Chapter 15), as well as that of the verbal affixation system.

5. A variety of previously undescribed interactions of morphosyntax with pragmatic factors were discussed. In chapter 13 I showed that a variety of discourse factors must be taken into account to adequately describe the meanings and use of the two main passive structures. Just as I concluded in chapter 3, the data accord well with the view that one cannot make a principled distinction between semantics and (linguistic) pragmatics (cf Wierzbicka 1989).
Appendix A

Semantics of phonologically 'optimal' morphemes: control sample

To check the hypothesis that there is a regular relationship between non-optimal phonological shape and 'expressive' semantics (chapter 3), a control sample of semantic classes of morphemes predicted to have the most 'neutral/central' possible structure was analysed (3.2.3). This involved a sample of 209 morphemes of shape C₁V(N)C₂VC₃, where C₁ was one of the three commonest initial consonants /p, t, s/ and C₃ from the commonest final consonants /ŋ/ and /h/. A 'random' sample was made: every third item in the list generated by the Fiesta computer program was chosen, aiming for a total sample of at least 30 tokens for each combination. Any (to my knowledge) potentially marginal phonotactic combination was excluded, including: (i) morphemes where C₂ is homorganic with either C₁ or C₃ (hence morphemes with otherwise very common initial /k/ were ruled out) (ii) morphemes where either syllable onset contained a filled L slot (iii) morphemes with intervocalic /r/, since morphemes with it are liable to be expressives (iv) and other shapes suspected or demonstrated to be non-optimal: morphemes where any of C₁, C₂ or C₃ is not filled, where C₂ is a nasal consonant, and so on.

In analysis of the control sample, an attempt was made to err on the side of 'over-assignation' to expressive semantic classes, to avoid charges of bias in the analysis: the higher the level of expressives in the control sample, the more difficult it would be to prove the hypothesis of interaction between semantics and expressiveness. The overall results were as follows: of a total sample of 209 tokens, 120, or 57%, were classifiable as expressives or loans), and the remaining 89 tokens (43%) were not immediately classifiable in one of the 'expressive' classes (Table A1).

<table>
<thead>
<tr>
<th>C₁...C₃</th>
<th>SENSE</th>
<th>PL/AN</th>
<th>BAD</th>
<th>LOANS</th>
<th>Other PERIPH</th>
<th>UNCLASS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p...ŋ/</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>21</td>
<td>40</td>
</tr>
<tr>
<td>/p...h/</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>14</td>
<td>37</td>
</tr>
<tr>
<td>/t...ŋ/</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>35</td>
</tr>
<tr>
<td>/t...h/</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>33</td>
</tr>
<tr>
<td>/s...ŋ/</td>
<td>4</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>3</td>
<td>11</td>
<td>29</td>
</tr>
<tr>
<td>/s...h/</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>12</td>
<td>29</td>
</tr>
</tbody>
</table>

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>totals</td>
<td>23</td>
<td>17</td>
<td>30</td>
<td>28</td>
<td>22</td>
<td>89</td>
<td>209</td>
</tr>
<tr>
<td>= % 'age</td>
<td>9%</td>
<td>9%</td>
<td>10%</td>
<td>13%</td>
<td>12%</td>
<td>46%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1 Semantic classes of morphemes in control sample

319
The detailed analysis of the data now follows. In checking for loanwords, principal works used were Gonda (1973) for Sanskrit; Tim penyusun kamus (1988) Malay/Indonesian; Poerwadarminta (1976) Malay/Indonesian; Mardiwarsito (1981) Old Javanese, Sanskrit; Prawiroatmojo (1981) Javanese. Note that I classify morphemes belonging to the the High/Honorific lexis together with other loanwords: I show in Clynes (1989) that this component of the lexicon is borrowed almost 100% from (Old) Javanese. An equally valid analysis would have been to set up an independent expressive class just for these items, just as Uhlenbeck (1950) does for Javanese. 

1. /p... ...ng/ first 40 items, 21 (52%) Unclassified:
   
<table>
<thead>
<tr>
<th>SENSE</th>
<th>BAD 3</th>
<th>PLANT/ANIMAL 4</th>
<th>OTHER PERIPHERAL 4</th>
<th>LOANS 5</th>
<th>OTHER PERIPHERAL 4</th>
<th>OTHER PERIPHERAL 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>SENSE</td>
<td>/p... ...ng/</td>
<td>/p... ...ng/</td>
<td>/p... ...ng/</td>
<td>/p... ...ng/</td>
<td>/p... ...ng/</td>
<td>/p... ...ng/</td>
</tr>
<tr>
<td>BAD 3:</td>
<td>paling 'disoriented'; pandung 'steal'; pendeng 'not give food to'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/ANIMAL 4:</td>
<td>padang 'grass' (generic); pantung 'k o prawn'; pelang 'k o tree'; petung 'k o bamboo'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER PERIPHERAL 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOANS 5:</td>
<td>paing 'day of 5-day week' Jav.; penting 'important' Malay; pacang 'will.HI'Jav.; pisang 'banana.HI' Jav./Mal.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER PERIPHERAL 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 5:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/ANIMAL 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER PERIPHERAL 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOANS 5:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER PERIPHERAL 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 6:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/ANIMAL 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER PERIPHERAL 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOANS 5:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. /p... ...h/ 37 items; 14 Unclassified:
   
<table>
<thead>
<tr>
<th>SENSE</th>
<th>BAD 5:</th>
<th>PL/AN 4:</th>
<th>SENSE</th>
<th>LOANS 6:</th>
<th>OTHER 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>SENSE</td>
<td>/p... ...h/</td>
<td>/p... ...h/</td>
<td>/p... ...h/</td>
<td>/p... ...h/</td>
<td>/p... ...h/</td>
</tr>
<tr>
<td>BAD 5:</td>
<td>padah 'blunt'; pedih 'angry'; pelih 'wrong' BAD; pich 'not properly developed (of a limb, due to sickness); pisuh 'hurl abuse at';</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 6:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 7:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 8:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 9:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 10:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 11:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 12:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 13:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 14:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/AN 4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. /t... ...ng/ 35 items, 20 unclassified:
   
<table>
<thead>
<tr>
<th>SENSE</th>
<th>BAD 4:</th>
<th>PLANT/AN 3:</th>
<th>SENSE</th>
<th>LOAN 3:</th>
<th>OTHER 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>SENSE</td>
<td>/t... ...ng/</td>
<td>/t... ...ng/</td>
<td>/t... ...ng/</td>
<td>/t... ...ng/</td>
<td>/t... ...ng/</td>
</tr>
<tr>
<td>BAD 4:</td>
<td>timang 'unevenly loaded'; tebong, --né 'you' (insulting); talang '[return] empty handed; without any load'; timpang 'not fitting, inappropriate' (Loan? =Mal.);</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td>tambang 'large coconut shell used as container'; teling 'k o flower';</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENSE 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENSE 3:</td>
<td>tayung 'swing (arm)'; tambung 'fly high and far'; timbeng 'totally absorbed';</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOAN 3:</td>
<td>tabung 'save (money)' Mal.; tumpeng 'conical offering made of rice Jav.'; timbang 'weigh' Mal.;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 5:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 6:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 7:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 8:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 9:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 10:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 11:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 12:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 13:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 14:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 15:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 16:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 17:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 18:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 19:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAD 20:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLANT/AN 3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER 2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

20 UNCLASS: tabeng 'screen (n)'; tabing 'wall'; taling 'letter of alphabet, /e/'; tambeng, tameng 'dam (v)'; tawang 'know'; tawing 'wall'; taweng 'protect'; tembang 'tune, song'; tembing 'k o game'; tempung 'join in'; tepeng 'cook (rice)'; tepung 'flour';
timbung 'cook in a bamboo tube'; timpung 'cook to eat communally'; tubung 'add to'; tulung 'help'; tumpang 'level, roof (on pagoda)'; tumpung 'k o basket'; tulang 'bone'; tulung 'help'; tumpang 'level, roof (on pagoda)'; tumpung 'k o basket';

4. /t... ...h/ 35 items, 12 (37%) unclassified, 22 (63%) expressives & loans
BAD 5: tawah 'strange'; tebah 'speak impolitely (use the wrong 'level')'; teluh 'magic'; tempuh 'fine [punishment]'; tulah 'accursed';
PLANT/ANIMAL 4: tabih 'base of tree trunk which divides into several board-like root-segments'; tuluh 'egg'; tabuh 'k o tree'; tubuh 'coconut tree nearly ready to bear fruit';
SENSE 7: tabah 2 'fresh (not salty, of water)'; tabeh 2. sound beaten to indicate time; tabuh 'intonation; tune'; timpah 'resting one on another (eg thighs of people sleeping)'; timpah sit with legs folded to one side'; tulih 'look around behind one', tempuh 'touch';
LOANS 3: tabah 'brave; resilient' Mal; tambeh 'add to.HI'; teluh 'cool, not hot' Mal.;
OTHER 5: tampah 'slaughter' TRUE PATIENT; tebah 'stab' TRUE PATIENT; tehuh 'split (v)' TRUE PATIENT; tuyuh 'finished, used up'; tempah 'place an order';

5. /s... ...ng/ 33, 20 (61%) expressives & loans, 13 (39%) unclassified
BAD 4: selang 1. suspect (vt); sepeng 'angrily rebuke'; simbing 'subtley criticise';
sumbung 'arrogant';
PLANT/AN 2: sepang 'k o tree'; sumpang 'flower worn in hair';
SENSE 9: sampang 'k o kemenyan (incense)'; sayang 'love (eg of family members for each other)'; sepung 'fall (eg leaves)'; sileng 'gaze up'; simbang 'stand with hands held up'; suling 'flute(?)'; simbang 'quiet; silent'; sawang 'meditate, concentrate mind to make contact with gods or ancestral spirits'; simping 'dancer's ornament';
LOAN 4: selang 'borrow.HI'; sumpang 'call in.HI, visit' Mal.; seling 'add (comment, thing)' Mal.; sumpang 'donate' Mal.
OTHER 3: sumping 'k o CAKE'; selang 'depression between collar-bone and chest' BODY PART; saling 'reciprocal morpheme' GRAMM;
UNCLASS 11: tabuh 'reinforcement on shoulder-pole'; tamah 'hoe'; tampih 'fold (v)'; tapih 'women's under-cloth'; tabah 1. span (unit of measurement); tabah 4. weigh up an issue; tabah 5. sometimes win, sometimes lose (in gambling); tehuh 'finished, used up';
temah 'place an order';
telah 'unchanged, in original state'; tunguh 'grow';

6. /s... ...h/ 29 items, 11 (38%) unclassified, 18 (62%) expressives & loans
BAD 2: sayah 'drought'; sulah 'bald';
SENSE 5: sepah 'spit out'; simbah 'spray liquid from mouth'; sembah, sumbah 'worship, holding hands and body in respectful posture'; silah, tuara ~, 'invulnerable';
LOAN 7: salah 'wrong' Mal., Old Javanese; selah 'hit ball (in game of kasti) Dutch slaag; sumbah 'take an oath' Mal.; sapuh HI 'broom'; silah 2 'reciprocal morpheme HI';
simbah 'HON dripping (sweat)'; siwah 'badly matched pair (shoes etc) HI';
OTHER 3: sipah 'armpit' BODY PART; selah 'opportune occasion; correct way of doing something' GOOD; seluh 'scrape out contents (of coconut etc) TRUE PATIENT;
UNCLASS 12 (38%): salah change (of seasons); sembah 'name of ceremony'; sambah 'broadcast, spread out (seeds etc)'; sapuh 'bank of river, slope of mountain'; sapuh 'draw (in competition)'; sebah, sibeh 'frame (of painting etc)'; sepuh 'apply a layer of metal to surface'; silah 'borrow'; suluh 'lamp'; subuh 'container for holy water made from a small coconut shell'

1 The high proportion of expressives here may be due to the feature [+continuant] shared by /s/ and /h/.
Appendix B

Morpheme-internal interactions between consonants

Here the data and analysis in favour of the 'One place per morph constraint' (6.2.1) is presented. Morphemes containing three structural sequences were analysed for possible interactions between consonants: (i) \( /\text{C}_{\text{initial}}\ldots\text{C}_{\text{final}}/ \), (ii) \( /\ldots\text{C}(V)\text{C}_{\text{final}}/ \), and (iii) \( /\text{C}_{\text{initial}}(V)\text{C}_{\text{final}}/ \); the results are reported in subsections B1, B2 and B3 of this appendix, respectively. Data is presented in the endnotes. The data in Appendix J also support the conclusions drawn here.

For each 'suspect' combination a random sample of morphemes was examined (see individual data sets for techniques used for sampling), and each item was classified as either being a loanword, a grammatical morpheme, belonging to one of the expressive classes (BAD, SENSE, PLANT/ANIMAL etc, cf 3.2.2) or as 'unclassified', ie not obviously expressive, or with insufficient information available to be classify it. In taking samples items were generally avoided where they contained other phonotactic sequences known to be associated with expressiveness, such as syllables containing CL clusters, though in some cases, such as where there were not a large number of them, all the morphemes containing a particular structure were examined.


Appendix B1 /\text{C}_{\text{initial}}\ldots\text{C}_{\text{final}}/ interactions

Table B.1 shows the numbers of tokens of each possible initial and final consonant combination, for headwords of shape \#/C\ldots\text{C}/ listed in Wama et al (1990). Data includes (non-reduplicated) di- and multi-syllabic headwords. Monosyllables and underlyingly reduplicated monosyllables (only those of shape RED-C(L)VC) were not included in the data. These are discussed below. Footnotes, found at the end of this appendix, give semantic analyses of morphemes containing 'suspected' consonant combinations. The rest of this part of the appendix discusses each place of articulation in turn.

The constraint barring \#/\text{labial}\ldots\text{labial}/ is close to an absolute ban, with very few occurrences. For other places of articulation there are sometimes significant numbers of morphemes of shape \#/\text{C}_{\text{place}}\ldots\text{C}_{\text{place}}#/ (ie where initial & final consonants have
identical place specifications), but the morphemes predictably have expressive semantics. Based on this data alone, there is evidence for a constraint such as the following (but see 6.2.1):

(1) Disprefer \([C_{\text{place}} \ldots C_{\text{place}}]_{\mu}\).

<table>
<thead>
<tr>
<th>p</th>
<th>b</th>
<th>m</th>
<th>t</th>
<th>d</th>
<th>n</th>
<th>l</th>
<th>r</th>
<th>s</th>
<th>k</th>
<th>g</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>20</td>
<td>14</td>
<td>26</td>
<td>40</td>
<td>64</td>
<td>69</td>
<td>76</td>
<td>131</td>
<td>22</td>
<td>144</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>14</td>
<td>4</td>
<td>57</td>
<td>89</td>
<td>94</td>
<td>77</td>
<td>124</td>
<td>58</td>
<td>145</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>49</td>
<td>31</td>
<td>2</td>
<td>22</td>
<td>27</td>
<td>25</td>
<td>32</td>
<td>13</td>
<td>18</td>
<td>58</td>
<td>49</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>25</td>
<td>12</td>
<td>6</td>
<td>16</td>
<td>17</td>
<td>3</td>
<td>37</td>
</tr>
<tr>
<td>21</td>
<td>16</td>
<td>14</td>
<td>71</td>
<td>19</td>
<td>78</td>
<td>59</td>
<td>41</td>
<td>110</td>
<td>138</td>
<td>37</td>
<td>125</td>
<td>113</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>7</td>
<td>20</td>
<td>11</td>
<td>12</td>
<td>31</td>
<td>14</td>
<td>15</td>
<td>48</td>
<td>9</td>
<td>55</td>
<td>29</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>3</td>
<td>78</td>
<td>47</td>
<td>38</td>
<td>0</td>
<td>52</td>
<td>100</td>
<td>65</td>
<td>80</td>
<td>48</td>
<td>116</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>13</td>
<td>59</td>
<td>14</td>
<td>18</td>
<td>54</td>
<td>0</td>
<td>18</td>
<td>58</td>
<td>26</td>
<td>63</td>
<td>56</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>10</td>
<td>16</td>
<td>65</td>
<td>18</td>
<td>10</td>
<td>18</td>
<td>34</td>
<td>38</td>
<td>17</td>
<td>135</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>5</td>
<td>18</td>
<td>45</td>
<td>18</td>
<td>5</td>
<td>20</td>
<td>21</td>
<td>15</td>
<td>26</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>22</td>
<td>9</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>51</td>
<td>28</td>
<td>20</td>
<td>22</td>
<td>18</td>
<td>7</td>
<td>52</td>
<td>66</td>
<td>63</td>
<td>118</td>
<td>25</td>
<td>97</td>
<td>52</td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>24</td>
<td>101</td>
<td>6</td>
<td>18</td>
<td>64</td>
<td>89</td>
<td>89</td>
<td>66</td>
<td>28</td>
<td>93</td>
<td>26</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>9</td>
<td>4</td>
<td>10</td>
</tr>
</tbody>
</table>

Table B.1 \(#C_{\text{initial}} \ldots C_{\text{Final}}\) consonant combinations (data from Warna 1990, column = \(C_{\text{initial}}\), row = \(C_{\text{Final}}\)). (Figures in bold - combinations where interaction is suspected)

Further qualifications need to be made where the consonants are both apicals. For morphemes of structure \(#t\ldots C_{\text{apical}}\#, the semantic data indicates no constraint on cooccurrence, cf table B.2:

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>/#t\ldots ..t#/</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>/#t\ldots ..d#/</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>/#t\ldots ..n#/</td>
<td>30</td>
<td>15</td>
</tr>
<tr>
<td>Totals</td>
<td>53</td>
<td>36 (57%)</td>
</tr>
</tbody>
</table>

Table B.2 Semantics of /#t\ldots ..C_{\text{apical}}\#/ morphemes

However, when /d/ is initial, there are much tighter constraints against following apicals.
Morpheme-internal interactions between consonants

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>/#d... ...t#/</td>
<td>26</td>
<td>4</td>
</tr>
<tr>
<td>/#d... ...d#/</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>/#d... ...n#/</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td>Totals</td>
<td>53</td>
<td>9 (17%)</td>
</tr>
</tbody>
</table>

Table B.3 Semantics of /#d... ...C\text{apical}/ morphemes

That is, of 53 tokens sampled, only 9 were not classifiable as expressives, indicating the constraint is operating. An explanation assuming that coronals are underlyingly unmarked for place of articulation (cf Fudge 1968, Paradis & Prunet 1990 and references in both) doesn't account for constraints against /#d... ...C\text{corona}/ combinations.

Liquids may not cooccur at all, except /#l... ...r#/ sequences, which are tolerated, but always in loans or expressives. (In the sample of 25 items, only 2 were not immediately explainable in these terms).

As in the case of C\text{1}-C\text{2} interactions (cf Appendix B3), Table B.4 below shows that strong interactions occur between the laminal stops and the apicals, with 80% of tokens in the sample in the expressive classes, presumable because both carry the feature [CORONAL]:

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>&quot;non-expressives&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>/#c... ...t#/</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>/#c... ...d#/</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>/#c... ...n#/</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>/#/... ...t#/</td>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>/#/... ...d#/</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>/#/... ...n#/</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>Totals</td>
<td>99</td>
<td>20 (20%)</td>
</tr>
</tbody>
</table>

Table B.4 Semantics of /#C\text{laminal}... ...C\text{apical}/ morphemes

No similar interaction between /s/ and the apicals was found (Of a sample of 30 items 19, or around 63%, were not classifiable as expressives). On the other hand /#s... ...s#/ sequences were limited to a very few loans and historically complex numerals (function words), indicating a constraint barring */#C_{lam}... ...C_{lam}/ in neutral morphemes.

The dorsals (Table B.5) are particularly interesting in that though large numbers of morphemes with /#C_{dorsal}... ...C_{dorsal}/ sequences were noted (and particularly morphemes of shape of /#k... ...k#/), suggesting at first glance that the constraint simply does not hold for dorsals, these morphemes were nevertheless consistently expressive semantically:
### Appendix B

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>&quot;non-expressives&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>/#k... ...k#/</td>
<td>125</td>
<td>6</td>
</tr>
<tr>
<td>/#k... ...g#/</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>/#k... ...ŋ#/</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>/#g... ...k#/</td>
<td>33</td>
<td>5</td>
</tr>
<tr>
<td>/#g... ...g#/</td>
<td>30</td>
<td>4</td>
</tr>
<tr>
<td>/#g... ...ŋ#/</td>
<td>30</td>
<td>8</td>
</tr>
<tr>
<td>Totals</td>
<td>278</td>
<td>27 (9.7%)</td>
</tr>
</tbody>
</table>

Table B.5 Semantics of /#C_dorsal... ...C_dorsal#/ morphemes

I conclude in 3.2.5 that /#C_dorsal... ...C_dorsal#/ sequences are used productively to generate expressives, morphemes of shape /k... ...k/ in particular being of 2nd highest frequency of any combination of initial and final consonants listed in Warna (1990), and yet having 95% expressives in this count.

Moreover, and in agreement with the evidence from /C_initialVC.../ and /...VC_final/ sequences, there is evidence here to indicate that dorsals pattern phonologically with /h/ in a single 'back' phonological class. The proportion of morphemes of structure /k... ...h/ and /g... ...h/ is much higher than that in the control sample (3.2.5), \( Z = 3.24, p < .0006 \):

<table>
<thead>
<tr>
<th>/C_1...C_3/</th>
<th>SENSE</th>
<th>PL/AN</th>
<th>BAD</th>
<th>Other LOANS</th>
<th>PERIPH.</th>
<th>UNCLSS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>/k... ...h/</td>
<td>11</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>8</td>
<td>14</td>
<td>60</td>
</tr>
<tr>
<td>/g... ...h/</td>
<td>12</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>-</td>
<td>8</td>
<td>30</td>
</tr>
<tr>
<td>total PLANT/ANIMAL, BAD, SENSE</td>
<td>48 (53%)</td>
<td>Z = 3.24, p &lt; .0006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total Expressives &amp; loans</td>
<td>68 (76%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total Unclassified</td>
<td>22 (24%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.6 Semantics of /#C_dorsal... ...C_back#/ morphemes

The above discussion is based on the data in Table B.1, which omits monosyllables, since these are consistently peripheral (see Appendix G).\(^{33}\) As a result, where they are lexical morphemes monosyllables will be expressives or loans, whether they contain homorganic initial and final consonants or not.

The conclusion of this part of the data is thus that, except in expressives and loan words, Balinese does not allow morphemes to begin and end with phonemes of the same place of articulation. The degree of tolerance of exceptions to this constraint varies with place of articulation: for labials, only very few tokens are found, for dorsals, there are many, but they are consistently expressive.
Morpheme-internal interactions between consonants

Appendix B2 \#...CVC#/ restrictions

Interactions were checked between the penultimate and final consonant, Ci and C_coda, in morpheme-final syllables of shape CiVC_coda. Data from Warna 1990 was used. The number of tokens occurring of each possible Ci-C_coda combination is given in table B.7, with the detailed data and analysis given in footnotes. As with /#C_... ...C#/ and /#C_1(V)C_2.../ sequences (see other the subsections of this appendix), the only significant interaction noted is a 'dislike' for /...Ci(V)Ccoda#/ consonants being of the same place of articulation. These patterns can be described with this constraint (but see 5.2.1):

(2) Disprefer \[...Ca.placeVCa.place#]\m

The constraint barring labial-labial combinations is again close to an absolute ban, with only 13 occurrences (in a data base of more than 14,000 items). Underlyingly, 7 of these 13 are reduplicated %VC_labial% sequences such as epep 'hide (in closed hand)', ebeb 'force in (intending to cause pain)'. For other places of articulation, while there are sometimes significant numbers of morphemes of shape /...CiplaceVCiplace#/, the morphemes predictably having expressive semantics.

Where both consonants are obstruents or nasal stops interact there is strong interaction (though of the nasals only /n/ and /ŋ/ occur in large numbers morpheme-finally). Manner of articulation seems important only for the liquids /l/ and /ɾ/, which seem again to form a subclass of their own (cf below).34

The rest of this section discusses each place of articulation in turn.

<table>
<thead>
<tr>
<th>p</th>
<th>b</th>
<th>m</th>
<th>t</th>
<th>d</th>
<th>n</th>
<th>l</th>
<th>r</th>
<th>s</th>
<th>k</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0</td>
<td>81</td>
<td>104</td>
<td>46</td>
<td>60</td>
<td>65</td>
<td>55</td>
<td>56</td>
<td>112</td>
<td>28</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0</td>
<td>87</td>
<td>21</td>
<td>60</td>
<td>50</td>
<td>68</td>
<td>46</td>
<td>66</td>
<td>46</td>
<td>145</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>82</td>
<td>36</td>
<td>46</td>
<td>21</td>
<td>29</td>
<td>43</td>
<td>27</td>
<td>8</td>
<td>48</td>
<td>44</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>19</td>
<td>4</td>
<td>3</td>
<td>8</td>
<td>26</td>
<td>21</td>
<td>12</td>
<td>7</td>
<td>33</td>
<td>31</td>
</tr>
<tr>
<td>19</td>
<td>17</td>
<td>15</td>
<td>58</td>
<td>11</td>
<td>62</td>
<td>82</td>
<td>94</td>
<td>67</td>
<td>165</td>
<td>55</td>
<td>160</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>14</td>
<td>61</td>
<td>4</td>
<td>41</td>
<td>79</td>
<td>82</td>
<td>46</td>
<td>102</td>
<td>85</td>
<td>163</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>2</td>
<td>15</td>
<td>9</td>
<td>20</td>
<td>24</td>
<td>35</td>
<td>4</td>
<td>70</td>
</tr>
<tr>
<td>36</td>
<td>14</td>
<td>34</td>
<td>66</td>
<td>45</td>
<td>75</td>
<td>43</td>
<td>65</td>
<td>64</td>
<td>112</td>
<td>36</td>
<td>134</td>
</tr>
<tr>
<td>26</td>
<td>33</td>
<td>103</td>
<td>32</td>
<td>44</td>
<td>3</td>
<td>81</td>
<td>97</td>
<td>57</td>
<td>127</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>11</td>
<td>11</td>
<td>58</td>
<td>14</td>
<td>16</td>
<td>37</td>
<td>38</td>
<td>7</td>
<td>93</td>
<td>18</td>
<td>78</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>3</td>
<td>36</td>
<td>1</td>
<td>19</td>
<td>23</td>
<td>45</td>
<td>5</td>
<td>42</td>
<td>42</td>
<td>80</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>7</td>
<td>16</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>3</td>
<td>13</td>
<td>11</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>10</td>
<td>66</td>
<td>7</td>
<td>30</td>
<td>37</td>
<td>53</td>
<td>6</td>
<td>64</td>
<td>22</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>5</td>
<td>15</td>
<td>4</td>
<td>21</td>
<td>6</td>
<td>12</td>
<td>3</td>
<td>27</td>
<td>11</td>
<td>48</td>
</tr>
<tr>
<td>28</td>
<td>22</td>
<td>9</td>
<td>113</td>
<td>34</td>
<td>45</td>
<td>92</td>
<td>62</td>
<td>66</td>
<td>9</td>
<td>22</td>
<td>106</td>
</tr>
<tr>
<td>44</td>
<td>22</td>
<td>65</td>
<td>35</td>
<td>32</td>
<td>31</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.7 Ci - C_coda interactions in final syllables (column = Ci, row = C_coda. (Figures in bold - combinations where interaction is suspected)
Apical stops: where the consonants are both apicals, the constraint again holds, as the following table (analysing semantics of only a sample of apical-apical combinations) indicates:

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>/...tVt#/</td>
<td>58</td>
<td>4</td>
</tr>
<tr>
<td>/...tVd#/</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>/...tVn#/</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>/...dVt#/</td>
<td>30</td>
<td>6</td>
</tr>
<tr>
<td>/...dVd#/</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>133</strong></td>
<td><strong>17 (13%)</strong></td>
</tr>
</tbody>
</table>

Table B.8 Semantics of /...VCapical VCapical#/ morphemes

Liquids: as with /#Cliquid... Cliquid#/ sequences, liquids barely cooccur in /...CVC#/ position. Only the sequence /...lVr/ (like /#...#t/) is tolerated, then mainly in loans or expressives: 74% of the sample fell into these classes.

Laminals: The only laminal occurring finally is /s/. As the figures in the table show, /s/ occurs freely morpheme-finally, except where the preceding consonant is a laminal. All the morphemes of structure /...C_laminalVs/ sampled were expressives or loans.

As in the case of /C_initial... C_final/ interactions (cf Appendix B1), strong interactions occur between the laminal stops and the apicals, with 90% of tokens in the sample classifiable in the expressive classes.

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>/...cVt#/</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>/...cVd#/</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>/...jVt#/</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td>/...sVt#/</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>/...sVd#/</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td><strong>Totals:</strong></td>
<td><strong>101</strong></td>
<td><strong>10 (10%)</strong></td>
</tr>
</tbody>
</table>

Table B.9 Semantics of /...C_laminal VCapical/ morphemes

For the dorsals, there are many /...Cdorsal VCapical#/ sequences, though (as with /Cdorsal... Cdorsal#/ and /Cdorsal VCapical.../ combinations), morphemes of this structure are strongly expressive (92% of the sample clearly expressives or loans), with many items belonging either to the 'BAD' or 'SENSE' classes.
Morpheme-internal interactions between consonants

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>/...kVk#/</td>
<td>99</td>
<td>5</td>
</tr>
<tr>
<td>/...kVg#/</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>/...kVŋ#/</td>
<td>30</td>
<td>3</td>
</tr>
<tr>
<td>/...gVk#/</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>/...gVg#/</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>/...gVŋ#/</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td><strong>Totals:</strong></td>
<td><strong>233</strong></td>
<td><strong>18 (8%)</strong></td>
</tr>
</tbody>
</table>

**Table B.10 Semantics of /...C\(_d\)orsal VC\(_d\)orsal#/ morphemes**

Semantic evidence supports the grouping of /h/ with the dorsals in a single 'back' phonological class, morphemes where /h/ occurs tauto-syllabically with dorsals being strongly expressive:

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>/...kVh#/</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td>/...gVh#/</td>
<td>30</td>
<td>4</td>
</tr>
<tr>
<td><strong>Totals:</strong></td>
<td><strong>60</strong></td>
<td><strong>6 (10%)</strong></td>
</tr>
</tbody>
</table>

**Table B.11 Semantics of /...C\(_d\)orsal Vh#/ morphemes**

My conclusion echoes that of the preceding section: except in expressives and loan words, Balinese strongly disprefers morphemes containing /...CVC/ sequences where both consonants are of the same place of articulation. The degree of tolerance of exceptions varies with place of articulation: for labials, only very few tokens are found, for dorsals, there are many.

**Appendix B3 Interactions between consonants in /\#C\(_1\) VC\(_2\) .../ positions, including inherent reduplications**

As with /\#C... ...C#/ and /...C(V)C#/ sequences, the only constraint noted on consonants in /\#C\(_1\) VC\(_2\).../ positions was again one barring consonants sharing (broadly) the same place of articulation, with again the liquids /l/ and /r/ behaving as a distinct sub-class, and the coronals (laminals and apicals) interacting. Interaction of /h/ with the dorsals could not be checked, since /h/ does not occur in syllable-onset positions.

Raw data is presented in tables 12 and 13. In general, the lowest numbers of combinations occur where C\(_1\) and C\(_2\) are of the same place of articulation (even allowing for the general low occurrence of nasals and glides in syllable onsets):
### Table B.12 C₁-C₂ combinations of Balinese (C₂ = non-nasal) (data from Warna (1990), column = C₁, row = C₂).\(^7^8\)

<table>
<thead>
<tr>
<th></th>
<th>mV</th>
<th>mp</th>
<th>mb</th>
<th>nV</th>
<th>nt</th>
<th>nd</th>
<th>rV</th>
<th>rc</th>
<th>rj</th>
<th>ns</th>
<th>nV</th>
<th>nk</th>
<th>ng</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>7913</td>
<td>2</td>
<td>0</td>
<td>44</td>
<td>42</td>
<td>10</td>
<td>23</td>
<td>12</td>
<td>5</td>
<td>39</td>
<td>31</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>76</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>57</td>
<td>5</td>
<td>5</td>
<td>16</td>
<td>10</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>30</td>
<td>2</td>
<td>6</td>
<td>31</td>
<td>19</td>
<td>4</td>
<td>13</td>
<td>5</td>
<td>5</td>
<td>17</td>
<td>16</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>123</td>
<td>73</td>
<td>90</td>
<td>22</td>
<td>6</td>
<td>31</td>
<td>3</td>
<td>5</td>
<td>16</td>
<td>10</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>27</td>
<td>20</td>
<td>3</td>
<td>13</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>25</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>n</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>l</td>
<td>60</td>
<td>56</td>
<td>43</td>
<td>14</td>
<td>39</td>
<td>22</td>
<td>15</td>
<td>19</td>
<td>14</td>
<td>21</td>
<td>31</td>
<td>32</td>
<td>22</td>
</tr>
<tr>
<td>r</td>
<td>20</td>
<td>32</td>
<td>23</td>
<td>14</td>
<td>28</td>
<td>18</td>
<td>7</td>
<td>12</td>
<td>2</td>
<td>13</td>
<td>15</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>c</td>
<td>16</td>
<td>23</td>
<td>4</td>
<td>20</td>
<td>21</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>36</td>
<td>12</td>
</tr>
<tr>
<td>j</td>
<td>7</td>
<td>24</td>
<td>14</td>
<td>12</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>32</td>
<td>13</td>
</tr>
<tr>
<td>ñ</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>16</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>78</td>
<td>58</td>
<td>57</td>
<td>47</td>
<td>48</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>14</td>
<td>50</td>
<td>54</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>y</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>78</td>
<td>43</td>
<td>20</td>
<td>35</td>
<td>31</td>
<td>25</td>
<td>34</td>
<td>32</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>31</td>
<td>23</td>
<td>31</td>
<td>26</td>
<td>50</td>
<td>39</td>
<td>0</td>
<td>14</td>
<td>26</td>
<td>13</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>ñ</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table B.13 C₁-NC interactions

Crucially, for /#/C₁(V)/C₂.../ sequences where the consonants share the same place of articulation, large numbers of forms only occur where C₁ is identical to C₂ (eg pipis 'money'; kuku 'finger nail', and so on). I present arguments in 8.8 that in such cases the surface identity derives from reduplication: underlyingly such morphemes are reduplicated monosyllables (eg pipis from %pis%, plus CV-reduplication) or disyllables (eg kekawa 'spider' from %kawa%, plus "C₁-reduplication". Additional statistical evidence is given in table B.14:
### Table B.14: Numbers of morphemes containing /#Ca VCa.../ sequences, by shape

| C₁ (¼ C₂) tokens | Reduplicated: (i) V₁ = V₂ (ii) C₁-RED (iii) 'Non-reduplicated' forms (V₁ ≠ V₂) |
|------------------|-------------------------------------------------|---------------------------------|
| p 56             | 40 11                                           | 5                               |
| b 66             | 36 21                                           | 9                               |
| m 33             | 26 3                                            | 4                               |
| w 16             | 9 2                                            | 5                               |
| t 46             | 37 1                                           | 8                               |
| d 43             | 29 7                                            | 7                               |
| n 13             | 9 1                                            | 3                               |
| l 71             | 43 17                                          | 11                              |
| r 47             | 29 2                                           | 16                              |
| c 36             | 33 0                                           | 3                               |
| j 31             | 23 6                                           | 2                               |
| ny 16            | 15 0                                           | 1                               |
| s 55             | 43 9                                           | 3                               |
| y 5              | 4 0                                            | 1                               |
| kc 72            | 44 17                                          | 11                              |
| g 36             | 30 4                                           | 2                               |
| ng 6             | 6 0                                            | 0                               |
| **TOTALS** 648   | **456 + 101**                                   | **91**                          |

**Figure 14:** Numbers of morphemes containing /#Ca VCa.../ sequences, by shape (data from Warne 1990)

---

Table B.14: for lexemes of shape /#Ca V₁CaV₂.../ gives the numbers of items where (i) V₁ = V₂ (ii) V₁ is [a], followed by a disyllabic base (eg kekupu, butterfly), and so analysable as C-reduplication, and (iii) V₁ ≠ V₂, and so the lexeme is possibly not a reduplication. Thus of the 648 items listed in Warne 1990 where C₁ is identical to C₂, 557 or 86% can be analysed as inherently reduplicated on this evidence. Only 91 (14%) are not immediately analysable as reduplicates, and some of these no doubt contain reduplication plus vowel overwrite (8.7.1). These figures alone suggest a constraint against simple morphemes containing /#Ca VCa.../ sequences.

Once these underlying reduplicates are excluded, it becomes clear that very low numbers of underlying /#Ca VCa.../ sequences occur. The same constraint as with /...CₐVCₐ#/, and /#Cₐ... Cₐ#/ sequences is seen to hold: essentially only expressives and loans occur.

Detailed analysis of the semantics of morphemes of shape /Cₐ... Cₐ/ and /...CₐVCₐ/ were given in Appendices B1 and B2. A similarly detailed analysis for /CₐVCₐ.../ sequences is not presented here. Instead, a brief survey of apical-apical and dorsal-dorsal interactions is presented below, with similar data on some labial-labial interactions given in the relevant footnotes of the above tables.

For apicals, 88% of all forms of shape /#tvd.../ and /#dvtd.../ were clearly expressives or loans (see relevant footnotes of table B.12 above for data and analysis):
Similarly, a sample of 58 forms containing dorsal-dorsal sequences included 93% expressives or loans:

<table>
<thead>
<tr>
<th>structure</th>
<th>Tokens sampled</th>
<th>unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>/kVk.../</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>/kVg.../</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>/gVk.../</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>/gVg.../</td>
<td>(doesn't occur)</td>
<td>2</td>
</tr>
<tr>
<td>Totals:</td>
<td>58</td>
<td>4 (7%)</td>
</tr>
</tbody>
</table>

These results, with the evidence for reduplication presented above, are sufficient to show that the same constraint against homorganicity in /#C_place... ...C_place#/ and /...C_placeVC_place#/ sequences is found here.

The rest of this appendix consists of the endnotes to the tables presented above.

Data analysis for Table B.1 (Appendix B1) is given in footnotes 1 to 32, that for Table B.7 (Appendix B2) in footnotes 36 to 69, and that for Tables B.12 and B.13 (Appendix B3) in footnotes 70 to 77 and 78 to 82.

Data analysis, Table B.1, footnotes 1-31:

1/ p... b/ 1 token: parab 'kingly name, title.HON' loan Jav.
2/ p... m/ 14 tokens, 10 expressives & loans:
HI/HON: padem 'die, pass away.HON+1sg'; peram loan Jav. 'sleep.HON'; pendem 'bury.HI'; pahum, parum 'meeting.HI'
BAD: palem 'greasy ingrained dirt on body'; COMPLEX: p-igum 'discussion, meeting' listed in Warna under igum)
LOANS: parrem 'kris' loan Old Jav.; plagam 'trophy' loan Malay plem 'film' loan Dutch UNCLASS: 4 items: palam 'y(v, n) crush (meat) to a paste & cook'; paum 'meeting' (and HI variants parum, pahum); pedalem 'feel pity for' probably historically complex *pe-dalem, cf edalem 'shy, reluctant'; pedem 'sleep (of animals)'
3/ b... p/ 3 tokens bengap 'shameless' BAD; and two loans bioskup 'cinema' Dutch; balap 'race(v), overtake' Malay
4/ b... m/ 14 tokens UNCLASS: 3 items; bogem, bugem 'large (wooden stand) with lid, used as a table for food'. UNCLASS; béréam 'k a black cloth' UNCLASS;
PLANTS and ANIMALS: bacem-bacem 'k o forest PLANT'; bayem 'k o spinach-like PLANT'; beem 1. 'fruit/nut of the lontar palm' 2. 'dried blood' BAD, BODY PART; bekasem, bukasem 'preserved duck egg', with a salty-sour taste historically COMPLEX, cf asem 'sour'; bekem 1. 'name of small sea-fish
SENSE: bungkem, 'close mouth tightly and refuse to speak')
LOANS: baem 'jaw-bone' ~ warak '...of rhinoceros' LOAN OLD Jav. BODY PART; baem HON 'jaw-bone' Old Jav. BODY PART; OTHER: balem 'calm, patient' GOOD; b(l)ekem 'hold, restrain (Indonesian: tahan) (true patient verb); beram 'rice wine' (Plant product);
The following 5 morphologically complex forms are listed in two places in Warna and are therefore not included: m-inab 'perhaps.HI' COMPLEX of listed also under inab, other forms inab-ang. Peguyangan often [mainab] which shows analysed as -inab; me-sib 'look like, be similar to' listed under both mesib and sib; mirib 'like; perhaps' COMPLEX listed also under irib; m-unteb 'bright & festive looking (meriah)' (eg of offering) COMPLEX listed under both muntab and untab; m-urub 'shoot up high (of flames)' COMPLEX (listed under both murub and urub).

Plus the following expressives: maen, maam 'eat, childish term' SENSE'; m-esem 'sour' COMPLEX also listed under asem; merem COMPLEX, cf peram; mrim 'cannon' LOAN, Malay; metaram 'k.o. flower' PLANT; miempem 'unpleasantly soft and chewy (of food which should be crunchy)' BAD, loan Jav.; mесem 'smile sneeringly' BAD.

7[t... ...t] UNCLASS at most 6 items out of 16:

timbut 'hole made by animal in ground eg rat, used only as emergency in danger'; topot 'kind of rice cooked in bamboo leaf'; torot me~---- lined up neatly in a row (SENSE?); teket pi-teket 'advice of parents etc to children' UNCLASS; tutut 'follow; obey'; tuut 'follow; imitate'.

Plus the following expressives and loans; tutak 'afraid' BAD; tambet HON (deprecatory) 'stupid' BAD; tenget 'magically dangerous (of place)' BAD; tengkeput 'startled' BAD historically COMPLEX, LOAN Malay ter-kejut; tepet in kali tepet, tangai tepet 'precisely midday' VERY; tengit a~ 'a little bit' SENSE/VERY cf a-kedik 'id.' also with identical initial and final conscientious; tebut 'bubble up (of water)' SENSE; tаit 'flash of lightning' SENSE; LOANS: tungut 'press a claim' LOAN Malay; turut 'obey, follow advice' LOAN Jav. (tutu native term)

9[t... ...t/d] 14 out of 17 forms UNCLASS, plus two not included in count (cf below).

UNCLASS 14: tangai 'clear up and put away (household things)'; tаked 'scales, weighing implement (Bangli)'; tаleb 'base, support (ug cloth on table, leaves used as base for offerings)'; tаmpad 'cut away (grass, on rice-field terraces); tаmpep 'store, clear away'; teket 'arrive at, reach (destination)'; tepud 'go all the way through, soak through; be handed down over several generations'; tild, tulad 'use as (good) example, copy, imitate'; tinking 'wit, mental resources; turn'; tungked 'walking stick'; tued 'base (eg of tree trunk, of neck); tukad 'river'; tuld 'n tool used to push rubbish or to smooth earth eg in ricefields); tangked 'reef, hidden sea-bed rising to near the surface of the water' BAD; VERBS OF PUSING (SENSE 2); tanggiled 'lean (elbow, knee) against sthg'; tamed 'push, apply pressure to eg with stick'.

OTHER, not counted: (The two forms beginning with ti- in Warna are expressive, with perhaps 50% with BAD meanings): ticead, cicedad 'mean, miserly' BAD; t‌emed 'tread on'; RED forms (also not counted): tudtud 'worn out' BAD; tidad 'dribble' BAD; tildad 'in dribs and drabs' BAD; tadrad 'carry in hand' SENSE ...t... ...t/d Sample of the first 30 forms. 16 out of 30 forms UNCLASS; UNCLASS 16: tаgung 'build (vt)' UNCLASS; taban 1. 'bamboo bed' 2. (vt) keep in stall (of animals); tabun 'light a fire using rice husks'; tagen 'be ready and waiting'; tagon tetagon 'fiance(e)'; takon 'ask'; talen 'cutting board'; tаman 'park, garden'; tаmbun 'collect, pile up'; tampin 'roll betel-leaf etc to form a quid'; tаpan 'rolled banana leaf used as drink container in ceremony' (SENSE? cf tampin); taun 'year'; tаwan 'hold prisoner'; tēbēn 'upstream'; teen 'press against'; tegen 'carry on shoulder' HI/HON 5; taen 'ever (aspectual).HI' LOAN?; takon 'ask.HI' LOAN Jav.; tаken te~ 'walking stick.HON'; tаten 'right hand side.HI'; tedun 'descend.HON'.

GOOD/BAD QUALITIES: tagieng 'steadfast in suffering'; tangēn 'have sharp hearing; be easily awakened'; taen 'bear (suffering); suffer (illness etc)'; SENSE? 1; talen 'slowly and carefully' SENSE? PLANT/ANIMAL 3; tamban 'k o sea fish' ANIMAL; tenggulun 'k o plant' PLANT; tawon 'wasp' LOAN from Jav. cf tabanu native Balinese term ANIMAL GRAMMATICAL 2; te'ēn 'ever' aspectual; tekēn 'with' *COMPLEX? *tēka-in? tēka 'come, arrive' 10[t... ...t] 26 items, only 3 UNCLASS: laet, dat, 'dry land'; dapet 'meet, find'; dewet 'k o drink made of coconut milk, rice flour and palm sugar'. VERY: dekot selam ~ 'very black'; daat, dahat 'very.HI' Jav. LOAN; diat, diut 'petegn ~ 'pitch dark'; dugkent 'very short'; True patient verbs: doet 'hit hard' VERY; dukit 'put in a closely spaced line eg prau' SENSE; dalet 'hit'; dalit 'v' glue, close tight' ZV; dampet, damprat, demprat 'yell loudly at' VIOLENCE ZV; dau 'pull (by magical force) ZV; deket 'stick tightly to (v) ZV'; BAD: démnpet 'stuck together (eg fingers, bananas)' BAD; demit 'mean, avaricious' BAD; diut HON me~ en 'goan' BAD.

LOANS: dawat 'ink' Jav.; dompèt 'wallet' LOAN Malay; dresaksat 'like, as though' COMPLEX, LOAN cf sakatsa same meaning, from Sanskrit PLANT: dukut 'k o grass'.

11[t... ...t/d lok 'seawards' COMPLEX d-lod 'LOC-seawards'
Appendix B

18 PLAN/T/ANIMAL: kacwak 'cockroach'; ka tik 'kind of banana'; kalian wak 'kind of tree'; keciplak, kecipluk 'kind of shrub'; kapuk 'kapok'; karuk 'kind of plant'; katak 'frog'; kati k 'stall (of plant)'; kikir 'kind of duck'; kelli 'kind of bird'; kepok pok 'kind of mango'; ketot tok 'kind of plant'; klopak, klopek 'sheath surrounding flowers/fruit (kelopak)'; klutuk 'kind of spirit/animal'; keduduk 'kind of plant'; klampwak 'kind of tree'

6 LOANS/HI/HON: kaduk 'even though'.HI; kedik 'small amount.HI'; kekerik LOAN 'frill.HON'; kekosok 'rice flour, white or yellow, used in a ceremony'.Jav.; kibik 'cubic'.Dutch; kidik 'a small amount'.HI;

6 OTHER: kadék 'title of third-born child' NAME; kampék 'small bag' SMALL; kensi 'very (silent, lonely (of place), deserted)'; kabak 'boy-/girl-friend' FAMILIAR; kencak 'clever, diligent' GOOD; a contents'; a conviction'; a appearance?;

26/k ...... g/ 25/k ...... g/

kecipluk 'kind of duck'; kampwak 'kind of tree'; kecipluk 'kind of shrub'; kapuk 'kapok'; karuk 'kind of plant'; katak 'frog'; kati k 'stall (of plant)'; kikir 'kind of duck'; kelli 'kind of bird'; kepok pok 'kind of mango'; ketot tok 'kind of plant'; klopak, klopek 'sheath surrounding flowers/fruit (kelopak)'; klutuk 'kind of spirit/animal'; keduduk 'kind of plant'; klampwak 'kind of tree'

26/k ...... g/ 25/k ...... g/ Sample of first 30 items (katu-lajeg, kali-manteg not counted, assumed to be complex forms). All 30 were expressives, 28 SENSE, 2 BAD:

SENSE 28: kangeg 'suddenly bought to a halt'; kanag 'of chest' hunched forward'; kaplug, kaplig, kaplug 'crash into' BAD; kebyeg 'be split (water)'; kecig 'jump, hop about'; kokog 'hop, jump'; keplag 'sound of explosion'; keplog, keplug 'fall out in one solid lump (of rice, from steamer)'; kepok 'collide', BAD; kereg 'sound of heavy blow'; kerug 'sound of thunder'; kese yek 'make a syeg sound (of marching feet etc)'; keteg 'beat (of heart)'; ketog 'hit (hollow container) to extract contents'; keteg 'beat fast (of heart)'; ketig 'sound of loom'; kirig 'move backwards?'; kiteg 'walk in a way to plump child'; klanyig 'playful'; kleseg 'sound of heavy steps'; kleteg 'feeling, inner conviction'; klipong 'throw wood at'; kokog 'bathed in sweat'; kelag '1. bend (wood); 2. make a rooster angry by rubbing its thighs; 3. fuck (crude) BAD; kerig 'do (eg, reach puberty)' VERY/SENSE;

BAD: kecig 'be accidentally left behind, forgotten' BAD; kemig 'talk too much' BAD;

26/k ...... g/ Sample of 30 in (1 8) Unclassified at most 4 items, expressives 26;

UNCLASS: kangsang 'a post-death ceremony'; kerang 'strong'; kléwang 'sword' (of frightening appearance?); kalang 'harvest moon (surrounded by large halo)' SENSE;

SENSE: kadang 'family' (used in expressions suggesting close-knittedness, warmth); kempleng 'hit on the head'; kebing 'open and shut repeatedly eg chicken's anus'; kembang 'lie without flowing (water)'; kemprong 'sound of factory bell'; kene ng 'totally absorbed (eg in play)' VERY/SENSE; kepeng 'be a vague rumour'; kenting 'make something sound of an empty drum being bumped'; kuyayang 'move anxiously' BAD;

PL/AN: kaping 'crab's claw (Bl dialect)' ANIMAL; katang 'k o creeper' PLANT/ANIMAL; kentingkong 'k o spider' BAD PLANT/ANIMAL; konsyong 'k o tree' PLANT; kloncing 'k o plant, leaves used in cooking'; klabang 'woven coconut leaves used as wall etc';

BAD: koréng 'scab, skin disease' BAD; kung 'not enough' BAD; keeng 'grumble' BAD;

HI/LOAN: kekulung 'HON throat'; kunambang 'poetic metre' LOAN OJ; koping 'ear' BODY PART;

27/k ...... h/ occurs in 128 different words. Sample of first 60 forms:

14 UNCLASSIFIED (27%): kampyah 'thatch roof extension at the end of a balé, which slopes into the roof and may not join up with the main roof' BAD?: kancuh 1. bucket used on prau to bail 2. use ~ ZV; kantih 'outtrigger'; kapah 'rarely, only occasionally'; karah 'iron ring used eg on cask'; kawah 'dead center of volcano'; kayeh 'bathe vi'; kayuh 'reach, hug'; kebah 'wear for the first time; grow occasionally';

46 expressives (73%):

BAD 11: kamph 'stranded on beach (boat)'; kaash 'suffer for a long time a very serious sickness'; kahiy 'creep up on and suddenly make start'; kaach 'carelessly, thoughtlessly (speak, work)'; kakhah, kakah 'rough, hard (of skin)'; kalah 'be defeated' LOAN?; kambuh 'recur (of sickness)'; kamuh 'unclear'; kocah 'talk non-stop' N—'eat betel non-stop'; kejéh 'afraid, worried'; keleh 'unchanging, the same old state of affairs';

SENSE 11: kaayiah 'hit with a broom' VIOLENCE SENSE; kailacah, kailalah 'rainbow' SENSE; katulayah 'lying stretched out unable to move' SENSE, BAD; kengeh 'look from side to side' SENSE; kebe 'fly wings' SENSE; kebyuh 'suddenly remember' ^an scandal SENSE; kebayah 'suddenly sore as though stung' 2. suddenly remember SENSE; kecu 'sput (v)' SENSE BODY?

BAD?: kekéh 'stiff' SENSE; kenuh 'gale' SENSE; PLANT/ANIMAL 7: kejoh 'kind of grasshopper' ANIMAL; kai neyah 'kind of centipede' ANIMAL; kampuh 'kind of fish' ANIMAL; kai raih 'kind of PLANT'; kecambah 'bean sprout (n) PLANT'; kadoi 'kind of bird' ANIMAL; kencak k.o. PLANT

OTHER 17: kangeh HI LOAN 'considered to be' complex?; kebabé 'left over' LOAN; kalih LOAN HI 'two'; kampuh LOAN asi 'waist-cloth'.Jav.; kantah LOAN ami 'hock (vt, ZV)'; karuh 'have sex CRUDE', of ANIMALS; kash 'affection, love between friends, family members' historically complex, Malay (k)ash; kalah LOAN 'many', HI; kath 'classifier for plants eg bambu' GRAMM; k-auh 'to the
Morpheme-internal interactions between consonants

west' COMPLEX; kedeh LOAN HI "full of hope, desire"; kedih LOAN HI 'ask for'; keeh... (= kaah-RED); kehuu 'somen' BODY PART; kekuulu 'holy water' (GOOD); kekoh 'scratch with nails'; kencah 'broadminded' GOOD;

28/g.../k Sample of 33 items [1 in 3]

At most 5 items UNCLASSIFIED (27%): gakuk 'combine'; geek 'busy (working in the rice fields)'; getak 'lead (cow) vi' SENSE?; gocek 'cause to fight (cock)' (SENSE?); gorok 'sweep (vi) (SENSE?)

25 items 'expressives' (73%):

SENSE 18: gaak 'sound of a crow'; galek 'furious'; gauk 'take greedily'; gelek 'swallow (vi)'; giak 'call out, shout'; gimpkek 'Jembrana dialect'); glendok 'hang swaying'; glanuk 'sway back and forth when walking'; glintik 'lying stretched out on the ground'; gluntuk 'lying stretched out on the ground'; glékonk 'move (body) very slowly'; gletak 'lying stretched out on the ground'; gontek 'rub one's finger against so's body'; gotok 'hammer, beat' (sound); grasak 'attack in a group'; grinkek 'make a pinching/massaging motion'; grisik 'twist, move (body) very little'; gecék 'tick';

BAD 2: gundik 'kept woman' BAD LOAN Malay; géék 'yell angrily', BAD/SENSE;

PLANT/An 5:

UNCLASS GOOD;

31 /g......h/

Sample

32

nation)'

GOOD;

BAD, SENSE;

336

nglotok 'exactly like' VERY; ngluntuk 'roll over and loud' SENSE;

ngrepayak

30/g.../k Sample of 30 [1 in 4 of 1st 120-plus items, forms beginning with gl- & gr- were not included since nearly all appear to be expressives]

UNCLASS 8 items: gedang 'green' SENSE?; gamping 'lime (mineral)'; ganjing 'bamboo stake used when measuring water channels'; gelung 'dancer's crown'; gembung 'pot made of wood'; gibun 'eat together in communal meal'; gêébén 'bundle (of corn, garlic etc)' GRAMM (classifier); gentong 'large pot used to hold water etc' SENSE?

Expressives (73%):

SENSE 8: gangsing 'spinning top'; galing 'slim'; gandong 'carry on back'; gantung 'hang'; gaung 'sound of badly sick or frightened person'; gincang 'curly (of hair)?'; goang 'wind-harp attached to a kite?'; gulung 'cook by turning on a spit; cylindrically shaped (of pillow, cucumber)';

PLANT/An 5: gongdong 'trunk of banana plant' PLANT; gerdang 'fon bamboo' PLANT; gumpang 'padi stalk' PLANT; guntung 'outer part of coconut leaf stalk' PLANT; gerang 'dried fish' ANIMAL;

BAD 7: gendang 'crazy' BAD; gendang 'quarrel' BAD; gëéning 'diminish (eg wealth)'; goléng 'badly proportioned' BAD; gônëng 'shake, sway'; gondrong 'have too-long hair' BAD LOAN Malay; gonteng 'torn earring-hole on lobe of ear' BAD;

OTHER 2: gipang 'k o CAKE'; gônëng 'sway' BAD MALAY;

31/g.../h Sample of 30, every second item of 1st 80 or more on list (forms beginning with gl- & gr- were not included since nearly all appear to be expressives)

UNCLASSIFIED 9 items (33%): gahud 'name of a Pasak clan; gahul 'opportunity'; gavah 'name of offering made from pig's bones'; geduh 1. move (of baby in womb) SENSE? 2. "deep with mud (ricefield)"; gelah 'possession'; gelih 'choose'; genjah 'go'; godoh 'fried banana' CAKE?; goloh 'loose' SENSE?; gubah 'pork dish'

EXPRESSIVES & loans 21: gangsuuh 'gedé' - 'tall and well-built' SENSE; gejah 'elephant' ANIMAL, loan Malay?; genjih 'wobbly, unstable' BAD; gêmeh 'almost stale' BAD; géréh 'thunder' SENSE; geruh 'rough (of skin) BAD; gathah 'sad' PLANT; gidih 'geli benaafsü SENSE, BAD; gileh 'move back and forth (eyes)' SENSE; gimeh 'do with enthusiasm' SENSE; gipih 'do busily, energetically' SENSE; goah 'weep loudly' BAD; goroh 'extravagant' BAD; gupuh 'energetic & busy' GOOD, SENSE; gédoh 'have a large stomach (because in late pregnancy)' SENSE; gingisih 'still in one piece (eg of harvested padi carried home) 2. 'clean' (of padi) GOOD; galuh 'princess (in folk tales) LOAN Old Jav.;

gambuh 'poetic metre' LOAN Old Jav.; gebuh 'easy to till (soil)' GOOD; gemuh 'flourishing (of a nation)' GOOD;

32/g.../k Two UNCLASSIFIED, 7 expressive:

nragcrap 'get an equal share' UNCLASS; ngrasak 'harvest ceremony' UNCLASS

naak 'sound of goose' SENSE; ngalak 'thirsty from heat' BAD; ngetaktak 'suddly weep loudly' BAD, SENSE; nglotok 'exactly like' VERY; ngluntuk 'roll over and loud' SENSE; ngrepayak

336
Appendix B

A detailed study of interactions of /l/ and /r/ with other consonants was not done.

Multi-syllables (ie tri-syllabic or more) ending in /an/ are here considered to bear the suffix -an, and are not counted.

*************************************************************

Data analysis, Table B.7, footnotes 36 to 69

36/...pVp/: 2 loans and 2 %VC%-reduplications: apap 'make flat, level' ZV, SENSE; epep 'hide (in closed hand)' SENSE.
37/...pVm/: apem 'k o CAKE' (plus mlempem 'stale & chewy' (BAD) from Jav., underlyingly reduplicated N-mlempem)
38/...bVb/: No native disyllables. 3 expressive reduplicated %VC%-RED: abab 'stir up', ubub 'thrust into (ashes, mud etc)', eebub 'force in (intending to cause pain)'; plus bobab 'lying, deceitful' (underlyingly probably a reduplicated monosyllable -bab, with vowel alternation) and two Arabic loanwords, via Jav.: rebab 'k.o. stringed instrument', srebab 'ginger drink'.
39/...mVm/: Both reduplicated %VC%-RED: amam 'unenthusiastic' BAD, umum 'general, public' LOAN Malay, plus (apem 'k o CAKE' (plus mlempem 'stale & chewy' (BAD) from Jav., underlyingly reduplicated N-mlempem)
40/...mVm/: Both reduplicated %VC%-RED: amam 'unenthusiastic' BAD, umum 'general, public' LOAN Malay, plus (apem 'k o CAKE' (plus mlempem 'stale & chewy' (BAD) from Jav., underlyingly reduplicated N-mlempem)
41/...tVt/: Includes 48/56 (92%) expressives and loans, plus 4 tentatively neutral/unclassified (though several are good candidates as expressives).

UNCLASSIFIED 5: petet kind of comb with large widely spaced teeth SENSE?; potet 'attendant in gambuh dance'; setut 'mask; thing used as a mask' 2.'prop up, support' SENSE?; bontot 'cooked meat dish'

VERY 7: getet 'VERY clay-ey'; gutit me- 'i(work) diligently and energetically'; bintit '(run) VERY quickly, eg in fear; kotot 'VERY [black]; rentet 'VERY [closely packed together]'; bentet 'exactly, precisely' a round one thousand'; untit '(do) very energetically'

SENSE 16: catat 'steal, take by snatching away from holder'; petit 'small trompong [kettle cymbal]'; atat 'pull sharply, forcefully'; catot, ceto 'peck (of bird); bite (of snake); cetét (i) 'take quickly'; bentet (ii), jetet 'clean cotton by pulling at it very hard with an implement'; bentet 'pull, stretch eg slingshot'; entut 'fart'; kretet 'sound of load carried on shoulder-pole'; petet 'stretch'; pantet block up by pushing something into; pitet 'control (one's lust)'; cantet 'whip vt'; cantet 'hang up (eg on hook, key)';

ANIMAL/PLANT 11: atat 'cockatoo'; bétét 'kind of parrot'; gumitit 'small creatures [insects, small evil spirits] of all kinds'; kotot 'banana-leaf/pineapple-fibre'; bantot male buffalo; kutat 'name of a tree'; pepetat 'kind of small lizard'

BAD 8: bantat 1. 'naughty'; bantat 2. 'abscess'; gantut 'too short (of clothes, rice plant etc)'; kretet 'give not of one's free will'; pantet 'eat CRUDE'; plintut 'untidy; of string, tangled'; plintat-plintut 'be wishy-washy'; ketet 'difficult to get eg water supply'

LOANS/HI/HON 12: ngotot Indonesian/Jav. 'stubbornly persist in a demand'; catat Bl'note'; cantut 'vt pluck, defraud; n tweezeers'; rutut 'order, series' Malay; rénté LOAN (Malay) 'in a row, line'; tuntut 'demand' (Malay); atut, patut [Jav.] 'fitting, suitable (HI)'; cutet 'brief, to the point (HI)'

OTHER 6: gétét, gantet, gétét all grammatical words, classifiers/count words for PLANTS 'bunch (of corn, of coconuts tied together)'; TRUE PATIENT VERBS: buntut 'cut off the tail (of dog)'; catet 'cut off the tip of something';

The following 4 monosyllables and presumed reduplicates were not counted: tatit 'lightning'

42/...tVd/: 2 unclassified, 9 expressive.

UNCLASS: keted 'firm, not muddy (of soil)' (SENSE/GOOD?); letot 'v, produce buds'

SENSE: ated 'fluid, orderly manner of speaking'; lontot 'stick out eg eyes'; lantot 'dribble, hang (eg snot)'; sendot 'lean on another's body, using shoulder or arm'; sántot 'lean on one's elbow'

BAD: sinted 'angrily speak to'; lambot 'be blocked, meet obstacles (of movement)';

OTHER: entut 'knee' BODY PART;

43/...tVn/: Of a sample of the first 30 disyllables of this shape, 5 were unclassified and 25/30 (83%) expressives and loans.

UNCLASS: eñen 'marry', batañ, baten 'below' GRAMM?; eñen 'wake up' (SENSE?)

orten 'sketch'
EXPRESSIVES: anten 'HI younger sibling' LOAN Jav.; antun 'join together by twisting (rope etc) SENSE; banten 'offering' said to be historically complex, a HI form from bal LOAN from Sanskrit; beton 'concrete' LOAN Dutch; binten 'full of joy' SENSE; coton 'take before invited' BAD; gantén 'of (HON) used betel quid'; genten 'pure, 'unstained' (eg of offerings, virginity) VERY, GOOD; ginten, jinten 'a creep PLANT'; janten 'HI definite, clear'; kanten 'HI visible'; kantén 'close friend' SENSE; koton 'HI visible'; ketan 'glutinous rice' PLANT; keten --- 'only realise after the event' BAD; ketén 'unit of 100,000' COMPLEX from keti '100,000'; kontan 'pay in cash' LOAN Malay; kénént 'like that/HI; metén 'pavilion, on mountainwards side of house compound'; nénten 'no.HI'; onton 'birthday' COMPLEX LOAN Old Jav.; panten 'fall in status.HI' BAD; pantun 'padi.HI'

44/...dVv: sample of 30; (1 in 2 of list, not counting those containing #CL...), since these are predictably expressible. 24 expressives (80%, including 17 (57%) SENSE) & loans; 6 (at most) unclassified.

UNCLASSIFIED 6: bandut 'rope used when carrying things on shoulder'; cendet 'brief, to the point' --- 'in short, in sum' SENSE; endut 'mud'; sedat 'shine on (eg sun)' UNCLSS; sodet: nyama ~ 'true, nuclear sibling' UNCLSS ('emphatic?'); indit 'carry in' ~ saput (probably SENSE, not recognised by informants).

EXPRESSIVES/LOANS: adat 'custom LOAN Malay (Arabic); bedét 'pull on, stretch out' SENSE; bedótt 'pull body part (eg person's hair, ear) SENSE, BAD; cadot 'not reach destination; not have opportunity to' BAD; cédót 'sound of coconut falling' SENSE; gedét, gotot 'cut up into large pieces' SENSE, ZV; gendét N ~ mata --- 'stay up, not go to sleep' SENSE, ZV (mata 'eye'); kadut 'carry folded in sarong' SENSE; kedat 'open eyes (eg after sleep)' SENSE; kédít 'tug (v)' SENSE; kendét 'be heatwave' BAD; kédité 'tie the left foreleg (of cattle etc) SENSE; kédité 'credit' LOAN; madat 'opium' PLANT; mendét 'not blink (of eyes)' SENSE; pédité 'dwarf, small in growing' SENSE/BAD; endét, pendét 'not hand back things to rightful owner (Kersten) BAD; pudget 'plump and firm (eg of body)' SENSE; pundat HON 'carry (honoured thing, person)' SENSE; randát 'very slow (eg of walking)' SENSE; redut 'not flow smoothly (eg, discussions)' BAD; sangkodot 'carry on both hands' SENSE; sedut 'suck' SENSE; sindét 'k o grub, lives in ground' ANIMAL; sudat 'cross-eyed' SENSE, BAD; nyédit 'isit straight up, in the middle' SENSE

45/...dVv: adat 'hit with thrown ball, stone eg person'. All are verbs of violence, and 3 out of 4 are VC-RED

46These figures are for bases of form /.../VVC/, where the /l/ is unambiguously syllable initial; the 216 bases of form /.../CVC/, are not analysed here.

47/...IVv: All three VC-RED: aal 'wait a long time' BAD; olol 'swallow (vt) CRUDE/ANIMALS; élél N ~ 1. wait a long time BAD 2. tongue hang out and pant (of dog) SENSE, ANIMAL; Not counted: /élél/ 'not sell (wares)' BAD (assumed to be underlyingly monosyllabic).

48/...IVv: In sample of the first 30 tokens, 8 were UNCLASSified (26%), and 22 (74%) were expressives/loans.

EXPRESSIVES 22: 74% (inc 11 (50%) BAD, 5 SENSE), and loans: alur, galur 'in spate (of river)' BAD; balér 'HI mountainwards'; belar 'be badly behaved (of child)' BAD; belur 'BADly grazed (skin) BAD; celir 'go quickly, clear out' SENSE; celor 'unfaithful' BAD; dolor 'scar on leg from skin disease' BAD; dulur ' supposedly LOAN Jav.; galar 'loose (eg key) because damaged' BAD; galur 'eull out, holler' SENSE; gelor 'come without bringing anything' BAD; gilur 'allowed turn' LOAN Malay; golur 'gently sway' SENSE; iler 'want sthg very much' VERY, SENSE; jalar 'improper' BAD; kelir 'screen used in shadow puppet performances' LOAN Jav.; kelor 'grazed (of skin) BAD; kulur 'very (raw) VERY; kélélor 'k.o. tree' PLANT; mélñor 'cart pulled by two horses' LOAN (= Jav., Malay); palur 'floor' LOAN English; pilur, pilur 'twist (eg ear of naughty child) SENSE; pulir 'form a spiral' SENSE UNCLASSIFIED: palar 'worth, usefulness'; galur 'bamboo slats forming base of a bed'; balur 'lump, piece cut (of meat, jackfruit etc) SENSE'; galur 'family relationship' (Bii dialect); gelur 'capital, finances'; jaldó 'trowsers'; kalor 'groove in a frame'.

49These figures are for bases of form /.../VrVC/, where the /r/ is unambiguously syllable initial; the 148 bases of form /.../CrVC/, are not analysed here.

50/...IVv: delir 'drill (k o cloth)' LOAN; obiral 'sale in shop) LOAN Malay; obrol 'chat' LOAN, Malay

51/...IVv: mirir 'gentle (of breeze)' GOOD, LOAN, complex, Old Jav. irr 'blow (of wind)', a reduplication.

52/...cVv: sample of 25; 23 (92%) expressive, 2 UNCLASS: ancut 'roast well' GOOD; ancut 'isolated place' BAD; becat 'very quick [in walking, running]' VERY; buncat 'pull up, tear out' true patient verb/VIOLENCE; encat 'liquid faeces' BAD; kacat 'mosquito wriggler' ANIMAL; kancat 'part of male loin-cloth which dangles down' SENSE; kecat ' (very) sour' BAD; kincat HI 'a little' (SENSE; koçot 'sound of swaying thing eg [squeaking?] bamboo bed' SENSE; koçot 'piglet' ANIMAL; lancat 'rocking of high waves' SENSE; lecit, licit, licat 'wet through, soaked (eg by rain)' SENSE BAD; mecat 'begin to flow (saliva, sweat)' SENSE; occt --- an 'sound of gender instruments' SENSE; pacat 'flat (of buttocks)' SENSE/BODY PART; pacīt 'k o small bird' ANIMAL;
pincat 'fall (because out of balance)' BAD; pucet 'too pale (of skin)' BAD; gancét 'bunch eg of grapes, of fruit (coconuts, corn etc) etc tied together to make a bundle' GRAMM classifier; UNCLASS?: kecit 'k o cloth, of cotton, with floral motifs or coloured, used for women's blouses'; kénécut 'carved decoration on pillar of building' 53/...c/vd: All expressive, with one possible exception, mincid 'tip of rice grain; fine pieces of broken rice' (PLANT); EXPRESS: enced 'cracks and other damage caused by subsidence' BAD; indec 'broken because wound too tight (of rope) BAD; kecud 'too pale' BAD; kleced 'sound of sthg breaking' SENSE; kéced 'limp' BAD; lándéd 'thick (of liquids); jet [black]' VERY; panced 'very sad BAD; not soft (of cakle) BAD'; planced 'eat repeatedly (more than usual)' SENSE; plinced 'wind (v)' SENSE; pénécad 'quick witted and intelligent; eloquent' GOOD; tlecéd 'mean, miserly' BAD; pascad 'sharp witted' GOOD (so/ sequence suggests LOAN) 54/...c/vs: Two inherent expressives: kecis 'jump about' SENSE; kecos 'jump, hop' SENSE, plus 5 loans: buncis 'k o bean' PLANT LOAN Dutch; nécis 'snazzy, snappy (dress)' LOAN Dutch; karcis 'ticket' LOAN Dutch; percis 'precisely' LOAN Dutch; prancis 'France/French' LOAN Dutch 55/...v/t: In a sample of the first 30 items (none containing CL... sequences), at least 28 (93%) were expressive: UNCLSS: lejt 'sprout, first begin to grow (eg plants)'; rujj 'cut (hair) short' (both sense?) EXPRESS: anjut 'lifted up because out of balance' SENSE/BAD; bebinjat, binjat 'basil, illegitimate child' BAD; beijt 'run around with delight' SENSE; bénjot 'swollen on one side (eg head)' BAD; enjat 1. ~~ 'move up and down' SENSE; light (eg lamp, with a flame) SENSE?; enjot 'move up and down' SENSE; enjut 'set fire to' SENSE of enjít; enjút 'move up and down' SENSE; genjot 'heel BODY PART'; genjotó 'move heel down (eg when pedalling, when kick-starting a motorcycle)' SENSE; genjotó 'speak too quickly, yelling at people' BAD; enjot 'shake, sway (eg house in earthquake)' SENSE; enjút 'speak very quickly, speaking very quickly' SENSE/BAD; pl écéd 'mean, miserly' BAD; pascad 'sharp witted' GOOD (so/ sequence suggests LOAN) ....v/dt: A sample of 25 (1 in 2, not including CL...) contained 23 expressives and loans, only 2 UNCLSS: unclSS: longsút 'rope used as a sling or as a noose/trap'; pusút 'small hand-bore' EXPRESS: asat 'HON dry'; basút 'unit of two hundred' GRAMM; bosót 'exclamation of surprise etc' SENSE; dresaksat 'like, as though' LOAN (Sanskrit draśa 'visible', saksaśa 'clear, visible'); gesut 'tangled' BAD; kaisset 'k o PLANT'; késat 'stiff, difficult to close (of eyelids) BAD; kesat 'tear up into small pieces' ZV; kesát 'tear (vt) ZV; kosot 'rub' SENSE? ZV; langset 'k o fruit' PLANT; lesat 'fly off and disappear' SENSE; lesánt 'stretched out stiffly (of limbs)' SENSE; kejat, kejüt 'stretched out stiffly (of limbs)' SENSE; kejat 'speak very quickly, speaking very quickly' SENSE/BAD; kejet, kejút move repeatedly and stiffly' SENSE; kejít 'raise one's eyebrows SENSE; koját, pejat, pejót 'lame' BAD; lonjat 'speak very quickly' SENSE; lanjút 'tail with white tip (of cows)' SENSE; pijét, pijít 'massage' SENSE?; tengkejüt 'startled' BAD; 66/...v/dt: bejad 'busy' 67/...v/t: Of which kisd 'move place' UNCLASS (SENSE?) Plus 4 items relate to rubbing/grazing, hence UNCLASS: kosed 'graze, surface wound (n) BAD; kosod 'rub'; plés séd 'rub backsides' SENSE; usu'd 'rub'; Also: kosed 'not tough, pleasant to eat (of meat)' GOOD; punsesd 'navel, belly button' BODY PART 59/...vs: Includes 3 VC-RED: aas 'pick in a disorderly way until all gone' BAD, ZV; isas 'not completely dry (of washing)' UNCLSS SENSE?; usus 'intestines' LOAN Malay; besus 'quick to anger BAD; kursus 'course' LOAN Dutch; 60/...k/vk: All 99 were examined: 90 were expressives, 4 loans, at most 5 unclassified; 38 of 99 (38%) had BAD semantics, 17/99 (17%) were PLANT/ANIMAL; 17 SENSE: these 3 classes make up 72% of the total. UNCLASS: akik 'coloured stone, used as jewel' FEEL? LOAN? (= Malay); cangkok 'base, support' (eg, of a ring, a lamp); cekak, pokok 'cut a groove in (eg tree trunk, as foothold)' ZV/sense?; congkok make a hole in (the earth, to plant sthg) ZV; congkok 'stick, pole' LOANS 4: mangkok 'small bowl' Malay; pokok 'capital, finance' Malay; sengkék (BLL dängkék) name of a playing card Malay; tekuk 'fold something which resists eg rattan' Malay EXPRESSIVES: BAD 38 only forms counted, 4 homophones ignored: anglik 'thin and sickly'; bangkuk 'arrogant'; bangkluk 'not sell (of goods)'; bengkek 1. 'short (of body) 2. 'unclear (of sound)'; bengkuk 'stoopeed', SENSE; bengkék 'out of tune' SENSE; bengkuk 'large boil' BODY PART; bungkuk 'stoopeed, hunch-backed' BODY; congkik 'trick by using ambiguous words'; congkok house- or yard-wall which leans inwards in a U-shape; cekak N 'too tight of clothes; cekok (also cekok) 'have a lot of holes eg road';
Morpheme-internal interactions between consonants

cekuk, cekik 'strange' ZV, SENSE; crongkak (also crokak, crokék) 'full of holes eg road'; cékak rude, impolite; dengkak 1. 'parched, dry (of earth) 2. sense name of a shape of kite; dengkak, tengkak 1. 'catch and overpower eg thief 2. short (of body); dengkik, dengkék, tengkék 'scream at, shout at' ZV; dokok 'stupid'; engkok 1. uneven (of earth); 2. arrogant; ikiik 'skin disease' BODY; mangkok 1. not firm, dense enough (eg of heaped soil) 2. imposter; mokak lie, untruth; mrekak speak arrogantly and untruthfully; rekrek, rungkuk N- 'remain silent because sick' HUMAN; slekak N- 'be caught in the throat; sit uncomfortably'; slekék slekak-~ slekék; repeatedly blocked, in fits and starts'; sukak 'have a bone caught in one’s throat'; sengkok 'mis-shapen (of arm)'; tungkak N- 'not yet finished (of work), left unfinished'.

SENSE 17: angkuk 'move the lower body forward and backwards, as though in sexual intercourse'; bikak 'sit with legs spread apart'; brekek, brekekek 'hug tightly' ZV; cangkok, cangkeng 'high pitched (of voice)'; cengkok, cengkeng 'high pitched (of voice)'; cengkang 'unit of measurement, distance from one’s bottom' ZV; 'logical'; cek 'order given to cow when ploughing'; cekak 'short for pekak grandfather' ADDRESS; cekag, cekag 'cracked, split' WALL; cekah 'kind of bird'; cekak-kékak 'roam about, show off; depicting mantis which makes a noise with its wings at night' SENSE, ANIMAL;

rokak 'break by squeezing (eg fruit)'; engkek 'afraid and turns'; l6kok 'untruthfully; rengkek, rungkuk N- 'remain silent because sick'; mangkok 'mis-shapen (of arm)'; tungkak N- 'not yet finished (of work), left unfinished'.
jangkong 'waist, hips' BODY PART; pikang 'groin' BODY PART; pokeng 'cut up into pieces (meat)' VERB of violence

63/...kVh: sample of 30, every 2nd of first 60 (no CL...), at least 28 (87%) of this 30 were expressive or loans:

angkeh  'pant' SENSE; bengkêh 'having a small body and large stomach' BAD SENSE; cekoh, cekêh 'sound of coughing' SENSE; dekêh 'hold breath while straining' SENSE; lingkêh 'sleep with knees pulled up against chest' SENSE; lukuh 'curve backwards (of horns)' ANIMAL/SENSE; lekêh 'circular support for pot' SENSE (like others noting roundness);
cengkêh 'cloves' PLANT LOAN Malay; cenikh 'k o snake' ANIMAL; ikuh 'tail (of ANIMAL)'; jukêh 'sea snail' ANIMAL; pengkâh 'k o ghost with a large stomach' ANIMAL/BAD

Verbs of violence:true patient verbs: cengkêh 'scape out' ZV; cekêh 'cut thickly' ZV; lekêh 'burst, overflow, break open' ZV; bongkah, bungkêh 'pull down, pull apart'

cêkoh 'eat raucously' BAD; jangkêh 'stiff (waist)' BAD; katekêh 'grumble' BAD; okeh 'disturb' BAD; pangkêh 'arrogant' BAD; jungkêh, kekêh  -- 'very (stiff)' VERY; ikuh 'big (of tubers)' GOOD; mokêh 'fat' FEEL/SENSE; akêh 'Hi many'; rengkåh 'Hi work at communally'; langkåh 'pace, stride' LOAN (= Malay, cf native tindak)

UNCLASS: angkuh 'behaviour'; dukuh 'k o priest'; pukuh 'base (eg of wing), basis (of story)'  
64/...gVk: sample of 1st 30 items (not including CL...) 4 UNCLASS, 1 loan, 25/30 (83%) expressive:

UNCLASS: cegak 'stick used as prop etc'; penggak 'small temporary food/coffee stall'; senggak 'proverb'; tagok 'rely on';

EXPRESS: anggak 'be suddenly made to stop' BAD/SENSE; agak 'bluster; boastful behaviour' BAD; aguk 'yell out' BLY SENSE; bêgak 'a great many' VERY; cêgak 'yell at using crude, sharp words' BAD; cêguk 'make?' the sound guk' SENSE; cêgêk 'sound of ducks' SENSE; inggêk 'sway (body)' SENSE; inggok 'strike (!?)' -- 'lame' SENSE; kanggêk 'be suddenly made to stop' BAD; kagêk 'be suddenly made to stop' BAD; kagok 'suddenly a loss and unable to reply' BAD; langgak 'arrogant, insolent' BAD; mëgak 'blustering, boastful behaviour' BAD; mogok 'go on strike; break down (car etc)' BAD LOAN Malay; mëngang 'turn, change directions' SENSE; ogoch 'urge on, repeatedly order (eg s o who avoids work)' SENSE; pegok, pégok 'turn, change direction' SENSE; pêgêk 'name of a neck-movement in dancing' SENSE; ringgêk 'sway' SENSE; seggak -- 'sound of [breathing of a person near death' SENSE; seguk 'sob' SENSE; singguk 'bump with elbow' SENSE; sigêk 'short (of s o's neck)' SENSE; sogok 'push from behind' (ZV, sense); tenggêk 'head (of insect, ANIMAL)'

65/...gVg: At least 21 of 22 expressive (tøgog 'statue' UNCLASS?)

/ BAD 12: agag 'open, gaping, eg wound'; begah 'a very big; canggih 'remain very big' VERY; canggih 'stay awake

all'; dekêk 'awkward' BAD; dekag 'sit with legs spread apart'; ogag, ogég 'large wound or tear'; ogig, ogug, ogég 'mischievous'; sigug 'easily offended'; srengag 'sleep with legs spread apart';

SENSE 6: bëgëg a-- 'a POINT large amount away'; jegegeg 'tall and large, visible from a long distance' BAD; jegög 'pretty' GOOD; jegog 'bamboo xylophone orchestra, very loud, very deep notes!'; pegeg 'assume a fighting attitude'; srengëg 'sit leaning against sthg'; tegeg 'HON firm, resolved';

VERBS OF VIOLENCE: ugar 'pull apart' ZV; uugug 'poke at sthg eg wasps nest, with a stick' both SENSE?

66/...gVg: sample of 30, every 2nd of first 60 (no CL...); 26 expresses (87%), 2 loans.

2 items at most UNCLASS: dagang 'merchant' UNCLASS (loan from Malay?); jagong 'stay awake all night, eating from time to time, usually when there is a death UNCLASS (SENSE?).

Expressions: BAD -- (46%); SENSE 8 (31%): agang -- 'awkward' BAD; agung 'great' VERY; canggang 'loose, not tight'; degeng 'quiet and orderly (of children)' GOOD; enggeng 'open (mouth)' SENSE; enggung 'k o toad' ANIMAL; langgang 'Hi eternal'; leging 'wave backwards at' BAD, SENSE; longgang 'toothless' BAD; léngég 'move very slowly 2. name of dance with graceful movements' SENSE; megeng 'remain silent like someone meditating' SENSE; negung 'for pitifully' BAD; ongang 'toothless' BAD; ogong '1. lacking in speed (reflexes, bodily movements) BAD 2. be unsuccessful' BAD; pøegag 'remain still, concentrating one's thoughts; not move at all' SENSE; punggung 'stupid' BAD; rønggeng 'HON lose, not tight'; rønggang 'broken and in a state of near collapse' BAD; røngeng 'stiff and spread, as though in revulsion (of fingers)' SENSE BAD; sigung COMPLEX? term of ADDRESS from (algung?); tanggung 'carry on pole, each end of which is supported by a different person' SENSE?; tinggang 'aslan (of pot etc incorrectly placed)' BAD; toqog 'not work (eg when others are) BAD; tunngung 'poke backwards out' SENSE

panggung 'stage (in)' (= Malay); senggeng 'spare time' LOAN Malay

67/...gVh: Sample of first 30 (not including CL...)

UNCLASSIFIED 4: anggeh 'buy on credit'; demanggâh 'half cooked (of tubers)' (BAD?); ënggêk 'broad'; cenggah 'fork, branch'.

EXPRESSIVES 26 (83%): agag -- 'walk slowly (of large-bellied person)' SENSE; anggih 'declare one's readiness to do sthg to God' SENSE; begah 'swollen (of stomach, from drinking too much water)' SENSE, BAD; begah 'a lot, very big' VERY; canggih 'like criticising the conversations of other people"
Morpheme-internal interactions between consonants

BAD; cenggeh 'oppose strongly' BAD/SENSE; cegah 'not know polite behaviour' BAD; curenghah pa- 'sprout branches (in great numbers)' PLANT; cégah 'disappointed' BAD; engghah 'Hi get up onto'; ingghih 'yes Hi'; jegah 'VERY full'; kangeh 'Hi considered'; kengghih twara ~ 'not care about' BAD; langghah 'insolent' BAD; legah 'flow fast and in great volume (river etc)' VERY; legah 'big and sticking out (stomach)' SENSE; BAD; lingghih 'HON seat, residence'; lungghah 'unit of counting for sirih/bettel leaves' GRAMM classifier; lungguh 'HON sit'; lēghoh, lēghah~ 'walk slowly because overweight' SENSE, BAD; manggeh 'Hi fixed, determined'; mengghah, minggha 'HON get up on; HON get angry'; oghah 'sway' SENSE; oghoh ~ 'status of evil spirit or monster paraded on Nyepi day' BAD/SENSE; SENSE, for semantic interactions.

Totals do not include underlying monosyllables, including %CLUVC% reduplications (eg cakcak, crucuk), or %CV-CVC% reduplications (eg kekeb); nor do they include tri-syllabics ending in i-an; They do include a small number of %VC% reduplications, such as ubub.

Data analysis, Table B.12, footnotes 70 to 77:

70/tVd.../: 19/23 expressives (83%), 4 unclassified:
SENSE 5: tedong, tedung 'graphic symbol of [a]'; tiding 'smooth and neat'; tuding 'point finger at'; tėdoh 'sit with legs to one side'
BAD 7: tadah 1. 'eat (of evil spirits, giants etc)'; tadah 2. 'boastingly claim to be (verb)'; tedak 'damaged because eaten by animals'; tēdik 'eat (of animals)' (also used as scornful jibe); tuduh 'leaky'; tēdah 'break promise'; tēdàng 'obstacle';
LOAN/H/HON 3: todaya 'see HON'; têdun 'descend (HON)'; tēda 'eat, (Hi, depreciating)'
OTHER GRAMM 4: tidong 'not [ + NPI] GRAMM; True Patient Verbs 3: todes, tudes 'slice' ZV, true patient verb, tudag 'slice open' ZV
UNCLASS 4: tadong 'prepare'; tadas 'clean' (SENSE? GOOD?); tuduh 'cool, calm' (GOOD?); tudung 'cover, lid' (Plus reduplicated monosyllables, uncounted: tadtad 'carry in hand' SENSE; tudud 'worn out' BAD.)
71/dVt.../: One unclassified, 10 loans or expressives
UNCLASS(7): data 'all of (something)'
BAD: data-data 'black magic' BAD; détê 'k. o. spirit' BAD
Hi: datang 'guest (Hi)'
LOAN: datu 'element' LOAN Sanskrit dhaktu 'element'; detik 'second, tick of clock' LOAN Malay; (a)ditia 'sun' LOAN Sanskrit aditàya-, redite 'Sunday LOAN Sanskrit (ld.), ditu 'there' GRAMM, dūta 'emissary' LOAN Sanskrit ducta; détia 'kind of giant' LOAN Sanskrit dātiya 'class of demons'
Plus these reduplicated monosyllables (not included in count) datat, dataté 'vainly dress up' BAD;
datet, diditit 'repeatedly tug' SENSE; 'repeatedly tug',
72/nVt...; all LOANS: nata 'ruler' Skt, natah 'household courtyard' OJav., natar 'household courtyard' OJav., natya 'polite.HI' OJ, niti 'statesmanship' Skt, nitya 'always' Skt, nota 'note' Du; notarisa 'notary' Du; notes 'notebook' Eng; netra 'eye' (literary) Skt.
73/nVd.../ 3 loanwords, one expressive: nada 'sound (literary)' Skt, nadi 'vein in lower arm' Skt, nadyan 'although' (variant of yadian) Skt, nedeng 'heavily laden with fruit (tree)'
74/kVv.../Sample of 20 tall 11 forms not potentially reduplications, plus 9 (taking every third item in list, from stak): 19 expressives, 1 Unclassified.
kekab 'lie face down' SENSE; kaki 'grandfather ADDRESS; kaku 'stiff' LOAN Malay SENSE?; kakung 'male person' LOAN Jav.; kakul 'snail' ANIMAL; kakur 'word used as an oath' BAD, SENSE; kakus 'toilet' LOAN Dutch, kakyasa 'shark' ANIMAL; kakwa 'tortoise' ANIMAL; kakyang 'HON grandfather'; kokila 'k.o bird' ANIMAL; kukih 1. 'interfere with anothers possessions' BAD; kuki 2. 'open little by little using fingers' SENSE; kokab 'sheet-like layer of sugar-palm fibre' PLANT; kalkal 'laugh aloud' SENSE; kekawa 'spider' ANIMAL; kekab 'stiff' SENSE; kekepuh 'k.o. tree'; kekes 'grow smaller from fear' SENSE, BAD
UNCLASS: kikén 'about to [verb]' UNCLASS
75/kVv.../: All expressives: kagok 'suddenly stopped' SENSE, BAD; kagok 'struck speechless' SENSE, BAD; kagyat 'HON startled', kagét 'it' GRAMM, plus 2 uncounted reduplicated monosyllables, both expressives: kígígik 'wander about lazily' BAD, kugkuogn 'k.o eagle' ANIMAL
76/gVv.../: 4 reduplicated monosyllables occur, all with expressive meanings: gekgek 'rhythmic sound of loud laughter, singing', gikgik 'sniffer', guguk 'laugh uproariously (Bl terpingkel-tingkel)', gégék 'sound of people laughing'
77/gVv.../: Sample of first 20 forms listed, 18 expressives, 2 unclassified:

342
SENSE 7: gegang 'of fingers) stretched wide'; gegar 'spread out wide (eg rice to dry)'; geded 'shiver'; gegel 'wavy' (of hair); gega 'shave small pieces off'; giging 'take with hesitating movements because of distaste'; gegeng 'strongly resolved (of will)'; PLANT/ANIMAL 2: gaga 'o. rice'; gagak 'crow'

BAD 3: geget 'stingy'; gegendong 'beggar'; gigihan 'swear or use indecent words';

Data analysis, Table B.13, footnotes 78 - 82:

78/kong...: Figures do not include underlyingly C1-reduplicated bases, or CVC-RED monosyllables. For morphemes like kongkong 'bark (of dogs)', dundun 'wake up (vt)', ie those of shape C1VNC1VC3 where N = C3, there are two possible structures: either as simple morphemes, or as reduplicated monosyllables. In all (nonloanword) cases I assume the latter analysis: kongkong is thus %RED-kong% underlyingly.

Stress patterns pattern this analysis: dundun 'wake up' ('dun'dun), tonton 'watch' ('ton'ton). As well, of the 29 forms of shape #C1VNC1V(C3) listed in Warna (excluding for now those where C1 = /s/) in 23 out of 29 cases, C3 is identical to the homorganic N. The six exceptions are all loans or realanalysed dialectal variants: tuntut 'demand' (=Malay), pimpin 'lead (vt) (=Malay), kangkat 'Hi may' (Jav.), ganggas 'big (Jav.)', or dialectal pumpun 'heat (vt)', and dendeng 'dry in sun (vt)' (dialectal variants of punpun and dengdeng respectively).

In other words in the native vocabulary generally, forms of shape #C1VNC1VC3/ must derive from reduplicated monosyllables. Only /s/-initial morphemes are exceptional in that of 15 items, only 4 are regular in being of shape /sVgVg/. The 11 exceptions are however predictably expressives or loans.

79/pVm...:/ All (but one?) expressives or loans:

paman 'uncle' ADDRESS (= Malay); pamit 'I beg your leave' (Jav.); pamor LOAN Jav. 'lime (mineral); pe-mangku 'temple priest' LOAN Jav. COMPLEX; pemada, pamada LOAN OJ 'punctuation sign', COMPLEX, from pada 'foot' Skt; pemaja 'eye of a crab' ANIMAL; pemash 'deed of village membership' COMPLEX, cf pe-pasih 'title' HON; pemating 'homeless' BAD;

LOANS: pemada 'middle element' from Sanskrit madhya, pemali 'curse' BAD, Malay; UNCLASS?: poma 'word used to finish a prayer' (SENSE?)

80/pVm...:/ pumpin LOAN Malay 'lead (vt)'; pumpun dialectal variant of punpun 'heat (vt)' true patient verb

81/pVmV.../ Loan but one loan and expressives: bama 'k.o PLANT; bamadéwa LOAN Skt 'another NAME for the god Siwa'; bima LOAN Skt NAME; bima LOAN Skt NAME; bumi 'earth, world' BAD, Skt (more commonly in dissimilated form: gumi); plus two mono-syllabics bom 'bomb' LOAN; bum 'boom (naut.)' LOAN Dutch;

UNCLASS?: bumara 'new (literary) COMPLEX? b-um-ara(?) cf mara 'just now, just then (do)'

82/kVNV.../ All 14 are expressive or loans:

kangeg 'suddenly stopped' SENSE, BAD; kangen 'feel moved, pitiful' SENSE; kangin 'to the east' historically complex k-angin 'to the wind'; kangol HON with difficulty'; kengeh 'look to left and right' SENSE; kengg 'Hi may'; kengis 'visible (of thigh, when wind disturbs kamen)' BAD; kingeh 'confused' BAD; kangka 'weak' BAD; kangkat 'Hi may'; kangker 'cancer' BAD, LOAN; kangkah 'Hi considered to be [verb]'; kangeek 'HON stopped momentarily'; kàngguk tuara -- 'not care about' BAD

These figures are higher than those in table B.12 above, since they include inherent C1-reduplications.

84 Morphemes are considered to have undergone inherent C1-reduplication (with insertion of epenthetic schwa) where they contain an initial /CsaC/ sequence preceding a disyllabic base, eg forms such as këgëndong 'tramp', kekuku 'butterfly' (chapter 8).

85 The constraint has lead to occasional dissimilation of borrowed forms, such as kompa 'pump' from Dutch pomp, gumi 'world' from Sanskrit bumi.
Appendix C

Morphemes with final /r/
Appendix D

Headwords in Warna et al (1990) of shape \((\sigma)\sigma[^{\text{RED}}-\text{CVC}]\)

Warna et al (1990) lists 29 tri-syllabic and longer headwords (which they thus implicitly analyse as unit morphemes), of shape \((\sigma)\sigma[^{\text{RED}}-\text{CVC}]\), ie containing a final reduplicated monosyllable. This list includes items assumed here to carry either of the the nominal or verbal suffixes of shape -an. I give the phonemic shape only of all items, as the meanings are not important to the purpose of this appendix (see 4.1, also Appendix F).

a) Morphemes of shape \(k\sigma[^{\text{RED}}]-[^{\text{CVC}}]\), 16 tokens:
   \begin{itemize}
   \item kebetbet; kebilbil; kecalcil; kecemcem; kemangmang; kenyongnyong;
   \item kepirpir; kepohpoh; kepokpok; ketestes; ketoktok; kedongdong; kedokdak;
   \item kedukduk; ngetabtab (%Nkatabtab %); ngetaktak (%Nkatakaktak %)
   \end{itemize}

b) Morphemes of shape \(k\sigma[^{\text{RED}}]-[^{\text{CVC}}]-\text{an}\), 3 tokens:
   kecélcélan; keségségan; kenyangnyengan

c) Morphemes of shape \(k\text{agt}-[^{\text{RED}}]-[^{\text{CVC}}]-\text{t-an}\), 2 tokens:
   \begin{itemize}
   \item katibangbung; kalimatmatan
   \end{itemize}

d) Morphemes of shape \(\text{C}_{\text{lam}}\text{gt}-[^{\text{RED}}]-[^{\text{CVC}}]\), 5 tokens:
   \begin{itemize}
   \item clekongkong; jrekangkang; jrekongkong; nyrekangkang; nyrekongkong
   \end{itemize}

e) Other, 2 tokens:
   \begin{itemize}
   \item kucalcil; mredigdig
   \end{itemize}
Appendix E

1. Semantic classes of inherently reduplicated disyllabic bases (eg kupukupu).

Data analysis, Table 4.1. Sample of the first 60 forms beginning from /1/ in Warna 1990. Only the base is listed for each morpheme, thus kupukupu would be listed as kupu.

PLANT/ANIMAL: labi 'k.o. seaweed'; lumba 'porpoise'; lungit 'k.o. plant'; mali 'k.o. seashell'; muluk 'kind of grass, good for pig food'; omang 'hermit crab'; opok 'k.o. shrub'; paci 'k.o. plant, leaves used as a medecine'; paet 'poisonous water grub'; peja 'kind of creeper used a pig food'; pici 'k.o. small water-snail' 2. 'group of kids' (jocular); pusut 'k.o. fish; k.o. snail'.

BAD: lapu 'always suffer misfortune (Bl terlunta-lunta)'; luu 'rubbish'; mait 'have a disorderly lifestyle'; mangu 'confused'; manyir 'be unreliable; break promise'; mema 'be speechless from nervousness'; minding 'confused' (2. move closer and closer to speaker); mretug 'angrily move hands and legs' ICONIC; murang 'wander about the countryside aimlessly'; neku 'pleadingly beg'; nerung 'walk on and on, in a pitiful state'; ngiuh 'anxious and sighing'; obed 'disorderly clothes or household items, unpleasant to see'; ocoh 'insincere' BD; ogan 'immoral'; ogoh 'paper mache monster'; orék 'aimless scrawl (Bl corat-coret)'; pénung me 'as long as or out of position, in a disorderly way'; puén 'not cared for in a fitting way (eg child)'; ron 'uneven'; rungah 'hurried, pressed for time'.

SENSE:
1. MOVEMENT/(BODY) POSITION: lengoh 'sway from side to side when walking'; lengok 'sit deep in thought [Bl termenung]'; liwa 'tighten a knot by twisting it over and over' movement; nebù 'save money little by little'; ongoh 'puff, pant'; pilì 'back and forth'; rampyak-rampyuk 'walk swaying as though about to faint'.

2. SOUND: nguéng 'sound of a cat'; ngueng 'sound of a beetle'; nguit 'sound of a piglet'.

VERY: namu 'mlecat ~' '[disappear mlecat] totally, without a trace'.

CAKE: ondè 'k.o. cake'; ongol 'k.o. cake/sweet'; orög 'k.o. cake';

SMALL IMPLEMENTS: puluk 'small pot made from clay'; pusi 'small wallet'(cf Malay = pundupundi).

ICONIC - Plural: preci 'all kinds of small things'; pungon 'rays of light which enter a room in the morning'.

LOANS: moga (Malay/Javanese) 'may it be that...'; mugi HI Javanese, (same source as previous); palén 'everyday needs' (Malay).

UNCLASSIFIED: lungka 'cushion'; obag 'doorway'; obéng 'loop in a rope'; pual 'waist' HON (Jav.?); pulapali 'carrying out of a ceremony (n)' probably LOAN from Old Javanese pali 'behaviour, virtue'.
Appendix E

2. Semantic classes of inherently reduplicated monosyllabic bases, of shape CVC

Data analysis, Table 4.2. Sample of the first 60 forms beginning from /l/ in Warna 1990

ANIMAL/PLANT: langlang ~ bé 'k.o. fisheating bird' (bé 'fish'); liklik 'half-ripe coconut'; mahmah 'k.o. civet cat (BL musang)'; linglang, juuk ~ 'k.o. citrus (juuk) which can bring creatures to life'

BAD: laglag 'thin (of hair)'; lahlah 'be contagious (disease)'; laplap, lélép 'be contagious (disease)'; lick (vt, of flames); lengleng 'be enraptured' 2. 'sad, moved'; lunlung 'broken, snapped'; lutlut 'flaking (skin, paint)'; lédéd me~ 'dribble out, slide out (of sticky fluids eg snot, intestines)'; léglog 'bent and twisted because weak (eg of a kris)'; léklok 'hang limply'; linglun 'weak and confused from desire'; liglig 'have no feathers; thin (hair)'; lhihi 'weak point; weakness'; lipip 'k.o. ghost'; longlongan 'hole (eg in a fence)'; lodlod me~ 'stick out (eg wick of lamp)'; loglog 1. blunt; léklé, loklok, loglog 2 'eat (crude)'; lottot 'flaking (skin, paint)'; lémiém, kembang ~ 'very pale (eg after sickness)'; ledled ~ 'magic potion'; lenlen ~ an 'stuck in throat when eating'; mémlé N~ 'do nothing but talk'; mingmang 'move in a restless, unsettled way (eyes, eg because one has a problem) ICONIC.; mihih 'chipped (of blade of a knife)'; mokmok, mukmuk N~ 'rave on in anger', ICONIC sound; molmol N~ 'grumble, moan'; mormor 'blurred, unclear (eg lighting, vision)'; mungmung 'very [deserted]'

VERY: leplep 'sound (asleep)'; pules ~; lumlum ~ 'gading' 'very yellow; bright yellow'; mekmek 'thick & black (of hair)' GOOD

CAKE: laklak 'k.o. cake'

LOANS: langlang Old Javanese 'travel about (pleasurable connotations)'

VERBS of FORCE, VIOLENCE: lebleb 'push underwater with hand'; lelék, lulkul 'force into'; lislis, lésès 'strip bare (bamboo, coconut leaves etc) 0 verb; pagpag 'dig up (field); pull down (wall) 0 verb; palpal 'split (bamboo, for rope)' 0 verb

UNCLASSIFIED: lamlam 'bathe for a long time in cold water' (iconic durative?); léhihé me~ 'open (of door)'; letlet 'k.o. saté' (a piece of meat, then covered with paste); luhluh 'mixture of earth and water for wall'; lingling N~ 'look at carefully; pay close attention to' 0 verb; monmon 'metal ring on the drinking-spout of a water-jug (kendi)'
Appendix F

Potentially monomorphemic lexemes of four or five syllables

The headwords in Warna et al (1990) are normally monomorphemic. In the database there were 229 headwords of four- and five-syllables.¹ Of these, 143 are clearly loanwords, mostly from Sanskrit; these are presented at the end of this appendix. There remain 86 items which are potentially non-loan/native; 7 of these have 5 syllables, and 79 have 4 syllables. The two subclasses are now discussed in turn. I argue that almost all are analysable as containing either complex formants, or more than one morpheme (cf 6.1.1). Only 4 were not thus analysable.

1. **Headwords containing 5 syllables**

There are 7 potentially non-borrowed headwords containing 5 syllables. They all end with the sequence /an/, and begin with /kali/, /kati/ or /tambu/:

(3) kalimatmatan; kalisasuan; katimumbulan; katiwawalan; katiwéwélan; tambulilingan; tamburérétan

I assume that these all consist of two elements: a root, plus a suffix (verbal or nominal) of shape -an. The roots clearly consist of concatenations of mono- or disyllabic complex formants (cf 4.1).

2. **Potentially monomorphemic native lexemes containing 4 syllables**

Of the 79 headwords of this shape, only 4 are perhaps simple morphemes (ie not complex morphemes, or bi- or poly-morphemic): *kapulaga*, *limandaru*, *sengaita* and *cumangkirang*. These are all marginal in various ways - either not in general use, possible loanwords, or both: 1. *kapulaga* 'k.o. medicinal plant' (a borrowing from Malay *kapulaga* 'id.?' ); 2. *limandaru* 'giant, ogre' (literary, not recognised by informants; borrowing?, cf OJ *liman* 'elephant's trunk', Skt *da:ru* 'wood'; and Sanskrit-derived *dewandaru* 'sandalwood' (Gonda 1973); 3. *sengait(a)* 'k.o. cake' (not recognised in Peguyangan; a speaker from Gianyar had tri-syllabic *sengait*); 4. *cumangkirang* 'k.o. smaller gong' (?compound *kirang* 'less.HI' OJ)

The remaining 75 4-syllabics are complex, consisting of concatenations either of complex formants or of morphemes.

1. **Lexemes containing complex formants** *kali-/kati/katu*, cf 3.3.1, 22 items:

¹A small number of polysyllabic morphemes listed in Warna et al, all loanwords, and including names of poetic metres and literary works, were not included in the database.
2. Lexemes assumed complex on phonological, semantic or formal/semantic grounds:

(i) Assumed to contain reduplicated sequences (cf Appendix J), 6 items:

(5) [lili]gundi 'k.o. plant'
    biang[lalah] 'rainbow'
    lalaména 'careless'
    tambul[lélé] 'elephant's trunk'
    suma[lulung] 'new-born chick'
    [dagdig]krana 'method of calculating propitious days'

(ii) Assumed complex on evidence of orthography/pronunciation mismatches (4 items):

(6) dumalada 'average, sufficient'
    kamaligi, kumaligi 'k.o. offering'
    kumaritig 'sound of rifles firing'.

These all are written with <a> in the antepenultimate syllable, even though the vowel is pronounced [ə]. This suggests that these words, and others like them, including ramayana, mahabrata and many other loanwords, are perceived as having a P-word boundary word-medially - except for words such as these, grapheme <a> only realises /a/ P-word-finally (see section 7.5).

(iii) Presumed complex on semantic grounds, 4 items:

(7) bangkebai 'k.o. cannon'
    majagau 'k.o. maja tree'
    basakuma 'intimate language'
    siamata 'k.o. water bird'

I assume these are morphologically complex in the following ways: bangke-bai 'k.o. cannon' (bangke corpse bebai 'evil spirit'), basa-k-uma 'intimate language' (basa 'language k-uma 'to the ricefields'); maja-gau 'k.o. maja tree'; sia-mata 'k.o. water bird' (sia 'nine' mata 'eye')

(iv) Containing recurrent affix combinations (39 items):

(8) a) 4 items of shape pa-[ROOT]-an:
    pangedangan padaringan penarangan penasatan;

    b) 13 items of shape [ROOT]-an:
    su-barengan, lumeténan, lekitikan, brahmantian, cilalongan, jemparangan,
    jempirangan, kukuluban, tambuyukan, tembrérétan, tembuyukan, temisinan,
    temuyukan,

---

2 watulampa, katilampa 'k.o. fruit with hard stone' lampa recurrent element in both, watu Javanese 'stone'.
Appendix F  Lexemes of four or five syllables

c) 22 items of shape ka-[ROOT]-an; ketekulan, kaceplakan, kalengisan, kecacungan, kecingan, kecipitan, kecucutan, kecelcélan, kemaranan, kemerakan, kemurugan, kemuugan, kepasilan, keponakan, kepupungan, kerayapan, kesimbutan, keségésan, ketaluan, keterangan, kewanditan, kenyangnyengan

Loanwords 143 items:

From Old Javanese/Malay (7 items):

(9) galaganjur  'k.o. gamelan tune' (Javanese)
guragada 'rude' OJ
kabinawa  'awesome' OJ
minakadi [-in-, kadi OJ], 'like, for example'
patikonang  [animals kept for slaughter, from OJ pati-ke-wenang]
siliganti  'reciprocal marker' cf Malay silih ber-ganti
tepisiring  'edge' OJ

Sanskrit forms (cf Gonda 1973) 136 items:

(10) alpaprana, anantaboga, anantasana, anresangsia, antaboga, antariksa, antawai, anugraha, anunasika, anusara, apramada, arimbawa, astungkara, aswalalita, atidreti, awiawahara, awignamastu, balasamar, bamadewa, bauandana, brahmacari, bratayuda, cakrawarti, canigara, caniscara, cikrabala, darapati, diatmika, dratikarma, dresaksat, duracara, duratmaka, durmanggala, indrajala, jaladara, jalantara, kajantaka, kalantaka, kalimosada, kantawali, karawista, kulagotra, kulawangsa, kulawarga, lascarya, manancika, mrecapada, nirupeksa, nirwikara, niskarya, nusantara, parajana, parijata, prabangkara, pragiwaka, pramadani, prameswari, pranagata, pranayama, pranidana, pratiksa, pratidina, pratipada, pratiwimba, prayascita, prayojana, predanggapi, predaksina, pretakjana, pula[sari], punarbawa, purnajiwa, purohita, rajasinga, rajawala, ramayana, recedana, sangaskara, samantara, samirana, samodana, sanikala, sanikaon, sanisacara, sanjawani, saraswati, sasadara, satawana, sawatara, sayawkati, sayuakti, senapati, singasana, singasari, sidagori, sidakarya, sidayogi, sidikara, sidiwakia, silangjana, silakrama, siratmaya, sitarasmi, suastistu, suwiwadani, sukasesi, sumanasa, sunantara, suranggana, suralaya, suraloka, surapsari, surudayu, swastistu, swayambara, sekapraya, setubanda, sewambara, tepasana, tigasana, ulantaga, unalatri, upadrawa, upanisada, uparengga, upasaksi, wanawasa, wianjana, wicaksana, widiadara, wijaksara, yadiapin, yadiastu, yadiastun, yogiswara
Appendix G

Analyses of mono- and poly-syllabic morphemes in Warna et al 1990

1 Monosyllabic morphemes

Semantic analysis of a sample of monosyllabic morphemes (cf 5.6, 6.1.1). The sample was generated by a Fiesta (Alsop 1990) search of Wama et al 1990. All 46 monosyllabics of shape V, CV, VC and CLV were analysed for semantic class, as well as a random sample of 145 of the CVC morphemes. Monosyllables of shapes V, VC CLV and CV included an unusually high proportion of functors: 25 out of 46, or 54%.

The 145 morphemes of shape CVC gave the following analysis: unclassified 15 (12%), expressives/loans 130 (90%); 59 (41%) belonging to the criterion classes BAD, PLANTS&ANIMALS or SENSE.

<table>
<thead>
<tr>
<th>SENSE</th>
<th>BAD</th>
<th>PL/AN</th>
<th>LOANS</th>
<th>OTHER</th>
<th>UNCLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>27</td>
<td>6</td>
<td>44</td>
<td>27</td>
<td>15</td>
</tr>
<tr>
<td>18%</td>
<td>19%</td>
<td>4%</td>
<td>36%</td>
<td>19%</td>
<td>12%</td>
</tr>
</tbody>
</table>

Semantic (etc) classes of a sample of 45 morphemes of shape CVC.

A. Morphemes of shapes V, CV, VC, CLV: 46 tokens, of which 2 unclassified, 44 expressives or grammaticals/loans.

UNCLASSIFIED 2 items: bo 'smell' SENSE?, lu 'pestle'

FUNCTORS (25 items): a 1. 3ACTOR pronominal a 2. 'irrealis' marker; i determiner, human nouns; -é GRAMM Definiteness marker, nouns; cí pro 'you'; ke 1 bound preposition ke 2 passive marker; di locative preposition; ké bound question marker; ma verb prefix; ni determiner, human female; nu aspect marker 'still'; né abbrev. of ané 'relative clause marker', 2 allomorph of -é 'definite'; pa 1. verbal, pa 2 nominal prefix pa 3. address abbrev 'father'; pi nominal prefix; pra- nominal prefix (loans from sanskrit); sa prefix 'one; si locative clitic; trï- 'three' Skt; nya (/pa/) 3pro; nyi (/ni/) 2pro, FEM (Badung dialect); -ang verb suffix 'causitive'; -in verb suffix 'locative'

SENSE (including 8 exclamations/expressive particles) 9 tokens: mu 'hold in mouth, eg a sweet' (~ emu); ah 'exclamatory particle'; ih 'exclamatory particle'; ja 'emphatic particle'; ko 'emphatic particle'; te 'emphatic particle'; uh, oh 'sound made when chasing away birds' 'yes'; ya 'emphatic particle'

ANIMAL 1: bé 'fish (generic)'

351
LOANS 6: mi 'noodles' Chinese/Malay; ri 'LOC' Javanese; ro 'two' Javanese; wé Javanese 'water; sun'; ér 'water' Malay; és Dutch 'ice'

ADDRESS (ABBREVIATIONS) 3: ji 'term of address to father' (abbrev of aji); mé 'term of address to mother' (mémé 'mother'); wa 'term of address to uncle/aunt' (~/ua/)

B. Morphemes of shape /CVC/: Sample of 145 items - unclassified 15 (12%), expresses/loans 130 (90%); UNCLASS 15 tokens: bek 'full' (SENSE?); boh 'tame'; goh 'accompaniment to food'; jot 'take gift/offering to someone'; jun 'small pot'; lén 'different'; nap 'subside (rain, epidemic)'; por 'part of loom'; rong 'room, compartment'; tuh 'dry'; tut 2. 'follow'; yag 'widespread' of news (?SENSE); nyat 'dry'; ngad 'bamboo strip cut to be razor sharp' PLANT?/SENSE?; ngeh 'be aware, realise'

BAD 27 tokens: ngab 'unpleasant smell or feeling'; ngeng 'shy, ashamed'; ngeb 'afraid'; nged malaria; ngel 'almost broken?'; ngés 'disappear, be lost on a journey'; bes, bas 'too, overly [state vb]'; beb 'fat and short'; bih exclam. of disappointment; buh 'malnutrition'; dél 'slow in growing (of rice)' PLANT; koh reluctant, lazy; lak exclam. 'noise made when chasing away dog' 2. 'suffering from extreme thirst and heat exhaustion' (of bird); lem 'surreptitiously steal'; lín 'weep'; med 'bored', 'scold' vb; met 'a bit too small (of clothes)'; mis 'rubbish'; mug 'rotten, flaky (of building materials)' - sép 'late'; wuk 'pus mixed with blood which comes out of a corpse'; nyah 'children' insulting term; nye'h 'afraid'; nyep 'afraid, scared'; nyék 'be unreturned (borrowed thing)'; loh '(do) in excess';

SENSE 26 tokens: bah 'fall lengthways'; ngeh 'tightly closed' (mouth etc); nges 1. pressed tightly together 2. 'many; abundant (of tricks/cunning)'; cuh 'sound made when spitting'; ding, dung 'notes of the scale'; gas 'scratch with leg' (of animal); gir 'sound of gong'; lop 'suck in the mouth'; nén 'sound of a bell'; ren 'echo'; seh 'rapid energetic movement when dancing'; sik 'scratch' of hen; sut 'draw lots with fingers'; tíng 'sound of steel when touched'; tét 'sound of cockatoo'; nyit 'set alight (vt, of small light eg candle etc)'; nget exclamatory particle '(do) suddenly'; ngid 'taste or smell like earth; ngon 'astonished'; ngek 'press very hard with the arm/hand'; nyud 'want something excessively (until one's desire can be seen by others)'; béh exclam. of disappointment; lak, sek 'noise made when chasing away dog'; péh exclam. of disappointment

PLANT/ANIMAL 6 tokens: ngas 'smell like that of sheep'; ngit 'small gnat'; bet 'shrubbery'; bun 'creeper, climbing plant'; kap 'kind of plant'; kem 'kind of plant'

Terms of ADDRESS, 9 (all but one are abbreviations): bik 'auntie' (abb. bibi(k)); ngah term of address (abb. nengah); cik address for chinese (abb. encik); l. cok address (abb. cokorda) 2. exclam (cockfight); gég address to young girl (abb. jegég); mék address 'mum' (abb. mémé(k)); nging affectionate address to a child (abb. cenég); tut l. address to fourth-born child abb. (ketut); yuk address for chinese
VERY 2: ngus 'nothing at all'; dil '(do) very often/much';

GRAMMATICAL elements 4: ben 'abbrev.; baan' marks actor in some 'passive' constructions; bin 'again'; duk 'when, subordinator'; sing 'no'.

TRUE PATIENT VERBS 7: geg 'force something into sthg'; leg 'bend'; lung 'break off part of'; mem 'soak vt'; ned 'bash, hit'; pag 'crush'; soh 'cut into (leaves)'.

OTHER 4: lir 'good luck charm: thing believed to bring good luck or happiness' GOOD; pék 'female genitalia' familiar term (abb. pépék); nyang 'not a single [...] VERY'; nged 'bountiful' (of fruit) GOOD.

LOANS 44 items: bang 'red' Javanese; bak Dutch 'sink'; bal 'ball'; ban 'tyre' Malay; bon Dutch credit; cat, cét 'paint' Malay/Chinese; dam 'dam'; dom 'dominoes (game)'; gul 'goal'; jeng 'foot.' HON Jav; jip jeep; kāng 'elder brother' Javanese; kul Dutch 'cabbage'; kén 'servant to a princess' Javanese; lat Du. 'late'; luk be bent (of a kris) Javanese; mal Du. 'model, pattern'; mur Du. 'bolt (n)'; pas 'exact' Malay; pir 'spring' Malay; pot 'pot'; pén 'pen'; rém 'brace' Malay; rug 'destroyed' Javanese; réng Malay/Javanese 'piece of wood to support tiles on roof'; sad 'six' Sanskrit; sél 'cell'; sor 'lower part (HI)' Javanese; sué 'long time HI' Javanese; sép 'chief' Dutch; sung HON Jav 'bestow'; sép Du. 'boss' 2. 'late' BAD; ta9 'pliers' Malay; tan OJav 'no' GRAMM; tong 1. GRAMM no 2. 'barrel' Malay; wang 'money' (Malay); wéh 'give' HI Javanese; woh 'fruit' Javanese; wul 'wool'; wér 'water' Javanese; wor 'mix' Jav.; wét Javanese 'because'; dwi Skt 'two'.

2 Polysyllabic morphemes/lexemes.

Below I analyse a sample of the first 195 headwords containing 3 or more syllables listed in Warna et al. 1990. Loans made up a very high proportion of the sample - at least 101 items, or 52%. Only 13% of the control sample (3.2.3) was identified as loanwords.

<table>
<thead>
<tr>
<th>SENSE</th>
<th>BAD</th>
<th>PL/AN</th>
<th>LOANS</th>
<th>OTHER</th>
<th>UNCLASS</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>12</td>
<td>45</td>
<td>101</td>
<td>16</td>
<td>15</td>
<td>195</td>
</tr>
<tr>
<td>4%</td>
<td>6%</td>
<td>23%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
</tbody>
</table>

Semantic (etc) classes of a sample of 195 trisyllabic morphemes.

Once loanwords are excluded from the count, the proportion of expressive in the remaining 95 items is high - 63 or 66% belong to the three criterion expressive classes, SENSE, PLANT/ANIMAL, BAD. In comparison, 70, or 39% of the 181 non-loans in the control sample belong to the three criterion expressives, Z = 4.55, p < 0.0000 (Median Test, 2.2.3). Again, a tiny probability that the difference is due to random effects.

PLANTS & ANIMALS 45 tokens:
Plants 21 tokens: ambawang 'kind of manggo'; ambulu 'k o plant'; andudu 'k o palm'; baingin 'banyan tree'; bedigul 'sawn-down trunk of tree'; bejadi 'coconut cream from
Mono- and Poly-syllabic morphemes

which the oil has been taken'; bentaro 'k o shrub'; bentawas 'k o tree'; bidara 'k o tree'; biluluk 'fruit of enau palm tree'; bintenu 'name of a tree'; bintawan 'kind of tree'; bintawas 'kind of tree'; blegiding 'k o plant'; bregiding 'kind of plant'; cempaga campaka, cepaga 'k o plant'; cempaluk 'young tamarind fruit'; cenana 'sandalwood'; cengurit 'pinang, areca nut';

ANIMALS 24 tokens: ambwati 'earthworm'; becica, bucica 'kind of bird'; bedawang 'mythical turtle'; bedigal 'maggot in cow dung'; beduka 'spleen of hen'; bintulu 'one-eyed giant'; candila 'stink bug' (from candala? Skt 'unclean'); cenggoret, cengore~cengkoret 'k o insect which makes a réét sound'; cenikih, ceningkih 'k o snake';

BAD 12 items: bejigar 'naughty, rude, habitually so'; bengkékok 'out of tune (note)'; bidugal 'rubbish'; blegandang 'abduct a girl against her will, with aim of marrying her'; blegándah 'tease, annoy so'; blegodah 'bits of padas stone of irregular shape'; brahmantian 'angry' (complex?); bucécér 'unable to keep a secret'; bulisah 'anxious';

BEDUGAL 'clotted blood on wound of a fighting cock'; cenidra 'jealous' (Skt? chidra- > OJ 'perfidious'); cengongol 'be lost in thought alone (bad)';

SENSES 8: bleganjur, galaganjur 'kind of loud martial gamelan'; cekékér 'sound of jungle cock (kékér)'; cekenting 'name of game: coin is dropped to floor, by flicking it with finger'; brekekek 'hug very tightly' cekepung 'song in Sasak language, accompanied by cries imitative of gamelan loan?'; andaru 'meteor, shooting star';

BEDOGOL 'large statue of ferocious guard'; bintulu 'carving [of one-eyed monster]' (Skt? cf bentala); cawiri 'decoration in form of evil spirit' (BAD) Skt?; ceniga 'decoration used in a ceremony, made from palm leaves'

TRUE PATIENT VERBS 2: bedagal, bedigal 'crash into, through';

CAKES, FOODS 8: bégalá mixture of spices used to cook suckling pig (complex?, bé 'meat'); bégina kind of CAKE; bekakak 'roast chicken used spreadeagled as an offering); bekasam, bukasem 'preserved duck egg'; bekayu 'k of CAKE; bukakak' pig roast on one side only, as offering' ANIMAL; béttu 'meat dish' (complex, bé 'meat')

LOANS: 101 items: acintia 'name of God' Skt; adnyana 'thoughts (HON) Skt/OJ; adwitia 'peerless' Skt; agama 'religion' Skt; agastia 'name of a saint' Skt; agorá 'another name of Siwa' (Skt); ahara 'food' Skt; ahimsa 'non-violent' Skt;; akasa angkasa 'heavens' Skt; akret name of a poetic metre' Jav.; akroda 'without anger' Skt; aksama 'request forgiveness' Skt; aksara 'grapheme' Skt; alpaka 'betray' Skt; alpaprana 'non-aspirated sound' Skt; ambara 'sky' Skt; amerta 'water of life' Skt; ampura Jav 'forgiveness, HI'; anamika 'little finger' HON Jav.; antaboga antaboga 'mythical serpent king' Skt; anantasana 'position of body when praying' Skt; andika 'say HON'
Appendix G

Jav.; **anggada** 'name of character in *Ramayana*' Skt; **anggara** 'Tuesday' Skt; **angkara** 'greed' Skt; **anjali** 'pray' Skt; **anesangsa** 'selflessness' Skt; **antaka** 'die' Skt; **antar** 'between' Skt; **antariksa** 'heavens' Skt; **antawali** 'k. o. creeper', Skt (cf Gonda); **antiga** OldJav. 'egg'; **anumana** 'compassion' Skt; **anugraha** 'grant' Skt; **anunasika** 'nasal sound, symbolised in *aksara* with *ulu candra*' Skt; **anugraha** 'nasal sound' Skt; **apitui** 'truly' Old Javanese; **atarawati** 'the fourth weda' Skt yg ke-empat'; **atawa** 'or' Skt; **atidreti** 'name of poetic metre' Jav.; **atita** 'past time' Skt; **awatara** 'incarnation of a God' Skt; **awéci** 'hell' Skt/OJ (cf OJ *awicinna* 'eternity'; **awesa** 'have power' Skt; **awielawa** 'always peaceful' Skt; **awignamastu** 'may there be no obstacles' Skt; **baita** 'boat' Skt; **bamadewa** 'Siwa' Skt; **bancana, bencana** 'disaster' Skt; **bencangah** 'genealogy' OJ; **bangkarna** 'red-eared hen' Skt/Jav.; **baudanda** 'court official' Skt; **begawan** 'priest of Ksatria caste' Skt; **bejana** 'dye water' (Skt 'pot') (transfer of meaning from container to thing contained); **belawa** 'name of Bima (character in *Ramayana*)' Skt; **bendéra** 'flag' Portuguese; **benégá** 'merchant' Skt; **bendésa** 'head of village' (desa 'village' Skt); **bengawan** 'title given to large river' Skt (assumed variant of *begawan*, Skt *bhagava:n* 'illustrious, divine'); **bentala** 'carving [of one-eyed monster]' Skt; **bérawa** 'Hindu sect' Skt; **bepala, cepala** 'insolent' Skt; **biapara** 'confused, in a mess' Skt/OJav.; **biapahara** 'deity' Skt; **bétara** 'purified spirit of ancestor' (< *betara* Skt?); **biakala** 'offering to destructive forces' Skt, *biakaon* 'id.' Javanese form; **biapara** 'confused, in a mess' Skt/OJav.; **biapahara** 'deity' Skt; **biapahara** 'purified spirit of ancestor' (< *betara* Skt?); **biakala** 'offering to destructive forces' Skt, *biakaon* 'id.' Javanese form; **biapara** 'confused, in a mess' Skt/OJav.; **bikupu** 'Dutch 'cinema''; **bisama** 'promise made to gods in return for help in difficult times' Skt; **biséka** 'royal title' Skt; **brahmacari** 'abstinence' Skt; **brahmana** 'caste title' Skt; **brahmanna** 'sun' Skt; **bratayuda** 'literary work' Skt; **bregedél** 'Dutch 'fricadel''; **bujangga** 'priestly sect' Skt; **bukpada** 'honorific address term' complex (*pada* 'foot' Skt); **busana** 'clothing' Skt; **cakrawarti** 'king of kings' Skt; **camura** 'dog' Skt; **candala, condala, centula** 'unclean' Skt; **cangkrama** 'pleasure trip' Skt; **canigara** 'flowering plant' Prakrit (Gonda); **caniscara** 'Saturday' Skt; **cempurit** 'kris sheath'; **cematu** 'dice'; **cekepit** 'small piece of bamboo used to close the tip of a guangan (wind-chime on kite)'; **cematu** 'dice'
Appendix H

CL Clusters

Analysis of a random sample of 120 morphemes containing a CL cluster (L = /r/, /l/) in the onset of the rightmost syllable (see 6.2.2). Sample is the first 120 items generated by a Fiesta (Alsop 1990) search of Warna et al 1990.

Of the 120 items, 91 (76%) fall into the 2 classes BAD, and SENSE (the two subclasses SOUND and MOVEMENT). Only 12 may not fall into one of the recurrent semantic classes.

BAD 31 items: děngklep 'remain silent with a moody appearance'; cangkrım k.o. skin disease; kaplèk 'too soft (of breasts)'; damprat, demprat 'yell at'; gébrot me- 'be obese'; sémprèt 'worn, tattered dress-cloth'; singlād, sèngglād, sènglād 1. 'awkward, not appropriate' 2. obstructed; bunglūn, bunglūh 'burnt on the outside, raw inside'; bunglūn 2 'albino'; bagler, baglūr 'crush (accidentally)'; baglug 'stupid'; gebrag 'yell at'; gebrig gebrag-~ 'move nervously, anxiously'; gébros me- 'visibly sulk'; koplar, koplēr 'blistered from heat (skin)'; saplīr 'miss the target'; sanglīr, tanglīr 'of different sizes (testicles)'; caplis 'VERY [old]' tua ~ 'old and feeble'; caplok 'eat like a pig, gobble up' (cf ambrés N- 'yell at'; banggras '(talk) loudly and harshly'; jébrès 'too hairy (brévok BI)'; kebrus kebras~- 'quick to anger and quick to calm'; kebros me- 1. 'spray from mouth' 2. 'suddenly angry'

SOUND 31 items: kaplet 'sound of object going into a hole'; kaplūt, kaplāk 'slap (vt)' (sound+movement); emprīt, emprēt '~~ kid's toy trumpet'; gebrūt, kebrūt, keprūt 'sound of fart'; kecrēt, kacrīt, kecrīt 'sound of spurting water'; kacrot, kecrot 'spurt (water, v)'; kaprit, kaprēt 'sprinkle (water) v'; keprīt 'sound of whistle'; keprēt 'sound of trumpet'; gedēblār, gedēblār 'sound of rifle'; gedēblūr 'sound of waves breaking'; gedē(br)āg, gedē(br)ēg 'sound of door slamming'; gedē(br)īs 'make a gebris sound when sneezing'; kebrīs 1. sound of sneezing, 2. 'laugh quietly and suddenly'; gedēbvlus 'sound of something being punched or falling'; gekelīs 'sound of boiled peanut being squeezed out of skin'; kecrēs 'sound of spraying from mouth'; keprīs 'hissing sound'; keprus 'sound of spraying from mouth'; plokplo~ 'sound of pigs eating';

MOVEMENT 25 items: sāgrēp 'pounce on prey [tiger etc]'; keblet 'suddenly go out (of flame)'; jumprīt 'roll over (usually, unintentionally)'; sempropt 'spray liquid onto (vt)'; tānggled 'rest both arms/elbows on sthng (body position); gēbrās, gēdēbrās 'stand up suddenly because angry'; ged(e)(de)brōs me- 'spray out of mouth strongly (eg blood); gedebrōs me- 'spray out of mouth'; ceplīs 'escape from a torn net'; ceplos, coplos 'come apart suddenly'; gedēblīs, keblīs, keblos 'suddenly come out, be born'; keplōg
Appendix H

'empty all rice out of container in a single movement'; **keplus**, **keplos**, **keplés** 'be born' (cf **keblos**); **keprés** 'sprinkle water'; **tembres** 'seep; drip'; **sanglapa~ -an** '[in the] wink of an eye'; **sunglap** (sulap BI) 'magic?/conjuring'; **undrus** 'let out (rope of a kite)'; **sungklit** HON/HI 'carry kris (slipped between sarong and body)';

PLANT 4 items: **amplem** 'mango'; **nyangkrut** 'k o sea urchin'; **cinglar** k o small bird; **tinglar** 'k o small bird'

OTHER PERIPHERAL 18 items:

LOANS 8: **amplop** 'envelope'; **coklat** 'chocolate'; **poplin** 'poplin'; **astrén** 'be given magic powers' (historically, Skt *astra* 'mystical syllable', +suffix); **komprés** 'dressing (on wound)'; **obral** Malay 'sale/sell cheaply'; **obrol** 'chat' Malay; **amplas** Malay 'sandpaper';

CAKE 2: **kaliadrem** 'kind of cake'; **semprit** 'k o cake';

VERBS of VIOLENCE/TRUE PATIENT VERBS 5: **sanglot** 'trap, snare vt (of dogs)'; **banglor, banglor, banglot** 'snare (vt)'; **sangkreb** HI? (example is HI) 'confiscate; take over by force';

VERY 2: **saplar** 'absolutely all gone'; **ambles** 'all gone'

GOOD 2: **paklis** 'appropriate, fitting (eg, speech)'; **bondres** 'comic segment in dance; humorous mask'

UNCLASSIFIED 8 items: **sanglup** 'sheath, wrapping, (salut BI)'; **cingkrem** 'compulsory savings by village members'; **sengklat** 'red lacquer; jomprot' 'piled up' (SENSE?); **sanglad ~-an** 'support for head or feet'; **sunglon** 'needle, used to make thatch'; **singlud** 'furthest to the rear' (SENSE?); **geblagan** 'name of offering made in cremation ground after ashes have been gathered'
Appendix I

Exceptions to \([+\text{ATR}]\) height harmony

In this appendix I give data to show that cooccurrence is dispreferred morpheme-internally of \([+\text{ATR}]\) vowels of differing height values (ie combinations of either of \(/o, e/\) with either of \(/i, u/\) in any order), cf 3.1, 6.3.1.

An analysis was done of all 162 headwords in the data base containing such a dispreferred combination, for six distinct patterns (see the following table). The overall figures are given in the table.

<table>
<thead>
<tr>
<th>structure</th>
<th>SENSE</th>
<th>PL/AN</th>
<th>BAD</th>
<th>LOANS</th>
<th>OTHER UNCLS</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>/...iCé.../</td>
<td>-</td>
<td>6</td>
<td>3</td>
<td>16</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>/...éCi.../</td>
<td>-</td>
<td>6</td>
<td>6</td>
<td>22</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>/...uCo.../</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>11</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>/...őCu.../</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>/...iCo.../</td>
<td>-</td>
<td>7</td>
<td>4</td>
<td>16</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>/...őCi.../</td>
<td>1</td>
<td>6</td>
<td>4</td>
<td>29</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Totals</td>
<td>1</td>
<td>25</td>
<td>17</td>
<td>96</td>
<td>6</td>
<td>17</td>
</tr>
</tbody>
</table>

As with polysyllabic morphemes (see appendix F), loanwords made up a very high proportion of the sample - at least 96 items, or 59%. By contrast, only 13% of the control sample (2.2.3) was identified as loanwords. Once loanwords are excluded from the count, the proportion of expressives in the remaining 66 items is high - 42 or 64% belong to just two of the three criterion expressive classes, PLANT/ANIMAL and BAD. Unexpectedly, only 1 item was classified in the SENSE category, but still there was a total of 65% in the 3 test classes. By comparison, 70, or 39% of the 181 non-loans in the control sample belong to the three criterion expressives. The Median Test (3.2.3) result again gives the result $p < 0.0000$ that this difference could be due to random effects.

The full data is given in the rest of this appendix.

/...iCé.../ all 28 items listed in Warna et al 1990

BAD 3: bisén 'attack unexpectedly from behind'; ilé 'dandy' BAD; ingé 'careless' BAD;

PLANTS 6: isén 'galanga plant'; kasiné, kesiné, kasinén'k o selasih'; kecirén'name of vegetable dish; katiwéwélan 'bud of nangka tree'
LOANS, including HIGH/HONORIFIC LEXIS 16: ilén ~ HI 'dance'; ipén HI/HON 'sibling-in-law' Javanese; biséká 'kingly title' HI HON Sanskrit (Gonda); pingé HI 'white' Javanese; wiréh 'because' HI Javanese; icén 'grant' Javanese, historically complex [ica+in]; bidé 'curtain' Malay bidai; tiré 'curtain' Malay tirai; présidén 'president'; redité 'Wednesday' Skt; silé 'razor blade' (<Gillette); résidén 'resident' Dutch; ribén 'sunglasses' (<Rayban); wijén 'sesame' Jav.; wisésa 'powerful' Skt.; wiwéka 'insight' Skt.

3 UNCLASSIFIED

kiré 'get ready'; kikén 'get ready' (probably *COMPLEX); pidén 'try, test' (probably *COMPLEX);

/...é_i.../ all 38 items in Warna et al 1990

BAD 6: kéris 'thin HI'; wéci 'dirty HI'; wégi 'enjoy annoying, mischievous'; béki; bebéki 'maliciously mischievous'; céti 'pimp';

PLANT 6: kasesí, sukasésí kind of PLANT 'water hyacinth'; kényéri PLANT; wéni, kwéni 'k o fruit tree'; déri 'derris, plant used to stun fish (loan? cf Balinese tuba)';

22 LOANS; béji 'spring' Old Javanese; cédí 'name of kingdom in Mahabarata'; déning complex Javanese, GRAMM; dési 'deci-'; détiá 'demon' Skt; déwi 'goddess' Skt; éling 'remember; be conscious' HI Jav; kemédi, komédi Dutch 'cinema'; krédi 'credit'; kénghi HI Jav; léci 'lychees'; méni 'paint used as undercoat' Malay méni; nécis Dutch 'snappy dresser'; nériti Skt 'south-east' (g); péni 'jewel' Old Javanese; présidén 'president'; résidén 'resident'; prési 'shield used in baris dance' Mal. perisai; téji 'race horse' cf Mal. téji; wésia 'name of a caste' Skt; yénghi 'if' Jav.

4 UNCLASS; apéní 'care for (eg flowers)' *complex? (cf -nin LOC); awéci 'hell' Skt?; béri 'k o gong' SENSE?; délign 'small statue of graceful person' GOOD?;

/...u_o.../ all 17 items in Warna et al 1990

LOANS (inc. HI) 11: cukong Malay 'chinese businessman'; kupon 'coupon'; nuroja Skt (Gonda); buron complex buru+an Javanese 'animal (lit. hunted thing)'; kulon Javanese 'west'; tunon 'cremation pile' Javanese; runon, sruron (se-ruron) 'work together', *complex cf Old Javanese runu 'look for'; dudon 'status, position' HI, tumon 'see' Jav id.; purohita 'priest' Skt.

MORPHOLOGICALLY COMPLEX 2; tludo 'odds of three (telu) to two (dua)'; kubon 'hut' *complex cf kubu (+an) 'hut';
Exceptions to [+ATR] height harmony

UNCLASS 4 (the four /...on/ forms are probably complex): pulo 'island' cf Malay pulau; dumon 'attack'; pungon-pungon 'rays of morning sun which enter a building'; upon-upon 'harvest' complex? *[upu-an] cf pupu 'be ready for harvest, from *p-upu?

/...o_u.../ Only 2 items in Warna et al 1990:
bolu Mal. gopura Skt

/...i_o.../ all 30 in Warna et al database

ANIMAL/PLANT 7: bekicot 'snail'; bligo 'kind of gourd'; kalimoko complex 'name of an edible fruit'; kalimoto 'vegetable dish'; limo k o lime; lirgilo k o rose; srikonta 'ko flower' (sri Skt);

BAD 4: linyok 'dishonest, deceitful'; ido, pido 'disbelieve, reject (eg one's parents)'; sipok 'arrogant and impolite';

LOAN 16: ilo 'green' (everyday word: gadang) Malay hijau Jav. ijo; ilon 'take someone's side in dispute' LOAN Javanese ilon 'follow'; kilo 'kilo'; kliwon Javanese 'day of week'; nilon 'nylon'; pitola 'k o cloth' Indian; sinom 'poetic metre' Javanese; wirosa 'angry' Skt; wirom 'fold of cloth at the front of loincloth' Jav.; wirosa 'cruel' Skt; ingon-ingon 'pet' complex cf Javanese ingon-ingon 'id.'; kepiong, pingon complex 'entranced in admiration' cf ngon 'astonished, amazed'; kalmosada 'science of medecine' complex, loan cf usada Skt; tiron 'derived (word)' (grammatical metalanguage) probably complex cf Mal. tiron 'false, not original'; patikonang 'domestic animal kept for slaughter' LOAN (cf Javanese pati 'death' ke-wenang'ruled over';

GRAMM 2: tidong, idong 'not' GRAMM

UNCLASS 1: iso (Karangasem dialect) 'tripe'

/...i_o.../ all 47 items in Warna et al 1990

BAD 4: boki 'worn out earthenware pot'; oding 'falsely artificial manner, behaviour'; ogig 'meddlesome'; okih 'lay claim to (bad)';

PLANT/ANIMAL 6: boni 'k.o. tree'; croring 'fruit tree'; medori 'k. o shrub'; soli 1. k.o flowering plant 2. k. o pine tree; somi 'straw'; kokila 'k o bird';

LOAN 29: doning 'by, because' Jav.; dori 'rear (n) Old Jav. wuri; goni 'k.o. sack' Mal. 'id'; joli 'palanquin' Skt; kodi 'kodi (unit of measurement) Mal. etc; komik 'comic'; kopi 'coffee'; kori 'gate' O Jav.; kosir 'driver' Mal., French cochere?; loji 'k o building' Mal. 'id.; losin 'dozen' Mal. 'id.'; polisi 'police'; ori 'rear (n) Old Jav. wuri; poni 'hair-style' Mal. 'id' (Eng. 'pony-tail?'; prayoga 'high official' Skt; puruhita 'priest' Skt; rodi 'forced labour' = Mal. (< Dutch?); roti 'bread' Ind.; sekoci 'k o boat' =Mal; sidagori 'k o medicinal shrub'; siddayogi 'accomplished yoga devotee' Skt; sloki 'small glass for drinking spirits' = Mal.;
somia 'calm' Skt; sopir 'driver' Dutch; topi 'hat' Mal.; yogia 'fitting' skt; yogini 'female priest' Skt; yogiswara 'yoga teacher' Skt; yoni 'symbol of female genitals' Skt
SENSE 1: tolih 'turn one's head to look about'
COMPLEX 2: kolih, polih see olih below
UNCLASS 5: olih 'get, obtain'; podi 'red jewel'; todia 'HON.see'; kloping 'dry flesh of coconut'; slopi 'k o basket from coconut leaves'
Appendix J

Internal reduplications: data set for morphemes containing non-initial /...IV₁IV₂.../ and /...rⅤ₁rⅤ₂.../

The figures in Table 8.1 are based upon the data presented below, generated by a search of the Warna et al (1990) database, using Fiesta (Alsop 1990) software.

1. Morphemes containing non-initial /...IV₁IV₂.../:

a. $V₁ = V₂$, 19 items: bianglalah; biluluk; blalak; bolong; bolok; biot; bloluk; cluluk; gliling; gliik; gluluk; plilit; plélén; slélég; sumalulung; tamblélé; tambulilingan; tambulélé; wlélé.

b. $V₁ \neq V₂$, 8 items: blalu; cilalongan (~ clalongan); culalong; culalé (~ clalé); klula; klélam; plilit; tulalé; plélos.

c. Loanwords (not included in counts): aswalalita Skt; blulang Mal. belulang; mléla 'k.o. sand' Mal. meléla id.;

2. Morphemes containing the sequence /...rⅤ₁rⅤ₂.../ (ie, where the sequence occurs non-initially):

a. $V₁ = V₂$, 24 items: ngeres; brarak; brorot; bréré; cret; criing; crot; créét; jiring; jirit; kiras; kerek; kérégan; prereng; priring; prurut; préréng; préréét; sririt; soro; sréré; sréré; tamburéréta; tembréréta;

b. $V₁ \neq V₂$, 8 items: brarut; brerong; crir; croring; priret; prirut; urirak; nrarit

c. Loanwords (not included in counts), 3 items (plus variant forms of those): sarira/sria Skt. krura Skt, [urirang'warirang'werirang' sulphur' Javanese],

I have assumed that 4 lexemes of the shape k(e)-...-an (kléran, kilangan, kilikan, krérégan) are morphologically complex, and so have not included them in counts; they have meanings consistent with this analysis, such as an 'adversely affected': kléran 'neglected, not looked after Mal'; kilangan 'have disease of genitals', kilikan 'have small boil on eye, stye'. Their inclusion would not affect the figures.
Appendix K

Diachronic Evidence for Reduplication

In this appendix (see section 8.8), I examine the shapes of morphemes reconstructed by Dempwolff (1938), and Blust (1980, 1983, 1986) for either proto Malayo-Polynesian (Dempwolff), or either of proto-Austronesian, proto Malayo-Polynesian, or proto-Western-Malayo-Polynesian (Blust).

1. In his 'Austronesian Etymologies I-IV' (Blust 1980, 1983, 1986), Blust reconstructs individual lexemes into one of three distinct sets: proto Austronesian, proto Malayo-Polynesian, and proto-Western-Malayo-Polynesian. I have treated these as a single data set for the purposes of this investigation. Blust reconstructs 40 morphemes of shape $C_1VC_1V(...)$, that is, where the onset consonants of the first two syllables are identical. Of these reconstructed forms, the vowels are also identical in 35 (87.5%) cases:

$$V_1 = V_2, \text{ 35 items:  } \text{CaCaw}, \text{ dapdap}, \text{ didis}, \text{ dudu}, \text{ duduq}, \text{ gagar}, \text{ kakak}, \text{ ka(s)kas}, \text{ ke(k)kek}, \text{ ki(k)kik}, \text{ kukud}, \text{ kukun}, \text{ lalatu}, \text{ lalaw}, \text{ lilin}, \text{ lu(q)luq}, \text{ ngel(ng)eng}, \text{ ngingi}, \text{ ngi(ng)ging}, \text{ ngu(ng)ngung}, \text{ ninih}, \text{ nunuk}, \text{ pa(m)pa}, \text{ pipit}, \text{ pupu(t)}, \text{ rara}, \text{ sasah}, \text{ sisi(q)1}, \text{ sisi(q)2}, \text{ susu(q)1}, \text{ susu(q)2}, \text{ tatadu}, \text{ tutub}, \text{ zizi, ziziq}$$

There are 5 'exceptional' reconstructions ($V_1 \neq V_2$): \text{DeDap, lalung, laluq, mamin, sisuq}

2. Dempwolff (1938) reconstructs 90 morphemes of shape $C_1VC_1V(...)$.

Of these, 75 (83%) have $V_1 = V_2$: \text{bab'a, babah1, babah2, babav, bibi, bib'i, bubu', bubuh, bubu[hl], dada', da(n)dang, DaDa', DeDem, d'ad'al, d'id[iL], gagah, gagak, gigi, gigit, gugut, juju', kaka, kekel, keket, kikik, kikil, kuku', kuku, kukuk, kukut', k'ak'ah, k'ak'a[ll], lala', lelet', lilin, lilit, luluD, lulun, lulug, lulung, [L]u[L]uh, mama', maham, mamit', nanah, nanat', nin[i'], nganga, papag, papah, papak, papan, pipi', pipit', pipul, puput, tatay, tatal, tetep, titit', tutu', tutul, [t]ut[t]ung, tutup, t'et'ed, t'it'ih, t'it'ik, t'it'[l], t'it'ip 1, t'it'ip 2, t'ut'u, t'ut'uk, t'ut'un, vavah


3. Combining the two corpuses, of the 130 morphemes reconstructed for either PAN, PMP, or PWMP, in 110/130 or 84.6% of cases, where the onset consonants of the first two syllables are identical, the vowels are also identical.
Appendix L

Other pronoun-avoidance strategies

Apart from those discussed in chapter 13, other strategies for the avoidance of 1st- and 2nd-person pronominals in everyday speech include:

1. Use of the 'DEFINITE' morpheme as a substitute for marking of 1st and 2nd (but not 3rd) person possessors in noun phrases:

   (11) a. sakit basang-é!
       sore stomach-DEF
       [My] stomach is sore

   b. suah- in = a bok-é
       comb-LOC = IRR hair-DEF
       [I]’ll comb [your] hair

   The influence of Style on the interpretation of ke-passives (3ACT readings being only possible in formal styles) is closely paralleled by the rules for interpretation of the definite morpheme: only in formal styles can it be associated with a 3POS interpretation, cf example (17), chapter 13, repeated here:

   (12) ka-gelek-ang poos-é
       ke-swallow-APP saliva-DEF
       [He] swallowed [his] saliva

   If the same phrase were uttered in (Low-style) conversation, only the 1POS reading, 'my saliva', would be possible.

2. Another avoidance strategy is the use of 3rd person pronominal forms with second person reference:

   (13) a. mani déén ingsap-in = a 'ba
       tomorrow just forget-LOC-3 already
       By tomorrow you’ll have forgotten about it!

   b. ke-cakcak tendas-né!
       ke-chop.up head(crude)-3
       [I]’ll chop up your head!

   Note that substitution of the definiteness marker -é for -né in the last example would produce a 1POS reading 'I’ll chop up my head', despite its use to mark 2POS in example (11b), and 3POS in (12). The rules governing the use and interpretation of these morphemes are evidently of similar complexity to those governing choice of bare-stem and ke-passives, just as in many ways they parallel them. They have not to my knowledge been described.


References


___ . A Linguistic key to the early Austronesian spirit world. University of Leiden (ms.). (Keynote address at the Third Eastern Conference on Austronesian Linguistics, Ohio University, Athens, Ohio, May 1983.)

References


Cohn, Abigail C. 1989. Stress in Indonesian and bracketing paradoxes. Natural language and linguistic theory 7:167-216

———. 1990. Phonetic and phonological rules of nasalisation. UCLA working papers in phonetics 76.


References


References


References


Joseph, Boy. n.d. Balinese. ms. Sydney University, Department of Indonesian and Malayan Studies.


Kirtya. [Indicates an unpublished typed manuscript from the folktale collection of the Gedong Kirtya, Singaraja].


References


References


372


373
References


