THE PHONOLOGICAL WORD IN STANDARD MALAY

A dissertation submitted for the degree of Doctor of Philosophy

DEPARTMENT OF ENGLISH LITERARY AND LINGUISTIC STUDIES
UNIVERSITY OF NEWCASTLE
NEWCASTLE UPON TYNE

TAJUL ARIPIN KASSIN

August 2000

Supervisor: Prof. Philip Carr

PAGE NUMBERS CUT OFF IN ORIGINAL

Acknowledgements

I would like to take the opportunity to thank wholeheartedly all those who have supported me, both practically and emotionally, in order that this dissertation be completed.

Firstly, I would like to extend special thanks to my very learned supervisor Prof. Dr. Philip Carr. His profound knowledge in the field of phonology together with his devotion, patience and encouragement provided me with the skills to carry out rigorous research and analysis. He also provided invaluable suggestions throughout the duration of the writing of this dissertation, which have been both immensely helpful and reassuring.

I would also like to thank the Universiti Sains Malaysia (Science University, Malaysia) for providing the scholarship, without which my postgraduate studies at the University of Newcastle could not have been realised.

I am grateful to Dr. S. J. Hannahs for detailed comments on a draft of Carr and Tajul Aripin (ms), which proved very helpful indeed.

I also would like to give thanks to my tutor, Professor Noel Burton-Roberts, who guided me through the academic and practical problems during the period in which I wrote up this thesis. Special thanks are due to the native speaker informants of SM who provided the empirical data for this research. Particular thanks are due to

Norasniza Sailan, Dzulhaliezad Iskandar, Maslita Abd. Aziz, Tuan Zalizam Tuan Muda, Rahmat Ismail, Mohd. Khalil Imran, Ismail Mahmood and Md. Nasir Daud. I also owe a great deal to the friends with whom I frequently discussed this study. Their suggestions and solutions to the problems I faced were both encouraging and enlightening. I would particularly like to thank Md. Nasir Daud, Marzuki Ibrahim, Mohamad Ikhwan and Patrick Honeybone who were ready at any time to discuss matters relating to my work. A further mention should also be made of David Williams who aided me by making intelligent criticism of this thesis.

Finally, I should not forget the great emotional support I received from each member of my family, especially my wife Rashadah Othman and my lovely children. I also would like to extend my warmest thanks to my mother, my father and my mother in law for their prayers, encouragement and love. Their love and care were a deep source of comfort and happiness during times of pressure and stress.

CONTENTS

Acknowledgements i

Contents iii

Abstract vii

INTRODUCTION 1

CHAPTER 1: OVERVIEW

1.0 The Malay language 4

1.1 Brief overview of previous studies on the morphophonology of Malay 8

1.2 Brief overview of the proposed alternative 13

1.3 Data Sources 16

CHAPTER 2: SYLLABLE STRUCTURE IN SM

2.0 Introduction 18

2.1 On Underlying Representations, Syllables and Storage in the Mental Lexicon 20

Evidence from Phonological Acquisition and Misremembered Words

Further evidence and argumentation

2.2 Standard Malay underlying syllable structure 28

2.3 Underlying and surface glides in SM 39

2.4 Summary 41

CHAPTER 3: THE STANDARD VIEW: EVIDENCE IN FAVOUR

3.0 Introduction 43

3.1 Prosodic phonology 43

3.2 The PW and Glide Formation 45

3.2.1 Root-internal glides 47

3.2.2 GF across a root-suffix boundary 57

3.2.3 GF is blocked across a prefix-root boundary 61

3.2.4 GF is blocked across a morphological word boundary 62

3.3 Gemination (Gem) 64

3.3.1 Gem operates across a root-suffix boundary 64

3.3.2 Gem is blocked across a prefix-root and morphological word boundaries 70

3.4 Summary 72

CHAPTER 4: THE STANDARD VIEW: COUNTER-EVIDENCE

4.0 Introduction 74

4.1 Nasalisation 75

4.1.1 Nasalisation in roots and across morphological boundaries 75

4.1.2 Nasalisation and morphological word boundaries 78

4.2 Nasal Obstruent Assimilation 79

4.3 Summary 84

CHAPTER 5: AN ALTERNATIVE ANALYSIS

5.1 Edge-based and non-edge based lexical generalisations 86

5.2 Non-Edge-Based processes 87

5.2.1 /k/ glottalling 87

5.2.2 Nasalisation 88

5.2.3 Resyllabification 89

5.2.4 Compensatory lengthening 97

5.2.5 /a/ Reduction 106

5.3 Conclusion 109

CHAPTER 6: WORD STRESS ASSIGNMENT IN SM

6.0 Introduction 112

6.1 Word stress assignment in SM 113

6.1.1 Stress in roots 114

6.1.2 Stress in roots with schwas 120

6.1.3 Stress in morphologically complex words 126

6.1.4 Stress in reduplicated forms 131

6.1.4.1 Doubling 132

6.1.4.2 Root reduplication 133

6.1.4.3 Partial reduplication 136

6.1.4.4 Rhyming and chiming 137

6.1.5 Stress in compounds 139

6.2 GSI and postlexical word stress adjustment 142

6.3 Conclusion 147

CHAPTER 7: GLOTTAL STOP INSERTION AS A POSTLEXICAL GENERALISATION

7.0 Introduction 148

7.1 Glottal stop in Arabic and Chinese loanwords 149

7.2 The postlexical status of GSI 151

7.3 Conclusion 160

SUMMARY AND CONCLUSION 161

References 166

Appendix I

Appendix II

Appendix III

Abstract

Previous analyses (Teoh 1994 and Zaharani 1998) have claimed that the phonological word (PW) in Standard Malay (SM) is best defined as a stem plus any suffix. This view gains support from the fact that the phonological processes of Glide Formation (GF) and Gemination (Gem) operate across a stem-suffix boundary but are blocked across a prefix-stem boundary. That is, they operate within the PW (thus defined) but are blocked by a PW boundary.

This view is undermined by regular phonological processes such as Nasal-Obstruent Assimilation (NOA) which operates across prefix-stem boundaries, but is blocked across stem-suffix boundaries. We claim that the PW is co-extensive with the morphological word in SM, and that the asymmetry between, on the one hand, GF and Gem and, on the other, NOA is best viewed in terms of a distinction between generalisations based on the right edge of the word and those based on the left edge.

The role of metrical structure in SM is also examined. Our observations show that SM Main stress is assigned from the right edge, while initial secondary stress is assigned from the left edge, thus supporting our distinction between left edge and right edge processes in the lexical phonology of SM.

We claim that Glottal Stop Insertion (GSI) is the default hiatus-avoidance process in SM and is an across-the-board postlexical rule that demands the second vowel be stressed, thus altering metrical structure postlexically.

As well as GF, Gem, NOA and GSI, we also provide analyses of Floating /r/ and Nasalisation (Nas) in SM. Floating /r/ and Nas, we claim, are lexical rules which operate across both prefix-stem and stem-suffix boundaries: unlike GF, Gem and NOA, they are not edge-based generalisations.

The research also examines a set of roots in SM whose syllabic status has been disputed in previous literature. We show that the non-application of GSI is unexpectedly blocked only within roots. We provide empirical evidence by focusing on a Johore onset-reversal language game. Evidence from this game shows that such roots are underlyingly bisyllabic, and we claim that they are phonetically bisyllabic. We also reveal that all such cases contain a sequence of a stressed low vowel followed by an unstressed mid vowel (or lax high vowel) and do not perceptually resemble hiatus sequences. This, we claim, explains the non-application root-internally of GSI as a hiatus avoidance strategy.

INTRODUCTION

In the literature on the morphophonology of Standard Malay (SM), it is widely acknowledged that certain regular phonological generalisations behave differently with respect to the presence of a prefix-root, as opposed to a root-suffix, boundary. Previous analyses of SM morphophonology, such as those of Teoh (1994) and Zaharani (1998), have sought to explain this fact by appealing to the Phonological Word (PW), defined, for SM, as a root plus any suffixes. This approach has its origins in Cohn's (1989) analysis of the closely-related language Indonesian, which in turn relies on the work of Nespor & Vogel (1986). The evidence that is adduced in favour of this hypothesis comes from the domain of application of the SM phonological generalisations Glide Formation (GF) and Gemination (Gem), both of which hold within the PW thus defined. That is, they hold across a root-suffix boundary, but are blocked across what is claimed to be a PW boundary, namely across a prefix-root boundary. They are equally blocked across the boundary between two morphological words (morphological words are the units concatenated in a syntactic phrase, between the two halves of a reduplicated root, or between the two halves of a compound see Carr & Tajul Aripin, ms). One might equally use the terms 'syntactic word' or 'morphosyntactic word for these units.

This hypothesis is falsified by the fact that an equally regular phonological generalisation, Nasal-Obstruent Assimilation (NOA), holds across a prefix-root boundary, but is blocked across a root-suffix boundary. Additionally, the regular

process of Nasalisation operates across *both* a prefix-root boundary *and* a root-suffix boundary, i.e. it is not blocked by the presence of the putative PW boundary.

The problem, then, is that, while there is robust evidence for defining the PW in SM as root (+ suffix, if there is one), there is equally robust evidence for defining the PW in SM as root preceded by prefix (if there is one). We propose a solution to this problem, in which the PW in SM is taken to be isomorphic with the morphological word, and certain morphophonological generalisations are said to be orientated towards a specific edge of the word.

The thesis is structured as follows. In Chapter 1, we examine the previous literature on SM morphophonology and provide a statement of the problem of defining the PW in SM, as well as offering a sketch of our alternative approach. In Chapter 2, we discuss the notion 'underlying representation' as used in generative phonology and allow for underlying syllable structure (following Kaye, Lowenstamm & Vergnaud 1985 and many others), and claim that SM has (C)V(C) syllable structure with underlying vowel-initial roots containing an empty onset position on the skeletal tier. In Chapter 3, we present the evidence supporting the standard view which takes the Phonological Word (PW) to be a root (plus-suffix). Counter evidence for the PW as a root (plus-suffix) is discussed in Chapter 4. Chapter 5 provides our alternative analysis in which the morphophonology of SM is seen partly as a function of edge-based generalisations. In Chapter 6, we seek to support our analysis by presenting evidence from word stress assignment in SM concerning main stress assignment as a right-edge generalisation and initial

secondary stress assignment as a left-edge process, with GSI readjusting stress assignment postlexically. In Chapter 7, we discuss a problem for our claim that GSI is a postlexical, across-the-board, generalisation, namely the non-application of GSI root-internally. We suggest that, although the evidence shows that these roots are bisyllabic their VV sequences are perceived as diphthongs, rather than hiatus sequences, which explains why GSI does not apply to them. Finally, we provide a summary of our main conclusions.

CHAPTER 1

OVERVIEW

1.0 The Malay language

The Malay language, previously known as *bahasa Melayu Purba* (ancient Malay), is believed to have originated from the great mainland area of Asia, arriving as early as 2500 B.C.; it is a branch of the Austronesian family of languages. Speakers of these languages occupied the coastal and lowland areas of Southeast Asia; some travelled even further eastwards to the Pacific Ocean while other groups went westwards as far as Madagascar and southwards to New Zealand (Nik Safiah 1995:1).

According to Asmah (1985), Nik Safiah, Farid, Hashim and Abdul Hamid (1987) and Nik Safiah (1995), the Austronesian language family can be divided into four branches, viz: the languages of the Malay Archipelago (or Nusantara), the languages of Polynesia, the languages of Melanesia and the languages of Micronesia. It has been claimed that the languages of Nusantara consist of some 200 - 300 languages, of which Malay has the greatest number of speakers. According to Asmah (1985), within the Austronesian languages, Malay is the most established[1] language that has been adopted as the national language for some of the countries in Southeast Asia such as Malaysia, Indonesia, Brunei and Singapore. It is recorded in a manuscript called *Sejarah Melayu* (The Malay Annals) that Malay flourished during

[1] By 'established, we mean that, during the centuries of the Malacca Sultanate, Malay became the language of the court, the administration, trading and culture. It is also believed that Malay played an important role in furthering the spread of Islam.

the dominance of the Malay Sultanates i.e. *the Malacca Sultanate* (1400 – 1511) as the centre of international trading in Southeast Asia with Malay playing the role of a *lingua franca.*

In Malaysia, Malay is known as *bahasa Melayu* (Malay). In Indonesia a variant known as *bahasa Indonesia* is spoken. Similarly *bahasa Melayu Brunei* and *bahasa Melayu Singapura* are used in Brunei and Singapore respectively. In addition, a further ethnic Malay language spoken in Southern Thailand is known as *bahasa Melayu Pattani* (Pattani Malay). According to Zaharani (1998), Malay is the national language of four of the Southeast Asia countries - the Republic of Indonesia (population 170 million), the Federation of Malaysia (16.5 million), the Republic of Singapore (3.25 million) and the Sultanate of Brunei (0.25 million).

In Malaysia, Malay is the mother tongue of about 45 per cent of the total population, most of whom are found in Peninsular Malaysia and the coastlands of Sabah and Sarawak. The remaining 55 per cent speak Malay as a second language which is learned formally in schools and universities. These include Chinese, Indian, Iban, Land Dayak, Melanau, Bisayah, Murut, Bidayuh, Kadazan, Temiar, Semai, Jah Hut, and other speakers (Zaharani 1998).

Indonesia has many Austronesian languages spoken as a first language. The most important of these languages are Javanese (60 million speakers) and Sundanese (20 million speakers). As well as these, more than one million speakers use their own languages such as Chinese, Batak, Minangkabau, Buginese, Makassarese, Madurese

and Balinese. Therefore, Malay is not widely spoken except in school. Only 7 per cent of the total population speak Malay as a first language, with the rest formally learning Malay as second language in school (Comrie, 1990).

The Malay language in Malaysia is characterised by a variety of dialects: the Kedah dialect is spoken in Kedah; the Perak dialect is spoken in Perak; the Minangkabau is spoken in Negeri Sembilan; the Sabah Malay dialect is spoken in Sabah; the Sarawak Malay dialect is spoken in Sarawak; the Johore-Riau dialect is spoken in Johore. It has been traditionally considered that the latter, which is predominantly spoken in the southern part of the Malay peninsular (Johore), is the standard dialect of Malay and is here referred to as Standard Malay (SM). It has been chosen the national language of Malaysia due to its long recognised role as the medium between the different ethnic groups which made up the population of Malaysia.

Much effort has been made to establish the Malay language in Nusantara. Principally, in 1967, Malaysia and Indonesia agreed to unite the languages. Initially a committee known as *Majlis Bahasa Melayu Indonesia-Malaysia* (MBIM) (Malay Committee of Indonesia-Malaysia) was formed to revise spelling, pronunciation and technical terms (Abdullah and Ainon 1994: 86). In Malaysia, the *Dewan Bahasa dan Pustaka* (DBP) (Language Planning Agency) was given a mandate to pursue this kind of goal. As a result, in 1972, *sebutan baku bahasa Melayu* (the standard pronunciation of Malay) was formally produced with the aim that spelling and pronunciation were standardised. In this study, it is referred to as Literary Standard

Malay (LSM). The reform of the language concerned not only spelling, pronunciation and technical terms but also, as the DBP was responsible for other reference materials such as magazines, journals and books to meet contemporary requirements, changes in these areas also took place. In 1984 and 1985 Brunei and Singapore joined the organisation, now known as *Majlis Bahasa Brunei-Indonesia – Malaysia* (MABBIM) (Malay Committee of Brunei-Indonesia-Malaysia).

In Malaysia, SM and LSM exhibit certain differences, largely confined to the pronunciation of orthographic <a> and <r> in word-final position. Firstly, in SM (in our opinion), orthographic 'a' in word-final position is not pronounced as a low back vowel [ɑ], instead it is pronounced as a schwa [ə]. By contrast in LSM, orthographic <a> in word-final position *is* pronounced as a low back vowel. Secondly, in SM, orthographic <r> is pronounced as a flap [ɾ], but is not pronounced except morpheme-internally. By contrast in LSM, it is produced as a velar fricative [ɣ] and must be uttered under all conditions (i.e. morpheme-internally, and across morphological boundaries and morphological word boundaries). Finally, in SM, orthographic <i> and <u> in stem-final closed syllables are pronounced laxed as mid vowels [e, o]. In LSM however the high vowels /i, u/ are not laxed in any position. These imposed spelling pronunciations in LSM, all of which seek to reverse historical phonological changes in SM, raise interesting questions about (a) the relationship between literacy and phonological knowledge and (b) the role of normativity in phonology, particularly with respect to the status of phonological knowledge in the current Chomskian conception of I-language. We do not pursue these issues in any depth here, but we will touch on them at several points.

SM has adopted terms from regional and ethnic dialects and from other languages (particularly Arabic and English); this is relevant for some of our analyses. In contemporary life, SM is the medium of all sectors of social, political, education and economic exchange. By contrast the English language is only used among the English educated sector of the population, which is a very small percentage of Malaysian society as a whole (Asmah 1977: 1).

1.1 Brief overview of previous studies on the morphophonology of Malay

An extensive overview of analyses of SM morphophonology in the previous literature would take too much space and would stall our discussion of our central problem. But we do need to briefly review here claims made in the existing literature concerning (a) the overall stucture of syllables in SM, (b) the shape of the underlying inventory of consonants and vowels and (c) the status of the PW in SM. The previous literature in question is: Yunus (1980), Farid (1980), Durand (1987), Teoh (1994) and Zaharani (1998).

Yunus's (1980) work is compiled from lecture notes, and was used as a main reference for undergraduate students at the University of Malaya during the late sixties and seventies. One of our concerns (in chapter 2) is to lay out the underlying inventory of SM vowels and consonants before proceeding to address our central question, namely the definition of the PW in SM. In this connection, Yunus claims

that the Malay phonemic inventory is comprised of 6 vowels - /i, u, e, o, ə, a/ and 19 consonants - /p, b, t, d, dʒ, tʃ, k, g, ʔ, s, h, m, n, ɲ, ŋ, l, r, j, w/, that is, including underlying glottal stop and two underlying glides. Yunus gives brief articulatory descriptions of the Malay segments, as well as their distributions within words in three environments (i.e. word initially, medially and finally). He claims that most SM words are disyllabic; monosyllabic and polysyllabic words are generally borrowed. He also claims that Malay is a language with a (C)V(C) syllable structure. We argue, in later chapters, that SM has no underlying glottal stops. We also argue that SM has underlying glides, but that these are high vowels in non-nucleus peak position in underlying representations. And we take 'vowel-initial' morphemes to begin with an empty nucleus slot. In as much as the (C)V(C) notation encodes this, we agree with it. But CV(C) could equally well be taken to express the same idea. It is unimportant for us which of these two notations is used; what matters are the claims that SM lacks underlying glottal stops and that 'vowel-initial' morphemes contain empty onsets. We return to the inventory of consonants and vowels, and to SM syllable structure, in chapter 2.

Farid's (1980) work falls within the framework of early generative phonology. It attempts to describe certain phonological and morphological alternations found in the language, including the generalisations we examine in this thesis. The regularities are captured and formalised as rules using an SPE-type formalism. Under Farid's analysis, glottal stop in SM is not regarded as one of the underlying phonemes, since its occurrence is highly predictable. We agree with this, and show why in chapter 2. With respect to syllable structure, both Yunus and Farid

agree that Malay is a (C)V(C) language. Again, in as much as this means that syllables may begin with an empty onset slot in SM, we agree.

Durand's (1987) analysis of SM phonology is couched in Dependency Phonology terms and argues that the phonology of SM does not require a category of underlying glides. They are, for Durand, simply high vowels in non-syllabic positions. In this respect, Durand (1987:98) points out that 'the majority pattern seems to be in favour of treating any high vowel as non-syllabic when preceding a non-high vowel'. Thus, he suggests that the output [hi.jas] is best analysed as underlying /hi.ias/ (i.e. two identical vowels in sequence), while [bja.sə] is underlyingly /biasa/. We will question the analysis of words like [hi.jas] below. Under Durand's analysis of SM syllable structure, a system of complex onsets and codas is allowed for, for example in the words /biasa/ [bja.sə] (CCV.CV) and /nack/ [najk] (CVCC) respectively. We will argue below that the first of these is correct but that words such as [najk] contain diphthongs.

Teoh (1994) abandoned the earlier linear representations of standard generative phonology in favour of a non-linear, feature-geometry approach. For instance, vowels and consonants are represented in the hierarchical model of Sagey (1986), and underlying segments are organised hierarchically into syllable structures built by an ordered series of basic syllabification rules in the style of Steriade (1982) and Levin (1984). Like Yunus (1980), Teoh (1994: 12 & 52) claims that Malay has 19 consonants and 6 vowels in its phonemic inventory. But, apparently contrary to Yunus (1980) and Farid (1980), Teoh (1994) claims that Malay basic syllable

structure is CV(C), which suggests that the requirement for an onset is obligatory. We suggest again that there is little at issue here other than the interpretation of the notations '(C)' and 'C': either may be interpreted as meaning that SM allows for empty onsets (which it does). The only other point of issue is whether SM has underlying glottal stops. We will argue that it does not. Importantly, Teoh claims that the PW in SM consists of a root + suffix. We query that claim at some length in the chapters that follow.

Zaharani's (1998) unpublished Ph.D. dissertation concerns the interface between phonology and morphology in prefixation and suffixation. One aspect he concentrates on is Malay reduplicated forms and root-reduplication: a process of copying the base root, most often in conjunction with prefixation and suffixation[2]. The work is based on the theoretical framework of Correspondence Theory (McCarthy & Prince 1994 and 1995), set within the constraint-based approach of Optimality Theory, where the relations between Input-Output Faithfulness and Base-Reduplicant Identity are formalised in terms of a set of formal constraints.

Unlike Teoh, Zaharani claims that Malay is of the (C)V(C) syllable structure type, as did Yunus (1980) and Farid (1980). As we have noted, this is perfectly reasonable if it means simply that SM syllables contain empty onsets. Like Teoh, Zaharani claims that a combination of root and suffix constitutes a phonological word (PW) in SM, and that the PW thus defined constitutes the domain for the

[2] This is the most productive reduplicated form in SM.

application of phonological rules. Additionally, he suggests that such a domain is not formed when a stem combines with a prefix (p. 164).

But this claim is flatly contradicted earlier in his thesis (p.107) where he states that

'generally, the phonology of suffixation reveals that the visibly active processes in the language are inapplicable in this particular domain, as if there was a barrier at the stem-suffix juncture blocking the application of the regular processes'.

In claiming that the root + suffix boundary block regular processes, Zaharani is referring to the fact that Nasal Obstruent Assimilation (NOA) does not hold at the stem-suffix boundary: the application of NOA is a mirror image of the application of Gemination (Gem) and Glide Formation (GF). But it is simply untrue that the visibly active processes in the language are inapplicable at stem + suffix boundary, and Zaharani's definition of the PW in SM rests on the fact that there are robust generalisations which hold across a root + suffix boundary (namely, GF and Gem).

The problem we seek to resolve is the one which gives rise to Zaharani's contradiction: on the one hand, there is evidence from robustly regular phonological processes (GF and Gem) that root + suffix forms a PW which constitutes the domain of those processes; on the other hand, there is equally robust evidence from another regular phonological processes (NOA) that prefix + root constitutes the domain of

application of those processes, and that they are blocked at root + suffix boundary, as if there were a PW boundary there.

To make matters more complex, there is yet another robust process (Nasalisation: Nas) which operates across both prefix + root and root + suffix boundaries. In short, the evidence does not point clearly either to a definition of PW as prefix + root or as root + suffix. We set ourselves the main goal of defining the PW in SM and explaining the differential behaviour of these processes. Accordingly, we offer an alternative analysis of the PW in SM: we claim that Nas, NOA, Gem and GF all fall within the domain of the PW, defined as being *isomorphic with* the morphological word (MW). By 'the morphological word', we mean any sequence of root plus affixes, if they appear. Under this definition, all of the following count as an MW: bare root; prefix + root; root + suffix. Reduplicated forms, we claim, are reduplications of MWs, and syntactic structures and compounds are concatenations of MWs. We claim that SM does not differentiate between PW and MW.

1.2 Brief overview of the proposed alternative

Given an analysis in which the PW in SM is isomorphic with the MW, the problem remains of how to account for the difference in the behaviour of the relevant generalisations with respect to prefix-root and root-suffix boundaries. A solution to that problem which takes SM word stress assignment algorithm into

account is presented here[3].

In order to account for the asymmetrical behaviour of Gem, GF and NOA, we claim that, in the lexical phonology of SM, GF and Gem are right-edge (of the PW/MW) rules, whereas NOA is a left-edge process. We also claim that right-edge (of the MW) rules in SM apply prior to left-edge rules[4]. We also claim that right-edge generalisations are blocked by left-edge affix boundaries (i.e. prefix boundaries), while left-edge generalisations are blocked by right-edge affix boundaries (i.e. suffix boundaries). That is, edge-based generalisations are limited in their scope by a type of locality constraint: they extend across no more than one morphological boundary from the relevant edge[5] (Carr & Tajul Aripin, ms).

We provide independent evidence for these claims by showing that primary stress assignment is a right-edge-of-the-MW process which can not, in principle, penetrate into prefixed material, while initial secondary stress operates from the left edge, and cannot affect suffixed material (see Cohn 1989). Thus, the application of right-edge effects prior to left-edge effects derives from the fact that main stress is assigned (of necessity) prior to secondary stress.

The other word-based generalisations (Nasalisation, Resyllabification, /k/

---

[3] A topic which has been largely ignored in the literature.

[4] The notion 'rule' is a useful notion in phonological analyses which do not purport to characterise on-line processing or production. The words 'rule' & 'generalisation' can be used interchangeably, since we take rules, like constraints, to be a species of generalisation.

[5] This limitation is similar to the subjacency constraint on cyclic rule application in earlier models of transformational syntax.

Glottalling, Compensatory Lengthening and /a/ Reduction) are, we claim, non-edge-based. These generalisations, we suggest, operate after the right-edge and left-edge processes[6]. By contrast Glottal Stop Insertion (GSI) is unlike the lexical generalisations mentioned above. We claim that GSI in SM is a postlexical and across-the-board process which applies to any sequence of two filled nuclei. The application of this demands that the second of the nuclei in question must be stressed: GSI may readjust word stress postlexically; in particular, the application of GSI may create stress contours which violate the lexical constraint Clash Avoidance.

There is a problem for our claim that GSI is a typical across-the-board generalisation: it is not attested root-internally, as it should be if it is, as we suggest, such a generalisation. In this connection, we examine a set of roots in SM whose syllabic status has been disputed in the literature; it seems that GSI is blocked only within roots. Evidence for this is elicited from native speakers of SM taking part in a SM onset-reversal language game. This evidence shows that such roots are underlyingly bisyllabic. They thus constitute serious counter-evidence to our view of the status of GSI.

However, given that all such cases contain a sequence of a stressed low vowel followed by an unstressed mid vowel (or lax high vowel), they are perceptually difficult to distinguish from monosyllabic roots containing either the

[6] Our analysis thus appeals to a kind of cyclicity, but not cyclicity as classically conceived, since we are not claiming that all three of the relevant rules operate on an edge-based cycle.

/ai/ or the /au/ diphthong, and thus do not perceptually resemble hiatus sequences[7]. This may explain the puzzling non-application of GSI in certain cases.

Our work differs from previous research on SM morphophonology in two main respects. Firstly, most of our data are empirically reliable since they come from tape recordings, made (in 1998) in Malaysia, of native speakers of SM. As a consequence, our findings differ from the previous literature in that they reveal inter- and intra-speaker variation. While we do not examine its possible sociolinguistic status, certain aspects of this variation back up our claim regarding GSI as a postlexical process which affects word stress patterns. Secondly, our analysis is, to the best of our knowledge, the first in the literature which offers a description of the word stress assignment in SM, and which integrates the stress assignment algorithm with SM morphophonology.

1.3 Data Sources

Our sources are:

a. Observations of casual conversation by native speakers of Johore Malay.

[7] Our definition of hiatus here is: a sequence of two filled nuclei, not separated by a filled onset. We appreciate that, for some, a sequence of two filled nuclei separated by a glottal stop constitutes a hiatus.

b. Observations of conversations including the Malay language reversal game by native speakers of Johore Malay (SM), Pahang Malay and Perak Malay.

c. Tape recordings of the pronunciation of a word-list given to Johore Malay native speaker informants.

d. Research into Johore Malay (SM) used in previous research, particularly Farid (1980), Teoh (1994) and Zaharani (1998).

e. The author's own observations and intuitions as a speaker of SM.

CHAPTER 2

SYLLABLE STRUCTURE IN STANDARD MALAY

2.0 Introduction

The analysis of the generalisations which we will be considering in this thesis often involve appeal to aspects of syllable structure. It is therefore important, before proceeding to the main topic of the thesis (the definition of PW in SM and related issues), that we set out our view of the nature of syllabification in SM. The aim of this chapter is to justify the idea of underlying syllabification and to provide an overview of underlying syllabification in SM. The analyses we provide in later chapters will presuppose the validity of what we say here. In what follows, we will use the terms 'UR', 'lexical entry' and 'lexical representation' synonymously, and use these terms to refer to real representations stored in the minds of real speakers, and accessed by them during acts of lexical retrieval.

Adopting a traditional derivational approach which allows for Underlying Representations (URs) and surface forms derived from them, we claim that syllable structure is present in URs universally. The generative phonology literature is divided on this issue; some, such as Kaye, Lowenstamm and Vergnaud (1985) allow for universal underlying syllabification, while others, such as Zaharani (1998: 22), claim that 'syllable structures are not present in the lexical representation, and are derived in the course of phonological derivation'. We therefore present evidence and argumentation in favour of underlying syllabification before proceeding to discuss

the overall underlying syllable shapes of morphemes in SM and the status (underlying or derived) of glides and glottal stops in SM.

This chapter is structured as follows. In section 2.1, we discuss the notion 'Underlying Representation' (UR) as appealed to in generative phonology. We argue there that work on child acquisition of phonology shows that mental representations of syllables precede representations of segments in the course of development: mentally stored phonological representations contain syllable structure from an early stage in development. We also argue that psycholinguistic work on misremembered words shows that words are stored with their syllabification. We present further arguments in favour of underlying syllabification by suggesting that there is inconsistency in the generative literature which postulates URs stripped of syllabification; we suggest that this inconsistency is obviated under our approach. Section 2.2 provides an overview of SM underlying syllable structure and the sorts of vowel and consonant sequence found in SM underlying representations; 2.3 outlines the status of glides in SM, claiming that SM has both underlying and derived glides; and section 2.4 provides a summary of our claims.

2.1 On Underlying Representations, Syllables and Storage in the Mental Lexicon

Evidence from Phonological Acquisition and Misremembered Words

There are many case studies within the child phonology literature which point to the syllable as a unit which is present in lexical entries, by which we mean, as noted above, the psychologically real representations stored by real speakers. One example, chosen from a large number of such cases, is an investigation by Vihman, Velleman and McCune (1994), who analysed the phonological development of two English-acquiring children in fine detail. They note that, as is known, in production, the child begins by uttering CV syllables in the canonical babbling stage. By around 10 months, individual differences in production emerge as the child develops vocal motor schemes which, crucially, reflect both the child's own pattern of vocal control (production) and phonetic patterns in the ambient language (gained via perception). We stress here that these patterns involve syllable shapes as well as specific feature configurations within those shapes. They postulate that, 'once some vocal motor schemes have developed, these patterns add to the salience of certain adult words that are, besides, prosodically highlighted, frequent, and inherently interesting to the child.' (11). They argue that storage need not be postulated at this developmental stage, but that, once the child's vocal forms are no longer embedded in a particular situation of use, they can be superimposed on the child's productions, such that the child uses them to form generalizations. It is at this stage, they argue, that the

beginnings of a phonological production system emerge: mentally stored phonological mental representations emerge at this point.

An example of this pattern of development comes from the child Timmy reported in Vihman *et al*'s paper. At 9 months, Timmy uttered [ba] in response to adult utterings of the words *ball* and *block*. By 10 months, he produced [ba] spontaneously in appropriate contexts for the uttering of those words. He also produced [ba] in response to a wider range of adult word utterances (*basket, bell, boat, book, button*). By 15 months, he uttered [ba] for *bird, brush* and *bunny*. From 11 months, Timmy uttered [ka] for *kitty, quack-quack, ca,* and *key*. These are examples of vocal motor schemes. They are syllabic in nature, and they form the basis on which the child will build a phonological system.

At 14 months, Timmy begins to construct a system, uttering [ja] for *eye*, and then extending this to other words containing palatality (*light, ear, hair*). He also utters [βa] for the word *Ruth*, and then extends this to other words containing labiality and/or continuant friction (*fire, flies, flowers, plum*). That is, the initial vocal motor schemes [ba] and [ka] are extended to [ja] and [βa]; it is this extension that constitutes the emergence of a system. For Timmy at this stage, word and syllable are not distinct units in his system. Rather, his system is based on a [Ca] syllable template; the four different templates are differentiated in terms of the different autosegmental features they contain.

At 15 months, [na] enters the system, and at 16 months, [ga]; at this stage, Timmy begins to iterate syllables, so that [ba] for *block, peg, boat* contrasts with [baba] for *baby and bracelet.* At this point, it becomes necessary to postulate the word and the syllable as distinct units in Timmy's phonological system. Paradigmatic contrasts also emerge at this stage, with, for example, [nama] (*Simon*) contrasting with [gaba] (*goodbye*). At the beginning of the 16 month stage, [ta] and [ti] emerge, and later in that period, the system begins to expand, with [i] occurring after consonants other than [t]. The point we wish to stress is that Timmy's phonological system is built upon mentally represented syllable structures.

Vihman *et al* also report on another child, Alice, whose route into a phonological system is quite distinct in many respects from that of Timmy, and involves templates of the sort <CVCi>, <Vɪ> and <jV> imposed on the child's productions of adult target words. We do not report the full details of Alice's development here; our main point is that syllable and word shapes are central to the emergence of the child's production phonology: syllable structures are mentally represented from an early stage. It is hard to imagine how they could then come to cease to be represented in the adult, given their centrality to the child's mentally represented phonology. Additionally, as Vihman (1996) points out, the syllable is the child's path towards the segment as a unit in its phonological representations.

We also note that, if syllables are not present in adult representations of words, it is hard to explain psycholinguistic results such as those of Aitchison & Straf (1981), who show that both adults and children preserve syllable count and initial

consonant in misremembered words. Vihman (1996: 174) notes, interestingly, that, when this pattern breaks down, adults are more likely to retain the consonants while children are more likely to retain the syllable count. These results show that words are stored with their syllable structure by both adults and children.

It might be argued that the child data show that the syllable is present in lexical entries only in the child's *production* lexicon, but not in the child's *receptive* lexicon. That position is hard to sustain: as Vihman *et al* point out, and as we have seen above, there is an intimate relation between production and perception, with the child's production capacities directly influencing its speech perception capacities. If the syllable plays a role in production, it is also playing a role in perception. As Vihman (2001) has suggested, recent work on mirror neurons suggests a neural basis on which this intimate connection is based: mirror neurons are activated when the child hears (and sees, if sighted) someone else engaging in a vocal motor scheme which the child has established. As Vihman (1996: 227) puts it, 'familiarity with the articulatory pattern is what makes an auditory pattern memorable, not only for 1-year olds...and 2-year olds, but also for the older children of Aitchison and Chiat's study'.

As Lléo (1990: 275f), reported in Vihman (1996: 227), puts it:

> There is a certain reluctance to attribute a crucial role to the lexical item in phonological acquisition... based on the assumption that the phoneme and its oppositions play an exclusive role. But child phonology is committed to both, to oppositions and to patterns, that is, to segments, but to syllables and lexical items too. Adult phonology is also committed to both, although the segment plays a more important role than in child phonology. Within this framework, the transition from

child phonology into adult phonology...involves a quantitative rather than a qualitative step.

Given the external evidence, of which we have presented only a tiny proportion, it seems hard to deny that human beings store words with a specification of their syllable structure. To the extent that generative analyses omit such structure, they do not correspond to real mental representations of words; at best, they are indirect ways of modeling inductive generalizations over mental representations of words which contain syllable structure.

One might argue that, in a derivational model of phonology which postulates two levels of representation (URs and surface representations) speakers are accessing syllable count from their *surface* representations, not their URs (we remind the reader that the derived surface representations, known as systematic phonetic representations, are a species of mental representation under classical generative assumptions: see Bromberger and Halle (2000) for a restatement of this view). That would be to allow that surface representations, as well as URs, are stored. If that were the case, then one needs to ask what the rationale of a derivation might be: if the speaker stores surface representations, rather than creating them on-line, in the way that Bromberger & Halle (2000) envisage, then one needs to ask what cognitive work a derivation is doing. There seems to us to be no obvious answer to this question.

We suggest that the way out of these difficulties for a derivational phonology is to concede that URs contain syllable structure. We also note that, if the

idea of stripping out all redundant information from URs is taken to its logical conclusion, then it should lead to a view in which much of the linear sequencing of segments should also be stripped out of URs. Both Anderson (to appear) and Sauzet (1996) have followed the logic of that argument and now propose lexical entries in which there is little or no sequencing of segments at all. This approach has the merit of being consistent, whereas traditional derivational phonology is inconsistent in insisting on removal of syllable structure on the grounds that it is predictable, while not insisting on removal of linear sequencing of segments, despite its also being largely predictable in many cases (on the basis of sonority sequencing principles, for instance). We object to both approaches (consistently, we believe): morphemes are stored both with syllable structure and with linear sequencing of segments; this would have to be the case, otherwise one could not explain the role played by the initial segment in the misremembered words research cited above: in a model of the Anderson or Sauzet sort, there are no initial segments for the vast majority of morphemes, since there is no available notion of 'initial' in an unsequenced set of segments.

For the reasons given above, we follow Kaye and Lowenstamm (1984) in allowing for underlying syllabification but with the possibility of resyllabification during the course of phonological derivation. We have claimed that speakers store words with their underlying syllabification, and inductively arrive at generalisations concerning those stored representations. Thus, we do not deny that the speaker has access to generalisations concerning syllable structure; what we do deny is that these result in URs being stripped of syllable structure. Having presented external

evidence and argumentation (from areas outside of generative phonology) in favour of underlying syllabification, we now present internal arguments.

Further evidence and argumentation

Firstly, in certain languages, there is no alternative but to allow for underlying syllabification. For example, word stress in Modern Greek is arbitrary; a child acquiring Modern Greek simply must store the phonological form of a given word with its associated word stress. This is significant in two senses: firstly it shows that human beings are capable of storing the word stress patterns of their own language; additionally, since stress is a feature related to syllables (or sub-parts of syllables, i.e. rhymes or moras) languages such as Modern Greek must be said to have underlying syllabification: if the stress pattern is stored, so too is the syllable structure (we know of no word stress assignment algorithm that does not make reference to syllables or syllabic constituents).

This evidence does not show that stress or syllabification is underlying in all languages (as postulated by Burzio 1996), but it opens up the possibility. It is widely believed that word stress in English is not entirely arbitrary, but is stored underlyingly, and subject to word stress assignment rules operating on underlying representations to yield derived stress patterns (a tradition going back to SPE). However, following Burzio's claim that English words are stored with their associated stress patterns (and thus with their syllabification), it can be argued that

speakers inductively generalise over stored forms in the manner envisaged by Hooper (1972) and Hayes (1995). This argument allows stress generalisations to be interpreted as static inductive generalisations over stored underlying forms. We claim with respect to SM, that once underlying syllabification is allowed for, many of the *ad hoc* constraints postulated by Zaharani can be obviated. It might be argued that our argument (that stress assignment requires syllabification) does not go through unless stress is related to syllable structure. In response to this, and as noted above, we know of no language in which stress is not related to syllable structure one way or another, via syllable position and/or syllable weight.

A further argument for underlying syllabification concerns the characterisation of underlying and derived glides. It seems appropriate to allow in principle for both underlying and surface glides. Hannahs (1995a, b, ms) argues that Standard French has both, and we argue that this is the case for SM too. However, as Hannahs rejects underlying syllabification his approach will not allow for defining underlying glides in terms of syllable structure. It seems clear that the term 'glide' must remain constant in meaning in the phrases 'underlying glide' and 'derived glide', otherwise the conceptual distinction between underlying and derived glides cannot be formulated. Since derived glides are defined in terms of syllable structure, allowing for underlying syllabification to define underlying glides is essential, otherwise no consistent definition of 'glide' is available.

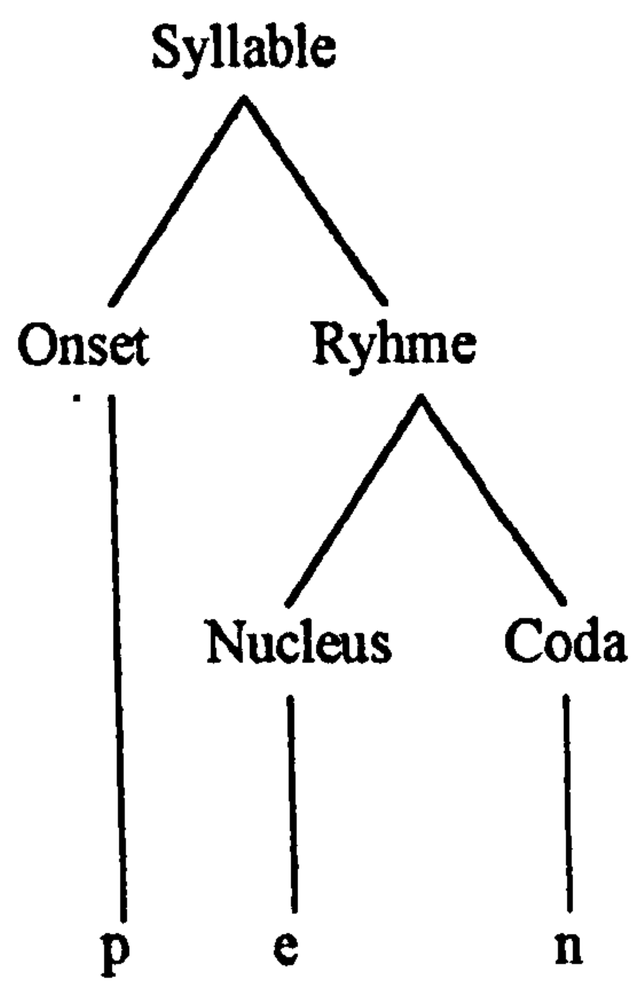
We claim that the same kind of inconsistency is evident in Spencer's (1996) discussion of underlying syllabification. Spencer claims that syllabification is

derived by algorithm, rather than being present underlyingly. But he considers pairs such as *aeon* and *yon* in English. Spencer (1996: 96-97), recognises that pairs such as these constitute a dilemma for the 'no underlying syllabification' approach, since, as he astutely remarks, 'we won't know whether the melody is to be interpreted as a vowel or a glide until we know whereabouts it appears in the syllable. But we cannot determine that until we know whether it's a glide or a vowel'. Hannahs' solution to this general problem is to provide a distinction at the underlying level, in the feature specification of glides and vowels. We argue, against this, that the greater degree of constriction in glides (as opposed to high vowels occupying a nucleus) *derives from* their place in syllable structure. Spencer suggests 'prespecification', such that the first vowel in *aeon*, but not in *yon*, is underlyingly specified as occupying a nucleus position. This, however, like Scullen's (1987) analysis of glides in French, undermines the 'no underlying syllabification' position, and results in the inconsistent claim that speakers store some words with underlying syllabification, but not others. Our position is, we claim, more consistent, while also allowing that languages do have syllabification generalisations.

2.2 Standard Malay syllable structure

We will assume a (widely, but not universally, adopted) conception of syllable structure in which a syllable contains an obligatory onset (which may be empty) and an obligatory *rhyme* which contains an obligatory *nucleus* followed by one or more optional *codas*. This kind of structure is shown in diagram (1).

(1).

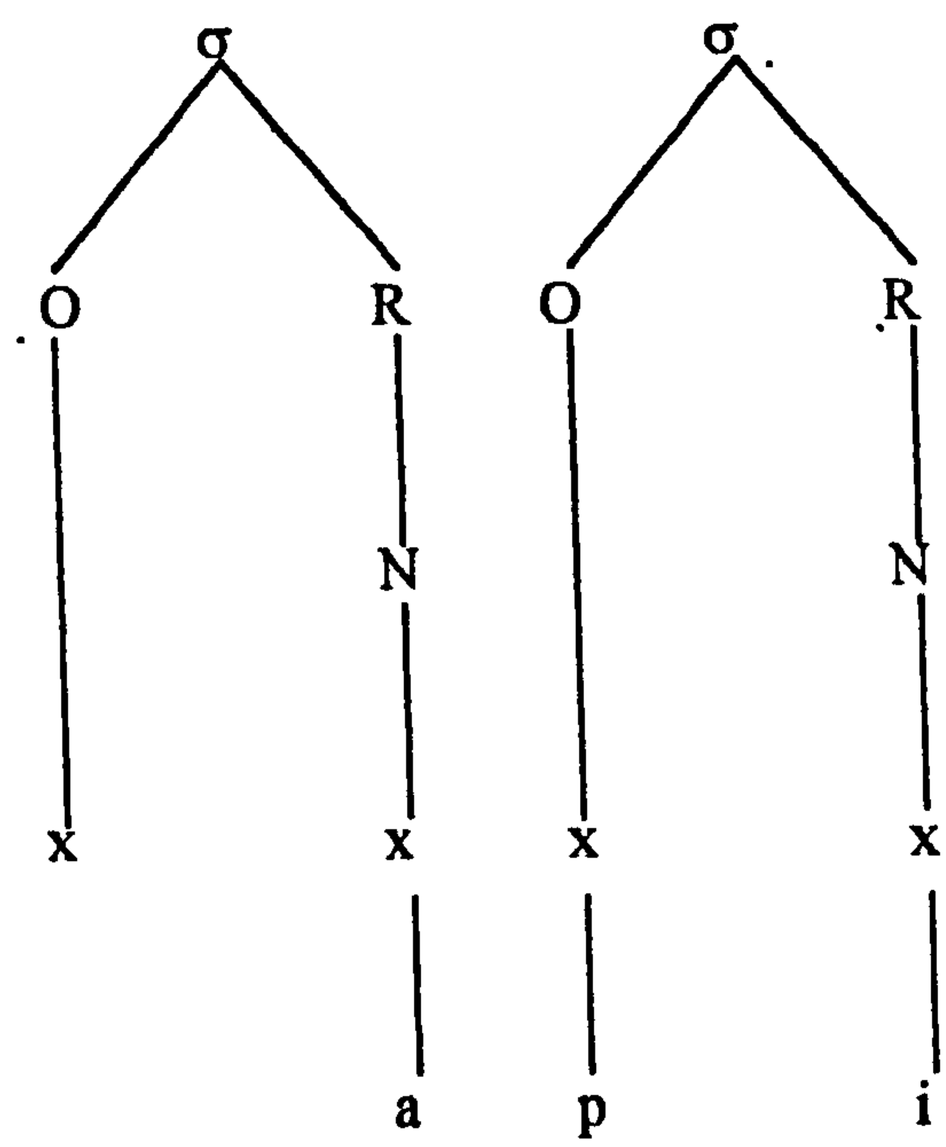


Assuming underlying syllabification, what generalisations can be made as to the overall underlying syllabic shape of morphemes in SM? The literature concerning this concentrates, as we have seen, on whether SM is basically (C)V(C) or CV(C). For example, Teoh (1994: 26) has argued that SM syllable structure is CV(C), but adds that 'Malay has empty X-slots in the vowel-initial suffixes' (p. 87). The claim that SM has empty onsets could be conveyed either by the notation 'CV(C)' or by the notation (C)V(C). We arbitrarily adopt (C)V(C). The notational issue is trivial; our claim is that all 'vowel-initial' morphemes in SM contain an empty onset.

On the related matter of whether SM has underlying glottal stops, we note that Teoh is inconsistent in his treatment of 'vowel-initial roots', representing them either with or without an underlying root-initial glottal stop, depending on whether they co-occur with a prefix, as in /tər + ʔambil/ ('to take unintentionally': p. 43) or not, as in /ikat/ ('to tie: p. 64). Teoh takes the underlying consonant system of SM

to contain a /ʔ/ phoneme (p. 8), and often represents it as occurring root-initially in underlying representations, while simultaneously asserting that 'the absence of glottal stop in word-initial position in Malay makes no semantic difference to the word' (p.59). It is not clear from Teoh's discussion whether his use of '/ʔ/', is intended as an informal means of representing an empty skeletal slot: his use of [ʔ] in vowel-initial roots suggests absence of root-initial underlying /ʔ/. (For example he uses [ʔ] rather than /ʔ/ in the compound word 'train', which he represents as /kəreta + ʔapi/). We have obviated this inconsistency by suggesting that the initial '(C)' in (C)V(C) be said to convey the idea that syllables in SM contain an obligatory skeletal slot onset position underlyingly, so that the syllable-initial 'C' position may be empty. This skeletal slot is the position that GF spreads into, where glottal stops are inserted, where floating /r/ is syllabified, and into which coda consonants may be resyllabified. We adopt the following sorts of representation for morphemes containing empty onsets, illustrated here by the word /api/:

(2)



We agree with Zaharani that SM does not have a glottal stop phoneme: all glottal stops in the native vocabulary are predictable, and thus the notation '/ʔ/' is inapplicable in the case of SM. SM does, however, exhibit phonetic glottal stops ([ʔ]), almost all of which result from the application of GSI (the remainder are substitutions for [ʕ] in Arabic loanwords).

In SM, some vowels occur in all positions in a root whereas others depend on whether the syllable in question is open or closed. We agree with Farid (1980) and Teoh (1994) that there are 6 underlying vowels in Standard Malay, viz., /i, e, ə, a, u, o/. We may depict the underlying vowel inventory of the language in the following chart:

(3) Vowel Inventory

High	i		u
Mid	e	ə	o
Low		a	

In open syllables of bisyllabic roots, the first syllable may contain any vowel, while the vowel in the final syllable is restricted to phonetic [i, ə, u], as in the words /ini/ 'this' [i.ni], /tiga/ 'three' [ti.gə] and /itu/ 'that' [i.tu]. Underlyingly, the final vowel in a bisyllabic root must be /i, u or a/. That is, only simplex vowels, in the sense of Harris (1994), may occur in that position: vowels containing a single element. An analysis of this phenomenon falls out with the aims of this dissertation, but we note that this kind of pattern is common in the world's languages, and may well be related to SM's trochaic stress patterns in bisyllabic words, which make up much of the native vocabulary.

In this connection we note that the /a/ of /tiga/ ('three') laxes to [ə] in word-final open syllables, while there is no productive synchronic process of high vowel laxing in root-final closed syllables; rather, such syllables contain underlying mid vowels, /e/ and /o/ as in /pileh/ 'choose', phonetically [pi.leh] and /batok/ 'cough', phonetically [ba.to?].

Another question that arises with respect to the range of SM underlying sequences is whether SM has root-internal V + V sequences, i.e. hiatus sequences of

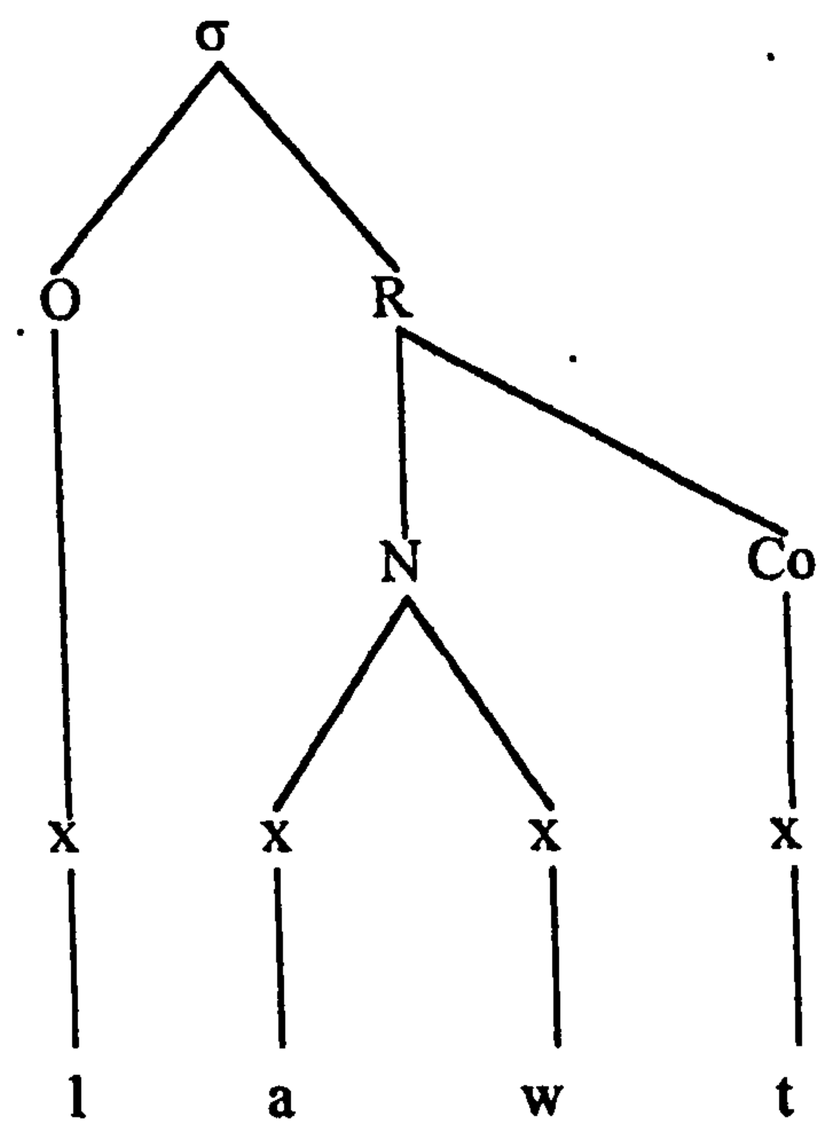
vocalic nuclei. Zaharani claims that it does, in roots such as [pisaw] ('knife'), postulated to be underlyingly /pi.sa.u/ by Zaharani, and [pakaj] taken to be /pa.ka.i/ underlyingly. These postulated underlying representations are driven by Zaharani's view that glides do not exist. We discuss this below.

In contrast to Zaharani, Teoh (1994) allows for root-internal diphthongs in cases such as [pisaw] ('knife') and [pakaj] ('to wear'), and suggests that other cases such as [lawt]/[laot] ('sea') and [kajn]/[kaen] ('cloth') actually occur with an intervocalic glottal stop, as [laʔot] and [kaʔen]. Zaharani denies this, and our phonetic observations support Zaharani's observation that there is no discernable glottal stop in such cases.

The question then is whether the latter cases are (a) phonetically or (b) underlyingly bisyllabic in nature. If they are considered to be underlyingly bisyllabic with a hiatus sequence that is repaired by a hiatus-avoidance strategy of diphthongisation, in which case a loss of a syllable from the underlying representation will occur (in these cases, it is the first, rather than the second, vocalic segment which is most prominent). From observations of tape recorded material we suggest that these cases are phonetically monosyllabic[8]. If this is the case, we may represent them as follows:

[8] Although it can be difficult to distinguish, whether the respective case is a diphthong whose head is the first element, or a bisyllabic sequence in which the first syllable is stressed. Where stress is made on a given sequence of two vocalic segments in the first segment.

(4)



In order to decide whether such words are phonetically monosyllabic, we need to consider whether they are underlyingly monosyllabic or bisyllabic. To do this, Zaharani's objections to Durand's (1987) analysis[9] need to be considered.

Zaharani's cites two principal objections to Durand's analysis. Firstly, he notes that it 'contradicts the general view that the basic structure of the Malay syllable is simplex' (p. 67). This claim is also supported by the work of Yunus (1980) and Teoh (1994)[10]. Secondly, Zaharani states that such sequences are parsed heterosyllabically, rather than tautosyllabically. Zaharani provides evidence from a Malay syllable reversal language game, showing that it applies to forms such as [nãẽʔ], yielding [ẽʔnã]. However, this may only show that the *phonetic*

[9] Durand claims they are underlying monosyllabic forms that contain diphthongs.

[10] This view may be merely an assertion of differing views among four analysts, rather than a claim based on evidence and argument.

representations of such forms are bisyllabic; it does not *show* that they are *underlyingly* bisyllabic[11]. These considerations would be puzzling if SM did indeed have a diphthongisation strategy (which resolved putative underlying root-internal hiatus sequences) since the language does not otherwise permit reduction in the number of syllables as a hiatus-resolving strategy.

An analysis which takes these forms to be underlyingly bisyllabic need not suggest they are phonetically monosyllabic (and thus that a process of diphthogisation, resulting in a phonetic monosyllabic, operates on them). For example, Zaharani (p. 68) allows for tautosyllabification of underlying sequences of two vowels as a hiatus avoidance strategy, and argues that such cases are both phonetically and underlyingly bisyllabic. He sees this as the only case in SM where an underlying hiatus remains unresolved by either GSI or the putative tautosyllabification process. For further discussion on this issue, see Chapter 7.

Another aspect of underlying vowel and consonant distribution in SM concerns floating /ɾ/. We assume that postulating an underlying floating segment is a plausible way of characterising the sorts of consonant/zero alternations found in phenomena such as French Liaison (Encrevé 1988 and Charette 1991 both postulate floating consonants for French latent consonants, for instance) and in the patterns of non-rhoticity found in SM, where floating coda /ɾ/ surfaces only if there is an empty onset position for it to anchor into, as in /tampaɾ + an/ ('smack') → [tam.pa.ɾan] *vs*

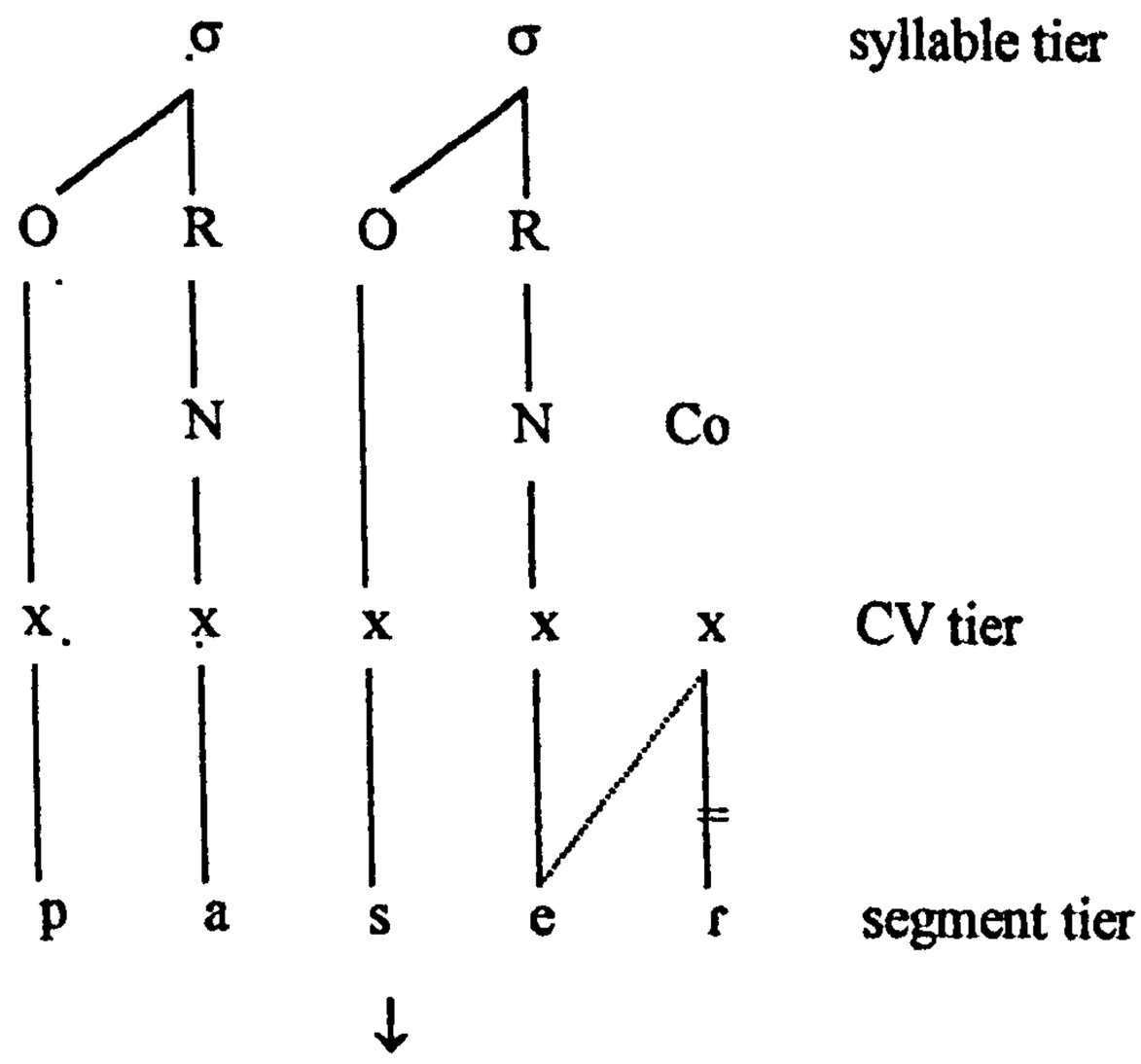
[11] It is of importance to note that the language game applies equally well to monosyllabic forms as to bisyllabic forms.

/tampar + kan/ → [tam.pa.kan], and /tər/ + /bakar/ → [tə:baka:] ('burnt down'), where both a prefix-final and a root-final /r/ fail to be realised since they are not followed by an empty onset. The idea of floating consonants has been applied to non-rhotic accents of English by Harris (1994), but we note that such accents are unlike SM in exhibiting 'intrusive r', which is unattested in SM, and which perhaps suggests that, for those accents, an insertion rule if preferable to an analysis appealing to floating consonants.

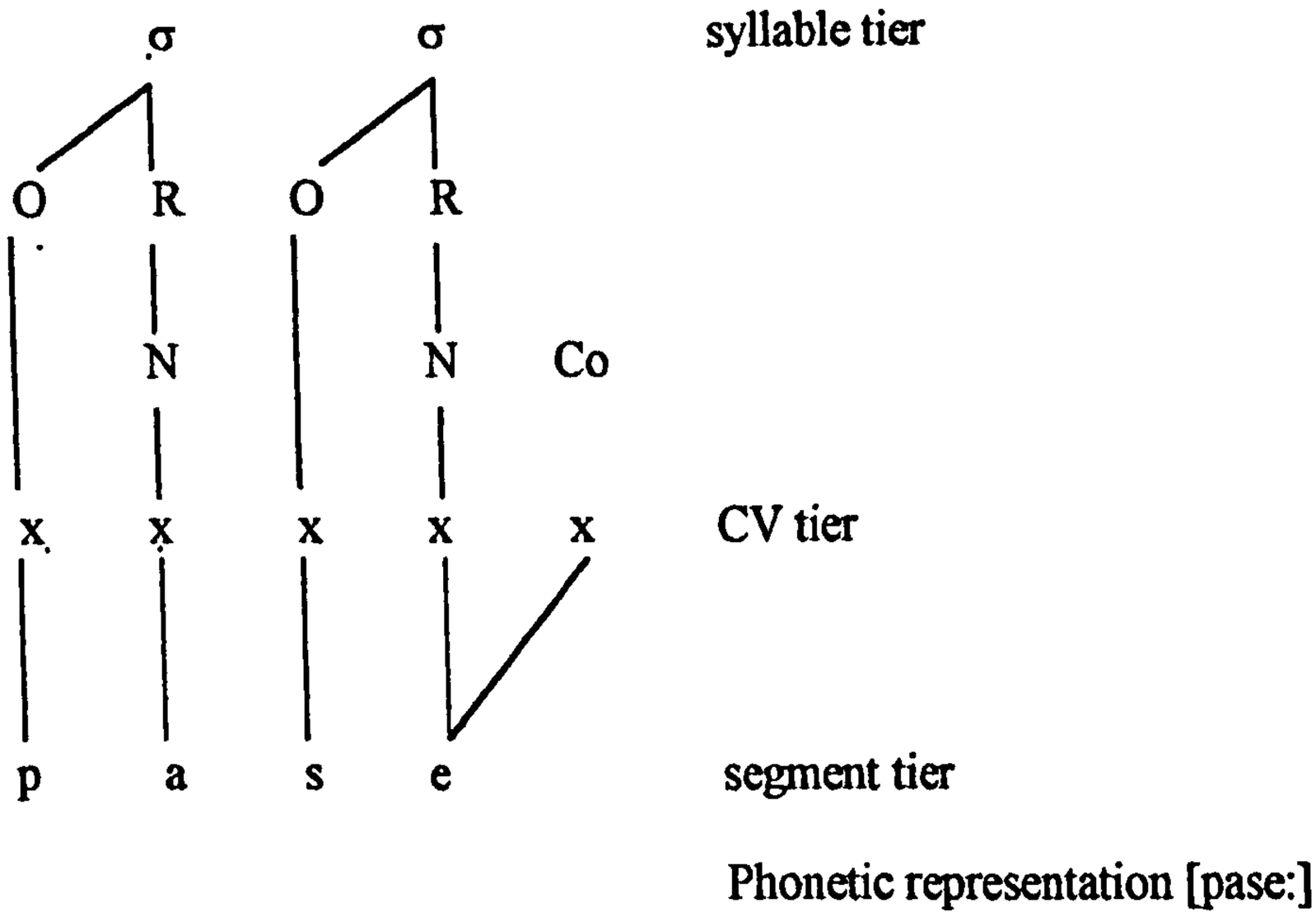
Floating /r/ is delinked from its skeletal point when it has no empty onset to anchor into, with subsequent lengthening of preceding /ə/ and /a/ in [tə:baka:]. Other examples are the words /paser/ ('sand') → [pase:] and /tabor/ ('to scatter') → [tabo:]. We represent the delinking process in (5):

(5) Vowel lengthening as a result of delinking of floating /r/

a.



b.



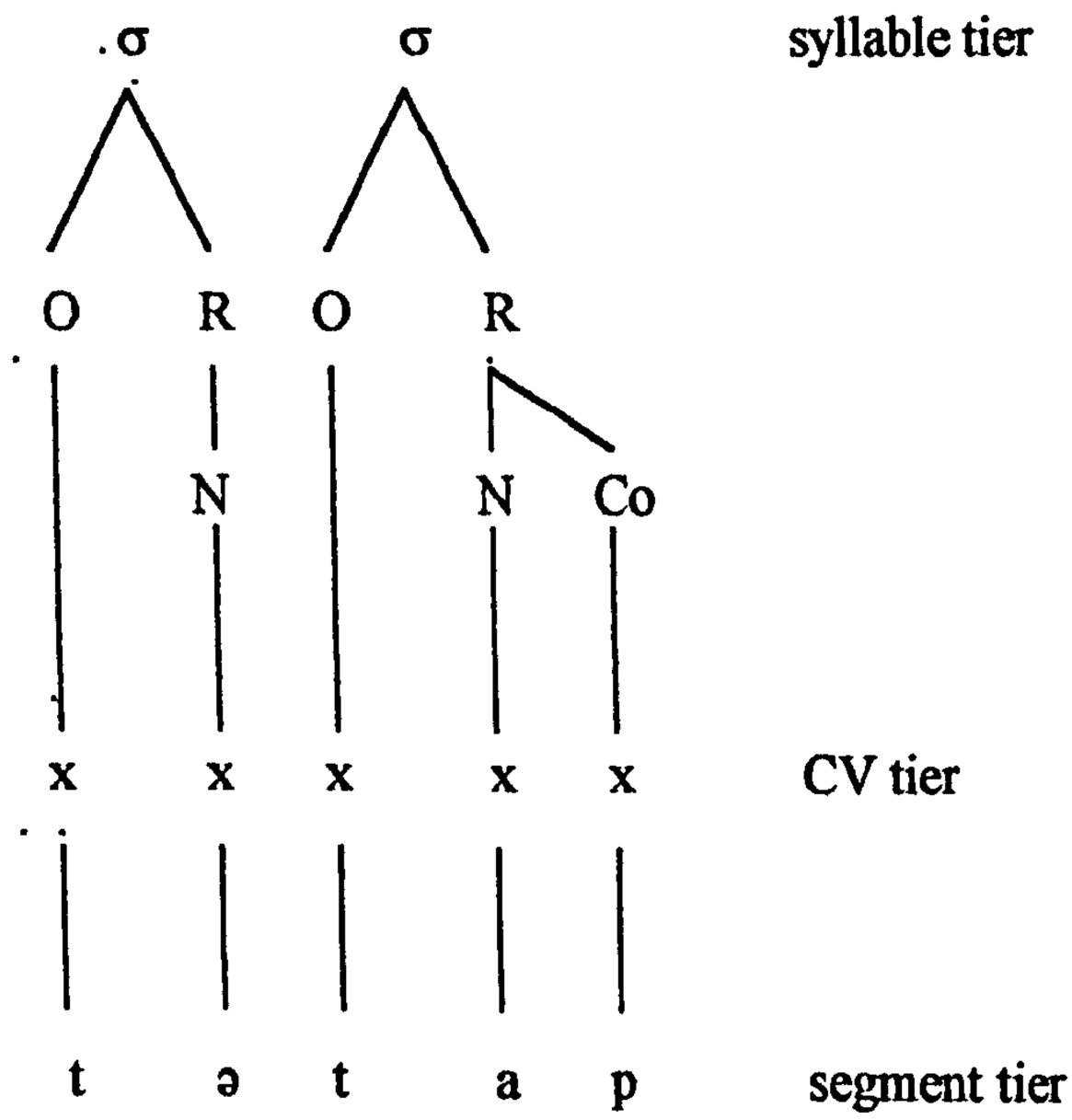
The vowel /a/ in a root-final open syllable does not undergo laxing when it is attached to a suffix: it is realised as [a] in surface representation, as in the words /suka+kan/ 'like' and /kə+suka+an/ 'hobby' (becoming [sukakan] and [kəsukaʔan][12] respectively).

In SM, the voiceless stops /p, t, k/ are unreleased when word-final or followed by another consonant. Examples of the latter are /tətap/ 'permanent' [tətapˀ], /kətat/ 'tight' [kətatˀ], and /kakak/ 'sister' [kakaʔ][13]. The syllabic context for unreleased realisations of /p, t, k/ in word-final position can be represented as follows:

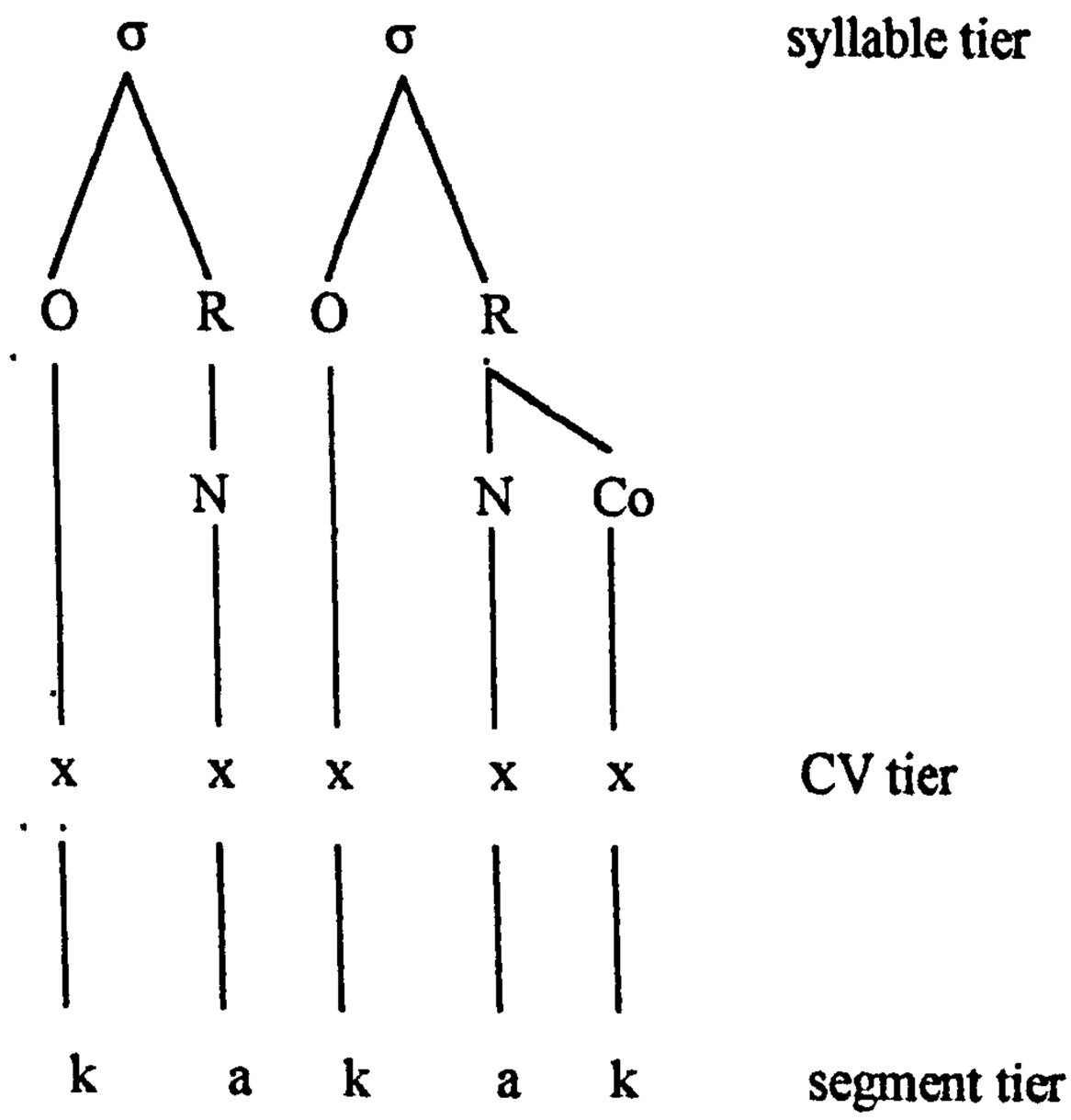
[12] As we have noted Glottal Stop Insertion (GSI) in vowel sequences is one of the phonological processes that resolves hiatus strategies. This will be explored in more detail in Chapter 7.

[13] The voiceless velar stop in SM is realised as a glottal stop [ʔ] when it is in syllable final position.

(6) a. /p/ in the coda position as in the word /tətap/ 'permanent'



b. /k/ in the coda position as in the word /kakak/ 'sister'

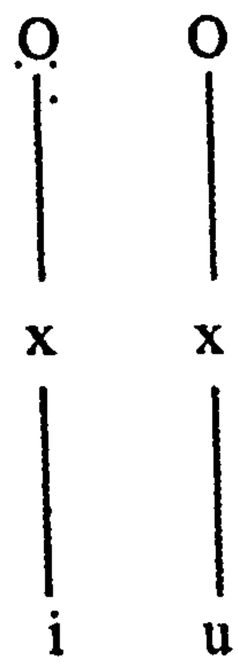


2.3 Underlying and surface glides in SM

Having adopted underlying syllabification, we claim that SM has both underlying and derived glides, the latter resulting from the operation of GF. We note, in passing, that GF in SM may not (as it does in other languages) result in reduction in the number of syllables. For instance, /tari + an/ always results in a trisyllabic [tarijan] form, and never in a bisyllabic [tarjan] form. Additionally, SM never exhibits the widely attested hiatus-resolving strategy of vowel deletion, which results in a decrease in the number of syllables in a derived word. We also claim that the underlying representation for the prefix /məŋ/ is specified both for its syllabic structure and for the place of articulation of the final nasal (i.e. it is velar, and in coda position). In cases where the /məŋ/ prefix co-occurs with a root beginning underlyingly with a voiceless bilabial stop, such as [məmakai] (from /məŋ/ plus /pakai/), we allow for both resyllabification and nasal assimilation.

In SPE, it was claimed that {j, w, ʔ , h} form a class of glides; this was disputed early on by Lass (1976); we agree that the glottal sounds are not readily conceived of as glides, but we focus here on the relationship between high vowels and glides. Hannahs (ms: 1) notes that languages containing both high vowels and glides have several relationships between them: they may be unrelated, such as when /i/ and /j/ are separate phonemes, surfacing as [i] and [j]; they may be allophonically related such as in instances of [j] deriving from /i/. We claim that SM has both underlying and surface glides; we represent onset glides as follows:

(7)



The difference between phonetic [i] and [j], [u] and [w] is their position in the syllable: the high vowels [i] and [u] occupy a nucleus peak position, whereas [j] and [w] do not. A high vowel in nucleus peak position may spread into an adjacent empty onset. So too may glides which are high vowels in nucleus offglide position – we refer to this spreading process as Glide Formation. Having allowed for underlying glides in SM as high vowels in non-nucleus peak position, we show the underlying consonant inventory of the language in (8):

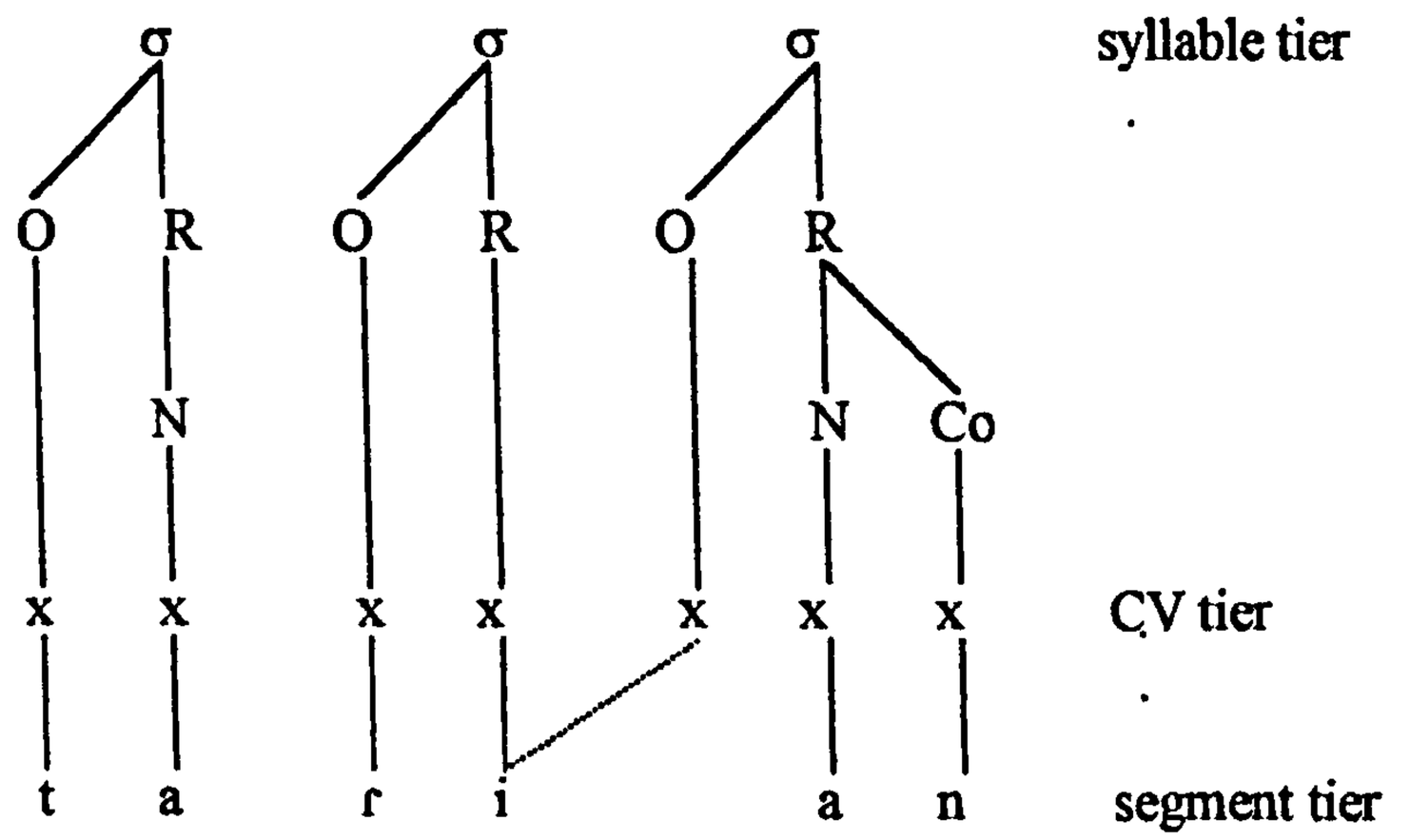
(8) Consonant Inventory

stop	p b	t d		k g	
affricate			tʃ dʒ		
fricative		s			h
liquid		l r			
nasal	m	n	ɲ	ŋ	

In traditional SPE phonology, the formulation of GF generalisation in SM is problematic as this generalisation concerns the *spreading* of features rather than the

changing or inserting of features. Non-linear phonology shows that GF can easily be depicted as the spreading of high vowels from a nucleus to an onset position. This is shown in (9):

(9) A glide as high vowel in non-nucleus peak position as in the word /tari+an/.



Phonetic representation = [ta.ri.jan]

2.4 Summary

We have claimed here that morphemes are stored with underlying syllabification. This is not a novel claim, but since it is not universally accepted, we have presented evidence and argumentation in its favour. We have set out the range of underlying sequences which occur in SM and the processes that accompany them. We have also claimed that SM has both high vowels and glides, and that they are often phonologically related via GF: the only major difference between them is the position in the syllable; differences in degree of constriction follow from the

syllabification. Thus the vowels [i] and [u] occupy a nucleus peak position, while the glides [j] and [w] occupy an onset position, or the dependent position (offglide) in a nucleus.

Given underlying syllabification, it is seems clear that SM belongs to the class of languages which Clements and Keyser (1983) refer to as the type IV group: it has four syllable structures, V, VC, CV and CVC. This can be abbreviated to (C)V(C), where the initial '(C)' should be taken mean that SM morphemes contain either a filled or an empty onset position. Postulating such empty onset positions is central to the set of SM hiatus-avoidance strategies (Glide Formation, Glottal Stop Insertion and Floating /r/ Realisation), which we will be discussing in the following chapters.

CHAPTER 3

THE STANDARD VIEW: EVIDENCE IN FAVOUR

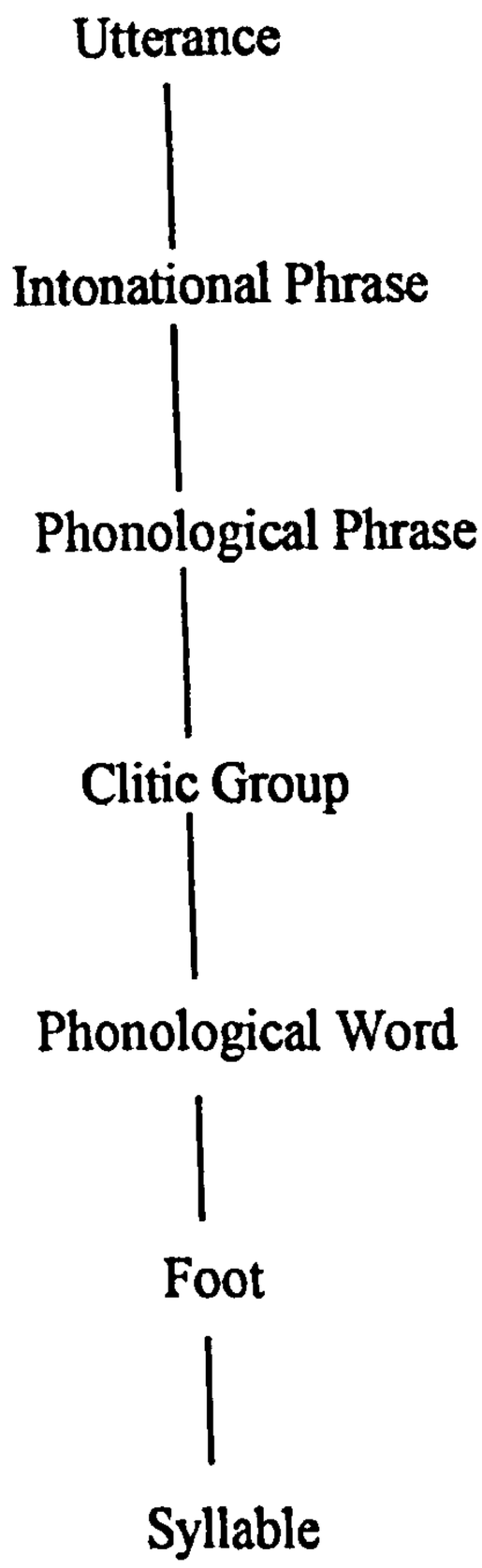
3.0 Introduction

This chapter presents the available evidence in favour of what we will call the standard view, a view which we question in the next chapter. This is the view adopted in previous analyses, namely those of Teoh (1994) and Zaharani (1998), who have sought to explain SM morphophonology by appealing to the PW defined, in SM, as a root plus any suffixes. For them, the phonological processes of GF and Gem in the language operate within the PW, thus defined. The Standard View is based on assumptions made within the theory of Prosodic Phonology, which we now give an outline of.

3.1 Prosodic phonology

The theory of Prosodic Phonology is a theory about the representation of phonological domains. This theory postulates a mapping procedure from the syntax to the prosodic phonology as proposed by Selkirk (1981), Nespor and Vogel (1986), and Hayes (1990). The theory states that phonological constituents are arranged in the form of a hierarchy. Each hierarchy set contains phonological domains that arrange themselves from the smallest to the largest as in (1) (McHugh 1990 and Inkelas 1996).

(1)



Hannahs (1995a: 126)

It remains controversial whether or not the units in the prosodic hierarchy arrange themselves according to the Strict Layer Hypothesis (Selkirk 1984). According to that hypothesis, within the prosodic hierarchy, each level must consist only of units from the next level down in the hierarchy. Another controversy concerns the existence of some of the postulated units, such as the clitic group. Others question the existence of specific units within specific languages. For instance, it is a moot point whether Standard French contains prosodic feet; Rose (2000), for instance, assumes that it does; Delais-Roussarie (1996) argues that it does not. These points of controversy need not affect our discussion of SM

morphophonology, however, since we are focussing on the PW. For our purposes, the adopting of this kind of model offers a possible way of describing the domains of the processes which characterise the lexical phonology of SM, to which turn shortly. Additionally, the PW has been widely accepted as a real unit in phonological organisation, although theories vary as to how exactly it is to be defined. It is the lowest constituent of the prosodic hierarchy that makes substantial use of morphophonological notions. According to Nespor and Vogel (1986:109),

> "The interaction between the phonological and morphological components of the grammar will in either case take the form of a definition of PW that maps morphological structure onto phonological structure. Within the PW domain we can then proceed either to readjust the syllables and feet when necessary or to construct them in accordance with universal principles and possible language-specific constraints."

We will allow that the definition of the PW in a particular language may differ from its definition in another language, and we now turn to our central question of how the PW is to be defined in SM.

3.2 The PW and Glide Formation

In the following, we illustrate the behaviour of GF in SM by referring to the prosodic constituent PW defined as under the standard view, where PW is taken to

be a root (plus suffix). GF applies within a root and between a root plus suffix. These environments are illustrated in (2).

(2) a). GF is often said to apply in the following cases:

i). High vowel becomes glide in word-initial onset position when followed by a vowel

/iə/root	[jə]	'yes'
/uaŋ/root	[waŋ]	'money'

ii). High vowel becomes glide in branching onset

/diə/root	[dijə]	'he/she'
/buah/root	[buwah]	'fruit'

(iii). Root-final High Vowel undergoes GF when followed by a vowel-initial suffix. (i.e. GF applies in the context [[...$V_{[+high]}$]root[V...]suffix])

[/aku/root/an/suffix]	[akuwan]	'confession'
[/bəli/root/an/suffix]	[bəlijan]	'buying'

We claim that, of these cases, only (iii) constitutes a genuine case of GF; the glides in the other cases are, we claim, underlying; we return to this point shortly.

b). GF is blocked:

When a prefix ending in a high vowel is followed by a vowel-initial root

(i.e. [[...$V_{[+high]}$] prefix [V...]root])

[/di/prefix/ambel/root]	[diʔambel]	'taken'
	*[djambel]	
	*[dijambel]	
[/dʒuru/prefix/acara/root]	[dʒuruʔacara]	'presenter'
	*[dʒurwacara]	
	*[dʒuruwacara]	

The blocking of GF also occurs at morphological word (MW) boundaries: between members of a compound, reduplicated forms and syntactic phrases. This process is also observed in French by Hannahs (ms:31) where GF is blocked when the edges of the PW fall between a prefix and root, between members of a compound, and between syntactic words.

3.2.1 *Root-internal glides*

Much discussion in the literature has focussed on the question of whether SM has derived or underlying glides in monosyllabic roots containing vowel sequences. As indicated in Chapter 2, the question that needs to be examined here is whether the phonetic glides [j] and [w] root-internally are derived glides or

underlying, or either, depending on the syllable structure of the word. We present relevant data in (3).

(3) (a). *Monosyllabic roots*

/iu/	[ju]	'shark'
/ia/	[jə]	'yes'
/iaŋ/	[jaŋ]	'which'
/uaŋ/	[waŋ]	'money'

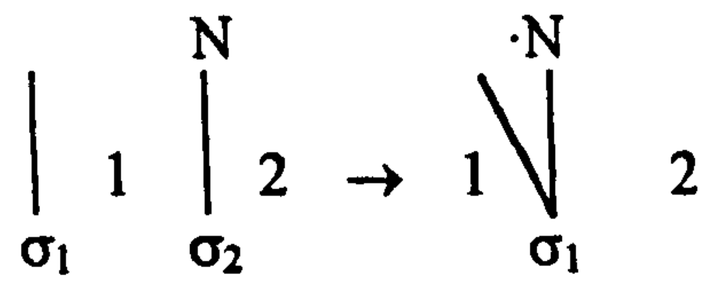
(b). *Bisyllabic roots*

/iakin/	[ja.kin]	'confident'
/iuran/	[ju.ran]	'fee'
/laiu/	[la.ju]	'to wither'
/saiaŋ/	[sa.jaŋ]	'love'
/siaŋ/	[si.jaŋ]	'day'
/uaiaŋ/	[wa.jaŋ]	'movie'
/uaŋ̃i/	[wa.ŋ̃i]	'fragrant'
/lauan/	[la.wan]	'compete'
/kahuen/	[kah.wen]	'marry'
/leuat/	[le.wat˺]	'late'

How is one to analyse the vocalic sequences in (3)? One analysis would be to take them to be diphthongs (namely /ia/, /iu/, /ua/ and /ui/), with the second vowel constituting the peak. That seems implausible, since there are no occurrences of words containing such putative diphthongs preceded by an onset consonant.

Another approach would be to define them as two sequences of vowels in separate nuclei, constituting a hiatus sequence, as in /i.u.ran/ 'fee'. This is the analysis adopted by Teoh (1994: 30), who suggests that the phonetic glides are the result of a process of Devocalisation (another name for Glide Formation). This putative process is formalised in (4).

(4). Teoh's (1994) Rule of Devocalisation



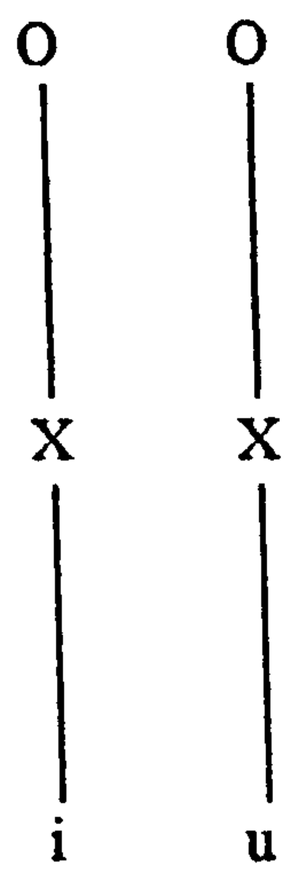
(Teoh, 1994:30)

The claim is that a high vowel in a nucleus position associates itself to an onset node and is realised as a glide (Teoh 1994). Teoh's analysis is inconsistent, we argue. On the one hand, he postulates underlying glides /j/ and /w/, for cases such as [dʒawap] 'to answer' and [wudʒot] 'to exist', said by Teoh (p. 53) to be /dʒawab/ and /wudʒud/ underlyingly (rather than /dʒauab/ and /uudʒud/). However, monosyllabic roots such as [ju] 'shark' and [jaŋ] 'which' (see p.30) are treated by Teoh as underlying /iu/ and /iaŋ/ respectively. The contradiction is that Teoh describes non-alternating glides as arising via GF/Devocalisation in some cases *and* as being represented underlyingly in others. It is difficult to see the justification for this inconsistency.

Another reason for rejecting Teoh's analysis is as follows. Forms such as [tarijan], from putative /tari/ (root) plus /jan/ (suffix) are analysed as having an underlying glide in the onset of the suffix. But that suffix will not contain a glide if the preceding root is consonant-final, as in /masak + an/ [masaʔkan] (Gem and Glottaling apply here - see below). Thus, in cases such as [tarijan], the glide *must* be derived, and in the case of words such as [ju], [waŋ] and [jaŋ], there is no evidence that they are derived. Given that words such as this are phonetically bisyllabic, it is difficult to see the justification for resyllabifying an initial nuclear vowel such that it comes to occupy the onset position of the following syllable.

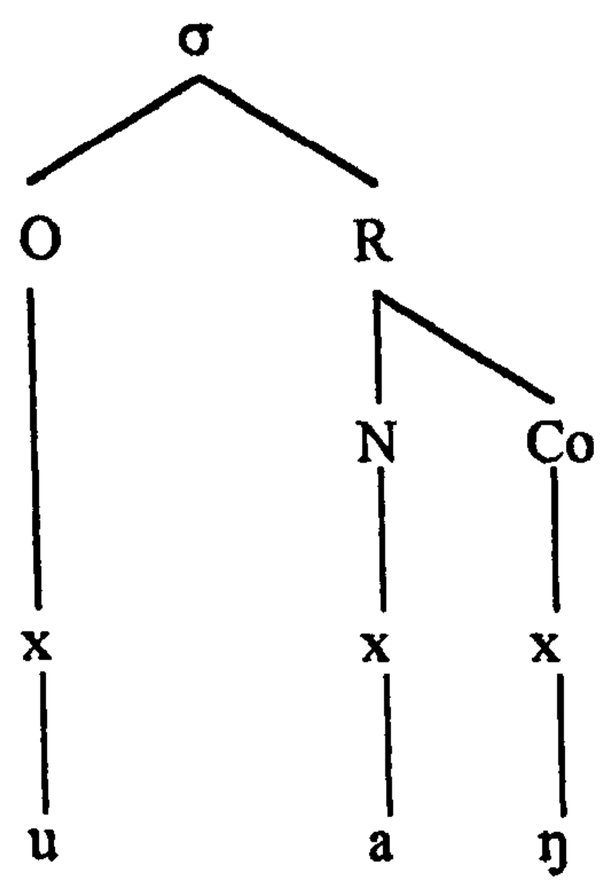
We suggest that the simplest analysis, previously proposed by Durand (1987), consists in taking all of the cases in (3) to contain a high vowel in onset position underlyingly. Given our definition of a glide (a high vowel in non-peak position), all such words contain an underlying glide. Thus, for us, the monosyllabic roots are underlyingly /iu/, /uaŋ/ and /iaŋ/, and are parallel to, for example, /ru/ ('casurina tree'), /raŋ/ ('code of law') and /dan/ ('and') in that they contain a filled initial onset underlyingly. Fuller representations showing the syllable slot occupied by these underlying high vowels in onset position (i.e. glides, on our definition) are reproduced here in (5).

(5).



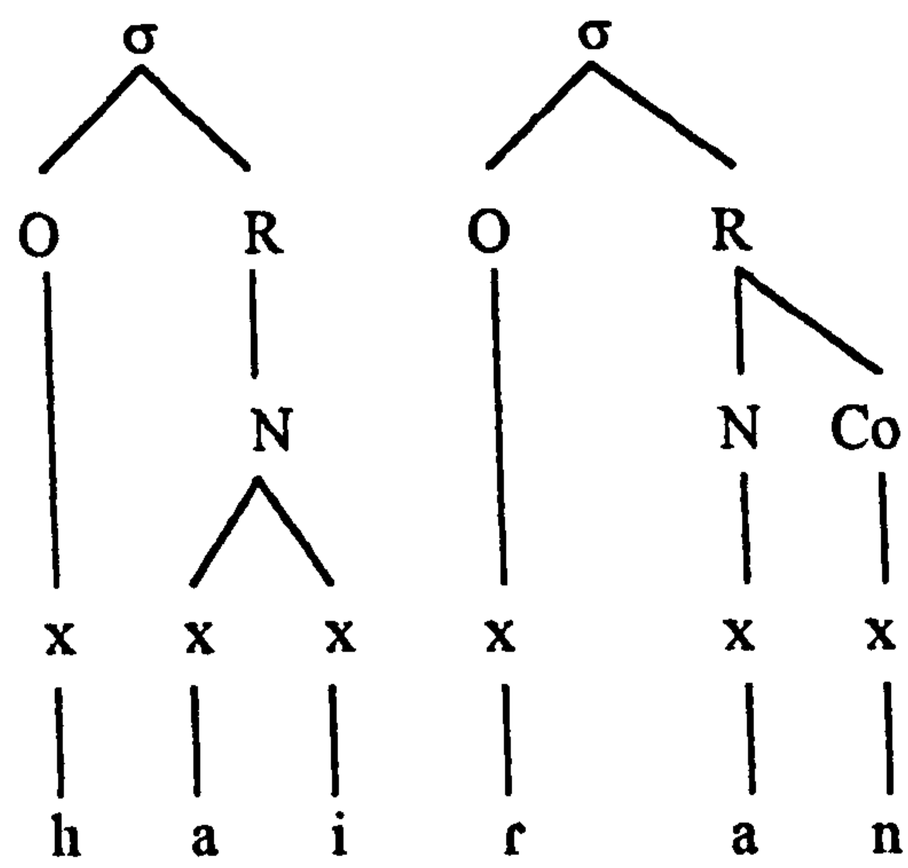
Similarly, we will refer to the glide of *'wang'* as an underlying high vowel in onset position, and represent it as follows:

(6)



We take the underlying diphthongs /ai/ and /au/ to contain /i/ and /u/ in non-nucleus peak position, as in /hai.ran/ 'wonderment' and /ka.lau/ 'if'. In this case, we represent the underlying diphthongs as follows:

(7) High vowel as non-nucleus peak position in diphthong

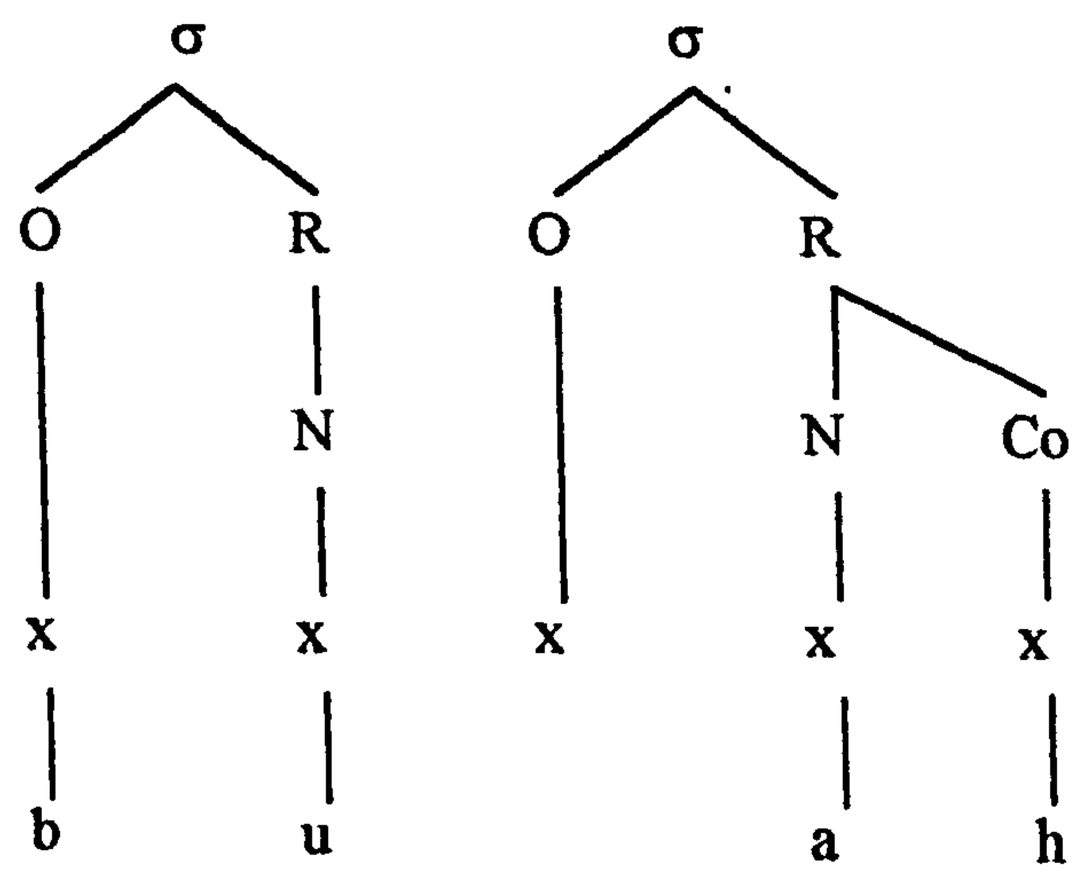


Once again, a glide is a high vowel in non-nucleus peak position, and in all of these cases, the glide is underlying. This analysis obviates an entire set of constraints postulated by Zaharani (1998: 37) who considers that such cases contain an underlyingly *heterosyllabic* /ai/ hiatus sequence, and postulates a hiatus-avoidance strategy of diphthongisation, in which the high vowel is syllabified into the same nucleus as the /a/. We see no reason to add this unnecessary complexity to the grammar of SM. A parallel, and equally unnecessary, analysis would be to argue that the English [ai] diphthong in words such as *buy* results from the application of a hiatus-avoidance strategy of diphthongisation, applied to an underlyingly bisyllabic form /ba.i/. We suggest that SM, like English, has underlying diphthings in cases such as this, and thus that the offglide in the diphthong is underlying. It seems to us that our definition of 'glide', and our claim that these occur underlyingly in both the [waŋ] and [hajɾan] cases is clear and consistent and obviates the unnecessary postulating of phonological processes to account for facts which are not rule-goverened.

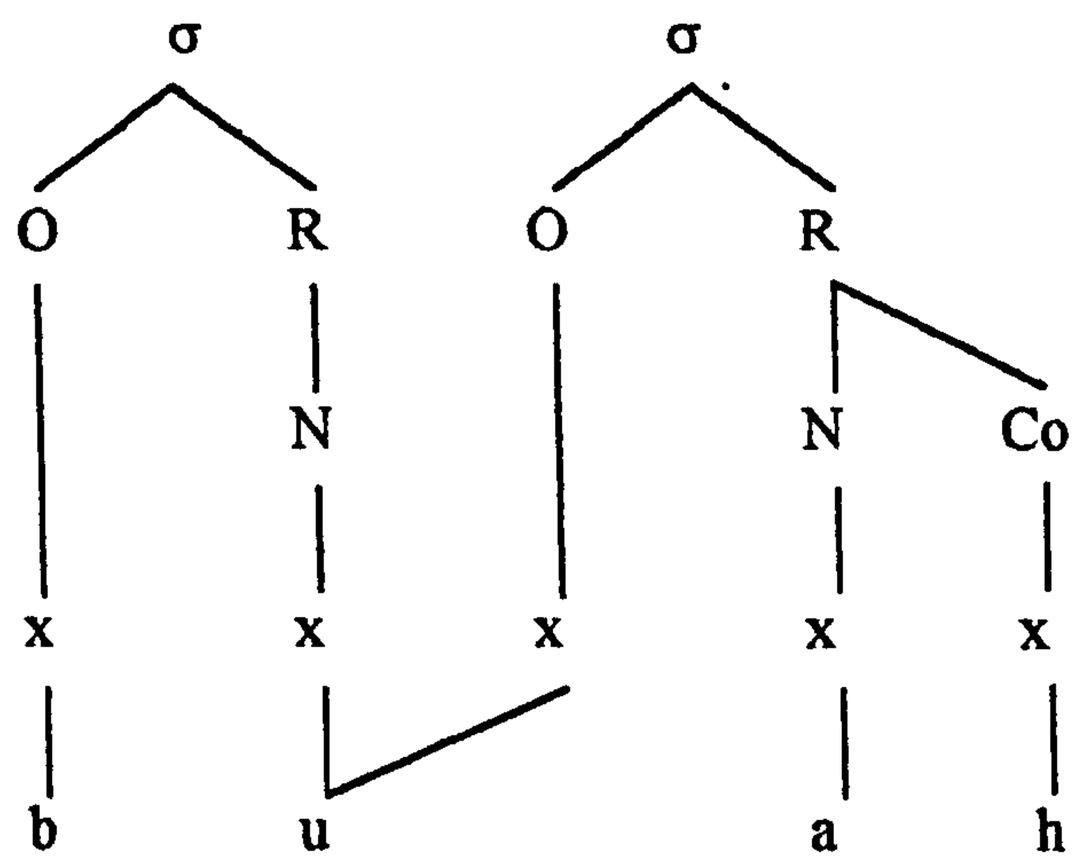
We note, in passing, that, when GF applies in Standard French, nuclei containing a high vowel may become onsets, and consequently, a syllable may be lost. Durand (1987) argues that GF in SM works this way too. In contrast, we claim that GF in polysyllabic roots in SM results in the vowel-plus-glide sequences [ij] and [uw], with no reduction in the number of syllables in the word. The difference is that GF in SM is a hiatus-avoidance phenomenon as it creates an onset segment between what would otherwise be a sequence of two vocalic nuclei but at the same time are separated into two syllables [i.j] and [u.w] such as in the roots /tiaŋ/ [ti.jaŋ] 'pole' and /buah/ [bu.wah] 'fruit' respectively. We take these cases to have underlyingly empty onsets and derived glides, on the grounds that the glides in [ij] and [uw] are always predictable. We present this phenomenon as in (8) below.

(8) GF without loss of syllable

a.



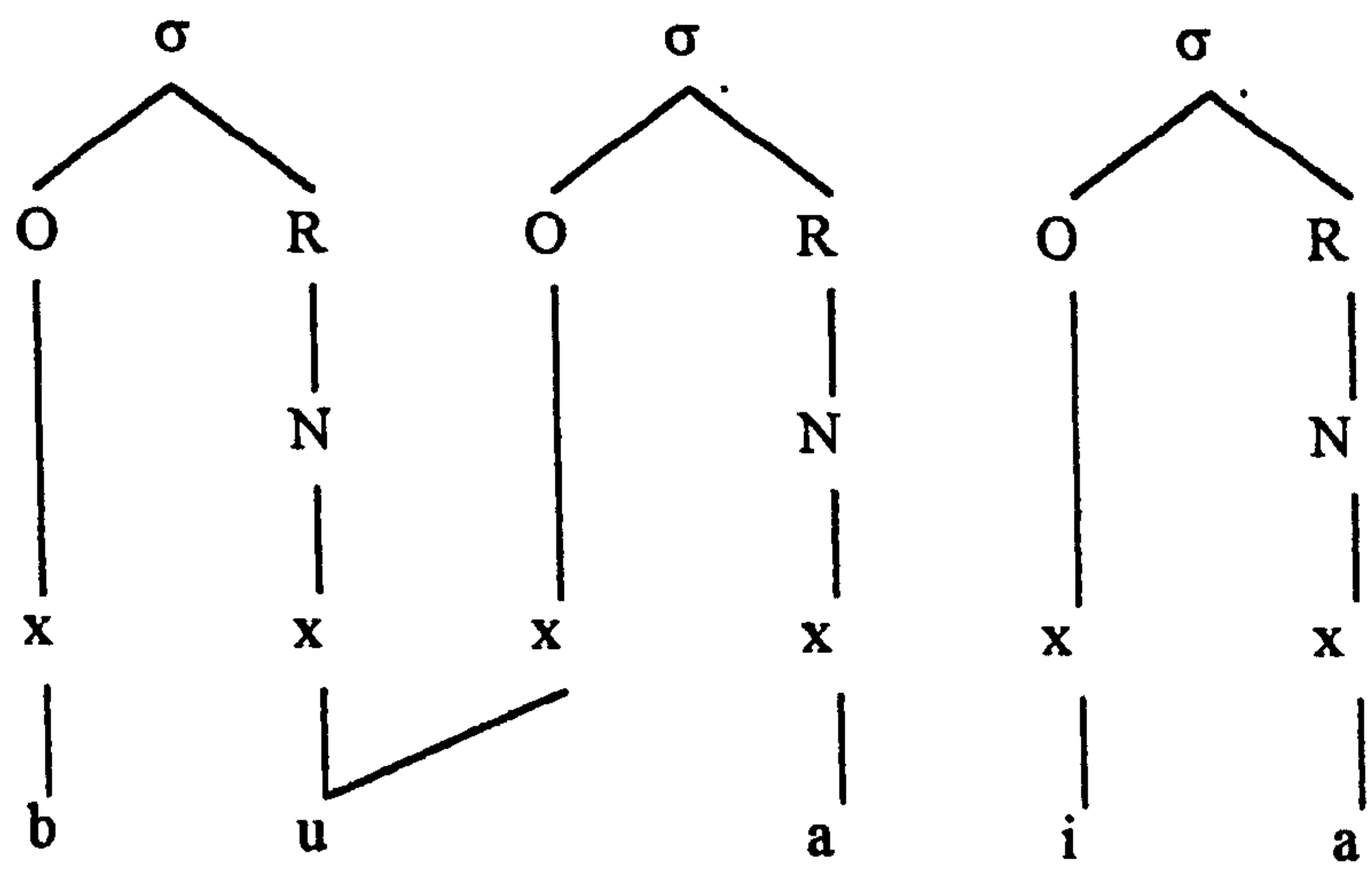
b.



Phonetic representation =[buwah]

In cases such as [bu.wa.jə] 'crocodile', the [w] glide is derived via GF which is similar to the case of [bu.wah], while the [j] glide is underlying. The informal representation of this phonological form is /buaja/, where the /j/ is shorthand for /i/ in the non-nuclues peak (onset) position. Thus, we represent /buaja/ (after the operation of GF) as in (9):

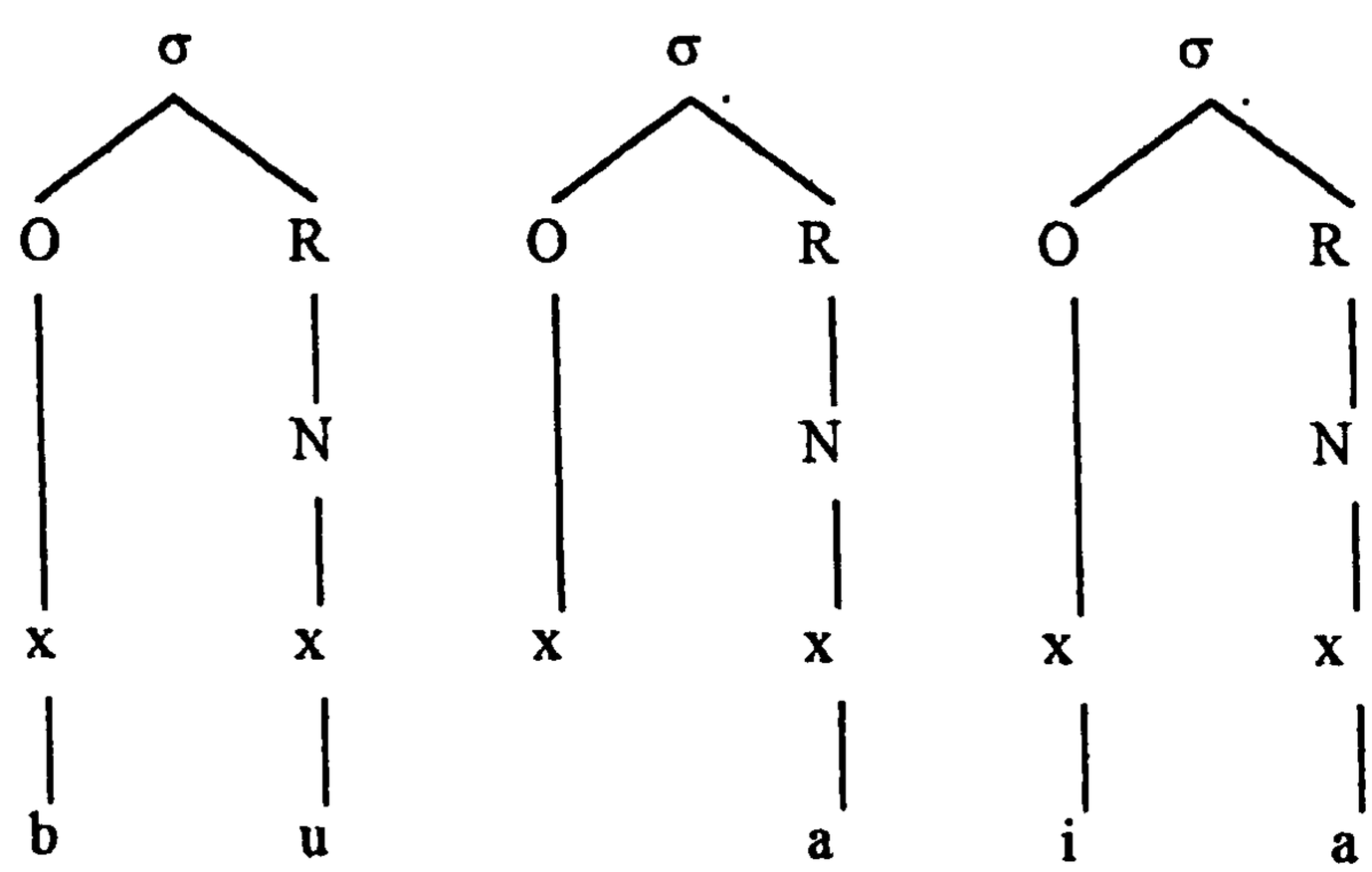
(9)



Phonetic representation = [buwaja]

We will assume that, in derivational terms, roots such as this have the following underlying syllable structure

(10)



Phonological representation =/buaia/

In (9), the application of GF breaks up any underlying hiatus.

As we have seen, SM, like Standard French (Hannahs 1995a, b, ms), has both underlying and derived glides[14]: the occurrence of all glides in Malay is *not* entirely predictable. For example, the glides in [waŋ], [kuju] ('half closed eye') and [waŋi] ('fragrant') are cited by Zaharani as being predictable. We claim that, with glides defined as high vowels in non-nuclesu peak position, these glides are not predictable. It might be objected that there is perhaps an air of circularity about this claim, so let us spell out exactly what we mean. It is an entirely arbitrary fact that words such as /uaŋ/ ('half-closed eye') begin with a /u/, just as it is an entirely

[14] Like Durand (1987), SM, we do not consider [ʔ] and [h] to be glides. Although glottal obstruents pattern with [j] and [w] in SM, it does not follow that the four segment types constitute a class of glides. Rather, what the two sets of segment types have in common is that neither of them constitute an oral obstruent.

arbitrary fact about /raŋ/ ('code of law') that it begins with a /r/. The /u/ here is pronounced [w] since it occurs in a non-nucleus peak position. That it is so pronounced is, of course, predictable. But that phenomenon is not the result of the application of a rule of Devocalisation in the sense proposed by Teoh and Zaharani, under which a vowel occupying a nuclear position is shifted into an onset position. Under Teoh and Zaharani's analysis, there are two phenomena here: shifting a high vowel into an onset position, and pronunciation of the onset high vowel as [w]. It is the former that we object to. The latter, we claim, results automatically from the position of the high vowel in a non-nucleus peak position; it does not result from the application of a phonological rule[15].

We now sum up our treatment of root-internal glides. We have claimed that the phonetic glides in cases such as [waŋ] ('money') and [kuju] ('half-closed eye') are underlying. So too, we claim, are the offglides in cases such as [haj.ran] 'wonderment' and [ka.lau] 'if': these words have underlying diphthongs. Both glides in a word such as [wa.jaŋ] ('movie') are underlying: the UR is /ua.iaŋ/. Words such as this do not have underlying diphthongs or sequences of adjacent vocalic nuclei. The glides in words such as [si.jaŋ] ('day') and [lau.wan] ('compete') are derived glides, resulting from the application of GF, operating on URs of the form /si.aŋ/ and /lau.an/. Finally, the [w] glide in cases such as [buwaja] ('crocodile') is derived from the application of GF, whereas the palatal glide is underlying, the UR being /bu.a.ia/.

---

[15] We also note, in passing, that Zaharani appears to believe that deriving glides from high vowels is a novel idea; in fact, the idea is over 40 years old (see Jakobson, Fant and Halle 1952).

In cases such as [si.jaŋ] and [buwaja], GF can be seen as a hiatus-avoidance strategy, similar in function to the process of GSI, under which a glottal stop [ʔ] is inserted between a sequence of two vocalic nuclei, as in the case of Arabic loan words in SM, e.g. /maaf/ [ma.ʔap] 'forgive' and /saat/ [sa.ʔatˤ] 'second'. GF in cases such as [lau.wan] is slightly more complex: GF can be seen as a hiatus-avoidance strategy, since the sequence in question is an offglide within a diphthong followed by a vowel in nucleus position. But this latter case differs slightly from the others in that the high vowel in question is nuclear, but not in nucleus peak position; in other words, in this case, it is an underlying glide (an offglide) which undergoes GF. We now turn to the application of GF in polysyllabic words, where it operates across a root-suffix boundary.

3.2.2 *GF across a root-suffix boundary*

SM has three suffixes, two of which are vowel-initial: /-an/ and /-i/[16]. These two suffixes possess an empty onset (i.e. /-Xan/ and /-Xi/) in which the slot acquires melodic material when followed by a vowel-final root. The following vowel-final roots consist of a non-high vowel with a glottal stop inserted in the empty slot by default. This is Glottal Stop Insertion (GSI)[17]. In other cases, where the final vowel

[16] The verbal suffix, which signals the semantic roles of goal and benefactive: see Teoh (1994: 61). SM has a third suffix, /-kan/, a verbal suffix which signals either transitivisation, or one of the semantic rolesof accusative or benefactive.

[17] The occurrence of GSI in SM will be discussed in more detail in Chapter 7.

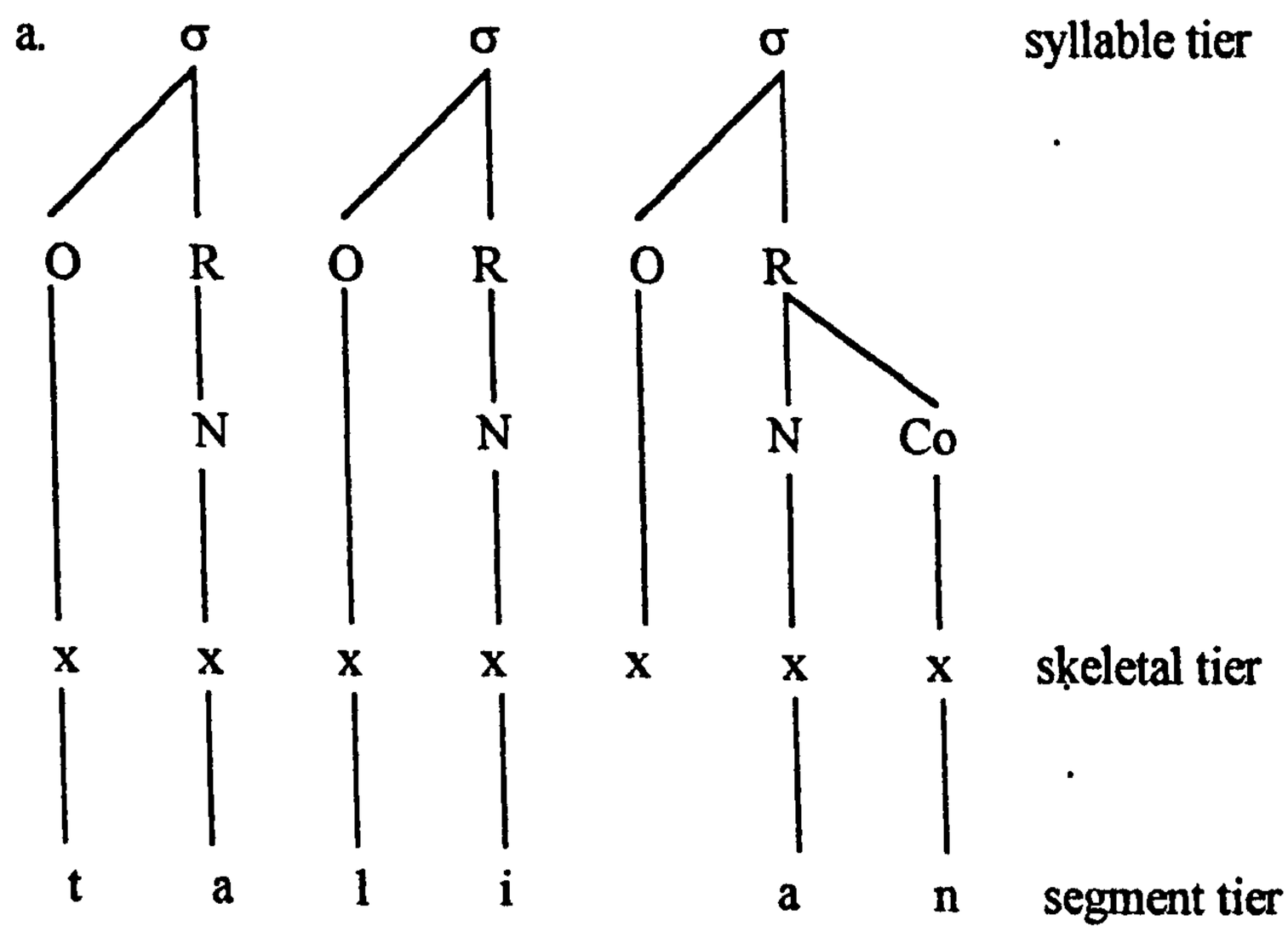
in the root is a high vowel, the empty slot (i.e. onsetless syllable) will be filled by a glide. The occurrence of a glide can thus be seen as a derived glide *via* GF[18]. Examples of such GF are shown below and can be represented as in figures 11 & 12.

(11) A [+high] root-final vowel + onsetless suffixes

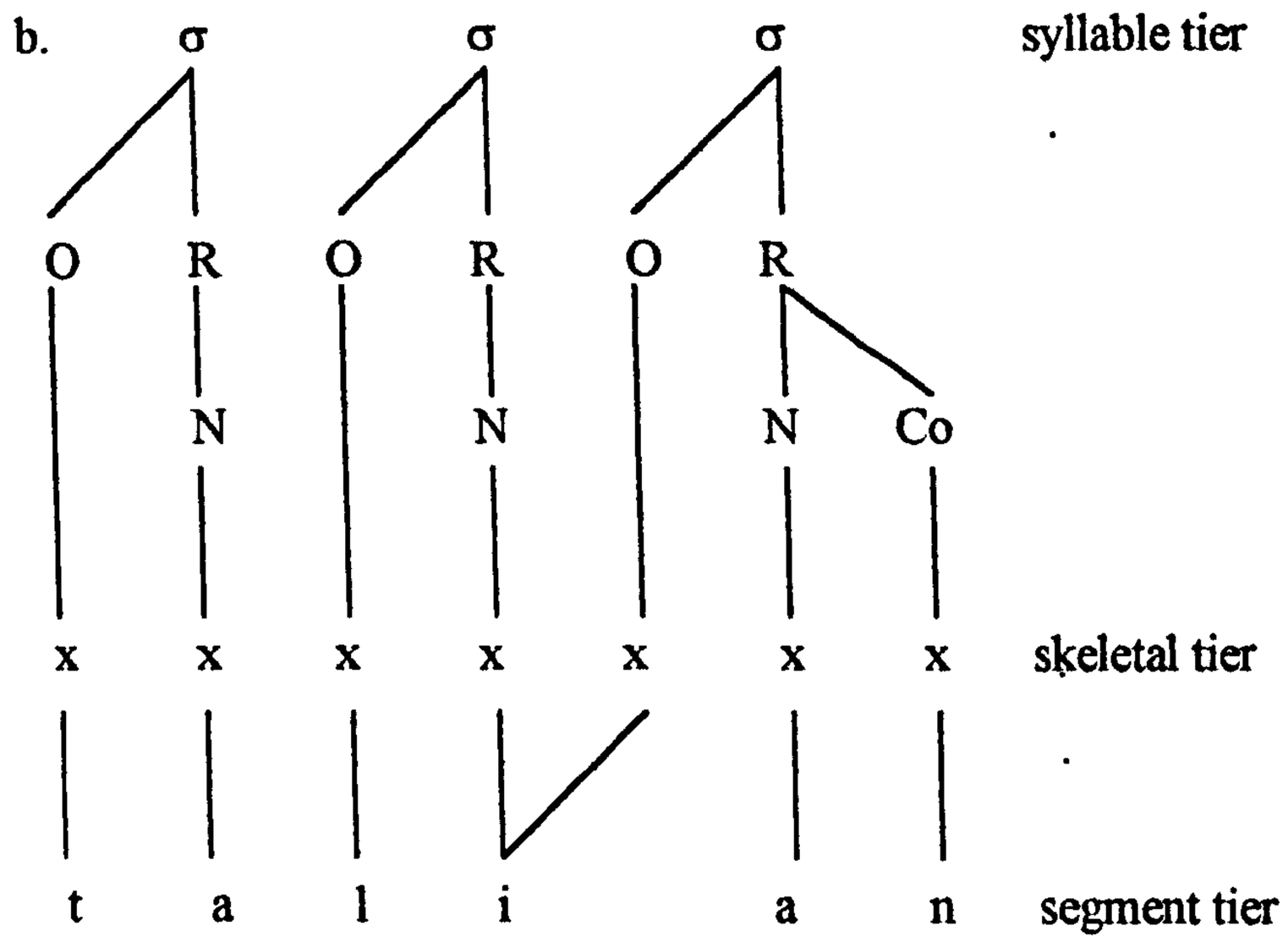
/tali + an/	[talijan]	'line'
/tari + an/	[tarijan]	'dancing'
/bəli - an/	[bəlijan]	'buying'
/buru + an/	[buruwan]	'hunting'
/dulu + i/	[duluwi]	'leading'
/aku - i/	[akuwi]	'confess'

We assume the following sorts of representation for such cases:

(12) GF across a root-suffix boundary in word /ta.li.jan/



[18] In Zaharani's (1998) analysis, GF is called V-gemination.



Phonetic representation =[ta.li.jan]

It is important to note that GF in SM occurs between a diphthong-final root (with ending in a high vowel, that is /ai/, /oi/[19] or /au/) and an onsetless suffix. Examples of the application of GF in this context are:

(13) Stems with diphthongs + onsetless suffixes

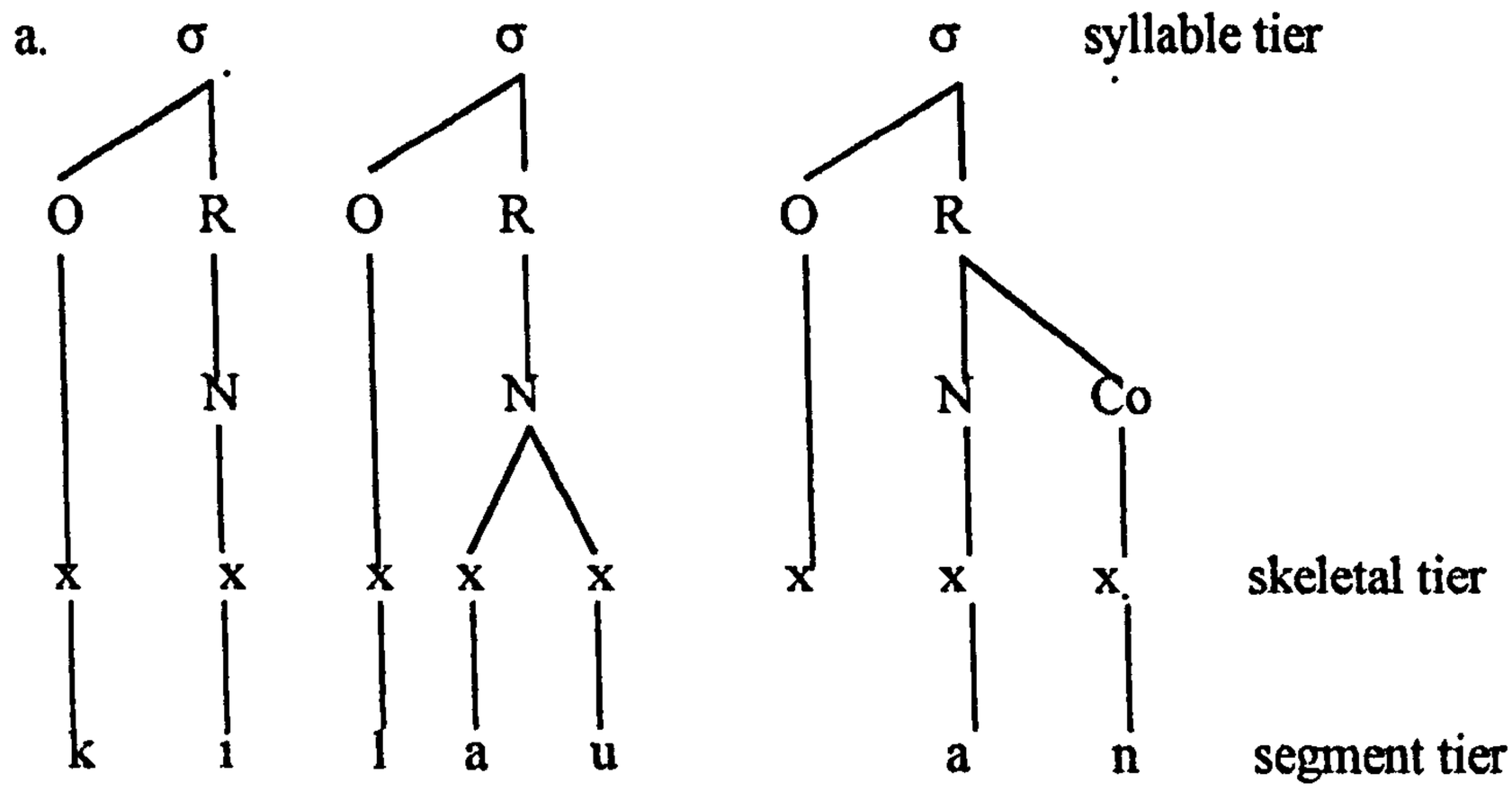
/lambai + an/	[lambaijan]	'waving'
/kilau + an/	[kilauwan]	'shining'
/məŋ + lampau + i/	[məlampauwi]	'extreme'

[19] There are a few monomorpheme words in SM ending with diphthong /oi/, for example /dodoi/ 'lullaby' and /kaloi/ 'a fish name'. However, it is impossible to add the word ending with diphthong /oi/ with suffix /-an/ or /-i/.

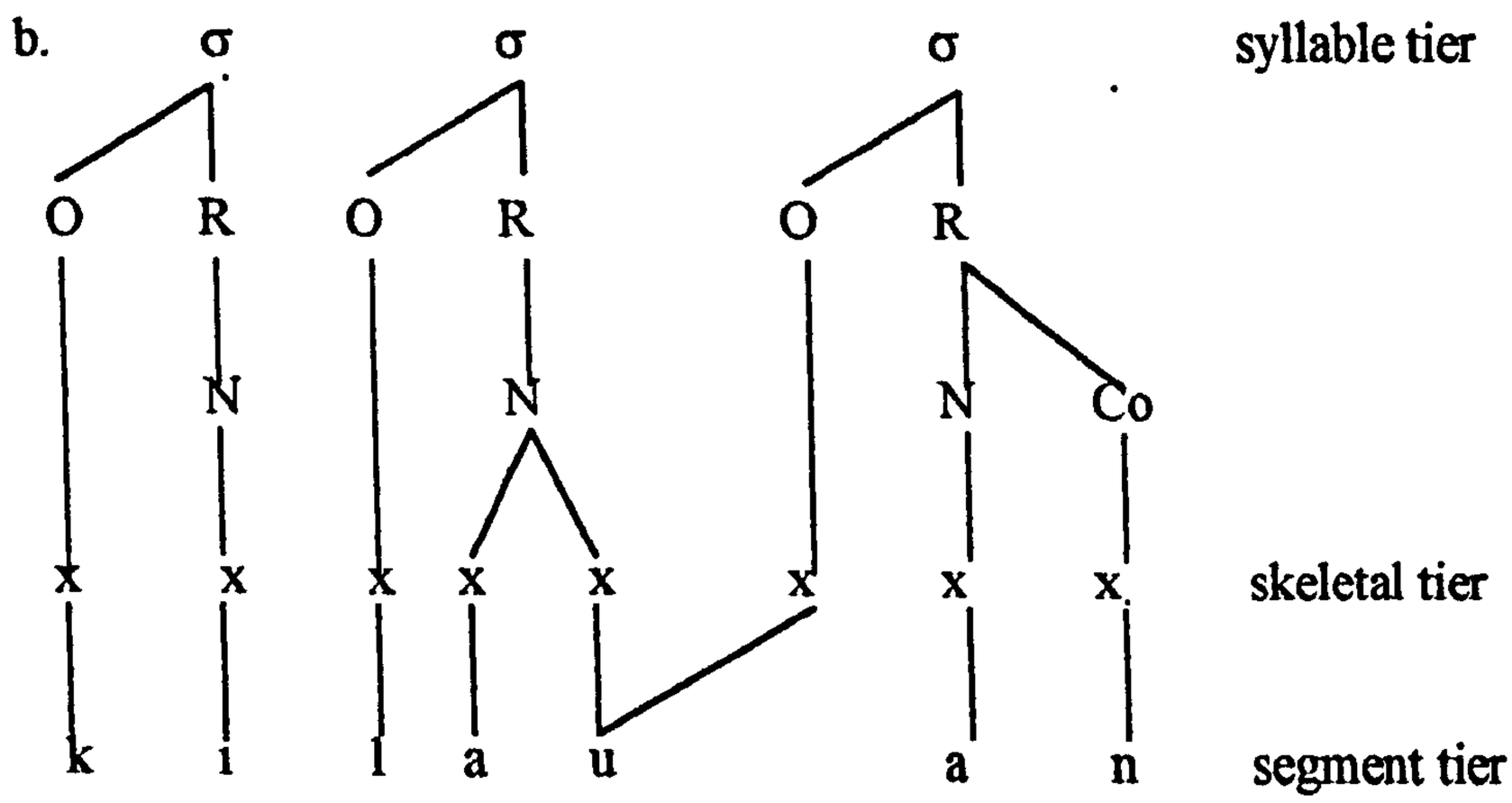
We will represent these diphthongs as a sequence of /a/ plus high vowel within a nucleus. (14) illustrates how GF applies to a non-peak high vowel of a diphthong.

(14)

a.



b.



Phonetic representation =[ki.lau.wan]

The case shown in (14) indicates that the high vowel is in the same nucleus as the /a/ (but not in nucleus peak position), with GF forming a derived glide in the empty onset of the suffix /-an/. Our view that GF may apply to the offglide in a diphthong differs slightly from Zaharani's (1998) view that the vowel /a/ is in the nucleus position, whereas the high vowel /u/ is in the coda position. We concede that it is difficult to provide convincing evidence either way as to whether the offglide is located within the nucleus or within coda position. And many historical changes suggest that a segment may switch from one position to the other, perhaps passing through a stage where the segment is not clearly located within either the nucleus or the coda. One example is the nasalisation of preceding vowels by a coda nasal, resulting in coalescence of the nucleus vowel and the coda nasal, as in Standard French. Another is the historical transition from sequences of vowel plus coda /r/ in non-rhotic accents of English, such as RP, resulting in schwa offglides in the centring diphthongs, where the schwa offglide in the nucleus seems to be the remnants of a historical coda consonant.

3.2.3 *GF is blocked across a prefix-root boundary*

In SM the vowel-final prefixes are /di-/ and /dʒuru-/ (with high vowels) and /ke-/ and /se-/ (which contain schwa). When these vowel-final prefixes attach to vowel-initial roots; GF is blocked, even if the prefix ends in a high vowel and the root is vowel-initial. Furthermore, the sequences of the two underlying vowels /V +

V/ must be syllabified heterosyllabically as [V.V] separated by a glottal stop. Glottal Stop Insertion (GSI) is, we claim, a postlexical, 'across-the-board'[20] rule.

(15)

a. Vowel-final prefixes with high vowel

/di+ambel/	[di.ʔam.bel]	*[di.jam.bel],	*[di.am.bel]	'taken'
/di+ubah/	[di.ʔu.bah]	*[di.wu.bah],	*[di.u.bah]	'moved'
/dʒuru+ukor/	[dʒu.ru.ʔu.ko:]	*[dʒuruwuko:],	*[dʒu.ru.u.ko:]	'surveyor'
/dʒuru+atʃara/	[dʒu.ru.ʔa.tʃa.rə]	*[dʒu.ru.wa.tʃa.rə],	*[dʒu.ru.a.tʃa.rə]	'presenter'

b. Vowel-final prefixes with schwa

/sə + iras/	[sə.ʔi.ras]	*[sə.i.ras]	'resemble'
/sə + utoh/	[sə.ʔu.toh]	*[sə.u.toh]	'as strong as'
/kə + ada + an/	[kəʔadaʔan]	*[kə.a.da.ʔan]	'situation'
/kə + əmas + an/	[kə.ʔə.mas.san]	*[kə.ə.mas.san]	'golden'

3.2.4 *GF is blocked across a morphological word boundary*

Morphological Word (MW) boundaries in SM can be seen in compounds, reduplicated words, and syntactic phrases. When vowel-final roots combine with vowel-initial root in MW boundaries, the hiatus is resolved by GSI. GF is blocked in

[20] For further discussion of GSI, see Chapters 5 & 7.

all of these contexts, even for a sequence of high vowel being followed by another vowel. Examples of GF being blocked in these contexts are:

(16) GF is blocked in compounds

/ibu/ plus /ajam/	[i.bu.ʔa.jam]	*[i.bu.wa.jam]
('mother') ('hen')	'pimp'	
/səgi/ plus /əmpat/	[sə.gi.ʔəm.pat']	*[sə.gi.jəm.pat']
('shape') ('four')	'square'	

(17) GF is blocked in reduplicated forms

/anai/	[anaiʔanai]	*[anaijanai]	'termite'
/alu/	[aluʔalu]	*[aluwalu]	'welcome'

(18) GF is blocked in syntactic phrases

Hari ini ada kuliah [hariʔiniʔadəkuliyah] *[harijinijadəkuliyah]
'Today there is a lecture'

Dia ada ibu angkat [dijəʔadəʔibuʔaŋkat']*[dijəʔadəʔibuwaŋkat']
'He/she has an adopted mother'

The above data make it clear why the domain of application of GF is taken to be the phonological word as defined under the Standard View. GF in SM can thus be

characterised as a domain span rule without requiring ordered lexical levels as proposed by Mohanan (1986) and Pulleyblank (1986).

3.3 Gemination (Gem)

3.3.1 *Gem across a root-suffix boundary*

In SM, when a vowel-initial root concatenates with a consonant-final prefix, the onsetless root always acquires its onset from a preceding consonant through Resyllabification[21]. But Resyllabification does not apply across a root-suffix boundary (Abdullah 1974, Teoh 1994 and Zaharani 1998). Instead, Gem applies. Examples of Resyllabification may be found in the words /bər+aŋkat/) [bə.ɾaŋ.kat] 'to depart' and /məŋ+aŋkat/ [mə̃.ŋãŋ.kat]) 'to lift' *vs* Gem in /lətop + an/ [lətop.pan] 'explosion' and /sambut + an/ [sambut.tan] 'reception'. Teoh, Zaharani and Abdullah all agree that a consonant only geminates across a root-suffix boundary[22], when a consonant-final root attaches itself to a suffix with an empty onset, i.e the suffixes /-an/ and /-i/. Gem is like GF in that it only occurs across a root-plus-suffix boundary. According to the Standard View, Gem, like GF, is a process which occurs within the PW.

---

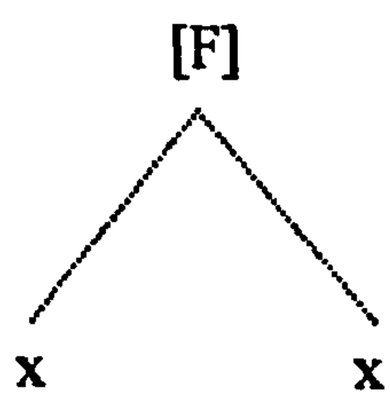
[21] In this analysis, the resyllabification process does occur in root-plus-suffix boundary. This involves Floating /r/ root final position plus vowel-initial suffix, such as in /tampar + an/ → [tam.pa. ran] * [tam.par.ran]. For further details see Chapter 5.

[22] Consonant geminates are also seen as not occurring in monomorphemic words such as in /sapu/ 'sweep' *[sap.pu] and /kota/ 'castle' *[kot.tə] vs. [sa.pu] and [ko.tə] in SM.

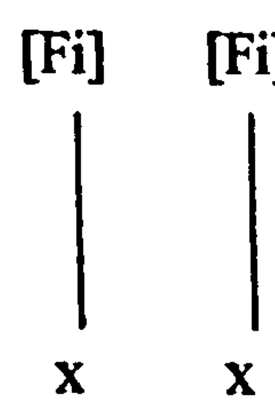
As far as the representation of geminates is concerned, Hayes (1986), Steriade (1987), and Teoh (1994: 68) all claim that Gem is represented as a multi-linked segment. A distinction between true geminates, as in [lətop.pan] 'explosion' [sambut.tan] 'reception' and false geminates, as in /masak + kan/, can be postulated for SM. We contrast /masak + an/ and /masak + kan/ in (19b). We follow Teoh in adopting the following means of representing true and false geminates.

(19). Distinction between true and false geminates

(a) monosegmental geminate (true geminate: derived in SM)



(b) heteromorphemic geminate (false geminate)



(19a) represents the kind of case in which an underlying root-final consonant is geminated, as in /masak + an/, which, we claim → masak + kan, and then the coda velar stop undergoes a further process of glottal reduction to become [masaʔkan]). (19b) represents cases such as masak + kan, where a sequence of two velar stops arises via suffixation. Glottalling occurs whether the sequence is a true or a false geminate: both /masak + an/ and /masak + kan/ → [masaʔkan]. Under a traditional rule-based derivational model of phonology, Gem applies prior to /k/

Glottalling. We unfashionably suggest that this derivational, rule-ordered account of the interaction of the generalisations is perspicuous.

Not all of the SM consonants undergo Gem. When a Floating /r/ in a root attaches to a vowel-initial suffix, it resyllabifies as in /tampar + an/ 'smack' [tam.pa.ran] *[tam.par.ran]; we remain agnostic as to what a geminated /r/ might sound like, but [d] and [r] are plausible candidates[23]. All of our informants uttered /r/ in this context. This evidence is supported by previous findings that show that /r/ does not geminate in SM but undergoes resyllabification (see Ismail 1996). Most other authors (e.g. Teoh 1994, Abdullah and Ainon 1994 and Zaharani 1998) claim that /r/ does undergo Gem but they appear to have made observational mistakes as all of the informants in our study have a single tap in the relevant cases, as in [tam.pa.ran] from /tampar + an/ 'smack'. Examples of geminated forms of the full range of consonants which undergo Gem is given in (20).

(20) Gem: Consonant-final roots + vowel-initial suffixes

/lətop + an/	[lə.topˀ.pan]	'explosion'
/sambut + an/	[sam.butˀ.tan]	'celebration'
/masak + an/	[ma.saʔ.kan][24]	'cooking'
/tanam + an/	[ta.nam.man]	'plants'
/atas + an/	[a.tas.san]	'superior'
/kaseh + an/	[ka.seh.han]	'pity'

[23] It has been suggested to us that Spanish /r/, contrastive with /r/, arose from a historical process of gemination. If so, it is the only consonant that underwent this process in the history of Spanish.

[24] /k/ is realised as [ʔ] in coda position *via* /k/ Glottalling.

/kəsal + an/	[kə.sal.lan]	'regret'
/ubat + i/	[u.bat'.ti]	'treat'
/hadap + i/	[ha.dap'.pi]	'face'
/məŋ + masok +i/	[mə.ma.soʔ.ki]	'to enter'
/faham + i/	[pa.ham.mi]	'understand'
/kaseh + i/	[ka.seh.hi]	'causing to love'
/səsal + i/	[sə.sal.li]	'causing to regret'
/atas + i/	[a.tas.si]	'overcome'

Non-gemination of /r/ is shown in (21).

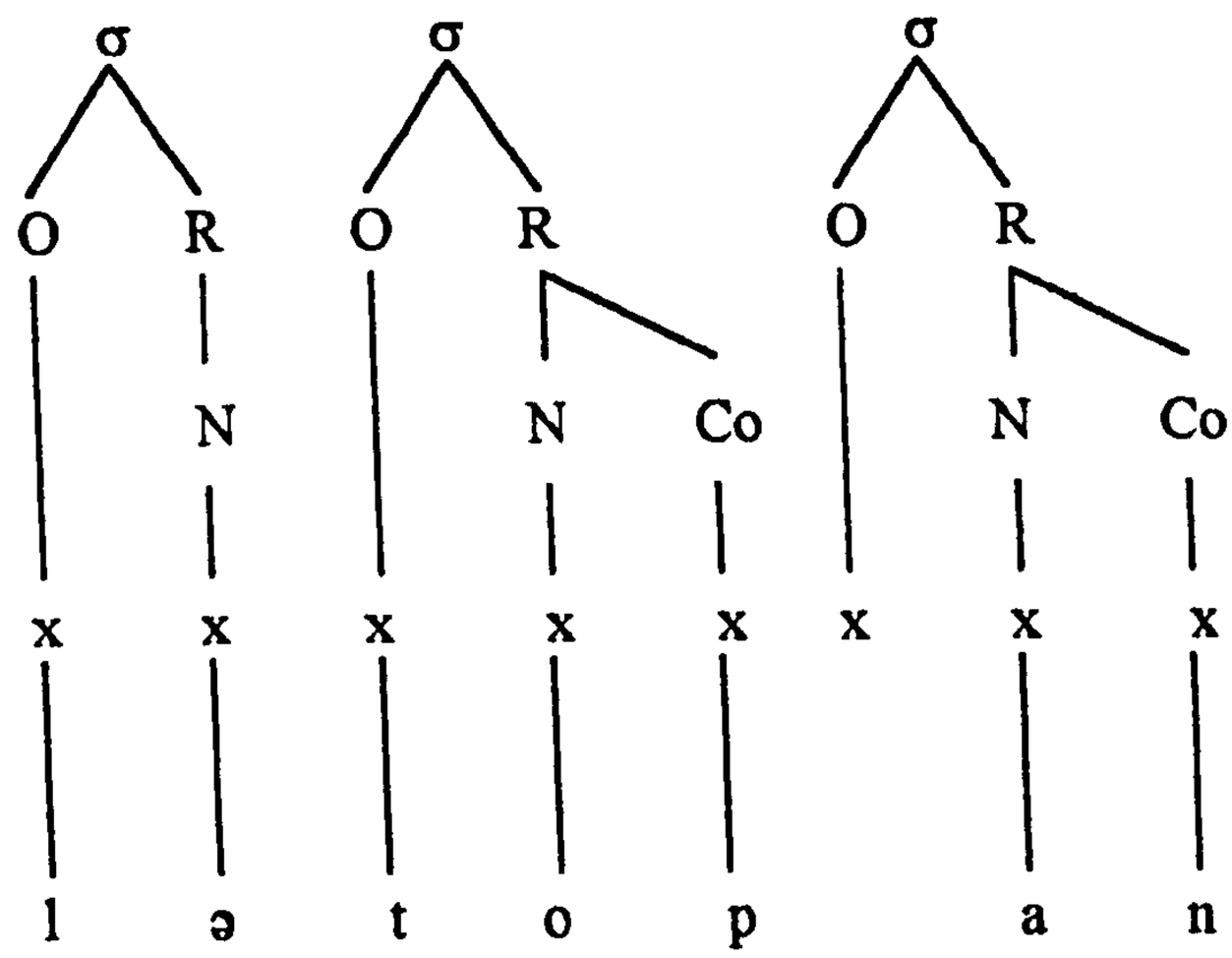
(21). Floating /r/-final roots + vowel-initial suffix

/tampar + an/	[tam.pa.ran]	*[tam.par.ran]	'slap'
/adʒar + an/	[a.dʒa. ran]	*[a.dʒar.ran]	'teaching'
/tʃabar + an/	[tʃa.ba.ran]	*[tʃa.bar.ran]	'challenge'
/pər + tukar+ an/	[pə:tu.ka.ran]	*[pə:tu.kar.ran]	'alternation'

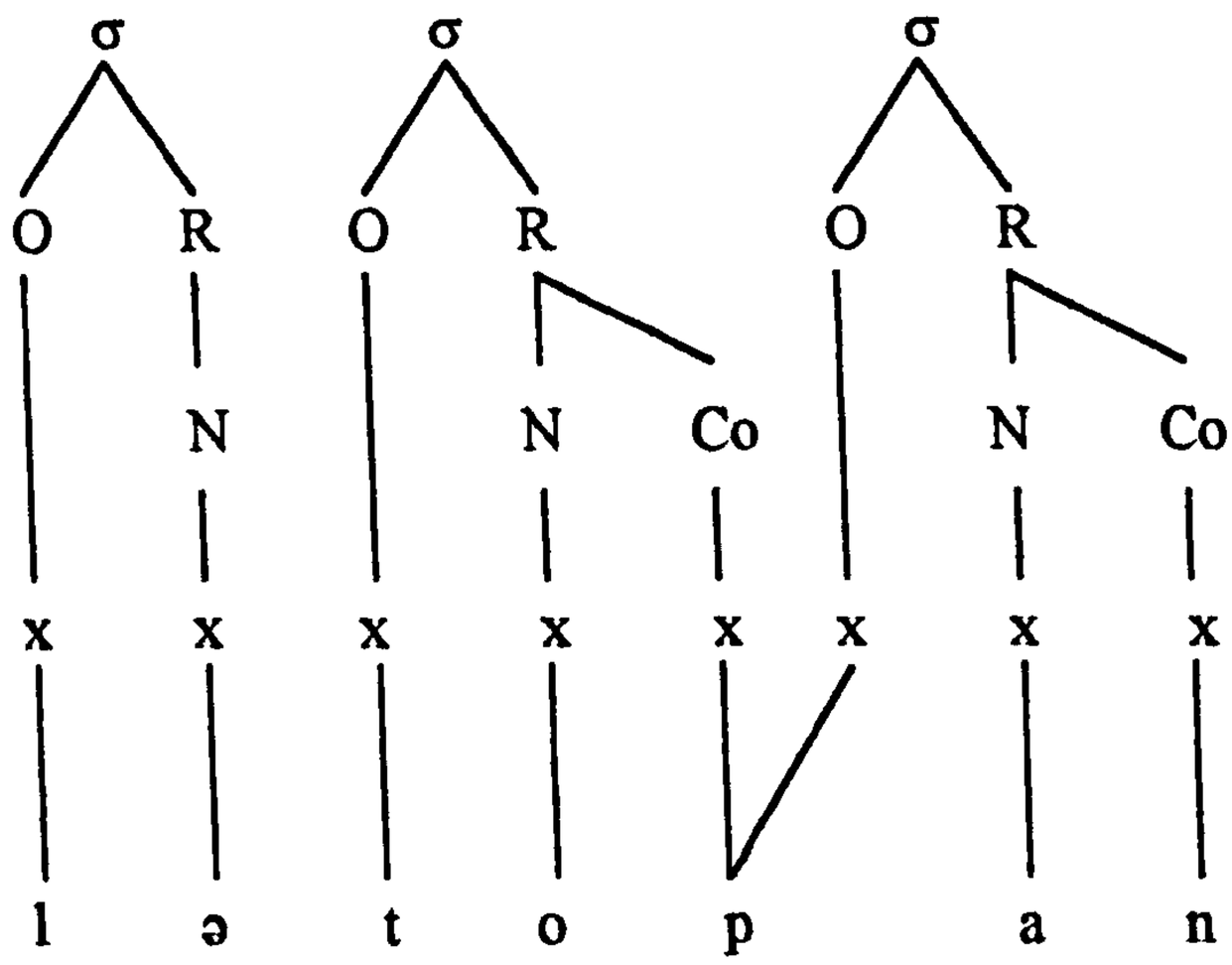
The geminates in (20) can be represented as in (22).

(22) Gemination –segmental linking of the root-final consonants

a.



b.



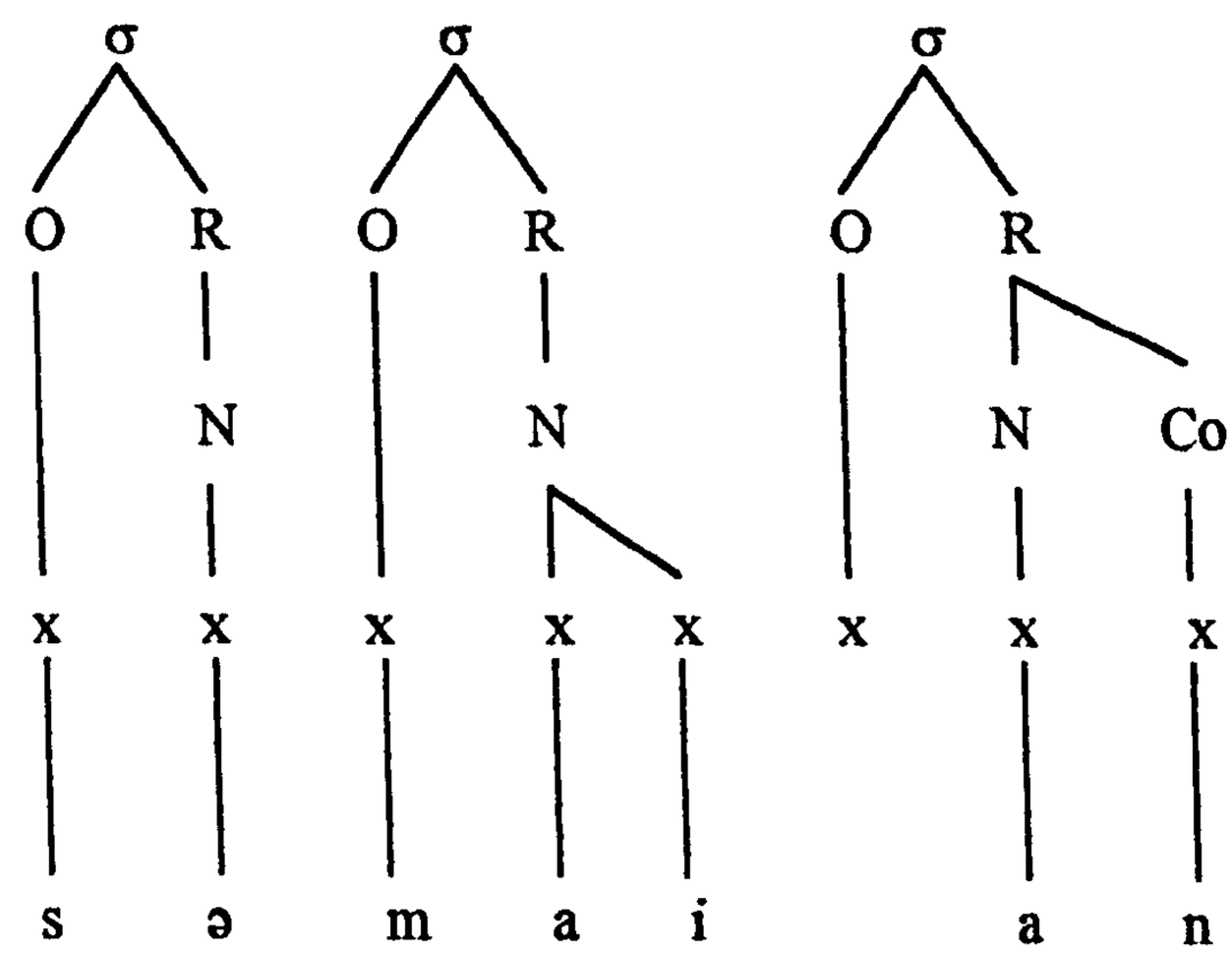
Phonetic representation (=[lə.topˀ.pan])

In the diagram above the segment /p/ (a root-final consonant) in the coda position of the word /lətop/ associates to the onset of the empty x-slot of the suffix /-

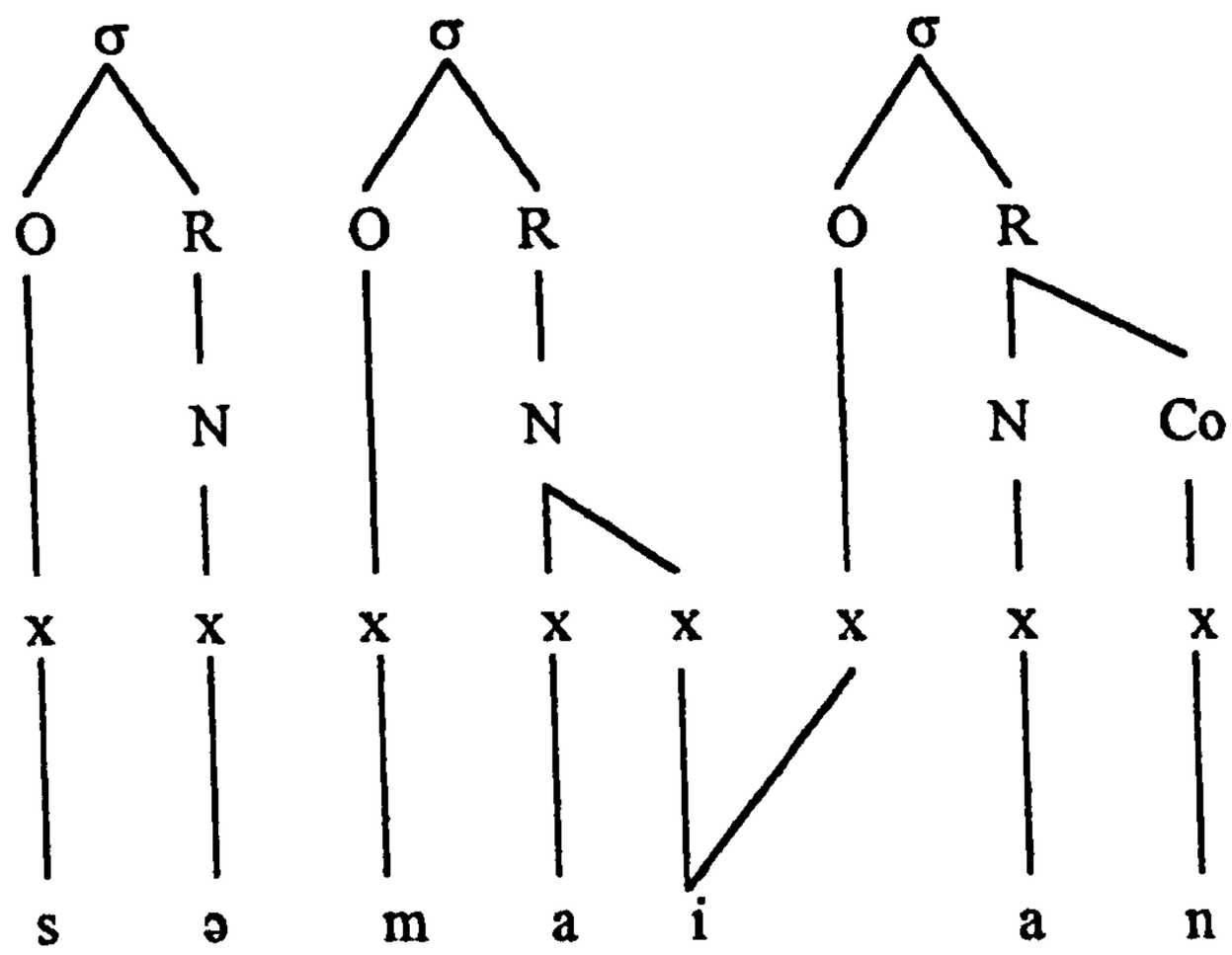
an/. As a result, the consonant /p/ becomes a geminate [p˺p]. The process of gemination shown in (22) shows how the association line of the consonant-final root to the empty x-slot of the onsetless suffix closely parallels the 'vowel geminates' case of GF in Section 3.2. To illustrate this parallelism, the vowel gemination of GF is shown as a tree diagram in figure (23).

(23) GF: linking of root-final high vowel root onto an empty x-slot of the suffix /-an/.

a.



b.



Phonetic representation =[sə.mai.jan]

In SM, Gemination is observed to take place when a consonant-final root (in coda position) attaches to an onsetless suffix only. By contrast, a consonant in the onset position cannot geminate with a preceding syllable, i.e. Gem in SM occurs from left-to-right but is blocked from right-to-left. Examples are the words /lətop + an/ [lə.topˀ.pan] 'explosion' vs /nama + kan/ 'to be named' [na.ma.kan] *[na.mak.kan].

3.3.2 *Gem is blocked across a prefix-root and morphological word boundaries*

Like GF, Gem is blocked across a prefix-root boundary and a word boundary. For example:

(24) Prefix + root

a. *Consonant-final prefixes + consonant-initial roots*

/məŋ + ŋaŋa/	[mə:ŋa.ŋə]	'to gape'
/tər + rakam/	[tə:ra.kam]	'recorded'

b. *Consonant-final prefixes + vowel-initial roots*

/məŋ + adʒok/	[mə.ŋã.dʒoʔ]	'to tease' (active)
/məŋ + ubah/	[mə.ŋu.bah]	'to change'
/məŋ + ikut/	[mə.ŋi.kot]	'to follow'
/ber + adu/	[be.ra.du][25]	'to fight'
/ber + ubah/	[be.ru.bah]	'changing'
/ber + ikot/	[be.rikot]	'the following'

(25) Compounds

/di + ambel/ plus /aleh/	[diʔambelʔaleh]	('to take over')
/di + anak/ plus /əmas + kan/	[diʔanãʔʔəmãskan]	('to make a favourite of)[26]

In (24a), prefix-final /ŋ/ and floating /r/ undergo deletion, reflecting a process of degemination. Additionally, this is accompanied by Compensatory Lengthening of the preceding vowel[27]. However, in (24b), /ŋ/ and /r/ *resyllabify* with the following vowel-initial roots. In (25), neither the /l/ of /ambel/ nor the /k/ of

[25] Some informants varied between prefix-final [r] in the onset position and prefix-final [r] syllabified as a coda consonant, followed by a glottal stop in the onset position. The latter phenomenon is a parallel of 'liaison non-enchaînée' discussed by Encrevé (1988) in his description of Standard French.

[26] The /k/ of /anak/ does not undergo Gem; the glottal stop in [anaʔ] arises from the application of /k/ Glottalling, rather than GSI .

[27] See Chapter 5 for Vowel Lengthening.

/anak/ undergo Gem, instead the empty onsets of vowel-initial roots are filled by a default glottal stop (See Chapter 5 for further discussion).

These data make it clear that, under the Standard View, the domain of application of Gem is the PW, defined as a root (plus suffix, if there is one).

3.4 Summary

We have claimed that a glide is a high vowel in a non-nucleus peak position. We also claim that glides in SM are both of an underlying and derived nature (are both predictable and non-predictable). GF in SM is a hiatus-avoidance phenomenon which creates an onset segment between what would otherwise be a sequence of two vocalic nuclei instead of a reduction in the number of syllables in a word as is possible in French (Hannahs *ms*). Gem is also a hiatus-avoidance strategy.

We have also observed that floating /r/ in the root final position does not geminate, as do other consonant-final roots when they are in combination with vowel-initial suffixes; instead it undergoes syllabification into a following empty onset position.

Under the Standard View, there are two phonological processes operating within the PW, thus defined: GF and Gem. We are about to question that view. /V + V/ sequences across a prefix-root boundary undergo GSI even where they contain a

high vowel followed by another vowel sequence. We also claimed that GSI in SM is a postlexical rule.

CHAPTER 4

THE STANDARD VIEW: COUNTER-EVIDENCE

4.0 Introduction

In this chapter, the processes of Nasal Obstruent Assimilation (NOA) and Nasalisation (Nas) in SM will be examined. The focus here will be on NOA, and we will be concerned mainly with discussion of the phonological process involving prefix-final velar nasals plus root-initial obstruents and root-final nasal stop plus suffix-initial obstruents. As noted by Farid (1980: 13) and Teoh (1994: 101), a prefix-final nasal is always homorganic with a following underlying root-initial obstruent. However, this process does not hold at the root + suffix boundary.

Nas operates left-to-right (progressive nasalisation), spreading nasality from nasal stops onto all following segments within roots and across word-internal morphological boundaries (i.e. prefix and suffix boundaries) but not across an MW boundary, as defined above (it is blocked by an obstruent with oral cavity constriction).

Section 4.1 discusses Nas in SM, focusing on root-internal morphological boundaries and morphological word boundaries. Section 4.2 discusses NOA, and section 4.3, provides a summary.

4.1 Nasalisation

All vowels in SM (i, e, a, o, u, ə) can be nasalised but do not contrast phonemically for nasality: there are no underlyingly nasalised vowels in SM.

4.1.1. *Nasalisation in roots and across morphological boundaries*

Despite there being no underlyingly nasalised vowels in SM, the language has a rule of Nasalisation that operates from left-to-right (progressive nasalisation), spreading nasality from a nasal stop onto all following segments within roots and across morphological boundaries (Teoh : 39 and Zaharani 1998: 192), subject to blocking by a consonant with oral cavity specification. The segments [j],[w] fail to block the spread of nasality since they are vocalic; [h] and [ʔ] do not block the spread of nasality since they are consonants without oral cavity specification. The remaining non-nasal consonants all act as blockers: {p,b,t,d,k,g, dʒ, tʃ, s, l, r}. All vowels exhibit nasality when they are preceded by a nasal segment. We show patterns of nasalisation in (1):

(1) Root-internal nasalisation:

/masak/	[mãsaʔ]	'cook'
/makan/	[mãkan]	'eat'
/nampak/	[nãmpaʔ]	'see'
/nanti/	[nãnti]	'wait'

/taman/	[tamãn]	'garden'
/tanam/	[tanãm]	'to plant'
/ŋaŋa/	[ŋãŋə̃]	'agape'
/taɲa/	[taɲə̃]	'ask'

(2) Nasalisation across a root-plus-suffix boudary:

/tanam + an/	[tanãmmãn]	'plants'
/makan+an/	[mãkannãn]	'cooking'
/kə+təraŋ+an/	[kətəraŋŋãn]	'statement'
/təman+i/	[təmãnnĩ]	'accompany'
/ulaŋ+i/	[ulaŋŋĩ]	'repeat'
/faham+i/	[pahammĩ]	'understand'

(3) Nasalisation across prefix-plus-root boundary:

/məŋ-asoh/	[mãŋãsoh]	'to educate'
/məŋ-elak/	[mãŋẽlaʔ]	'to avoid'
/məŋ-ikat/	[mãŋĩkat]	'to tie'
/məŋ-upah/	[mãŋũpah]	'to hire'
/məŋ-ia-kan/	[mãŋĩjãkan]	'to affirm'
/pəŋ-asoh/	[pəŋãsoh]	'maid'
/pəŋ-ikat/	[pəŋĩkat]	'one who ties'
/pəŋ-upah/	[pəŋũpah]	'payer'

As shown above, Nas in SM operates root-internally and across word-internal morphological boundaries (i.e. root-suffix boundary and prefix-root boundary) such as in (1, 2 & 3). These data show that Nas is a lexical process which

not only operates within the domain of the PW as defined under the standard view but also across a prefix-root boundary.

Examples of the blocking of Nas are set out in (4) below.

(4)

a. root-internally

/mati/	[mãti]	'die'
/malam/	[mãlam]	'night'
/nanti/	[nãnti]	'wait'

b. root plus suffix

/tanam+kan]	[tanãmkan]	'to plant'
/tərаŋ+kan/	[tərаŋkan]	'to explain'
/simpan+kan/	[simpankan]	'to keep'

c. prefix plus root

/məŋ-bəri/	[mãmbəri]	'to give'
/məŋ-dapat/	[mãndapat']	'to get'
/məŋ-gɔsok/	[mãŋgɔsok	'to rub'
/məŋ-dʒadi/	[mãɲdʒadi]	'to become'
/məŋ-lawat/	[mã:lawat]	'to visit'
/məŋ-rompak/	[mã:rompaʔ]	'to rob'

Examples showing non-blocking by of Nasalisation by glides and glottal segments are shown in (5).

(5)

(a) root-internally

/majaŋ/	[mãj̃ãŋ]	'stalk'
/mahal/	[mãh̃ãl]	'expensive'
/mewah/	[mẽw̃ãh]	'luxury'
/maaf/	[mãʔãp]	'forgive'

(b) root plus suffix

/kə+səni+an/	[kəsən̪ĩjãn]	'art'
/pər+təmu+an/	[pə:təmũw̃ãn]	'meeting'
/kə+mahu+an/	[kəmãh̃ũw̃ãn]	'wish'

4.1.2 *Nasalisation and morphological word boundaries*

In SM, nasality is blocked at a MW boundary even if there is no opaque consonant there to act as a blocker. This is illustrated in (6) – (8).

(6) Nas blocked across MW boundary in compounds

(a) /makan aŋin/	[mãkanʔaŋ̃in]	'vacation'
(b) /sama ada/	[samə̃ʔadə]	'either'

(7) Nas blocked across MW boundary in reduplicated words

(a) /anai-anai/	[anãĩʔanãĩ]	'termites'
(b) /uarna-uarni/	[warnãwarn̪ĩ]	'colourful'

(8) Nas blocked across MW boundary in Syntactic phrases

(a) /Sudah lama awak di sini?/ [sudahlamə̃ʔawaʔdisiñi?]

'Have you been long here?

(b) /kita sama sama jakin/ [kitəsamãsamãjakin]

'we are mutually convinced'

Clearly, Nasalisation operates across a prefix-root boundary. If the PW in SM is a root + suffix, (as claimed under the standard view) then, in a word containing a prefix, a root and a suffix, one would be obliged to postulate two Nas rules (or two applications of the same rule): a juncture rule when nasalisation operates over a prefix-root boundary and a domain-internal, non-junctural rule when it operates across a root + suffix boundary. Such an analysis implausibly suggests that there are two nasalisation generalisations, or two distinct applications of the same rule. It seems to us that there is only one generalisation here, and that its domain is the MW. The evidence from the application of Nas undermines the standard view.

4.2 Nasal Obstruent Assimilation

As previously mentioned, there are two prefixes in SM which end with a velar nasal: /məŋ-/ and /pəŋ-/. The prefix /məŋ-/ signals the active voice in verbs, while the prefix /pəŋ-/ signals the derived nominal. These velar nasals undergo assimilation when they combine with underlying root-initial obstruents (whether

voiced or voiceless). This process is referred as Nasal Obstruent Assimilation (henceforth NOA). The following examples illustrate the NOA process.

(9).

(a)	/məŋ-pudʒi/	[mə̃mũdʒi]	'to praise'
	/məŋ-taɲkap/	[mə̃nãɲkap]	'to catch'
	/məŋ-kojak/	[mə̃ŋõjã?]	'to tear'
	/məŋ-səru/	[mə̃ɲə̃ru]	'to summon'
(b)	/pəŋ-putar/	[pəmũta:]	'roller'
	/pəŋ-tari/	[pənãri]	'dancer'
	/pəŋ-karaŋ/	[pəŋãraŋ]	'author'
	/pəŋ-sapu/	[pəɲãpu]	'sweeper'

(10)

(a)	/məŋ-bəri/	[mə̃mbəri]	'to give'
	/məŋ-dapat/	[mə̃ndapat']	'to get'
	/məŋ-gɔsok/	[mə̃ŋgɔsok	'to rub'
	/məŋ-dʒadi/	[mə̃ɲdʒadi]	'to become'
(b).	/pəŋ-basoh/	[pembasoh]	'cleaner'
	/pəŋ-dapat/	[pendapat']	'idea'
	/pəŋ-gali/	[pəŋgali]	'digger'
	/pəŋ-dʒual/	[pəɲdʒuwal]	'seller'

In (9) and (10), NOA manifests itself in two phonological contexts: the prefix-final velar nasal plus a root-initial voiceless obstruent as in (9), and the prefix-final velar nasal plus a root-initial *voiced* obstruent as in (10). In (9), NOA applies, and is followed by a rule of Voiceless Obstruent Deletion in N + Obs clusters where the obstruent is voiceless. One way of looking at this is to say that NC̥ clusters are dispreferred, and that two strategies may be adopted to avoid them: voice the obstruent or delete it. SM opts for the latter, but only across prefix – root boundaries. This process is common to many Western Austronesian languages (Dempwolff 1934-1938) and the Philippine languages; it may be reconstructable as far back as Proto-Malayo-Polynesian (Kroeger 1988:5) and also accurs in some African languages (Rosenthall 1989: 50).

In (10) NOA occurs on account of a prefix-final velar nasal assimilating with a root-initial voiced obstruent, but here, no deletion takes place. Previous studies distinguish the two processes (see Kroeger 1988 and Teoh 1994), but here, since both involve assimilation of nasal plus obstruent, we assume that there is a single generalisation. Crucially for our argument, NOA does not occur at a root plus suffix boundary where a root-final nasal attaches to a suffix-initial obstruent (i.e. the suffix /kan/). Nor does VOD:

(11)

a. /tanam + kan/	[tanãmkan]	*[tanãŋãn] / [tanãŋkan]	'to plant'
b. /pindʒam + kan/	[pindʒamkan]	*[pindʒaŋãn]/[[pindʒaŋkan]	'to lend'
c. /simpan + kan/	[simpankan]	*[simpaŋãn] / *[simpaŋkan]	'to keep'
d. /saran + kan/	[sarankan]	*[saraŋãn] / *[saraŋkan]	'to suggest'

e. /təraŋ + kan/	[təraŋkan][28]	*[təraŋãn]	'to explain'
f. /sajaŋ + kan/	[sajaŋkan]	*[sajaŋãn]	'to love'

Note too that VOD fails to apply root-internally, as shown by forms such as [sampan] 'small boat', [tandə] 'sign', and [saŋkot'] 'hang'.

Finally, neither NOA nor VOD applies at an MW boundary:

(12) NOA and VOD blocked at MW boundary

Compounds:

a. /dʒam/ 'watch' /taŋan/ 'hand'	→ [dʒamtaŋãn]	'watch'
/dʒam/ 'watch' /taŋan/ 'hand'	→ *[dʒantaŋãn]	
b. /taman/ 'garden' /buŋa/ 'flower'	→ [tamãnbuŋə̃]	'flower garden'
/taman/ 'garden' /buŋa/ 'flower'	→ *[tamãnbuŋə̃]	

Reduplication:

/lintaŋ-pukaŋ/	→ [lintaŋpukaŋ]	'messy'
/lintaŋ-pukaŋ/	→ *[lintampukaŋ]	

Syntactic phrase:

a. /dalam/ 'in/inside' /guni/ 'sack'....	[dalamguñi]	'in sack....'
/dalam/ 'in/inside' /guni/ 'sack'....	*[dalaŋguñi]	

[28] The surface form of velar nasal [ŋ] is an output from underlying /ŋ/ and not via NOA.

b. /kəsan/ 'impression' /burok/ 'bad' [kəsanburok....] 'bad impression...'

/kəsan/ 'impression' /burok/ 'bad' *[kəsamburok....]

The domains of NOA and VOD are the mirror image of the domain of application of GF and Gem: they operate only across a prefix-root boundary and never root-internally or across a root + suffix boudary. What NOA, VOD, Gem, GF and Nas share is that they are blocked at an MW boundary.

One could claim, under the standard view, that prefixes constitute a PW in SM, and that VOD and NOA are juncture rules operating across the boundary between two PWs. But the standard view runs into serious difficulties with respect to NOA for the following reason. One might argue that Gem and GF apply within the PW as defined under the standard view, and that NOA is a juncture rule operating across the boundary between two PWs. The problem is that one could equally well define the PW in SM as a prefix + root, argue that suffixes constitute PWs, and take NOA to operate within the PW and argue that Gem and GF are juncture rules operating across the boundary between two PWs. We see no non-arbitrary way of choosing between these two approaches. Additionally, either choice forces us to claim that Nas operates both as a juncture rule and within a PW. We therefore suggest that both alternatives should be rejected, and that, instead, we take the PW in SM to be isomorphic with the MW.

4.3 Summary

To summarise, Nasalisation operates left-to-right and penetrates all glottal segments and glides [h, ?, w and j] either root-internally, at a suffix boundary or prefix boundary. It is however blocked at MW boundaries. We claim that Nasalisation in SM is a lexical process that appears root-internally, at suffix and prefix boundaries: its domain is the MW.

It is difficult to account for the application of NOA if one defines the PW in SM as a root (plus any suffixes), since NOA is blocked within a PW, thus defined, but operates across a PW boundary, according to the standard view. One would expect such phonetic assimilation, where it is blocked in the domain of a PW to be possible where no such boundary exists. Moreover, NOA can be seen as a phenomena that is the mirror-image of the application of GF and Gem. That is, a prefix-final nasal will assimilate in place of articulation to a root-initial obstruent, but a root-final nasal will not assimilate to a suffix-initial obstruent, or to an obstruent at the beginning of a following morphological word i.e. a root-final nasal plus root-initial consonant.

On the basis of this evidence, we claim that the PW in SM is isomorphic with the MW, and that the standard view is untenable. This claim leaves us with the task of showing why some processes are blocked at a prefix-root boundary, but not at a root-suffix boundary, while others apply in the mirror-image of those contexts, and while yet others apply across both sorts of boundary.

We provide a new analysis in the chapters which follow, an analysis that distinguishes edge-based and non-edge-based processes, and right-edge from left-edge processes.

CHAPTER 5

AN ALTERNATIVE ANALYSIS

5.1 Edge-based and non-edged based lexical generalisations

In Chapter 3 and 4, it was shown that the standard view concerning the Phonological Word (PW) in SM is falsified since an equally regular phonological generalisation, Nasal-Obstruent Assimilation (NOA) holds across a prefix-root boundary, but is blocked across a root-suffix boundary. The same is true of VOD. Additionally, the regular processes of Nasalisation and Resyllabification operate across both a prefix-root boundary and a root-suffix boundary. We claimed that the PW is isomorphic with the morphological word in SM, and that all of the following processes have the PW, thus defined, as their domain: GF, Gem, NOA, VOD, Nasalisation, and Resyllabification.

This analysis requires us to give an alternative solution to account for the asymmetrical behaviour of GF and Gem, on the one hand, and NOA, and VOD, on the other. We claim that the lexical phonology of SM can be subdivided into left-edge-based, right-edge-based and non-edge-based generalisations. By 'edge' here, we mean edge of a PW which, for us, is the same thing as the MW. The Edge-based generalisations are subject to a locality constraint such that they cannot operate across more than one morphological boundary from the relevant edge. We claim that GF and Gem are right-edge processes, whereas NOA and VOD are left-edge processes. We claim that the following generalisations, all of which have the PW

(=MW) as their domain, are non-edge-based processes: /k/ Glottalling, Nasalisation, Resyllabification, Compensatory Lengthening and /a/ Reduction. These lexical generalisations apply, we claim, after the edge-based generalisations. Each of these is considered in turn below.

5.2 Non-Edge-Based Processes

5.2.1 */k/ Glottalling*

In SM, all underlying coda /k/s are realised as a glottal stop, although the generalisation makes no reference to edges (the generalisation is thus not related to alignment phenomena, *pace* Zaharani 1998). However, Gemination applies when coda /k/s are suffixed with /-an/ and /-i/. If the coda /k/ is suffixed with onset /k/ of /-kan/, the sequences of /k/s which result are *false geminates* at a certain stage in the derivation (see Section 3.3). Examples of this phenomenon are:

(1)

a.	i.	/masak/	→	[mãsaʔ]	'cook'
	ii.	/masak + an/	→	[mãsaʔkan]	'cooking'
	iii.	/masak + kan	→	[mãsaʔkan]	'to cook'
b.	i.	/masok/	→	[mãsoʔ]	'to enter'
	ii.	/mãŋ + masok + i/	→	[mãmãsoʔki]	'entering'
	iii.	/masok + kan/	→	[mãsoʔkan]	'to put in/enter'

c. /di + anak/ + /əmas + kan/ → [diʔanãʔʔəmãskan] 'to make a favourite of'[29]

In figure (1a-iii, 1b-iii & 1c), the /k/s of /-kan/ are in onset position and do not undergo /k/ Glottalling on account of the segments singly attached to the different X-slot as fomalised in Section 3.3 – (19b). However, in figures (1a-i, 1a-ii, 1b-i & 1b-ii), all the underlying /k/s in root-final position undergo glottalling [ʔ][30] due to the fact that they are multiply attached association lines linked to the skeletal X as formalised in Section 3.3 – (19a). In these cases, it is necessary to state that Gem applies prior to /k/ Glottalling. From our analysis, this ordering follows from the fact that Gem is an edge-based generalisation and thus precedes /k/ Glottalling, which is not an edge-based generalisation[31].

5.2.2 *Nasalisation*

Nasalisation operates from left-to-right (i.e. it is progressive nasalisation), spreading nasality from a nasal stop onto all following segments within roots and across word-internal morphological boundaries in SM (subject to blocking, as

[29] The /k/ of /anak/ does not undergo Gem; the glottal stop in [anaʔ] arises from the application of /k/ Glottalling, rather than GSI: see below on the lexical status of /k/ Glottalling vs the postlexical status of GSI. *The second of the two [ʔ]s in the sequence is not derived from /k/ but from GSI occurring by default.*

[30] In some dialects of SM such as Kelantan and Trengganu, Glottalisation applies not only to velar stops, but also to all other voiceless root-final obstruents (see Teoh 1994: 65).

[31] For further discussion of /k/ Glottalling, from view of constraint interaction, see Chapter 8.

described above); we claim that it is non-edge-based process: its domain is the PW (as MW), but it penetrates both prefix-root and root-suffix boundaries.

5.2.3 *Resyllabification*

Like Nasalisation, Resyllabification in SM is also a lexical generalisation (i.e. holds within the PW as MW), which holds after the assignment of word stress and related edge-based processes. This is shown by the fact that it is blocked by the presence of a PW boundary. This can be demonstrated most clearly by the behaviour of floating (or 'linking') /ɾ/, which occurs both prefix-finally and root-finally in SM. Before we discuss the Resyllabification process in SM, it is important to consider the behaviour of /ɾ/ syllabification in both roots and prefixes.

SM is non-rhotic in the sense that it has underlying floating /ɾ/ which is realised only if an empty onset follows, in which case, it is syllabified into that onset position:

(2) Non-Rhoticity in SM

(a) Root-final floating /ɾ/ is not realised if only coda position is available:

/tuka(ɾ)/	[tuka:]	'change'
/baka(ɾ)/	[baka:]	'burn'
/sama(ɾ)/	[samã:]	'blur'
/koto(ɾ)/	[koto:]	'dirty'
/tuka(ɾ) + kan/	[tuka:kan]	'to change'

/tabo(r) + kan/	[tabo:kan]	'to spread'

(b) Root-final floating /r/ is realised if an empty onset position is available:

/uke(r) + an/	[u.ke.ran]	'carving'
/pasa(r) + an/	[pa.sa.ran]	'market'.
/koto(r) + an/	[ko.to.ran][32]	'dirt'
/mãŋ/ + /səda(r) + i/	[mã.ɲã.da.ri]	'realised'
/mãŋ/ + /gəma(r) + i/	[mãŋ.gə.mã.ri]	'to like'

(c) Prefix-final floating /r/ is not realised if only coda position is available:

/bə(r)/ + /maen/	[bə:mãẽn]	'to play'
/pə(r) + /kata/ + /an/	[pə:kataʔan][33]	'word
/tə(r)/ + pakai/	[tə:pakai]	'used'

(d) Prefix-final floating /r/ is realised if an empty onset is available:

/bə(r)/ + /aŋkat/	[bə.raŋ.kat]	'to depart'
/pə(r) + /aku/	[pə.ra.ku]	'to admit'
/tə(r) + ikut/	[tə.ri.kut˺]	'followed'

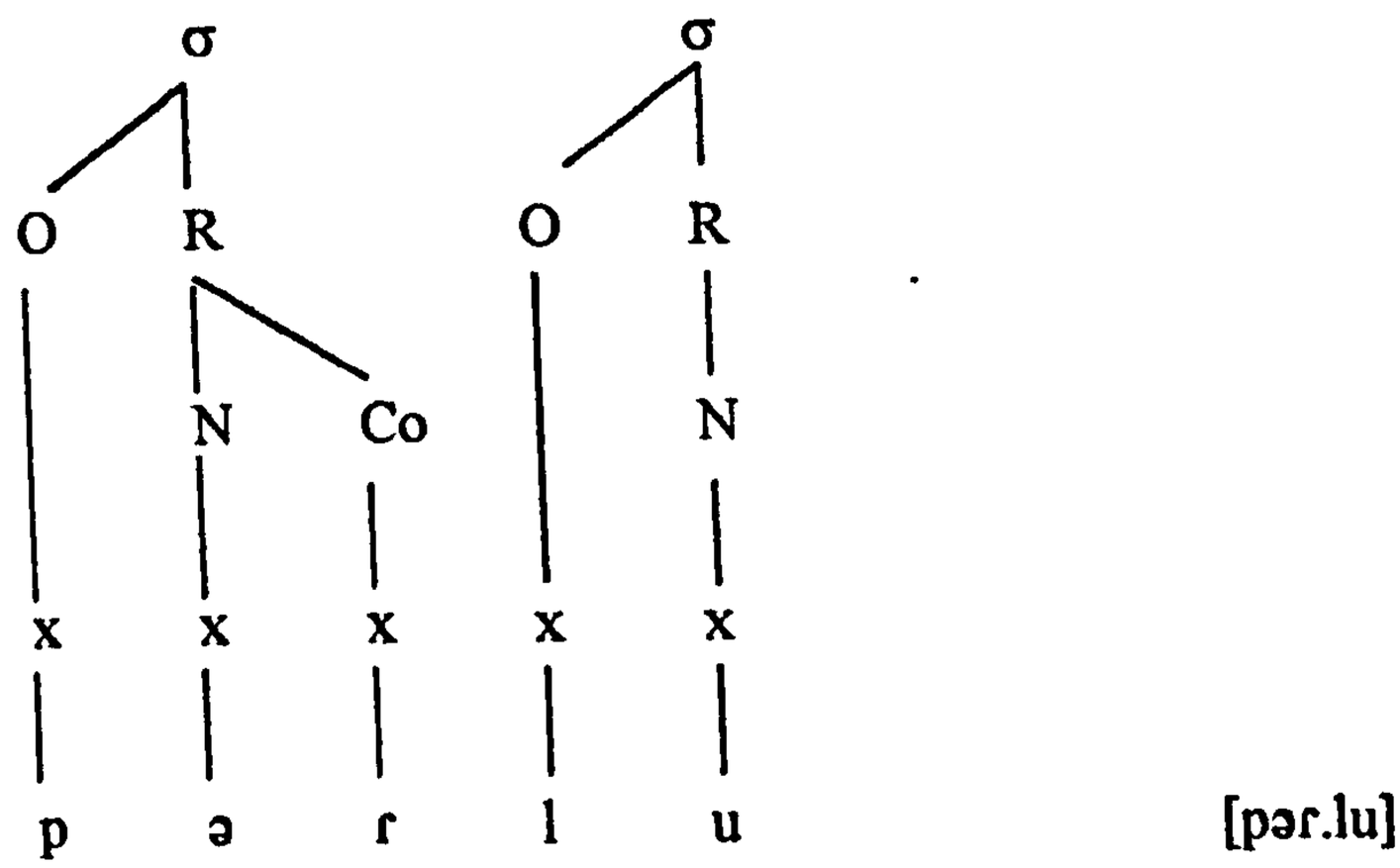
[32] It is unclear exactly what a geminated tap of the sort Teoh and Zaharani claim to observe might be: a [d], or a two-tap trill? Neither of these occurred in the speech of any of our informants.

[33] Our data show intra-speaker variation, in which floating prefix-final /r/, when followed by a V-initial root, may either be resyllabified, (taken to be the norm), or realised in coda position of the prefix, with a glottal stop preceding the root-initial vowel. This latter phenomenon is parallel to that of liaison non-enchaînée, as discussed by Encrevé (1988). This is discussed in the section on the effect of GSI on word stress patterns. There is also intra-speaker variation with respect to prefix-final floating /r/ when followed by a consonant-initial root. We take this to be due to the influence of Literary Standard Malay (see following footnote).

Root-internally, /r/ is realised both intervocalically and preconsonantally (in root-internal coda position), as in /tari/ [tari] ('dance') and /pərlu/ [pərlu] ('need').

The SM pattern of non-rhoticity is distinct from that found in non-rhotic accents of English, where /r/ may never appear in coda position. It is nonetheless a form of non-rhoticity, clearly. We suggest that the /r/ is syllabified underlyingly as a coda consonant in the cases of /pərlu/ and as an onset consonant in the case of /tari/. That is, /r/ is fixed root-internally, but floating in both root-final and prefix-final positions; a floating /r/ surfaces only if a following empty onset is available. This is justified since there are clear alternations available in the case of floating /r/, but no alternations at all in the case of fixed /r/. In taking this view, we are following the treatment of fixed and floating nasal stops in French, as postulated by Encrevé (1988). It also parallels Harris's (1994) treatment of non-rhoticiity in English, but we note that there is no equivalent of 'intrusive r' in SM, and that its presence in non-rhotic accents of English perhaps suggests that a 'floating r' analysis is questionable for those accents. Syllabified underlying /r/ as a coda consonant can be depicted as in (3).

(3) /r/ is syllabified underlyingly as a coda consonant



Resyllabification across morphological boundaries

In SM, when C-final prefixes[34] concatenate with V-initial roots, the empty onset position in the root is filled by the underlyingly preceding consonant: the prefix-final C becomes an onset to the second via resyllabification (Abdullah 1974, Farid 1980, Teoh 1994 and Zaharani 1998). We note in passing that these generalisations are true of SM, but not Literary Standard Malay[35]. Resyllabification reflects the operation of the Minimal Onset Satisfaction Principle (Roca 1994), assumed to be universal.

---

[34] In SM, only the consonants /r/ and /ŋ/ can be underlying in the final position of prefixes. The final /r/ and /ŋ/ in the prefixes /tər-/, /bər-/, /pəŋ-/ and /məŋ-/ all undergo the Resyllabification process. The final /r/ in the prefixes is referred to as 'floating /r/'.

[35] In Literary Standard Malay, prefix-final /r/ is fixed, rather than floating. It is thus realised both before a vowel-initial root and a C-initial root, as in /pər + atur + an/ → [pəraturan] (/pər + atur + an/ → [pəraturran]) /pər + main/ → [pərmãin] (/pər + main/ → [pərmãin]). This is a spelling pronunciation arising from one of the prescriptions of LSM, that 'all letters must be pronounced'. We observed that /r/ in LSM is realised as trill /r/ (See Ismail 1996).

According to Teoh (1994:11) and Zaharani (1998), Resyllabification only applies at a prefix boundary, and not at a suffix boundary (i.e. when C-final roots combine with V-initial suffixes). Instead, all C-final roots undergo Gemination, for example in the words /lətop + an/ [lətopˀ.pan] 'explosion', /sambut + an/ [sambutˀ.tan] 'reception' and /ukeɾ + an/ [ukeɾɾan] 'carving'. According to Zaharani and Teoh, Gem in SM applies when a root-final floating /ɾ/ combines with a V-initial suffix, as found in the putative pronunciation [ukeɾɾan]. However, none of our informant uttered geminates in such words; rather, the /ɾ/ was resyllabified. This shows that Resyllabification operates across a prefix-root and a root-suffix boubdary: it is a non-edge-based rule. Note that Resyllabification in SM not only applies to floating /ɾ/, but also occurs when a nasal prefix-final /ŋ/ combines with a V-initial stem. Examples of the application of Resyllabification are given in (4) and (5).

(4) Resyllabification across prefix-root boundary

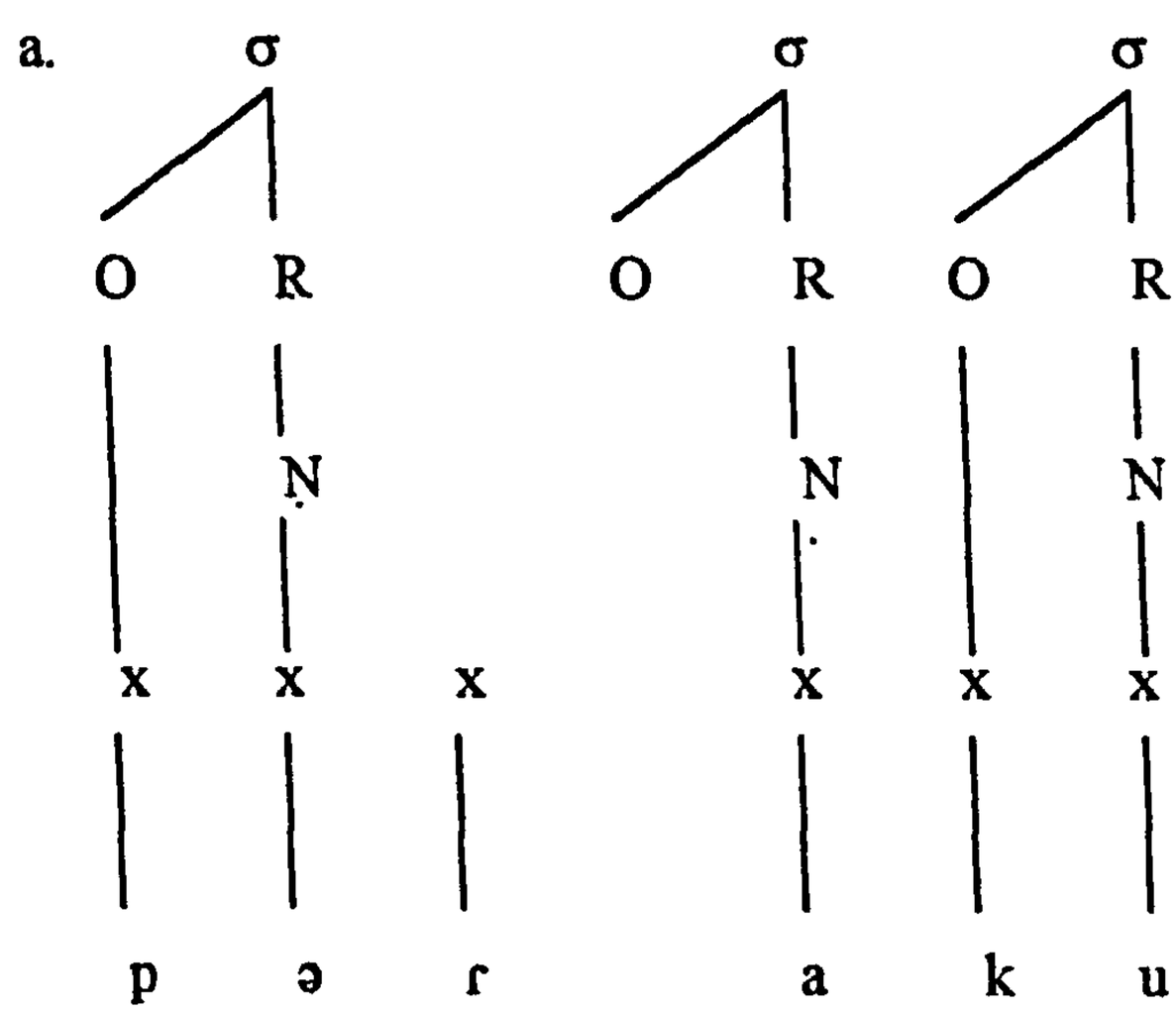
/bə(ɾ)/ + /aŋkat/	[bə.ɾaŋ.kat]	'to depart'
/pə(ɾ) + /aku/	[pə.ɾa.ku]	'to admit'
/tə(ɾ) + ikut/	[tə.ɾi.kutˀ]	'followed'
/mãŋ/ + /aŋkat/	[mã.ŋãŋ.kat]	'to lift'
/mãŋ + elak/	[mã.ŋẽ.laʔ]	'to avoid'.
/pəŋ/ + /awas/	[pə.ŋã.w̃ãs]	'invigilator'
/pəŋ/ + /ikot/	[pə.ŋĩ.kot]	'follower'

(5) Resyllabification across root-suffix boundary

/uke(ɾ) + an/	[u.ke.ɾan]	*[u.keɾ.ɾan]	'carving'
/pasa(ɾ) + an/	[pa.sa.ɾan]	*[pa.saɾ.ɾan]	'market'.
/koto(ɾ) + an/	[ko.to.ɾan][36]	*[ko.toɾ.ɾan]	'dirt'
/mãŋ/ + /səda(ɾ) + i/	[mã.ɲã.da.ɾi]	*[mã.ɲã.daɾ.ɾi]	'realised'
/mãŋ/ + /gəma(ɾ) + i/	[mãŋ.gə.mã.ɾi]	*[mãŋ.gə.mãɾ.ɾi]	'to like'

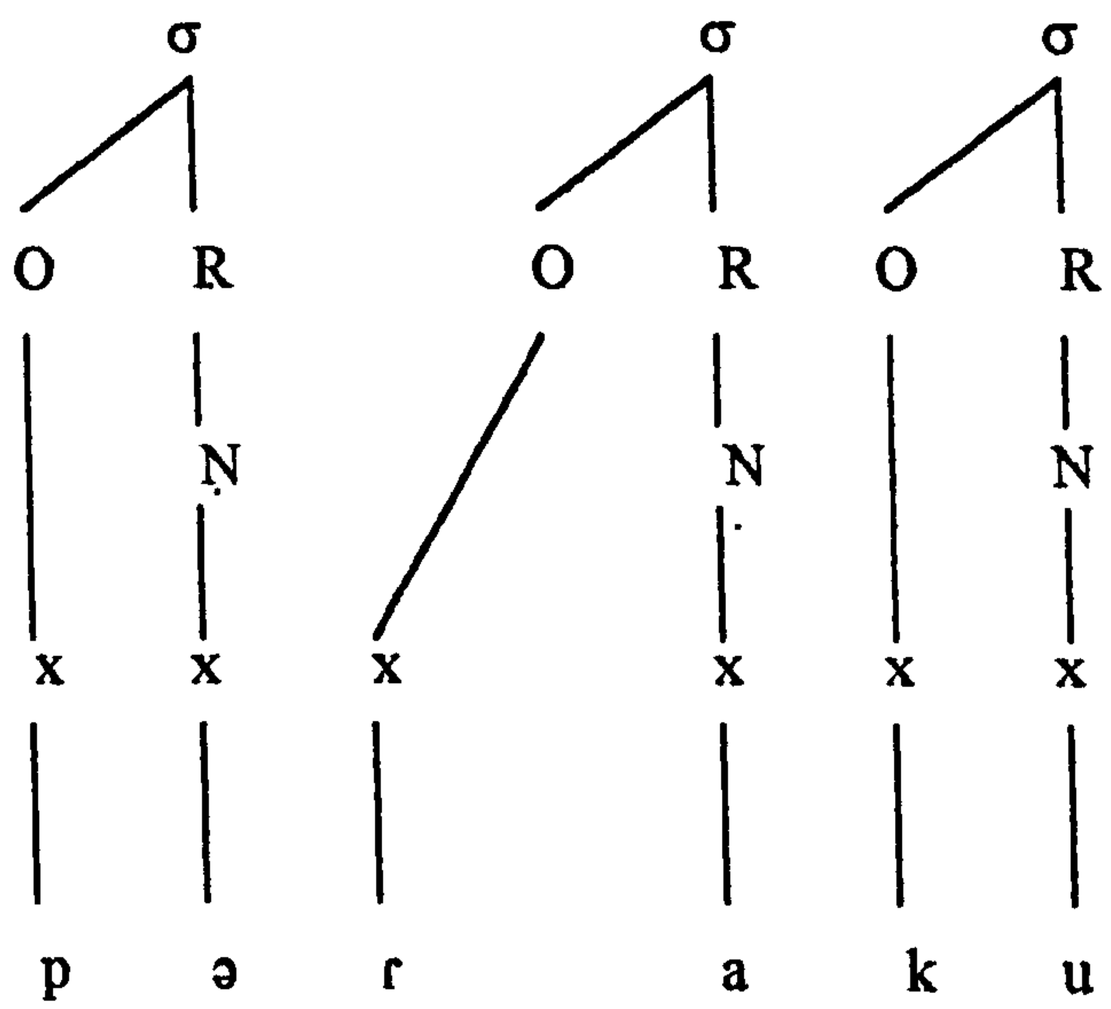
The behaviour of floating /ɾ/ is illustrated in (6) and (7).

(6). Resyllabification of prefix-floating /ɾ/



[36] It is unclear exactly what a geminated tap of the sort Teoh and Zaharani claim to observe might be: a [d], or a two-tap trill? Neither of these occurred in the speech of any of our informants.

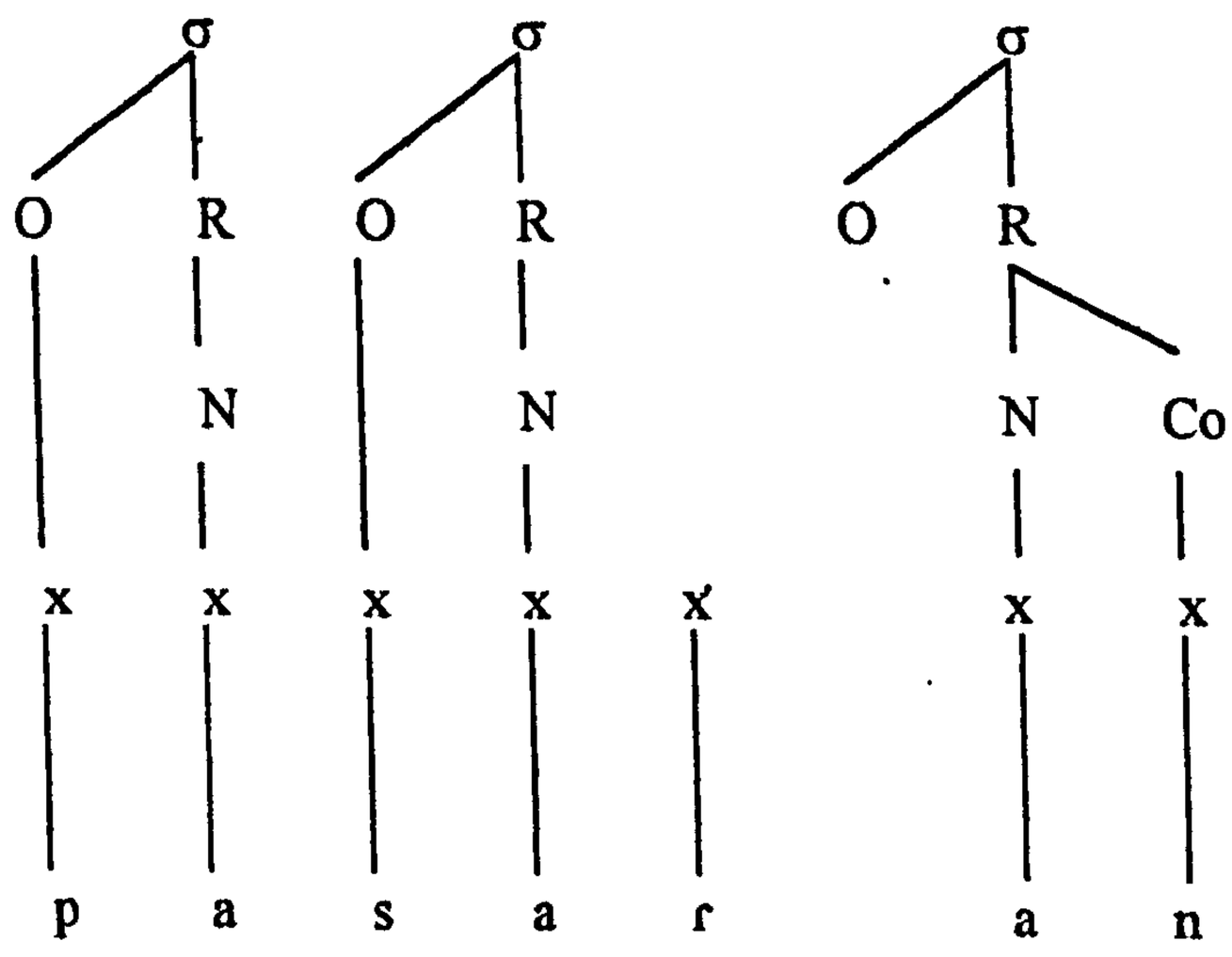
b.



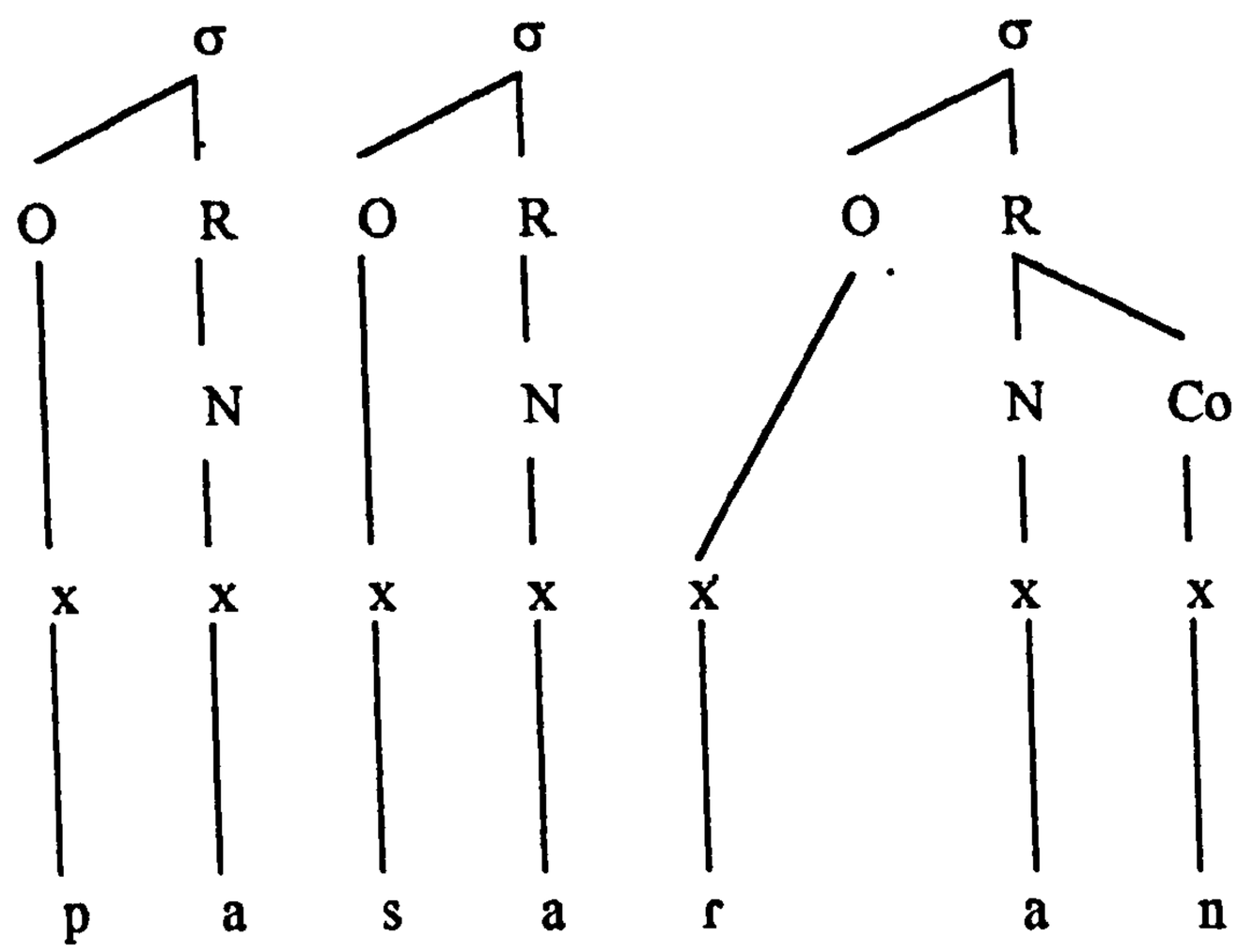
Phonetic representation [pə.ɾa.ku]

(7) Resyllabification of root-final floating /ɾ/

a.



b.



Phonetic representation [pa.sa.ran]

The behaviour of floating /r/ and prefix-final nasals serves to show that Resyllabification has the same domain of application as Nasalisation, that is, the PW (=MW): both processes operate over a prefix-plus-root and a root-plus-suffix boundary, but are blocked at a word boundary. Examples of this behaviour are:

(8) Resyllabification is blocked at MW boundaries

a. *Compound*

/ular/ + /aer/	[ula:ʔae:]	*[ularae:]
'snake' 'water'	'water snake'	

b. *Reduplicated forms*

/adʒa(r)-adʒa(r)/	[adʒa:ʔadʒa:]	'to teach repeatedly'
/aka(r)-aka(r)/	[aka:ʔaka:]	'roots'

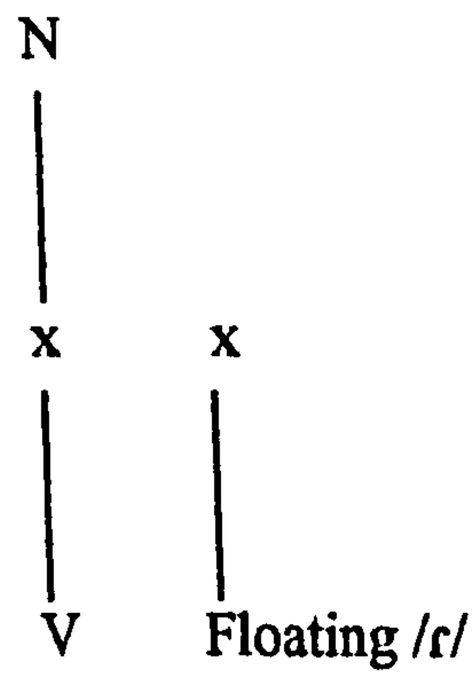
c. *Syntactic phrase*

/tolɔŋ/ +	/tukar/ +	/ikan/ +	/ini/	[tolɔŋtuka:ʔianʔiñi][37]
'help'	'change'	'fish'	'this'	'please change this fish'

5.2.4 *Compensatory lengthening*

In the case of PW boundaries (i.e. te boundaries in compounds, reduplicated forms and syntactic phrases), floating /r/ fails to resyllabify with the following V-initial word. Instead, floating /r/ remains unlinked and compensatory lengthening (CL) is triggered. CL of floating /r/ can thus be formalised as deletion as in (9) below.

(9) CL rule of floating /r/

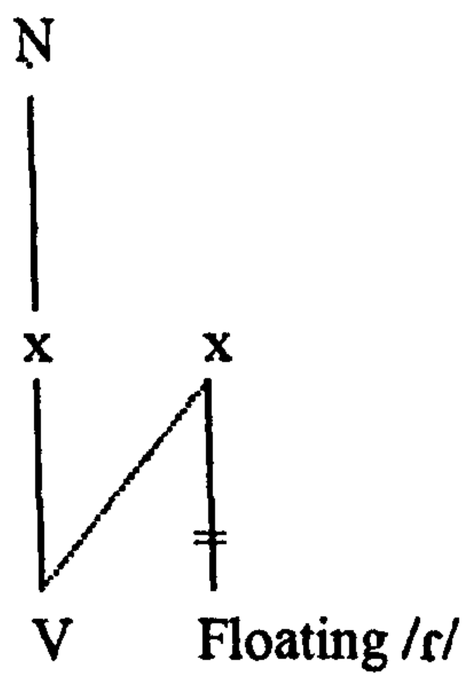


[37] Note that compensatory lengthening applies here.

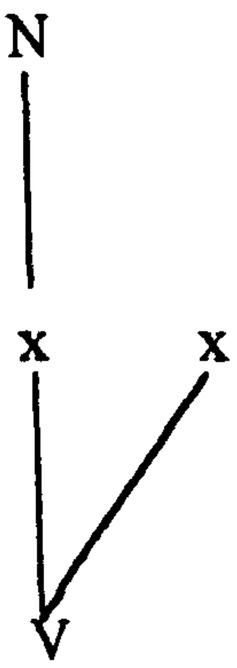
Hence, whenever /r/ is delinked, its tier docks (shown in dotted line)[38] onto the preceding vowel (i.e. in a rhyme), and this constitutes compensatory vowel lengthening as shown in (10) below.

(10) Delinking of Floating /r/ and Compensatory Vowel Lengthening

a.



b.



Phonetic representation = (V:)

CL operates at the right edge of a root and at the right edge of a prefix, as shown in (11 & 12) below.

[38] The dotted line shows that Floating /r/ leaves behind a branching rhyme when it is deleted. The x tier of the floating /r/ thus associates to a rhyme after its segment is deleted.

(11) CL PW boundary

a. *root-internally*

/tuka(r)/	[tuka:]	'change'
/baka(r)/	[baka:]	'burn'
/sama(r)/	[samã:]	'blur'
/koto(r)/	[koto:]	'dirty'

b) *PW-internal CL: root + suffix boundary*

/tuka(r) + kan/	[tuka:kan]	'to change'
/tabo(r) + kan/	[tabo:kan]	'to spread'

(12) PW-internal CL: prefix + root boundary

/bə(r)/ + /maen/	[bə:mãẽn]	'to play'
/pə(r) + /kata/ + /an/	[pə:kataʔan][39]	'word
/tə(r)/ + pakai/	[tə:pakai]	'used'

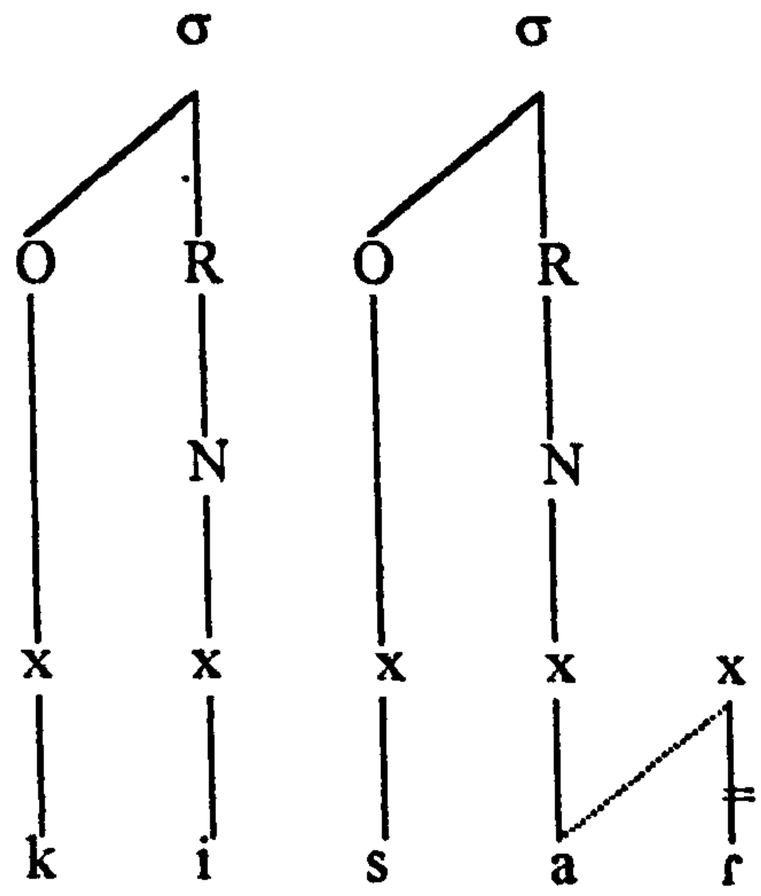
From a non-linear perspective, the occurrence of CL in (11) and (12) above is due to the delinking of a floating /r/ from the skeletal tier, since the preceding vowel associates with the following tier, simultaneously giving rise to compensatory lengthening (Harris 1994). This is shown in figure (13) below. Here, the vowel

[39] Our data show intra-speaker variation, in which floating prefix-final /r/, when followed by a V-initial root, may either be resyllabified, (taken to be the norm), or realised in coda position of the prefix, with a glottal stop preceding the root-initial vowel. This latter phenomenon is parallel to that of liaison non-enchaînée, as discussed by Encrevé (1988). This is discussed in the section on the effect of GSI on word stress patterns. There is also intra-speaker variation with respect to prefix-final floating /r/ when followed by a consonant-initial root. We take this to be due to the influence of Literary Standard Malay (see following footnote).

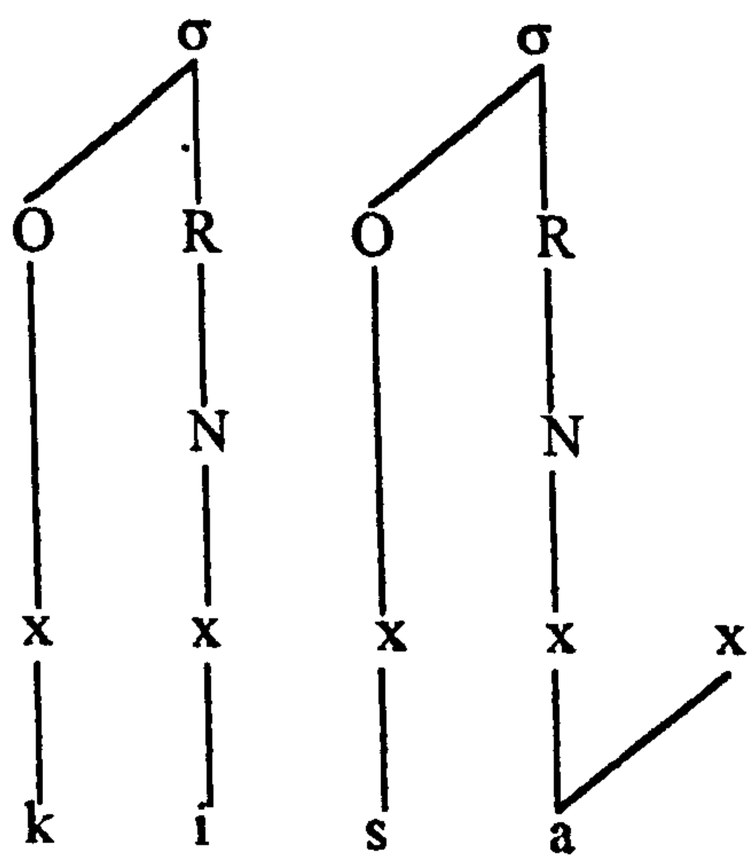
length is not phonemic but derived. In order to understand this process, consider the following tree structures.

(13) CL at PW boundary

a.



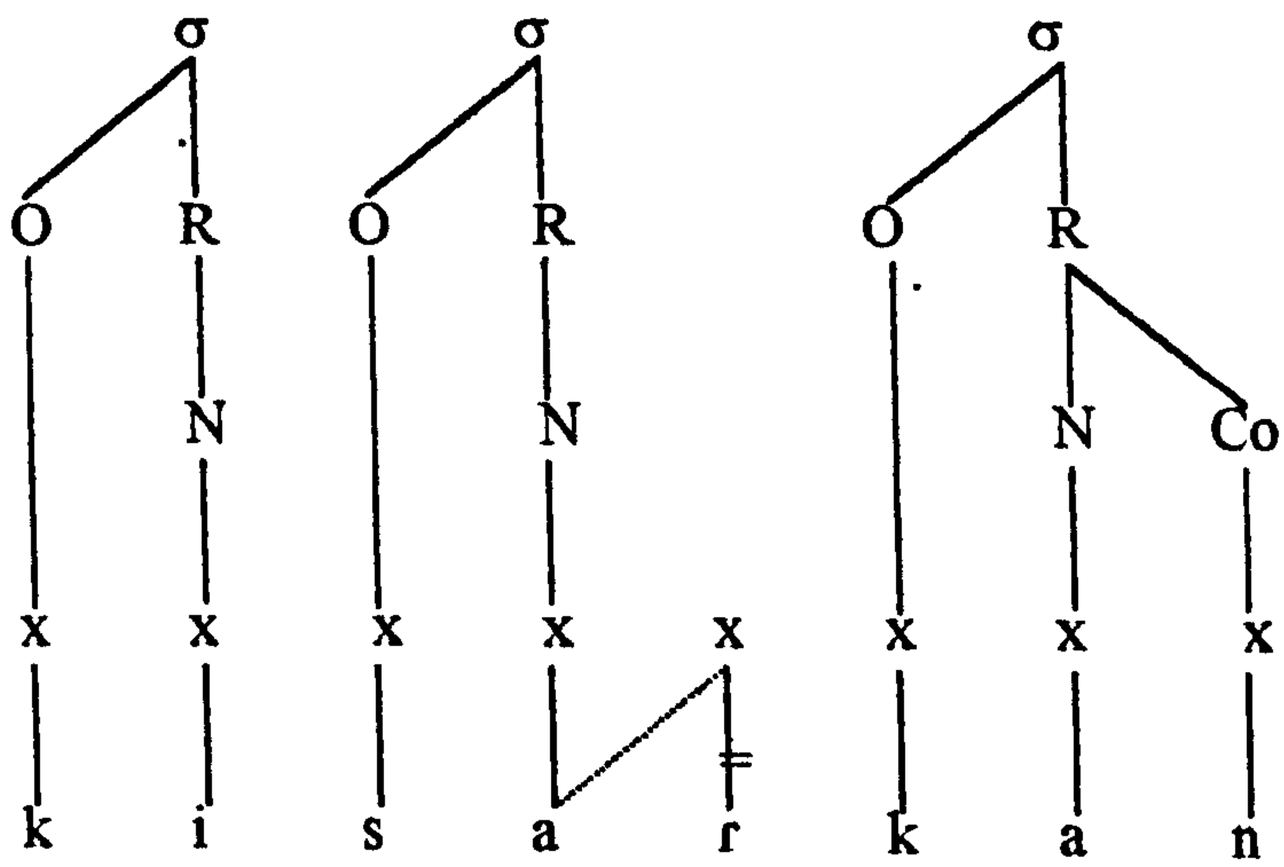
b.



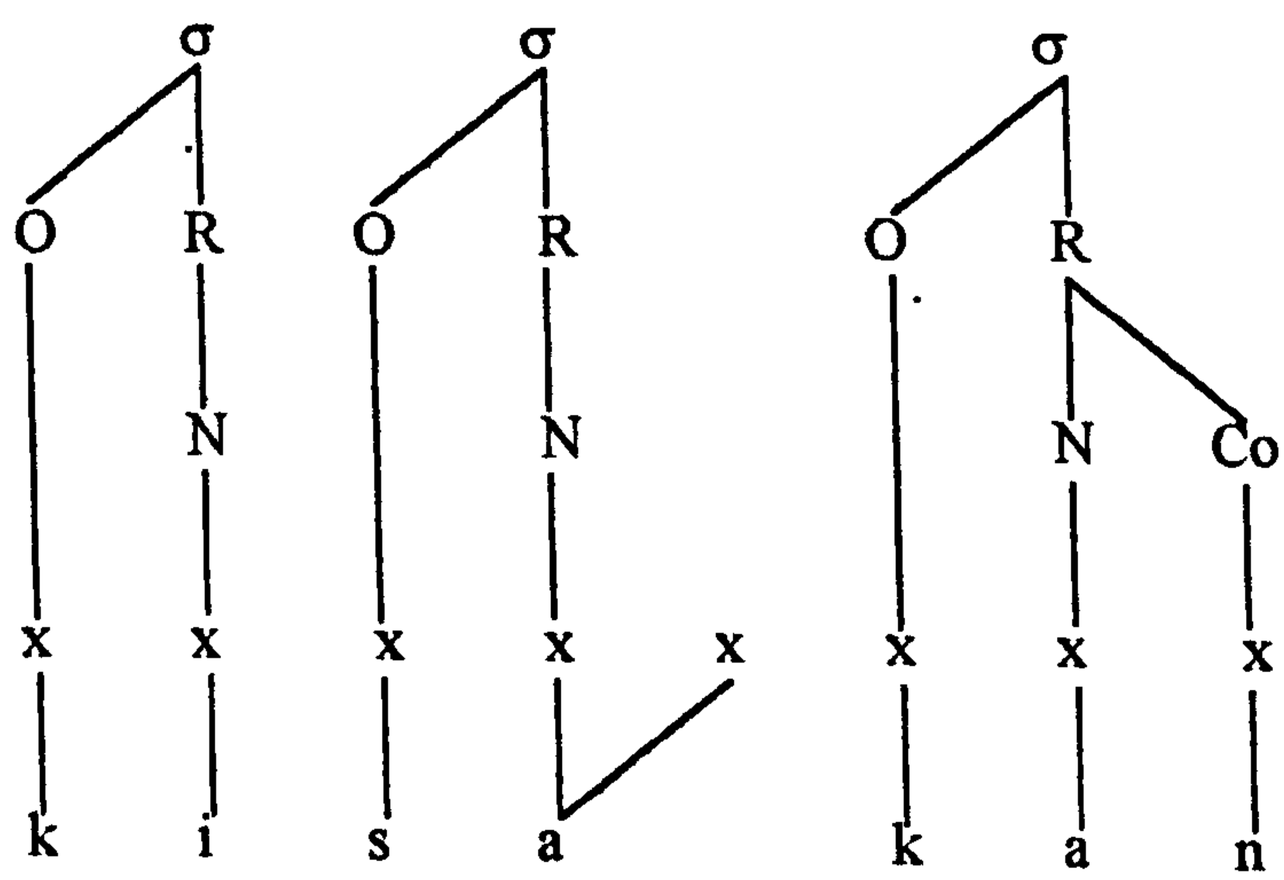
Phonetic representation = [ki.sa:]

(14) CL PW-internally: root + suffix

a.



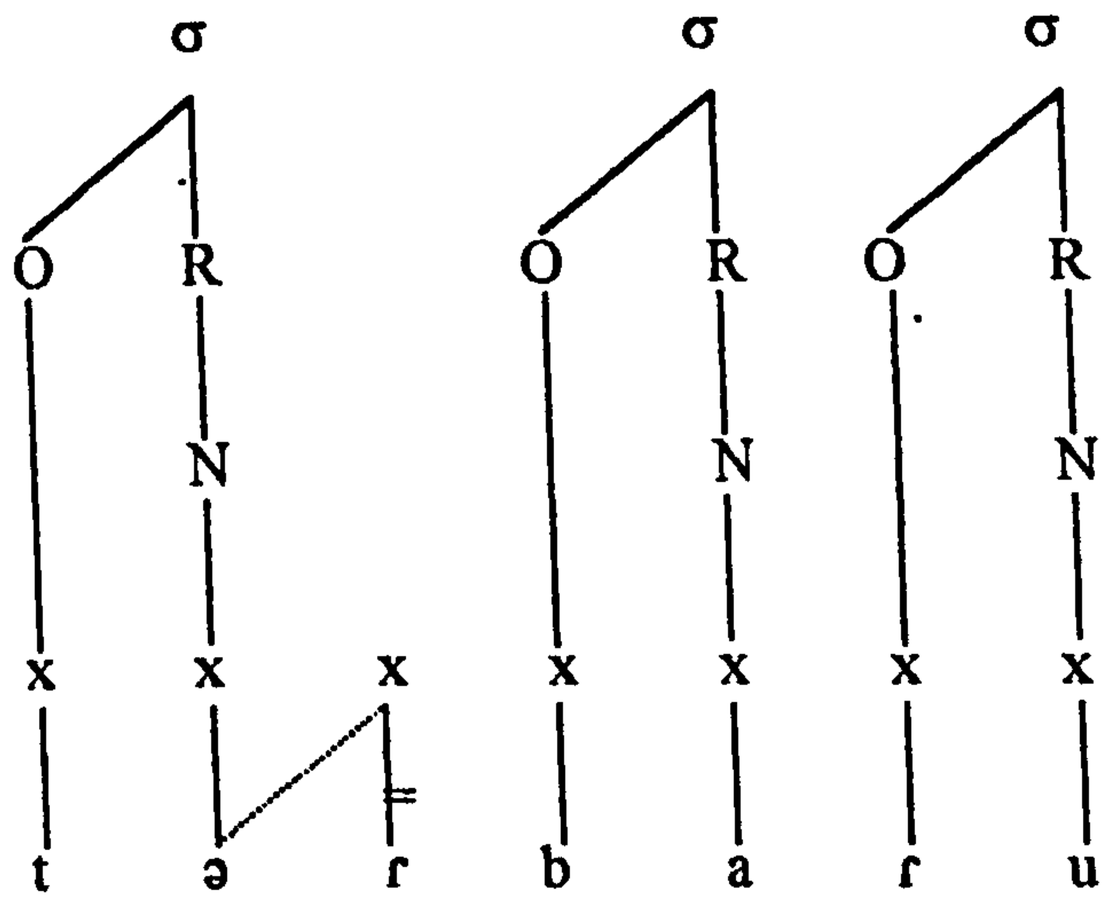
b.



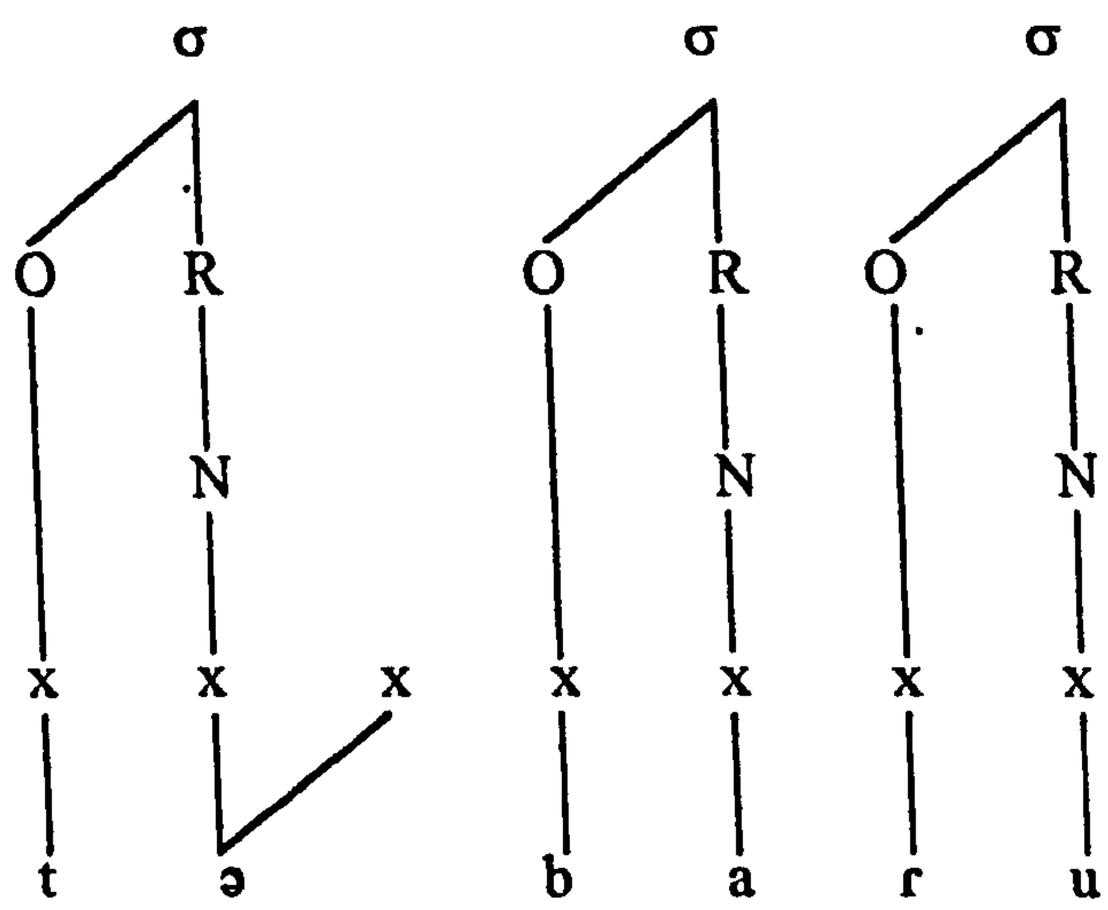
Phonetic representation = [ki.sa:.kan]

(15) CL PW-internally: prefix + root

a.



b.

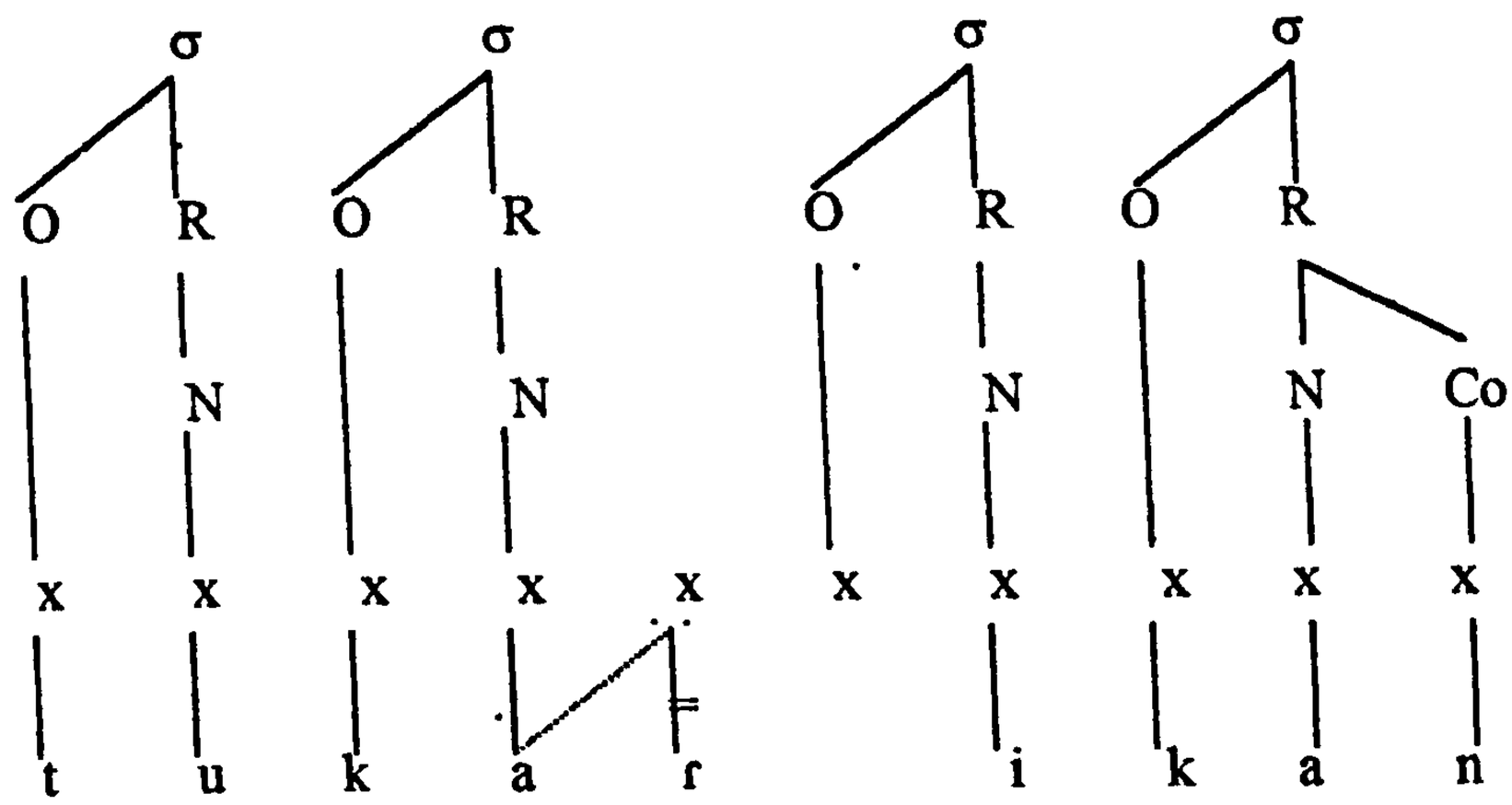


Phonetic representation = [tə:.ba.ru]

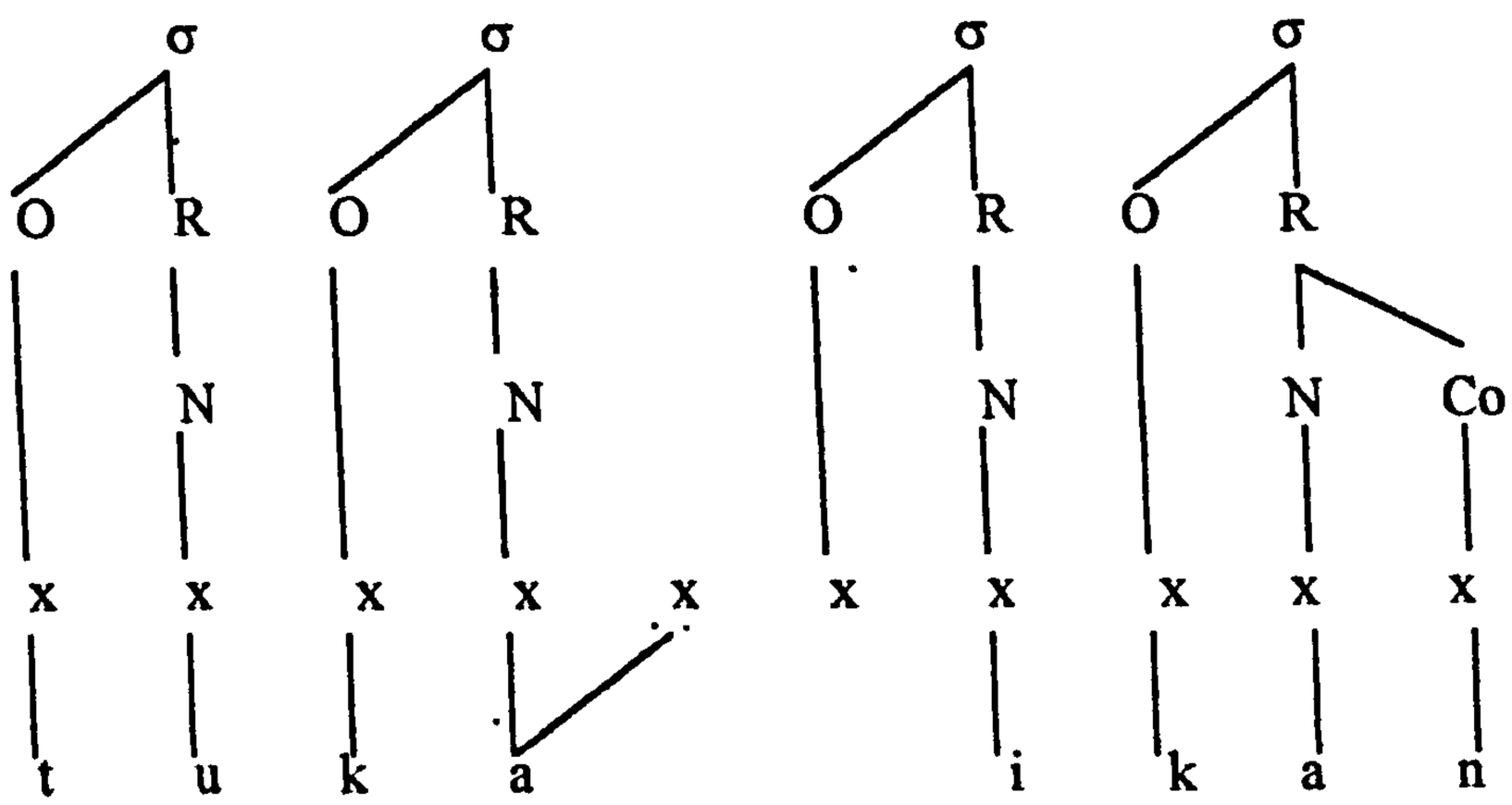
CL is clearly a non-edge-based process.

(16) CL at the PW boundary: syntactic phrase

a.



b.



Phonetic representation = [tuka:ʔikan]

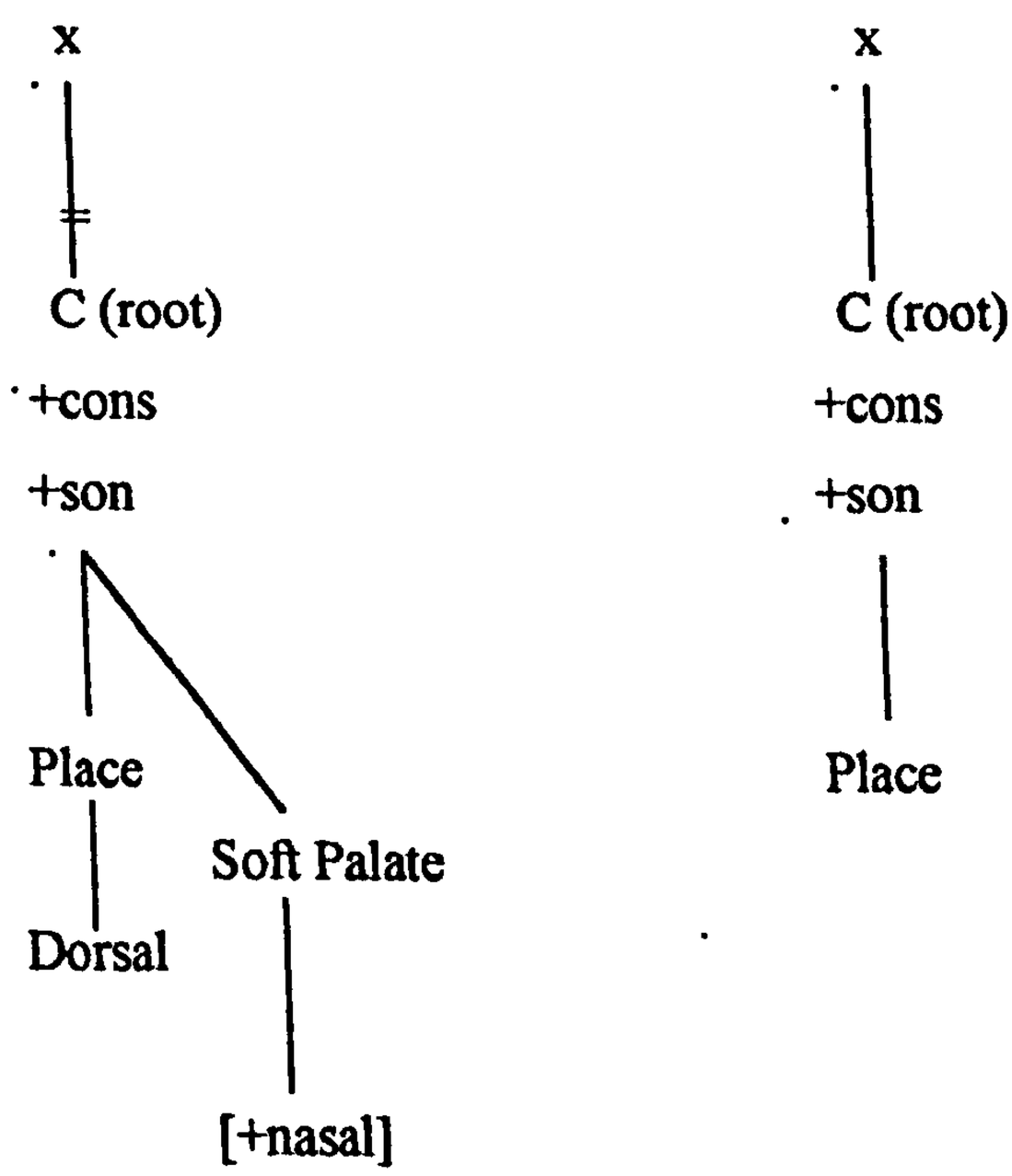
CL in SM can also be found where a nasal is deleted[40]. This occurs when a prefix-final nasal combines with root beginning with non-syllabic sonorants such as glides, liquids and nasals (i.e. w, j, l, r, m, n, ŋ, and ɲ) as in (17).

[40] Nasal deletion is distinct from NOA.

(17).	/mãŋ/ +/jakin/ + /kan/	[mã:j̃ãkinkan]	'to convince'
	/mãŋ/ + /ɾeka/	[mã:ɾekə]	'to design'
	/pəŋ/ + /waɾis/	[pə:waɾis]	'heir/heiress'
	/pəŋ/ + /lawak/	[pə:lawaʔ]	'comedian'
	/pəŋ/ + /malas/	[pə:mãlas]	'lazybone'
	/pəŋ/ + /nanti/	[pə:nãnti]	'person who wait'
	/pəŋ/ + /ɲaɲi/	[pə:ɲãɲũ]	'singer'
	/mãŋ/ + /ŋaɲa/	[mã:ŋãŋã]	'to gape'

We agree with Teoh (1994: 45) in that Nasal Deletion is a deletion of the root node of the nasal segment when followed by a non-syllabic sonorant. Given Halle's (1995) feature geometry, Nasal Deletion in SM is formulated as in figure (18).

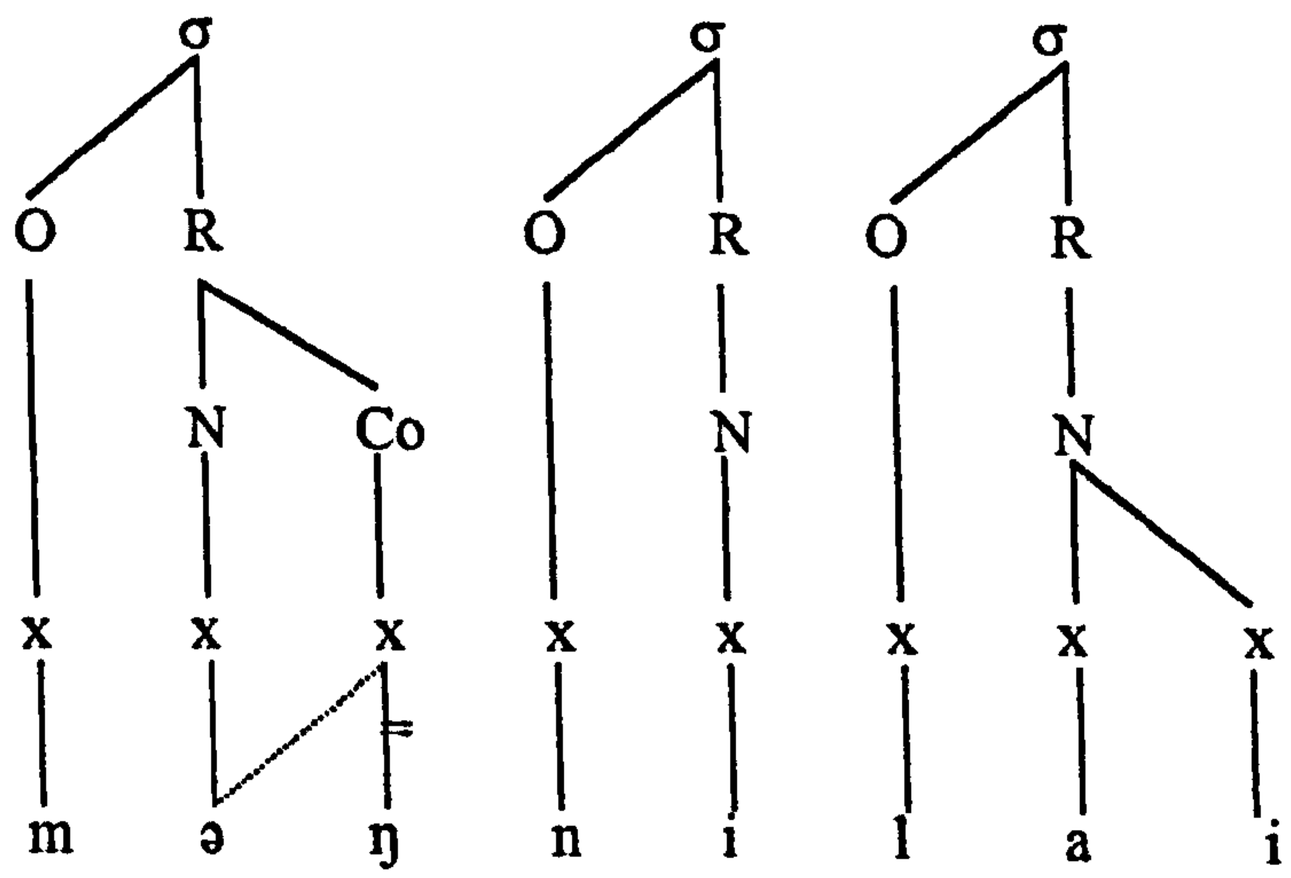
(18) Nasal Deletion rule as dinking of root node of nasal



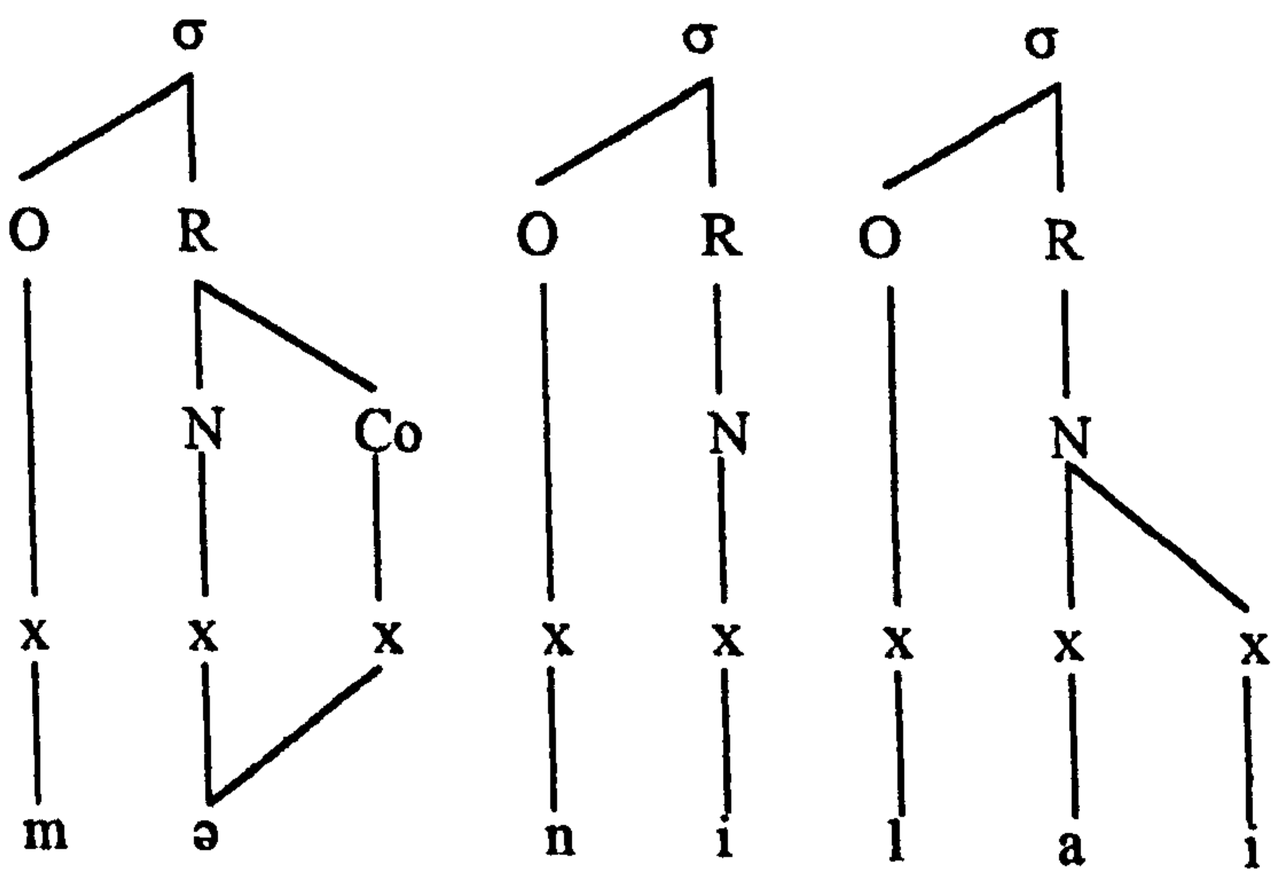
The rule in figure (18) shows that the prefix-final nasal will be deleted when combined with a non-syllabic sonorant. On account of this, the schwas of prefixes /məŋ-/ and /pəŋ-/ become lengthened when the root nodes of prefix-final nasals are deleted. This behaviour can be depicted as in figure (19) such as in the word /məŋ-nilai/ [mə̃:.ñi.lai] 'to assess'.

(19).

a.



b.



Phonetic representation = [mə̃:.ñi.lai]

In figure (19), the delinking of a root node of a nasal detaches everything that it dominates leaving a skeletal x slot. Concurrently, the preceding schwa relinks to the skeletal x tier.

5.2.5. */a/ Reduction*

The vowel /a/ in the right-edge of a word (i.e. in a final open syllable) in SM is always realised as [ə]. This is referred to as '/a/ Reduction'. It occurs root-internally; when the root attaches to any suffixes it does not lax See (20) below). Moreover, in morphologically complex words the vowel /a/ laxes to [ə]:

(20) PW (=MW): utterance final

a. /masa/	[mãsə]	'time'
b. /saja/	[sajə]	'I'
c. /bahaja/	[bahajə]	'dangerous'
d. /sahadʒa/	[sahadʒə]	'only'

Compounds

i. /sətia/ plus /usaha/	[sətijəʔusahə]	'secretary'
ii. /kərəta/ plus /api/	[kərətəʔapi]	'steam engine'

Reduplication forms

i. /dua dua/	[duwəduwə]	'both'
ii. /buka məŋ+buka/	[bukəməmbukə]	'open repeatedly'

Syntactic phrases

i. /Sudah/ + /lama/ + /awak/ + /di sini?/ [sudahlamə̃ʔawaʔdisiñiʔ]

'Have you been long here?

ii. /kita/ + /sama-sama/ + /jaken/ [kitəsamə̃samə̃jaken]

'we are mutually convinced'

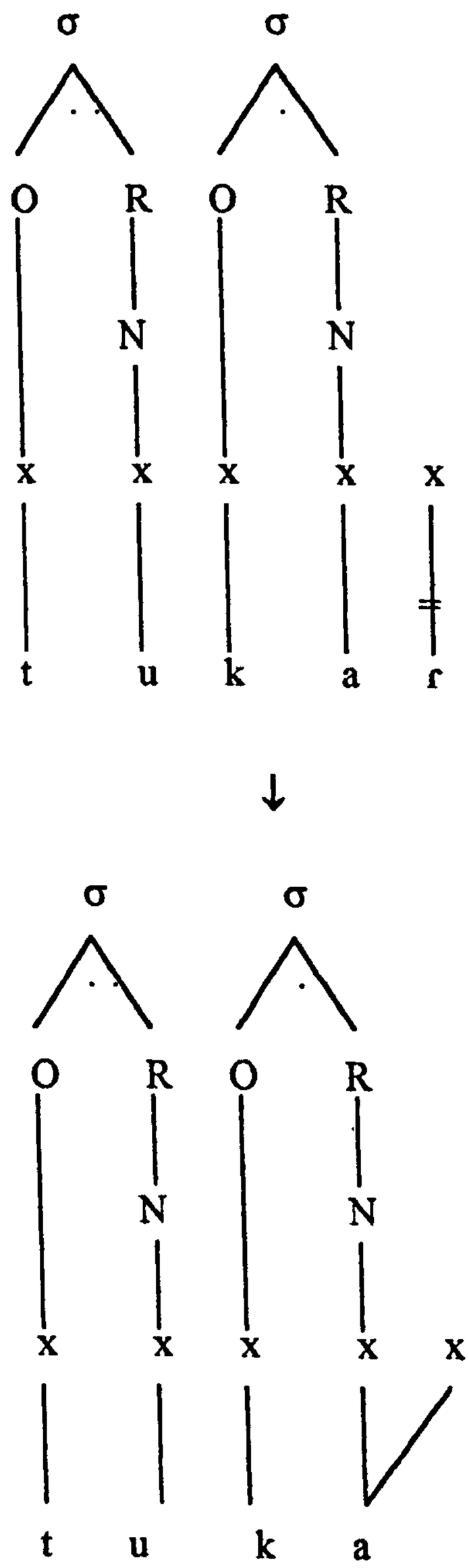
(21) Morphologically complex words

a. /mula + i/	[mulaʔi]	'start'
b. /tʃuba + an/	[tʃubaʔan]	'test'
c. /kata + kan/	[katakan]	'say'

On the basis of these examples, we claim that /a/ Reduction is a non-edge-based process, and /a/ followed by a floating /r/ in a closed syllable fails to lax to [ə]. Following Teoh's argument (1994: 49) the vowel [a:] fails to undergo vowel reduction because it is now long, i.e. attached to two-X-slots instead of one; the structural description of /a/ Reduction is not met. In a non-linear analysis, both vowel lengthening and vowel reduction can be illustrated as in (22) below.

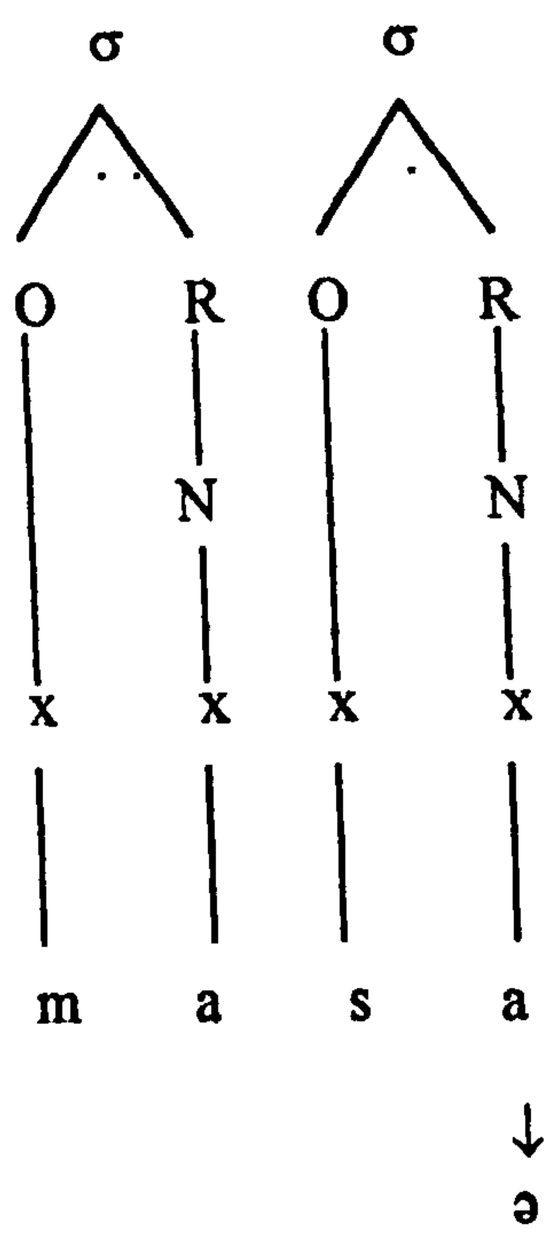
(22) Vowel Lengthening and /a/ Reduction

a. *Floating /r/ deletion incurs VL such as in the word /tukar/ 'change'*



Phonetic representation = [tuka:]

b. /a/ Reduction such as in the word /masa/ 'time'



Phonetic representation = [mãsə]

5.3 Conclusion

By postulating that the PW in SM is co-extensive with the MW, we have overcome the problems in assuming that the PW in SM is either a root + suffix (with prefixes forming PWs) or a prefix + root (with suffixes forming PWs). And out claim that SM exhibits both edge-based and non-edge-based processes provides an account of the asymmetrical behaviour of, on the one hand, GF and Gem (right-edge processes whose scope extends no further than one boundary to the left of the right edge of the PW) and, on the other hand, NOA and VOD (left-edge processes whose scope extends no further than one boundary to the right of the left-edge of the PW).

Other lexical generalisations, such as /k/ Glottalling, Nasalisation, Resyllabification, Compensatory Lengthening and /a/ Reduction, are non-edge-based processes. In illustration of this, we now give sample derivations in (36) involving the lexical generalisations of GF, Gem, NOA, Nasalisation, /k/ Glottalling, Resyllabification, Compensatory Lengthening and /a/ Reduction.

In the following chapter, we seek to provide independent support for our analysis by considering word stress assignment in SM.

(34)

U.R.	/tari+an/	/məŋ+dʒalan+i/	/koto(r)+an/	/masak+an/	/mula/	/koto(r)/
Right Edge GF Gem	tarijan	məŋ+dʒalanni		masakkan		
Left Edge NOA		məɲdʒalanni				
Post-Edge Nas Resyllab /k/ Glottalling Comp Length /a/ Reduction		mə̃ɲdʒalannĩ	kotoran	mãsakkan mãsaʔkan	mulə	koto:
Output of Lex Phonology	[tarijan] 'dancing'	[mə̃ɲdʒalanni] 'to undergo'	[kotoran] 'dirt'	[mãsaʔkan] 'cooking dish'	[mulə] 'start'	[koto:] 'dirty'

CHAPTER 6

WORD STRESS ASSIGNMENT IN SM

6.0 Introduction

We have now proposed an alternative definition of PW in SM, and an alternative analysis to account for the difference in the behaviour of certain generalisations with respect to prefix-root and root-suffix boundaries. Here we provide independent evidence to back up our analysis, evidence from the SM word stress assignment algorithm. Our claim is that word stress assignment in SM also shows edge-based effects, with the right-edge Primary Stress assignment operating prior to the left-edge-based initial Secondary Stress assignment. There are three main aims of this chapter: firstly, to claim that SM word phonology also exhibits edge effects in its stress assignment algorithm; secondly, to provide a full account of word stress assignment in SM (since this has not been done before, to the best of our knowledge, and also because such an account fills out our overall picture of the word phonology of SM); thirdly, to show how GSI, which is one of the main processes we discuss in this thesis, interacts with stress patterns.

The discussion here aims to integrate the stress assignment algorithm with SM morphophonology. In presenting the stress assignment algorithm we assume that stress assignment and related lexical generalisations hold over entire PWs. That is, we are not proposing a level-ordered account of word formation in SM. Rather,

we are claiming that there are phonological generalisations in the lexical phonology of SM which are interleaved with the stress assignment algorithm.

In considering the stress assignment algorithm, we also ask whether the stress contours affect the following phonological processes in SM: (a) GSI, (b) the schwa vowel, (c) affixation, (d) reduplication, and (e) compounds, and whether there are any significant changes of the stress patterns resulting from the application of these rules.

The structure of this chapter is as follows: section 6.1 covers word stress assignment in roots, morphologically complex forms, reduplicated forms and compounds. Section 6.2 is a discussion of GSI and postlexical word stress adjustment. In section 6.3 we provide a summary of our claims.

6.1 Word stress assignment in SM

A large number of proposals have appeared in the literature concerning the representation of word stress and word stress assignment algorithms, such as those of McCarthy and Prince (1986, 1993), Prince (1983, 1985), Hayes (1980, 1995), Selkirk (1980), Halle and Vergnaud (1987a, 1987b), Liberman and Prince (1977) and Liberman (1975). Such work has been variously based on metrical trees, metrical grids, principles and parameters and conflicting constraints, or some combination of these. We adopt an analysis of word stress assignment in SM which

is similar in many, but not all, respects to Cohn's (1989 and 1993) grid-based analysis of Indonesian (often taken not to be distinct from Malay). We believe that in previous studies the stress assignment algorithm in SM has been overlooked, or has only been described in term of primary stress (see Zaharani 1998), who also disregarded some of the consequences of morphophonological alternations for SM stress patterns.

The data shown in this study are more complex than those given in the previous literature as they come from our own tape recorded observations of native speakers of SM from different areas of Johore, Malaysia. Since we found close concurrence between the speakers, we feel confident that our observations are representative of the way speakers of SM (i.e. the Johore dialect) stress words.

6.1.1 *Stress in roots*

In SM, there are few indigenous roots longer than two syllables; in fact, the vast majority of roots in SM are bisyllabic, with stress on the penultimate syllable: the most common metrical structure is thus the bisyllabic trochaic foot. We believe that loanwords into SM are chunked up into trochaic feet precisely because the trochaic foot is the overwhelmingly most common metrical structure in the native vocabulary: it is the structure on which all of its word stress patterns are founded. This rendering of polysyllabic loanwords as sequences of trochaic feet constitutes their metrical nativisation. When examining the stress patterns of longer words, we

will mostly be considering such loanwords, borrowed largely from Arabic, English and Sanskrit (we will consider morphologically complex forms later in this chapter). Utilising these loanwords in a discussion of word stress in SM is not problematic as they are fully integrated into SM phonology; in fact, as we have just suggested, their nativised metrical structure helps shed light on the nature of word stress assignment in SM. We begin with stress patterns of roots in SM.

(1). Stress patterns of roots in SM

a.	σ	há?	'right'
		tʃám	'identify'
b.	σσ	bá.ku	'standard'
		sá.lah	'wrong'
c.	σσσ	da.hú.lu	'past'
		ha.lá.mãn	'courtyard, compound'
d.	σσσσ	bì.da.dá.ɾi	'angel'
		bì.ja.sís.wa	'scholarship'
		ò.lah.ɾá.gə	'athletics'
		sù.wa.sá.nə̃	'atmosphere'
e.	σσσσσ	dʒùs.ti.fi.ká.si	'justification'
		sì.la.tu.ɾá.him	'friendship'

f.	σσσσσσ	ì.den.tì.fi.ká.si	'identification'
		sà.ka.rà.tul.mấ.ũt	'near to death'
g.	σσσσσσσ	às.sa.la.mũ̀.ʔã.lái.kum	'may peace be upon you'
		ìn.di.vi.dù.wa.lís.ti?	'individualistic'

As indicated above, the most common metrical structure is the bisyllabic trochaic foot; it seems plausible, therefore, to analyse longer words as containing sequences of trochaic feet (*see* Spencer 1996, Kager 1996, Halle and Idsardi 1996, Halle and Vergnaud 1987, Prince 1983*)*. We assume that such feet are structured in both a left-to-right and a right-to-left manner, with Main Stress assigned from the right edge of the word, creating an initial trochaic foot, as in many bisyllabic and trisyllabic roots, illustrated in (2) and (3); the initial grid mark in (3) falls (we claim) outside of the foot which follows it. Round brackets indicate foot boundaries. The conventions for grid building which we adopt are: assign a grid mark to any syllable containing a full vowel (i.e. not schwa: see below on schwa and stress patterns); assign a further two grid marks to a primary stressed syllable; assign one further grid mark to a secondary stressed syllable.

(2) *Bisyllabic roots*

```
x
x
x  x
(ba ku)
```

(3) *Trisylabic roots*

```
      x
      x
  x   x x
  da (hu lu)
```

For many words containing four or more syllables, more than one trochaic foot may be constructed, as in (4) and (5), each of which contains two trochaic feet; the leftmost foot is created by initial Secondary Stress assignment, which operates left-to-right, as we will see below:

(4) *Roots with four syllables*

```
          x
  x       x
  x  x    x x
  (bi da) (da ri)
```

(5) *Roots with five syllables*

```
            x
  x         x
  x x  x    x  x
  (si  la  tu) (ra him)
```

Examples of words which contain three trochaic feet can be seen in (6) and (7):

(6) *Roots with six syllables*

```
                x
x       x       x
x x     x x     x x
(i.den.) (ti.fi.) (ka.si)
```

(7) *Roots with seven syllables*

```
                x
x       x       x
x x x   x x     x x
(as sa la) (mũ ʔa˺) (lai kum)
```

As these examples show, SM exhibits main stress on the penultimate syllable of a polysyllabic word. An initial secondary stress appears on the leftmost syllable in words of four or more syllables; a word consisting of three syllables will not be assigned a secondary stress on the leftmost syllable since this would result in stress clash. An additional secondary stress will be assigned where a root contains more than six syllables with full vowels in the surface form. Note that in words with five syllables, an additional secondary stress is not assigned either to the second syllable from the left, or to the third, since both would result in stress clash. This phenomenon is referred to as *Clash Avoidance*. Examples are dʒùs.ti.fi.ká.si 'justification' and sì.la.tu.ɾá.him 'friendship', where we would have the incorrect form if a secondary stresses fell on the second syllable, which is adjacent to a first syllable (secondary stress) i.e. *dʒùs.tì.fi.ká.si and *sì.là.tu.ɾá.him.

However, many of the initial dactyl cases discussed in the literature are ambiguous as to whether the stresses to the left of the main stressed (penultimate) syllable are due to left-to-right or right-to-left assignment[41]. Consider the example of stress assignment in cases such as 1(f) ì.den.tì.fi.ká.si 'identification'. In this case, stress assignment in SM could be either left-to-right or right-to-left. We argue that, to resolve this issue, words with seven syllables or more must be taken into account, such as às.sa.la.mǔ.ʔã.lái.kum and ìn.di.vi.dù.wa.lís.tiʔ 'individualistic' in 1(g) above. As pointed out by Cohn (1993), if a secondary stress surfaces two syllables to the right of the initial secondary stress, this suggests left-to-right assignment (σσσσσσσ), whereas, if a secondary is two syllables to the left of the *main* stressed syllable, this suggests right-to-left assignment (σσσσσσσ)[42]. SM words containing seven syllables show that additional secondary stresses in SM follow the second pattern: they are placed, right-to-left, on alternate syllables to the left of the main stressed syllable (i.e. the penultimate syllable)[43]: we do not find examples of the sort *às.sa.là.mũ.ʔã.lái.kum.

To sum up thus far, a primary stress is assigned to the second syllable from the right edge of the root, creating a trochaic foot; a secondary stress is assigned to the first syllable on the left edge of the root, also creating a trochaic foot. Any additional secondary stresses fall on alternating syllables to the left of the main

[41] For further detail, see Cohn (1993), Halle and Kenstowicz (1991) and Hayes (1991).

[42] Halle and Kenstowicz (1991) propose this sort of analysis for English, Italian and Polish.

[43] See Cohn (1986: 172) on the Right-to-Left Rule in terms of Grid Theory.

stress, also creating trochaic feet. Secondary stress on alternating succeeding syllables[44], subject to Clash Avoidance, prohibits a sequence of two adjacent stressed syllables within a root. It seems clear that SM main stress assignment is a right-edge process, and initial secondary stress is a left-edge process. But stress assignment is more complex in words containing schwa, which we now turn to.

6.1.2 *Stress in roots with schwas*

As mentioned in Chapter 2, SM has six underlying vowels, namely /i/, /e/, /a/, /u/, /o/ and /ə/, all visible to stress assignment, except /ə/ (as in Indonesian; but see below on roots containing only schwas). The vowels [a, e, i, o and u] are full vowels: schwas ([ə]) are non-full vowels. Given that the schwa is invisible to stress assignment in SM, stress assignment in bisyllabic and trisyllabic roots containing schwas are as follows:

Bisyllabic roots

If a schwa occurs in the first syllable, main stress falls on the final syllable, as in (8a and 8b); if a schwa is in the final syllable, then the stress falls on the initial syllable, as in (8c and 8d); bisyllabic roots in which both syllables contain schwas exhibit a trochaic stress pattern, as in (8e) and (8f): this kind of case is the only case in which a schwa may be stressed.

[44] We cannot agree with Zaharani (1998) that there are no secondary stresses in SM.

(8) *Schwa in bisyllabic roots*

a.	tə.(bú)	'sugar cane'
b.	sə.(ɾái)	'lemon grass'
c.	(sá.jə)	'I'
d.	(bá.wə)	'bring'
e.	(kə́.nə̃)	'strike'
f.	(bə́.lə)	'breed'

Trisyllabic roots

Main stress assignment is sensitive to the occurrence of schwas. A penultimate syllable containing a schwa fails to receive stress, as in (báh.tə.ɾə) 'ship'. If there are two full vowels, one on either side of a penultimate schwa, stress falls on the syllable to the left rather than right of the penultimate, as in pútəɾi 'princess'. If the syllable preceding the penultimate also contains a schwa, then main stress will fall on the full vowel of the final syllable, as in tʃə.mə̃ɾ.(láŋ) 'excellent' respectively. Trisyllabic words containing only schwas exhibit default trochaic patterns, as in tʃə.(də́.ɾə) 'injured'. We sum up these patterns in (9).

(9) *Schwas and main stress in trisyllabic roots*

a.	(báhtəɾə)	'ship'
b.	tə.(lá.gə)	'well'
c.	tʃə.(ɾá.mãh)	'speech/talk'
d.	(pú.tə.ɾi)	'princess'

e. tʃə.mə̃r.(lán) 'excellent'

f. tʃə.(də́.ɾə) 'injured'

Roots with more than four syllables

Stress patterns in roots with a schwa and four or more syllables are exhibited in (10):.

(10) Stress patterns in roots containing schwa(s) with four or more syllables

a. *Schwas and stress in four syllable roots*

gə.nə̃.(ɾá.si)	'generation'
sə.ɾi.(gá.lə)	'wolf'
(kò.pə.) (ɾá.si)	'cooperation'
a.(nū.gə.ɾah)	'reward'
(bà.ha.) (gí.jə)	'blissful/happy'

b. *Schwas in five syllable roots*

tʃən.də.ɾa.(wá.seh)	'bird of paradise'
kə.(sà.tə.) (ɾí.ja)	'warrior'
sə.(kù.la.) (ɾís.ma)	'secularism'

c. *Schwas in six syllable roots*

gə.nə̃.(ɾà.li.) (sá.si)	'generalisation'
(ɾè.vo.) (lù.si.) (jó.nə̃:)	'revolutioner'
pə:.(sò.ñi.fi.) (ká.si)	'personification'

d. *Schwas in seven syllable roots*

(ìn.tə.leʔ.(tù.wa.) (lís.mã) 'intellectualism'

As regards (10a), secondary stress is not assigned to the first syllable when a schwa intervenes between the penultimate stress and initial syllable. This is because the first syllable is itself a schwa: secondary stress must fall on a full vowel (it is only when all the syllables in a word contain schwas that main stress will fall on a schwa, for want of a syllable containing a full vowel. As a result, four syllable roots will not be assigned a secondary stress when a schwa intervenes between the main stress and the initial syllable. But if the intervening syllable has a full vowel, then a secondary stress will be assigned to the initial syllable, as in bà.ha.gí.jə 'blissful'. A secondary stress may appear on the left of the main stress in words of five and more syllables, if they (the main stress and the initial syllable) alternate with syllables containing full vowels, for example in sə.kù.la.rís.mã 'secularism' and in gə.nõ.rà.li.sá.si 'generalisation'. An additional secondary stress may also be assigned in six and seven syllable roots, if schwas do not intervene as in rè.vo.lù.si.jó.nõ: 'revolutioner' and ìn.tə.leʔ.tù.wa.lís.mã 'intellectualism' respectively.

Cohn (1989: 174) claims that, in Indonesian, schwas are invisible to stress assignment. This is because schwas are, on her view (but not ours), epenthetic: they are inserted after stress is assigned. We take schwas to be underlying and invisible to stress assignment except in default cases where there are only schwas in the word.

We express their invisibility by not showing them on the rhyme projection in the metrical grid: they lack a grid marker, as suggested above:

(11) *Schwas in bisyllabic roots*

a.

```
      x
    x
    x   -
   (ba wə)
```

b.

```
      x
      x
    -   x
   (tə bu)
```

(12) *Schwas in trisyllabic roots*

a.

```
      x
      x
      x   -   -
   (bah tə rə)
```

b.

```
    x
    x
  - x -
  tə (la gə)
```

c.

```
        x
        x
  -  -  x
  tʃə mər (laŋ)
```

(13) *Schwas in four syllable roots*

a.

```
        x
        x
  -  -  x x
  gə nə̃ (ra si)
```

b.

```
    x
    x       x
  x x  -    x
  a (nũ gə) (rah)
```

(14) *Schwas in five syllables root*

```
          x
          x
  -  -  x   x  x
  tʃən də ra (wa seh)
```

(15) *Schwas in roots with six syllables*

```
                x
        x       x
- -   x   x     x   x
gə nə̃ (ra  li)  (sa  si)
```

6.1.3 *Stress in morphologically complex words*

In SM, affixes may cause stress pattern changes, but only suffixes and prefixes with full vowels may take stress. The affixes which do so are as follows: (a) The derivational affixes, /-an/ and /kə-an/. The latter is a circumfix, but it is only the suffix part that takes stress, of course. (b) The verbal suffixes, /-i/ and /-kan/, which change the valence of verbs. They co-occur with the verbal prefixes /məŋ-/, /pəŋ-/ 'active', and /di-/ 'passive'. As mentioned, schwas do not take stress except by default, so the prefixes /məŋ-/, /pəŋ-/ are never stressed.

SM also has possessive suffixes, namely the enclitics /-ku/, /-mu/ and /-ɲa/. These possessive suffixes appear to the right of roots or to the right of other suffixes, i.e. /-kan/, /-an/ or /-i/. The final vowel of the suffix /-ɲa/ laxes to a schwa (giving [-ɲə̃]) in surface representation. As with the enclitics /-ku/ and /-mu/, the laxing of the vowel-final possessive suffix /-ɲa/ does not prevent main stress from being assigned on the penultimate syllable. These are illustrated in (16), (17) and (18).

(16).

/guru/	(gúru)	'teacher'
/guru + ɲa/	gu(rúɲə)	'his/her teacher'
/guru + ku/	gu(rúku)	'my teacher'
/guru + mu/	gu(rúmũ)	'your teacher'

(17)

/dʒuriga/	dʒu(ríga)	'suspicious'
/məŋ + dʒuriga + kan/	mãn(dʒùri) (gákan)	'to suspicious'
/məŋ + dʒuriga + kan + ɲə/	mãn(dʒùriga)(kánɲə)	'make him suspicious'
/məŋ + dʒuriga + kan + ku/	mãn(dʒùriga)(kánku)	'make me suspicious'
/məŋ + dʒuriga + kan + mu/	mãn(dʒùriga)(kánmu)	'make you suspicious'

(18)

/bahagia/	(bàha)(gíjə)	'happy'
/kə + bahagia + an/	kə(bàha)(gìja)(ʔán)	'happiness'
/kə + bahagia + an + ɲa/	kə(bàha)(gìja)(ʔánɲə)	'his happiness'
/kə + bahagia + an + ku/	kə(bàha)(gìja)(ʔánku)	'my happiness'
/kə + bahagia + an + mu/	kə(bàha)(gìja)(ʔánmu)	'your happiness'

Other affixes that may fall within the stress domain are the prefixes /di-/ and /dʒuru-/. Prefixes with full vowels in SM may be assigned stress when they combine with roots. Prefixes with full vowels are similar to the suffixes in that they are a part of the stress domain when they attach to roots. By contrast, in Indonesian, Cohn (1989: 204) claims that the prefixes with full vowels act as if they are not part of the

stress domain. In SM, since secondary stress falls on the initial syllable of a word, an initial secondary stress in the prefix /dʒuru-/ falls on the syllable '*dʒu*', i.e. the leftmost syllable of the prefix. However, in the case of the monosyllabic prefix /di-/ when attached to a monosyllabic root, main stress can not be assigned on the prefix as main stress is a right-edge process which cannot penetrate into prefixed material. For us, this is an important point: the prefix /di-/ when attached to a monosyllabic root could, in principle, form a trochaic foot structure, but it does not, for the reason just given. Examples of prefixes with full vowels that attach to roots and are visible to the stress assignment are:

(19)

a.	i.	/di+tʃat/	ditʃát	'painted'	
	ii.	/di+tʃat+kan/	ditʃátkan	'painted'	
b.	i.	/dʒuru+waŋ/	dʒùruwáŋ	'banker'	
	ii.	/dʒuru+waŋ+ɲə/	dʒùruwáŋɲə̃	'his/her banker'	
c.	i.	/ubah/	ʔúbah	'move'	
	ii.	/di+ubah/	diʔúbah	'moved'	(passive)
	iii.	/di+ubah+kan/	dìʔubáhkan	'moved'	(passive)
	iv.	/di+ubah+kan+ɲə/	dìʔubahkáɲə̃	'moved by him/her'	
d.	i.	/antara/	ʔantárə	'between'	
	ii.	/di+antara/	dìʔantárə	'in between'	
	iii.	/di+antara+ɲə/	dìʔantaráɲə̃	'in between him/her'	

e.	i.	/ atʃara/	ʔatʃárə	'event'
	ii.	/dʒuru+atʃara/	dʒùruʔatʃárə	'presenter'
	iii.	/dʒuru+atʃara+ɲə/	dʒùruʔàtʃaráɲə̃	'presenter'
f.	i.	/habasa/	bahásə	'language'
	ii.	/dʒuru+bahasa/	dʒùrubahásə	'interpreter'
	iii.	/dʒuru+bahasa+ɲə/	dʒùrubàhasáɲə̃	'his/her interpreter'

In (19), the main stress is always assigned to the penultimate syllable of the morphologically complex form except in the cases where a monosyllabic prefix is attached to a monosyllabic root. When secondary stress occurs, it falls on the left-edge of the morphologically complex form (i.e. the first syllable of prefix). However, an additional secondary stress is invisible on the second syllable from the initial syllable which alternates from penultimate stress on account of Clash Avoidance (as in (19c-iv, 19d-iii, 19e-ii and 19f-ii). In such cases, an additional secondary stress may assign on the third syllable as it is not subject to Clash Avoidance for both main stress (i.e. penultimate stress) and initial secondary stress. This is shown by the data in 19c-iii & 19d-iii: dʒùruʔàtʃaráɲə̃ 'his/her presenter' and 'dʒùrubàhasáɲə̃' 'his/her interpreter' respectively. Clash Avoidance (19c-ii) can be illustrated in terms of Cohn's (1986) metrical grid rule as in (20) (see Cohn 1986:171).

(20). Additional secondary stress and Clash Avoidance in the morphologically complex form /dʒuru + atʃara/ * dʒùrùʔatʃárə 'presenter'

```
            x
  x  *x     x
  x  x  x   x   -
  dʒu ru ʔa tʃa rə
```

In (20), the unmarked case a grid mark is not laid down if it clashes with one already laid down: an additional secondary stress cannot be assigned because it clashes with the secondary stress which is already assigned.

Given stress alternations in roots, as in cases such as gúru/gurúŋə and many others cited here, we do not assume that main stress is assigned to the first syllable of the root and then later switched to the second syllable after suffixation. Rather, stress assignment operates at the word level after affixation, with main stress being assigned to the penultimate syllable of a word, independently of whether the final syllable is part of a root or a suffix. Similarly, initial secondary stress is assigned to the leftmost syllable independently of whether that syllable is part of a prefix or a root. The only exception to this is the case of bisyllabic words consisting of a prefix and a root: in those cases, the penultimate syllable of the word is not stressed because of the locality constraint which we have suggested is imposed on edge-based processes in SM.

6.1.4 *Stress in reduplicated forms*

Reduplication in SM may take any of several forms to furnish a variety of semantic nuances signalling pluralisation, diversification in nouns, intensification of verbs, and the formation of adverbs from adjectives. As Zaharani (1998:207) puts it, 'various combinations of reduplication with suffixation and prefixation give rise to a number of significant phonological patterns'.

Generally, reduplication in SM can be categorised into two classes: primary reduplication; and secondary reduplication. Primary reduplication is productive reduplication in SM. Primary reduplication can be categorised into two classes, namely: (a) Root-Reduplication - the process of copying the root, often in conjunction with some other morphological process (prefixation and suffixation) and (b) Doubling - the process of a complete copying of the whole word (comprising affixation), most common in nouns, denoting the meaning of plurality and diversity. Secondary Reduplication comprises Partial Reduplication, Rhyming and Chiming.

As in Indonesian (Cohn 1989: 184), the reduplication process in SM is such that root reduplication occurs prior to affixation; doubling is reduplication following affixation. Reduplication must also be represented as a single entity as it may apply before affixation or after it, but never twice in the same form. This is true of all types of reduplication, even in highly lexicalised forms.

6.1.4.1 Doubling

By considering the SM reduplication process, we can see that Doubling in (21) is stressed differently from root reduplication forms as in (22). As doubling is reduplication after affixation, we therefore have a similar level of stress contours in both the root and reduplicant (comprises a combination of root and affix) assigned by two primary stresses. There is however no further application of the main stress rule, and the two halves surface with equal stress prominence. The main stresses in each half sound roughly equal, but the tendency is for the penult to sound a little more prominent. Doubling is shown below in (21).

(21) *Doubling*

Root	**Base**		**Doubling**
lári	pəlári	'runners'	pəlári-pəlári
táhan	tahánãn	'detainees	tahánãn-tahánãn
báŋon	baŋɔ́nnãn	'buildings'	baŋɔ́nnãn-baŋɔ́nnãn
bántu	bantúwan	'assistences'	bantúwan-bantúwan
ɾádʒə	kəɾàdʒaʔán	'governments'	kəɾàdʒaʔán-kəɾàdʒaʔán
bàhagíjə	kəbàhagìjaʔán	'happiness'	kəbàhagìjaʔán-kəbàhagìjaʔán
sájo:	sajó:ɾan	'vegetable'	sajó:ɾan-sajó:ɾan

The data in (21) shows that reduplication applies after the root is affixed. Doubling shows that there is no further application of the Main Stress rule and that the two halves surface with equal stress prominence. If we expect reduplication to apply before stress assignment, we would have an incorrect form such as *tàhanãn-

tahánãn 'detainess' and *bàntuwàn-bantúwan 'assistences'. This is parallel to the stress assignment of monomorphemic forms like ì.den.tì.fi.ká.si 'identification'. In respect of this we claim that whatever is being reduplicated is copied completely, including the whole metrical structure.

6.1.4.2 Root reduplication

Given that root reduplication is reduplication prior to affixation, we therefore show the analysis in (22) as follows: column I consists of root stress; column II consists of reduplicated root forms, and; column III consists of affixed root reduplication. In column I both root and reduplicant (or the two halves) surface with equal stress prominence. While in column III, the penultimate stress appears distinctly as the most prominent stress in the whole form. The Main Stress rule is thus applied such that the word may receive one main stress and be followed by secondary stresses. This shows that reduplication applies, then suffixation or cliticization causes main stress to reapply and give the observed output. Examples are presented in (22):

(22). Affixed *Root Reduplication*

Roots		**Reduplication**	**Affixation**
I		**II**	**III**
a. tʃát	'paint'	tʃát-tʃát	dìtʃat-tʃátkan
			tʃàt-mãŋãtʃát
			mãŋãtʃàt-ŋãtʃát

b. wáŋ	'money'	wáŋ-wáŋ	dʒurùwaŋ-dʒùruwáŋɲõ
c. úlaŋ	'repeat'	úlaŋ-úlaŋ	ùlaŋ-ʔuláŋɲãn
			ùlaŋ-ʔuláŋkan
			diʔùlaŋ-ʔúlaŋ
			diʔùlaŋ-ʔuláŋkan
			diʔùlaŋ-ʔùlaŋkáɲõ
			bərùlaŋ-rúlaŋ
			bərùlaŋ-rulá̰ŋɲãn
			mãɲùlaŋ-ɲúlaŋ
			mãɲùlaŋ-ɲũláŋkan
			mãɲùlaŋ-ɲùlaŋkáɲõ
d. táhan	'stop'	táhan-táhan	tàhan-tahánkan
			tàhan-mãnáñãn
			mãnàhãn- nãñãn
			ditàhan-táhan
			tə:tàhan-táhan
e. lámõ	'long'	lámõ-lámõ	làmã-kəlàmãʔã́n
			làmã-kəlàmãʔã́ɲõ
f. kədʒá:	'chase'	kədʒá:-kədʒá:	kədʒà:-mãɲõdʒá:
			mãɲõdʒà:-ɲõdʒá:
			bə:kədʒà: -kədʒáran

g. wilájah 'district' wilájah-wilájah wilàjah-wìlajáhɲə̃
dìwilàjah-wìlajáhɲə̃

In (22), we observe that the main stress assignment is a right-edge rule, which assigns stress to the penult, and that initial secondary stress assignment is a left-edge rule which assigns stress to the initial syllable. In addition, a further secondary stress assigns is assigned to alternately syllables from the main stress, that is from right-to-left, as in (22b, 22c and 22g) dʒurùwaŋ-dʒùruwáŋɲə̃, mə̃ŋùlaŋ-ŋùlaŋkáɲə̃ and dìwilàjah-wìlajáhɲə̃ respectively.

In the case of GSI applying, the second of the two relevant vowels must receive a main stress that is on the right-edge of the word such as in (22e) 'làmã-kəlàmãʔán'. Furthermore, the penultimate syllable that would be expected to receive a main stress is demoted as main stress has already been assigned on the final syllable. This shows that GSI creates a foot and can also affect postlexical readjustment of word stress. Meanwhile, a secondary stress falls on the left-edge of the word (i.e. the first syllable).

From examples (21 & 22), we can see that the basic difference between the two types of reduplication is how they are ordered with respect to affixation: root reduplication precedes affixation and doubling follows it[45]. However, in certain

[45] For further discussion contrasting stem reduplication and doubling, see Cohn (1989: 184).

circumstances, root reduplication behaves differently. For consideration of such forms, see data in (23) below.

(23)

a). i. tàhan-mənáñãn

ii. mənàñãn- náñãn

b. i. kədʒà:-mãŋãdʒá:

ii. mãŋãdʒà:-ŋãdʒá:

In (23), all outputs are derived from the roots *'tahan'* 'stop' and *'kədʒa:'* 'chase' respectively. These forms are ordered differently with respect to affixation and the relevant phonology. Considering cases (23-i and 23b-i), we expect that reduplication must precede prefixation and NOA whereas in the cases (23a-ii and 23b-ii), reduplication must follow prefixation and NOA.

6.1.4.3 Partial reduplication

There are three further secondary classes of reduplication, namely, Partial Reduplication, Rhyming and Chiming. Partial Reduplication is a process of copying an onset of a first syllable of the root in order to fill the framework of templatic morphology in which the onset (reduplicant) will be one member of a skeletal template CV. In order to fill the template with a nucleus the specified schwa [ə] is

inserted[46]. Partial reduplication applies on verbs and nouns in order to denote derived nouns and plurality (objects or animals) that resemble base forms. In partial reduplication, we expect reduplication to apply after stress assignment. Suffixation causes Main Stress to reapply to give the observed output. Some examples of partial reduplication are:

(24).

Roots		**Partial Reduplication**	
kúdə	'horse'	kəkúdə	'wooden horse'
láki	'guy'	ləláki	'guy'
dáon	'leaf'	dədáon	'leave'
		dədaónnãn	
póhon	'tree'	pəpóhon	'trees'
		pəpohónãn	

The data in (24) shows that the stress assignment of Partial Reduplication falls on the penultimate syllable (the right-edge of the reduplicated form). Inserted schwas between copied onsets and the left-edge of word however, do not affect stress assignment since they are unstressable.

6.1.4.4 Rhyming and chiming

Rhyming and Chiming reduplications are unpredictable phonetic changes. In Rhyming, reduplication occurs by copying one of the base syllables, either the initial

[46] In previous analyses (Asmah 1975 and Arbak 1981), the templatic morphology is obtained by copying the first syllable of the root, and reducing the reduplicated vowel to a schwa.

syllable together with the following consonant (onset of the final syllable) or the final syllable. By contrast, in Chiming, only the consonants are repeated, with the vowels in the reduplicant being unpredictable. These two classes of reduplicated forms are no longer productive and have a tendency to be attached by affixation although their utilisation is restricted. Zaharani (1998) claims that these forms are fully lexicalised in the language, and cannot be regarded as part of word formation. These classes of reduplication are shown below:

(25). *Rhyming Reduplication*

Roots	**Rhyming**	**Affixed Rhyming**
a. ánã?	ánã?-pínã?	bəràna?-pínã?
'child'	*'a large number of children'*	*'being generated'*
b. bátu	bátu-bátan	bə:bàtu-bátan
'stone'	*'assortment of stones and bricks'*	*'various of stones'*
c. búket	búket-búkau	dibùket-búkau
hill	*'various kind of hills'*	*'in the various kind of hills'*

(26). *Chiming Reduplication*

Roots	**Chiming**	**Affixed Chiming**
a. gúnõng	gúnõŋ-gánãŋ	bə:gùnõŋ-gánãŋ
'mountain'	*'range of mountains'*	*'various of mountains'*
b. gopoh	gópoh-gápah	tə:gòpoh-gápah
'hurry'	*'in hurry'*	*'do thing hastily'*
c. ásal	ásal-?úsol	bəràsal-?ùsol
'origin'	*'ancestor'*	*'originated'*

In SM, rhyming and chiming may be observed in two types of stress patterns. In (25 and 26), the reduplication forms in the 'Rhyming' and 'Chiming' columns show that the two halves surface with equal stress prominence. This shows that the stress patterns in Rhyming and Chiming are similar to Doubling. In columns 'Affixed Rhyming' and 'Affixed Chiming', a main stress is applied which assigns on the right-edge (i.e. the penultimate syllable of the reduplicant) and a secondary stress assigns on the left-edge of the base. This means that the application of affixation causes the Main Stress rule to be reapplied.

It is noteworthy that the stress assignment of SM, like that of Indonesian (see Cohn 1989) relies on reduplication or affixation being the most recent process. The surface contour is thus defined by the morphological process which applies last. By affiliating stress assignment rules with morphological processes, we claim that the surface stress pattern depends on which morphological process was most recent. However, in forms where the most recent morphological operation is affixation, the penultimate stress appears as the most prominent stress in the whole form. This shows that affixation triggers the reapplication of the Main Stress rule at the right-edge of the form. At the same time, the previous prominent stress that falls on the base is demoted to a secondary stress, which we claim appears at the left-edge.

6.1.5 *Stress in compounds*

Compounds in SM consist of two elements (words). These elements are monomorphemic, but the resulting compounds may undergo further affixational

morphology. In SM, each half of a compound receives its own stress pattern in the normal way, and the relative prominence contour is imposed on the compound as a whole. As in Indonesian, the stress pattern in a compound is largely similar to the reduplicated forms in which reduplication has last applied, (for example in the word **sətìja ʔusáhə** 'secretary'). Consider the following examples:

(27).

	Roots	Compounds		Affixed Compounds
a.	sətíjə + usáhə	sətìjaʔusáhə	'secetary'	sətìjaʔùsahájñə̃
				kəsətìjaʔusàhaʔán
b.	gúru + agámə̃	gùruʔagámə̃	'religious teacher'	gùruʔàgamájñə̃
c.	səgí + əmpát	səgíʔəmpat	'square'	be:səgíʔəmpat
d.	bátu + ápi	bàtuʔápi	'flint'	be:bàtuʔápi
e.	dúkə + tʃítə	dùkatʃítə	'sad'	bə:dùkatʃítə
				məndùkatʃitákan
				məndùkatʃitakánjñə̃

According to (27), we can see that compounds in SM have three distinct behavioural characteristics. Firstly, in a similar way to Indonesian, each element has its own metrical structure assigned. Secondly, in order to form metrical structure of a compound, both elements combine together with the right element more prominent to the left element. Like reduplicated forms, stress assignment of a compound depends on the application of affixation and causes main stress to reapply to produce the observed output. Thus, the formation of compounds precedes affixation. This is illustrated in (27) in the column 'Affixed Compounds'. When an affix is attached to

a compound, the regular stress rules apply, resulting a shifting of the main stress[47]. However, the secondary stress in the first element remains unchanged. For further illustration, we present the data in (27a) in metrical feet as follows:

(28).

a.

```
        x
    x   x   x
  - x x x x x -
  sə ti ja ʔu sa ha ɲõ
```

b.

```
            x
      x     x   x
  - - x x x x x x
  kə sə ti ja ʔu sa ha ʔan
```

Note that an additional secondary stress may apply on the second element which alternates from the main stress such as in (28a and 28b) sə̀tijaʔùsahá̀ɲõ 'his secretary' and kəsə̀tijaʔusàhaʔán 'secretarial line' respectively. What is interesting here is that the affixed compound in (28b) has a glottal stop in the final syllable. Given that where GSI applies, the second of the two relevant vowels must receive main stress, the main stress must fall on the rightmost side of the compound with any additional secondary stress falling to the left of the main stress alternately.

[47] The reapplication of main stress means that stress assignment undergoes Destressing rule. For further detail see Cohn (1989 and 1993) and Kager (1989).

6.2 GSI and postlexical word stress adjustment

Glottal Stop Insertion (GSI), unlike the lexical generalisations, is a late across-the-board postlexical rule which applies to any sequence of two filled nuclei (Further detail, see Chapter 7). In this section, we argue that its application demands that the second nuclei must be stressed. GSI may readjust word stress postlexically, and may create stress contours which violate the lexical constraint, Clash Avoidance. This phenomenon tends to cause phonetically-driven stress assignment in the language and may operate in morphologically complex forms, reduplicated forms and compounds in SM. Consider the following examples.

(29) *GSI in morphologically complex forms*

a.	/guna/	gúnã	'use'
	/guna + an/	gùnãʔán	'usage'
		*gunãʔãn	
b.	/pinda/	píndə	'amend'
	/pinda+an/	pìndaʔán	'amendment'
		*pindáʔan	
c.	/dʒaga/	dʒágə	'watch'
	/dʒaga + an/	dʒàgaʔán	'care/protection'
		*dʒagáʔan	
d.	/suka/	súkə	'happy'
	/kə + suka + an/	kəsùkaʔán	'happiness'
		*kəsukáʔan	

e.	/mula/	múlə	'start'
	/mula + i/	mǜlaʔí	'starting'
		*mũláʔi	

(30) *GSI in Reduplicated forms*

a.	/guna /	guna	'use'
	/kə + guna + an/	kəgùnãʔán	'usage'
		kəgùnãʔán-kəgùnãʔán	'usage repeatedly'
		*kəgunáʔãn- kəgunáʔãn	
		*kəgunáʔãn- kəgunàʔán	

b.	/lama/	lámã	'long'
	/lama-kə + lama + an/	làmã-kəlàmãʔán	'finally'
		* làmã-kəlamáʔãn	
		* làmã-kəlamàʔán	

(31) *GSI in compounds*

/sətíjə + usáhə/	sətíjəʔusáhə	'secretary'
/kə + sətia + usaha + an/	kəsətìjaʔùsahaʔán	'secretarial'
	* kəsətìjaʔùsahaʔan	
	* kəsətìjaʔùsahàʔán	

We noted above that primary stress is assigned to the second syllable from the right-edge of the word in SM. However, when a vowel-final root combines with vowel-initial suffix, GSI applies between the two vowels. In this case, the second of the

two relevant vowels must receive primary stress[48], while the penultimate syllable becomes invisible to stress assignment since primary stress is already assigned on the final syllable. This phenomenon can be illustrated in the metrical grid as follows.

(32)

```
        x
  x     x
  x  x  x
(pinda)(?an)
```

In (32) the bisyllabic root /pinda/ is stressed, in isolation, with primary stress on the first syllable. However, when the vowel-initial suffix /-an/ is added, GSI is triggered, and the second of the two vowels receives main stress. At the same time, the penultimate syllable which would receive a main stress is invisible to stress assignment and the previous primary stress is 'demoted', and receives secondary stress due to the fact that its syllable alternate from the main stress. This shows that GSI can affect postlexical readjustment of word stress. This SM phenomenon appears to be an extension of a commonly occurring phenomenon where GSI applies as long as the second of the two vowels is already stressed (cf Booij 1997 on Dutch).

Secondary stress can also be assigned to the first syllable on the left edge of the word, and on alternating succeeding syllables, subject to Clash Avoidance. This

[48] Note that only glottal stops resulting from the application of GSI trigger this process. Glottal stops arise from two other sources in SM: /k/ Glottalling, and cases where [ʔ] is substituted for [ʕ] in Arabic loanword as in [saʔat] ('second'). Stress remains on the first syllable in such cases.

prohibits a sequence of two adjacent stressed syllables within a word (examples of this are):

(33) Prefix /dʒuru-/ + root /atʃara/

```
                x
        x       x
        x x  x  x  -
        dʒu ru ʔa tʃa rə          'presenter'
```

(34) Prefix /dʒuru-/ + root /atʃara/ as clash avoidance

a.

```
                x
       *x x     x
        x x  x  x  -
        dʒu ru ʔa tʃa rə          'presenter'
```

b.

```
                x
       *x    x  x
        x x  x  x  -
        dʒu ru ʔa tʃa rə          'presenter'
```

In (33), the second syllable from the left edge does not receive secondary stress, despite the fact that it is alternate from the main stress. In this case an additional secondary stress is blocked by Clash Avoidance on account of initial secondary stress having already been assigned on the first syllable (see 34a). In

addition the third syllable from the left edge does not receive secondary stresses, despite the fact that it is the syllable containing a glottal stop. In this case, secondary stress is blocked by Clash Avoidance on account of the main stress having already been assigned to the penultimate syllable (see 34b).

Clash Avoidance may also be over-ridden (as in 35) in the word /sə+umpama/. In this case since the schwa is unstressable, the prefix /sə-/ is skipped over by the secondary stress assignment algorithm and primary stress falls on the second syllable from the right of the word. Secondary stress falls on the second syllable from the left edge of the word due to the syllable undergoing GSI. Since the main stress must assign on the rightmost syllable, the syllable containing a glottal stop is demoted by a secondary stress. This shows that the secondary stress adjacent to the main stress over-rides Clash Avoidance. Consider the metrical grid of over-ridden Clash Avoidence in (35).

(35) Prefix /sə-/ + root /umpama/

```
        x
    x   x
  - x   x   -
  səʔum pa mə̃
```

6.3 Conclusion

As shown in the analysis above, schwas in SM do not receive stress unless the word in question contains only schwas. Affixes form a part of the stress domain. A postlexical process GSI acts to readjust word stress.

It is interesting to note that GSI not only creates a foot, but can also create clashing stress. In this case, a syllable containing a glottal stop is more prominent than other stresses. Thus any penultimate syllable that precedes a glottal stop is demoted due to Clash Avoidance. Under certain circumstances Clash Avoidance may be over-ridden in order to prevent a secondary stress from destressing.

The main conclusion made here is that, in the lexical phonology of SM, main stress assignment is a right-edge-of-the-word process, while initial secondary stress a left-edge process. From this, it follows that the application of a right-edge-based main stress assignment process must be taken to precede the left-edge-based secondary stress assignment process. This conclusion is reached since main stress cannot penetrate into prefixed material and secondary stress cannot affect suffixed material. This, we claim, is evidence in favour of our edge-based analysis of certain aspects of the lexical phonology of SM.

CHAPTER 7

GLOTTAL STOP INSERTION AS A POSTLEXICAL GENERALISATION

7.0 Introduction

We have claimed that, in SM, underlying heterosyllabic /V+V/ sequences are always subject to the hiatus avoidance strategies of Glide Formation (GF) or Glottal Stop Insertion (GSI), with GF operating at the PW level, and GSI postlexically. In this chapter, we will examine what looks like counter-evidence to this claim: the non-occurrence of GSI in what look like underlying /a/ + /i/ and /a/ + /u/ sequences. The problem is this: if GSI really is a postlexical rule, as we claim, it ought to operate on such sequences, assuming that they really are heterosyllabic. One of our tasks will thus be to establish whether such sequences are heterosyllabic, or whether they are diphthongal nuclei. If they are diphthongs, then our claim about the postlexical status of GSI is unfalsified. If they are not, it appears falsified. In order to settle the issues, we appeal to evidence from an onset reversal language game. The evidence suggests that the forms in question do indeed contain heterosyllabic vowel sequences, which suggests that our claim is wrong. However, we suggest that GSI does not operate on such sequences since they are perceptually hard to distinguish from the diphthongs /ai/ and /au/. Speakers thus treat them as not containing a hiatus and thus not requiring the intervention of the hiatus-avoidance strategy of GSI.

Before we discuss GSI in detail, we consider the status of glottal stops in loan words, (section 7.1). In section 7.2, we consider the postlexical status of GSI.

The occurrence of heterosyllabic V + V sequences root-internally is discussed in this section and the issue of whether there are any root-internal contexts in which GSI could apply is discussed.

7.1 Glottal Stops in Arabic and Chinese loanwords

As we have seen, Yunus (1980), Farid (1980) and Teoh (1994) all claim that the glottal stop is a member of the underlying phonemic inventory of SM. But, as we have shown, there is a contradiction in the work of these authors, since they simultaneously argue that the glottal stop in Malay is *predictable*, and take it to be derived from either the velar stop /k/ or the velar stop /g/ (see Yunus 1980, Farid 1980 and Teoh 1994), as in '*budak*' 'child' [buda?] and '*ragbi*' 'rugby' [ra?bi]. Glottal stop in coda position in words such as /kakak/ [ka.ka?] 'sister' and /masak + an/ [ma.sa?.kan] 'to cook' are derived from the consonant /k/ and do not justify postulating /?/ as member of the SM phonemic inventory[49]. By contrast, a glottal stop [?] in vowel sequences such as in the word /di+ambil/ [di?ambil] 'taken' (passive) results from the application of GSI in the onset position: glottal stops act as default onsets in SM.. The case of /masak + an/ [ma.sa?.kan] 'to cook' contains a sequence of [?.k] resulting from a geminate kk sequence which arises in the course of derivation via the application of Gem.

[49] In SM velar /k/ in onset position is always realised as [k].

In SM, the Arabic voiced pharyngeal /ʕ/, when it occurs in Arabic loan words between two identical vowels root-internally, is replaced by a glottal stop [ʔ]; this replacement process is distinct from the GSI process, but resembles it to some extent in that the glottal stop is occurring in onset position intervocalically. The difference is that GSI is not a replacement phenomenon. Examples of replacement are:

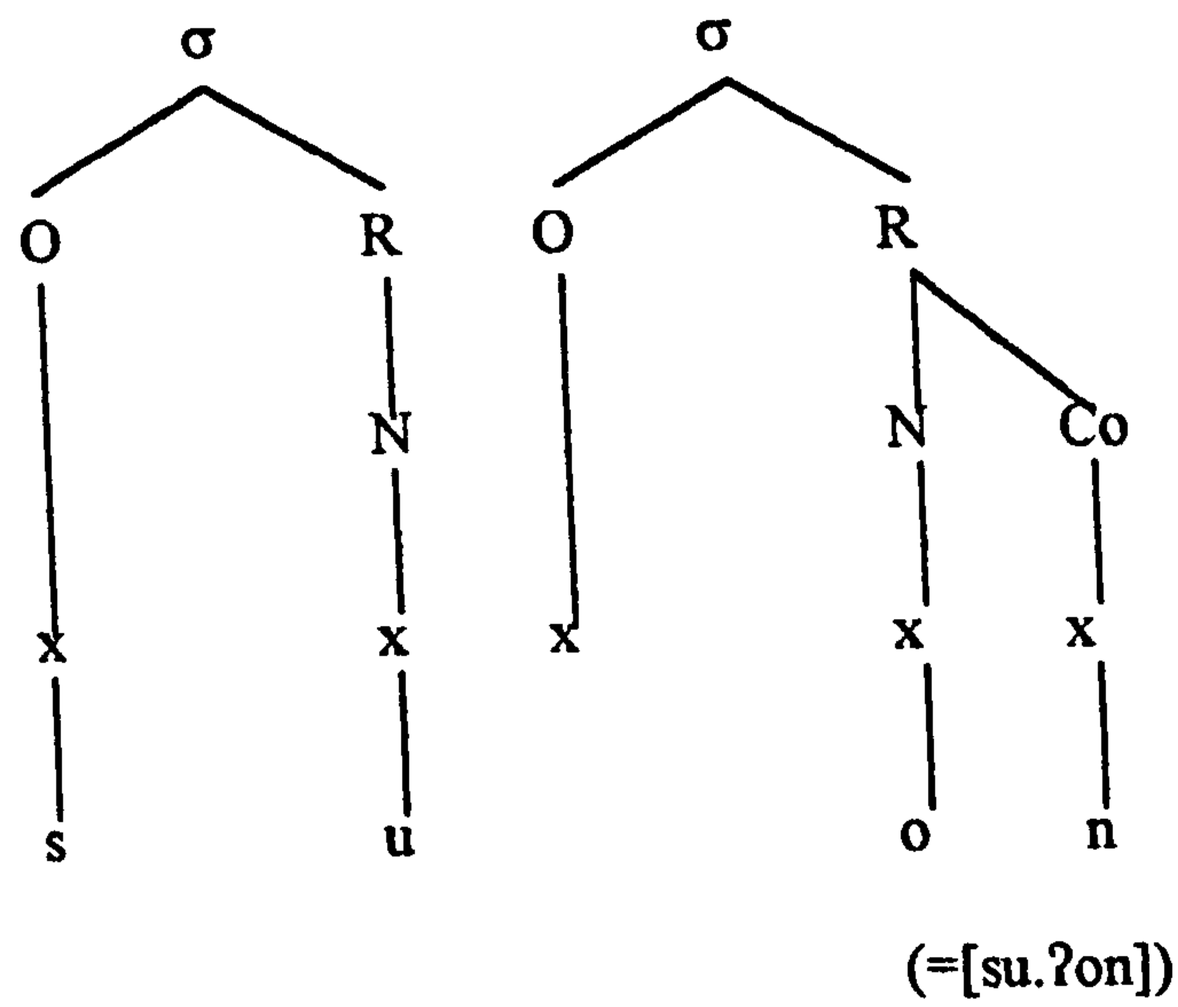
(1). [ʔ] in Arabic loan words

a. /saʕãt/	[saʔãt]	'moment'
b. /taʕãt/	[taʔãt]	'obey'
c. /peʕẽl/	[peʔẽl]	'behaviour'
d. /maʕãp/	[mãʔãp]	'forgive'
e. /tabiʕãt/	[tabiʔãt]	'habit'

In Chinese loanwords such as /suon/ [suʔon][50] 'rice noodle', the glottal stop appears via GSI. Such words are, we assume, lexically marked as failing to undergo GF. As a result, GSI occurs by default, as in (2):

[50] In SM, a loan word of root-internally from Chinese origin that involves GSI is very small number. This is one example that found in the study.

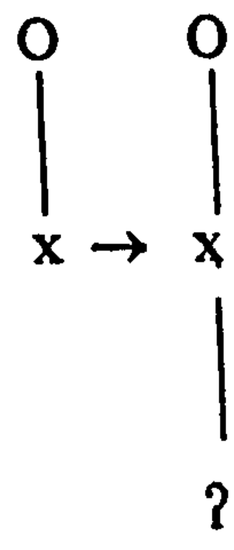
(2) Context for GSI in Chinese loanwords (GF blocked by lexical marking)



7.2 The postlexical status of GSI

We formulate GSI as follows:

(3) Default GSI



That is, a glottal stop is inserted to fill a vacant x slot dominated by an onset (O) node. This kind of filling glottal stop ensures that an onset is present between heterosyllabic vowel sequences. We claim that GSI is a postlexical process,

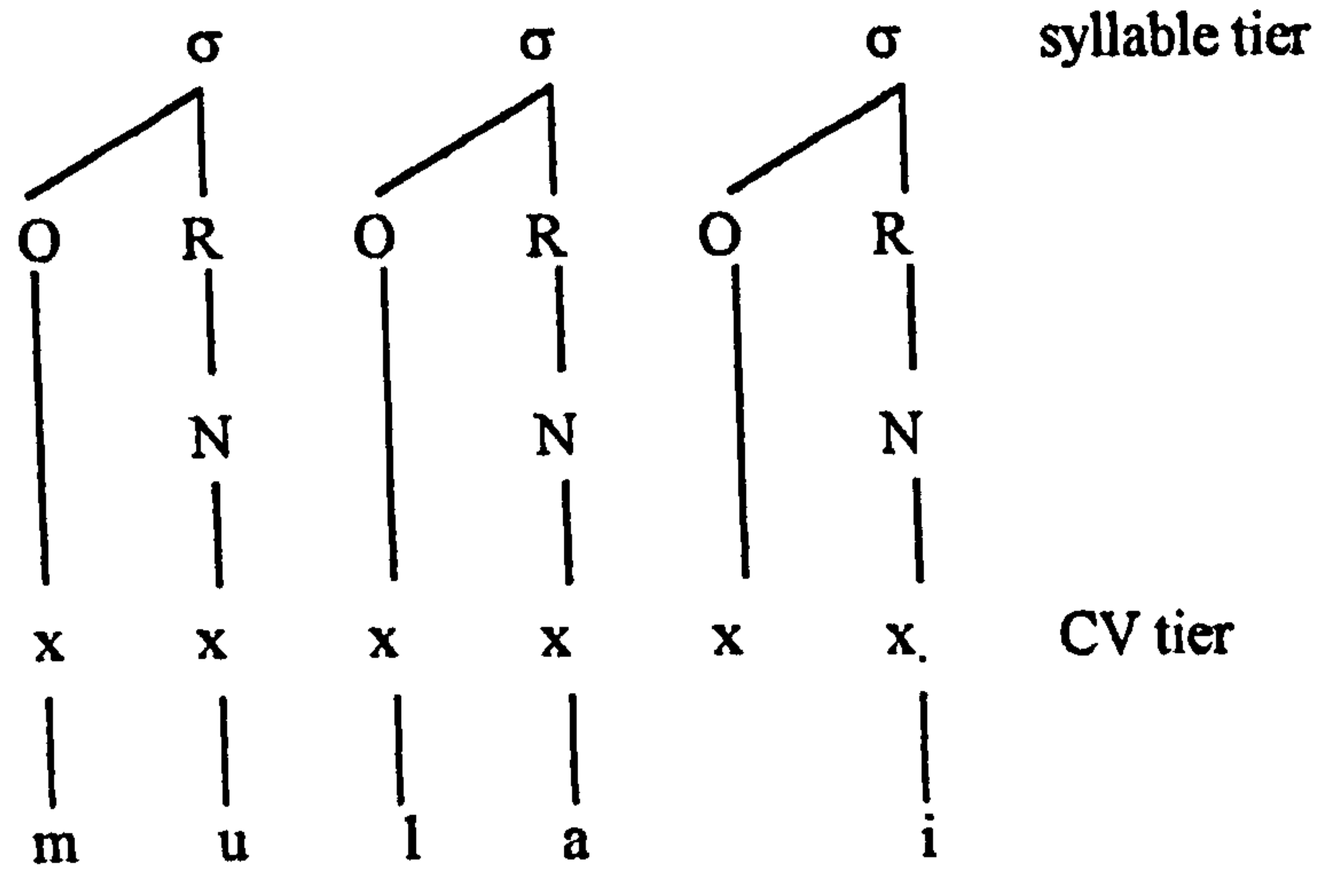
applying 'across-the-board', independently of the presence or absence of boundaries of any sort. The evidence in favour of this claim is strong: unlike all of the lexical processes discussed earlier, GSI can occur across the boundary between two MWs, as shown in 3.2.4 above. All of the lexical processes are blocked at precisely that boundary. This shows clearly that GSI is not part of the word phonology of SM. Further evidence that GSI is an across-the-board process, blind to the presence of word or affix boundaries, is that fact that it operates across prefix-root boundaries, as shown in 3.2.3 above, and across root-suffix boundaries, as shown in (4):

(4).

/tʃuba + an/	[tʃubaʔan]	'test'
/kə + guna + an/	[kəgunãʔãn]	'usage'
/kə + mulia + an/	[kəmũlijaʔan]	'glory'
/pər + kata + an/	[pə:kataʔan]	'word'
/mula + i/	[mũlaʔi]	'starting'
/noda + i/	[odaʔi]	'to stain'
/məŋ+ biaia + i/	[mə̃mbijajaʔi]	'financing'
/məŋ + kuasa + i/	[mə̃ŋũw̃ãsaʔi]	'to control'

We show the relevant syllable structures in (5):

(5) GSI by default across root-plus-suffix boundary



Phonetic representation = [mũla?i]

In (5), GF cannot apply across the suffix boundary because the preceding vowel is a non-high vowel. Since GF does not occur, GSI applies by default postlexically to fill the empty onset in between of vowel sequences.

As mentioned, GSI is not attested root-internally in SM, except in Chinese loan words. This appears to be a problem for our claim that GSI is a postlexical process. A major consideration here is whether there are any root-internal contexts in which it *could* apply, i.e. whether SM really has underlying heterosyllabic V + V sequences. If not, there is no problem: GSI is not attested root internally because there do not happen to be any root-internal sequences that meet its conditions. Central to this isssue is the status of a set of roots whos underlying structure is debatable, and which have been given a range of different phonetic transcriptions in the literature. As Zaharani (1998, p.67) shows, the following authors propose the following different underlying and surface representations for such roots:

(6a). Yunus (1980)

/kain/	[kaẹn]
/naik/	[nãẽ̞?]
/laut/	[laọt]

(6b). Farid (1980)

/kain/	[kaen]
/naik/	[nãẽ?]
/laut/	[laot]

(6c). Durand (1987)

/kain/	[kajn]
/naek/	[nãj̃k]
/laut/	[lawt]

(6d). Teoh (1994)

/ka?in/	[ka?en]
/na?ek/	[nã?ẽ?]
/la?ut/	[la?ot]

Teoh's claim that Malay has an underlying /?/ is unclear and inconsistent, as we have seen. As Zaharani shows, Teoh's transcriptions with glottal stop, such as [la?ot] and [ka?en], are likely to be observational mistakes. Our own observations of tape recordings of SM speakers show no discernable glottal stop in such words. (We

note in passing that Durand's transcription of a [k], rather than a [ʔ], in [nãẽʔ] / [nãj̃ʔ] also an error).

So the question is: what evidence might there be that would allow us to establish whether cases such as these are (a) phonetically and/or (b) underlyingly bisyllabic? If they are either underlyingly or phonetically bisyllabic (or both), then our analysis faces a problem. Our own auditory impressions are that these words sound monosyllabic, although it is not always easy to discern bisyllabicity in sequences of two vocalic segments, if the first is stressed and low, and the second is unstressed and non-low[51]. In this connection, we note that d'Andrade (1998), claims that, in Portuguese, a stressed vowel followed by a high vowel never counts as a hiatus sequence. But our auditory impressions alone are perhaps insufficient to resolve the issue. Accordingly, we now provide evidence bearing on this issue, from an SM onset reversal language game.

There are several versions of the reversal game data, depending on which Malay dialect is spoken by the informant. The generalisation is a simple reversal of the onsets in a sequence of two syllables. Crucially, where one of the onsets is empty, a glottal stop is inserted. For SM speakers, the game works as follows (we mark only the relevant syllable boundaries):

[51] This depends on whether one is dealing with a bisyllabic sequence or not.

(7) *Bisyllabic roots:*

/ba.uah/	[ba.wah]	→	[wa.bah]	'under'
/bu.ah/	[bu.wah]	→	[wu.bah]	'fruit'
/ua.iaŋ/	[wa.jaŋ]	→	[ja.waŋ]	'movie'
/ti.aŋ/	[ti.jaŋ]	→	[ji.taŋ]	'pole'
/si.ap/	[si.jap]	→	[ji.sap]	'ready'
/kər.tas/	[kər.tas]	→	[tər.kas]	'paper'
/a.tas/	[a.tas]	→	[ta.ʔas]	'top'
/i.tu/	[i.tu]	→	[ti.ʔu]	'that'
/pər.lu/	[pər.lu]	→	[lər.pu]	'need'
/pər.gi/	[pər.gi]	→	[gər.pi]	'go'
/sa.na/	[sa.nə̃]	→	[nə̃.sə][52]	'there'

(8) *Bisyllabic roots with diphthongs*

/pi.sau/	[pi.sau]	→	[si.paw]	'knife'
/dau.lat/	[daw.lat]	→	[law.dat]	'power'
/hai.ran/	[haj.ran]	→	[raj.han]/[raj.ʔan]	'amazed'

(9) *Root plus suffix*

/tari + an/	[tari.jan]	→	[rati.jan]	'dance'
/bəli + an/	[bəli.jan]	→	[ləbi.jan]	'buying'
/bantu + an/	[bantu.wan]	→	[tambu.wan]	'assistance'
/buru + an/	[buru.wan]	→	[rubu.wan]	'hunting'

The question is: how does the game affect (a) words in which the roots are unquestionably monosyllabic, and (b) roots such as /naek/? As far as the first type of roots are concerned, some SM speakers follow the pattern: [jaŋ] → [jaʔaŋ], [ru] →

[52] This form exhibits the rule of /a/ → [ə] in word-final position.

[ruʔu]. Others follow the patterns [jaŋ] → [jaŋʔi], [waŋ] → [waŋʔu]. Importantly, the (b) cases follow neither of these patterns. For our SM informants in this study, results were as follows:

(10)

[laot]	→	[halot]	'sea'
[bau]	→	[habu]	'smell'
[dʒaoh]	→	[hadʒoh]	'far'
[kaen]	→	[haken]/[jaken][53]	'cloth'

(11)

[mãẽn]	→	[jamẽn] (not [hamẽn])	'play'
[nãẽʔ]	→	[janẽʔ]/[hanẽʔ] (familiar form)	'ascend'

For SM speakers, these cases behave as if they were underlyingly bisyllabic, but have [h], rather than [ʔ] in the empty onset position: the problematic cases do not behave as though the roots are monosyllabic. Thus that Durand's (1987) description in (6c) above, describing them as underlying monosyllabic, is questionable.

If such roots were underlyingly bisyllabic, but monosyllabic at a derived level, then that would explain why postlexical GSI does not apply to them. A crucial question is thus whether reversal operates on underlying or derived

[53] The form with [j] is used when speaking to people with whom one is familiar.

representations. If the reversal game operates on underlying representations, then the data shows that such cases must be underlyingly bisyllabic, with a root-internal empty onset. However, if reversal operates *after* the operation of other rules in SM, then reversal does not operate on underlying representations. While the data exhibits the operation of several rules of SM lexical phonology, the only rules relevant to reversal are those of Nasalisation[54] and Nasal-Obstruent Assimilation.

As has been shown, the rule of Nasalisation operates left-to-right, after the edge-based processes, but prior to GSI, spreading nasality from a nasal stop onto all the following segments within a word, other than consonants with oral specification, which block the spread of nasality. Consider cases such as underlying /sa.na/ ('there'). In the non-reversed case, Nas applies to yield [sa.nə̃]: the onsets are reversed prior to the application of the rules of Nas and /a/ Reducation. In the reversed case, in which the output is [nã.sə], Reversal applies prior to Nas and /a/ Reduction. The data also supports the view that reversal operates prior to NOA assimilation (e.g. [bantuwan] → [tambuwan]). It therefore seems that reversal operates on underlying representations, and that the problematic cases are underlyingly bisyllabic.

Although the language game evidence suggests that these disputed cases are underlyingly bisyllabic, it is interesting that all of them involve an [a] followed by a [j]/[e] or [w]/[o] vowel, as in [kajn]/[kaen] ('cloth') and [lawt]/[laot] ('sea'). Given

[54] By 'relevant', we mean rules which could operate either before or after Onset Reversal. Rules such as /a/ Reduction and /k/ Glottalling are ignored.

that SM has only two diphthongs, in both of which the first vowel is /a/, namely /ai/ and /au/, it appears to be a coincidence that the problematic cases do not contain an underlying diphthong, but consist coincidentally exclusively of /a/ + /e/ and /a/ + /o/ sequences. We take this to be an arbitrary fact about the URs in SM phonology.

It seems therefore that Zaharani's (1998) ideas that root internal cases such as in (10) and (11) are bisyllabic both underlyingly and on the surface are correct even though they sound as if they contain a diphthong (i.e. they sound as if they are monosyllabic). It is important to note that stress in SM bisyllabic roots falls on the penultimate syllable and that the vast majority of roots are bisyllabic. Moreover, sequences of a stressed low vowel followed by an unstressed high mid/laxed high vowel (resulting from the operation of High Vowel Laxing) are perceptually similar to a diphthong. GSI thus does not operate on such sequences, we claim, since they are not readily perceptible as a hiatus sequence. The claim seems plausible: sequences such as those found in Chinese loanwords such as /suon/ [suʔon] ('rice noodle') are readily perceivable as hiatus sequences since such sequences do not resemble diphthongal sequences. In stark contrast, /a/ + /e/ and /a/ + /o/: sequences, in which the first vowel is stressed but the second is unstressed, are perceptually very similar to the /ai/ and /au/ diphthongs. We believe that SM is like Portuguese in this respect: heterosyllabic /a/ + /e/ and /a/ + /o/ sequences are not perceived as hiatus sequences.

7.3 Conclusion

GSI in SM is a postlexical default process in which a glottal stop fills in the onset position in a heterosyllabic sequence of two vowels. GSI in SM is an across-the-board postlexical process which applies after edge-based processes (see Chapter 5). There are, we claim, no root-internal hiatus vowel sequences which require the intervention of GSI as a hiatus-avoidance strategy, except in Chinese loan words.

With respect to the language game, NOA and Nasalisation, it would appear that reversal operates on underlying representations. We concluded that the problematic cases, all of which involve an [a] followed by a [j]/[e] or [w]/[o] vowel, as in [kajn]/[kaen] ('cloth') and [lawt]/[laot] ('sea') are perceptually very similar to monosyllabic words containing diphthongs, and that GSI fails to apply there because such sequences are not perceived as hiatus sequences.

SUMMARY AND CONCLUSION

We have claimed that SM, like Standard French, has both underlying and surface glides. We have argued that, since both underlying and derived glides are defined in terms of syllable structure, allowing for underlying syllabification is essential. We claimed that SM has both high vowels and glides that are often phonologically related, the only difference between them being the position in the syllable: the high vowels [i] and [u] occupy the nucleus peak while the glides [j] and [w] do not: the definition of 'glide' here is 'a high vowel in non-nucleus peak position'. Examples of derived glides are those in /buah/ [bu.wah] and /tari + an/ [tari.jan] 'dancing', whereas /iaŋ/ [jaŋ] 'which' and /uaŋ/ [waŋ] 'money' contain underlying glides. By contrast, we can see that the high vowels [i] and [u] occupy the nucleus peak position as in [bu.wah] and [tari.jan]. We agree with Zaharani (1998) who claims that the glides [j] and [w] are high vowels /i, u/ in non-nucleus peak positions. However, we do not agree with his argument that *all* glides in SM are predictable. Furthermore, Zaharani's definition of 'glide' cannot be seen as unique or novel since the idea of deriving glides from vowels is over 40 years old (see Jakobson, Fant and Halle 1952).

Allowing for underlying syllable structure, we argue that SM has CV(C) syllable structure, where this means that underlying vowel-initial morphemes contain an empty onset position on the skeletal tier. This is central to the set of SM hiatus-avoidance strategies: Glide Formation, Glottal Stop Insertion and Floating /r/ Realisation.

Previous research has sought to explain certain aspects of SM morphophonology by appealing to the Phonological Word (PW), defined, as a root plus any suffixes. This literature includes Teoh (1994) and Zaharani (1998). The evidence supporting this hypothesis comes from the domain of application of the SM phonological generalisations of Glide Formation (GF) and Gemination (Gem) that hold within the PW thus defined. Thus they hold across a root-suffix boundary but are blocked across a PW boundary (a prefix-root boundary or across the boundary between two morphological words), as defined under the Standard View.

The claim that the PW in SM constitutes a root plus any suffixes is falsified by the data. Instead, the PW in SM is, we conclude, isomorphic with the morphological word. To account for the asymmetry of application of, on the one hand, GF and Gem and, on the other, NOA and VOD, we claim that these generalisations are orientated towards a specific word edge: GF and Gem are right-edge processes whose scope extends no further than one boundary to the left of the right edge. However, by contrast, NOA and VOD are left-edge processes extending no further than one boundary to the right. Other lexical generalisations, such as /k/ Glottalling, Nasalisation, Resyllabification, Compensatory Lengthening and /a/ Reduction are not edge-based process. Edge-based generalisations are subject to a locality constraint which states that they only operate across one morphological boundary from the relevant edge.

We also claim that GSI is a postlexical process which necessarily applies after both edge-based and non-edge-based lexical rules. In order to explain the non-

application of GSI root-internally, we claim that its conditions are simply not met there because SM generally lacks root-internal heterosyllabic V + V sequences. The only instances of these are Chinese loanwords, where GSI does apply, and certain root-internal vowel sequences in underlyingly bisyllabic words, involving an [a] followed by a [e] or [o] vowel, as in [kaen] 'cloth' and /[laot] 'sea', which sound like monosyllabic [kajn] and [lawt] because the first syllable is stressed and the second syllable unstressed. This is connected with the fact that there are two diphthongs in SM, in which the first vowel is /a/, namely /ai/ and /au/. The unexpected non-application of GSI root-internally stems from the relevant sequences of filled nuclei being perceptually similar to the SM diphthongs, with each consisting of a stressed low vowel followed by a high or mid vowel. It is because of this that there is no resolution of hiatus in the cases.

In terms of the word stress assignment algorithm in SM, we assumed that stress assignment and related lexical generalisations hold over entire words. We claimed that there are phonological generalisations in the lexical phonology of SM which are interleaved with stress assignment algorithm. A main stress is assigned to the penultimate syllable and initial secondary stress assigned to the initial syllable. However, under certain conditions, the stress patterns in the language change when the words contain schwas and where GSI has applied. We claim that schwas are unstressable segments in SM, therefore syllables containing schwas are invisible to stress assignment. Syllables containing schwas are exempt from stress assignment (except in default cases where the word contains only schwas) and that the main stress and secondary stresses are as a result, assigned further to the left such as in the

words *pútəɾi* 'princess' and báhtərə 'ship'. Note that stress cannot be assigned to the final syllable of **putəɾí* since SM creates troachic feet.

SM stress assignment is an edge-based processes in which the right-edge process of main stress assignment applies prior to the left-edge processes of initial secondary stress assignment. Word stress assignment backs up our claim that right-edge processes precede left-edge-based processes. Stress assignment also shows that right-edge processes cannot penetrate into prefixed material such as in /di + tʃat/ [ditʃát] *[dítʃat] 'painted', and secondary stress cannot affect suffixed material. Non-edge processes apply after both right-edge and left-edge processes. This is evidence in favour of ordered, edge-based generalisations in the lexical phonology of SM.

Given that GSI causes stress pattern changes, we claim that the vowel following the inserted glottal stop is stressed and causes demotion of lexically assigned main stress. This demotion is due to Clash Avoidance. In certain circumstances however, this may be over-ridden when GSI applies and prevents secondary stress from destressing, for example /sə + umpama/ [səʔùmpámə] 'like'. However, words such as /dʒuɾu + atʃaɾa/ [dʒùɾuʔatʃáɾə] 'presenter' do not violate Clash Avoidance on account of secondary stress having already been assigned on the first syllable. This evidence shows that GSI in SM is a foot creator and readjusts stress assignment postlexically.

In assuming that stress assignment and related SM lexical generalisations hold over entire words, we do not propose a level-ordered account of word formation in SM but claim that there are phonological generalisations in the lexical phonology of SM which are interleaved with the stress assignment algorithm.

References

Abdullah Hassan. (1974). *The Morphology of Malay.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Abdullah Hassan and Ainon Mohd. (1994). *Bahasa Melayu untuk Maktab Perguruan.* Kuala Lumpur: Fajar Bakti.

Aitchison, J. & Straf, M. (1981). 'Lexical storage and retrieval: a developing skill?' In A. Cutler (ed.), *Slips of the tongue and language production.* Amsterdam: Mouton, 311-26.

Anderson, J.M. (to appear). 'A major restructruing in the English consonant system: the de-linearisation of [h] and the de-consonantaisation of [w] and [j]'. *English Language and Linguistics.*

Arbak Othman. (1981). *Tatabahasa Bahasa Malaysia.* Kuala Lumpur: Sarjana Enterprise.

Asmah Haji Omar. (1975). *Essays on Malaysian Linguistics.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Asmah Haji Omar. (1977). *The Phonological Diversity of the Malay Dialects.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Asmah Haji Omar. (1985). *Susur Galur Bahasa Melayu.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Booij, G. (1995). 'The phonology of Dutch'. In the series: *The Phonology of the World's Languages,* In Jacques Durand (ed.). Oxford: Oxford University Press.

Booij, G. (1997). 'Non-derivational Phonology Meets Lexical Phonology'. In Iggy Roca (ed.), *Derivations and Constraints in Phonology*. Oxford: Clarendon Press.

Bromberger, S. & Halle, M. (2000). 'The ontology of phonology (revised).' In Burton-Roberts, N., Carr, P. & Docherty, G. (eds.), *Phonological knowledge*. Oxford: Oxford University Press.

Burzio, L. (1996) 'Surface Constraints vs. underlying representation', in J. Durand and B. Laks (eds.), *Current Trends in Phonology*. Paris: CNRS/Salford: University of Salford Press.

Carr, P. (1993). *Phonology*. London: Macmillan.

Carr, P. (1999). *English Phonetics and Phonology*. Oxford: Blackwell.

Carr, P. (ms). 'The notion "underlying representation" in phonology'. University of Newcastle.

Carr, P. (ms). 'Avoidance of hiatus'. University of Newcastle.

Carr, P. and Tajul Aripin K. (ms). 'The phonological word in Standard Malay'. University of Newcastle.

Casali, R. F. (1996). *Resolving Hiatus*. PhD dissertation, University of California, Los Angeles.

Charette, M. (1991). *Condition on Phonological Government*. Cambridge: Cambridge University Press.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. Cambridge, MA: MIT Press.

Clements, G. N. and Keyser, S. J. (1983). *CV Phonology; A Generative Theory of Syllable.* Cambridge, MA: MIT Press.

Comrie, B. (1990). *The major languages of East and South-East.* Cambridge: Cambridge University Press.

Cohn, A. C. (1989). 'Stress in Indonesian and bracketing paradoxes'. *Natural Language and Linguistic Theory 7*: 167-216.

Cohn, A. C. (1993). 'The initial dactyl effect in Indonesian and the status of degenerate feet'. *Linguistic Inquiry 24*: 372-381.

d'Andrade, E. (1998). 'Glides and stress in Portuguese'. Paper presented at the 2nd CTIP conference, Royaumont, Paris, June 1998.

Delais-Roussarie, E. (1996). 'Phonological phrasing and accentuation in French.' In Nespor, M. & Smith, N. (eds.), *Dam Phonology (HIL Phonology papers 11).* Den Haag: Holland Academic Graphics.

Dempwolff, O. (1934). *Vergleichende Lautlehre des Austronesischen Wortschatzes.* Liechtenstein: Kraus.

Durand, Jacques. (1987). 'On the phonological status of glides: The evidence from Malay'. In John Anderson and Jacques Durand, (eds.), *Explorations in Dependency Phonology.* Dordrecht: Foris.

Durand, Jacques. (ms). 'Gliding along in Midi French'. Handout of paper presented at Durham University, February 1999.

Encrevé, P. (1988). *La liaison avec et sans enchaînement.* Paris: Seuil.

Farid.M. Onn. (1980). *Aspects of Malay Phonology and Morphology: A Generative Approach.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Goedemans, R. (1996). 'An optimality account of onset-sensitive stress in quantity-insensitive languages', *Linguistic Inquiry* 13: 33-48.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology.* Oxford: Basil Blackwell.

Halle, M. (1995). 'Feature geometry and feature spreading', *Linguistic Inquiry* 26 (1): 1-46.

Halle, M. and J-R. Vergnaud (1987a). *An Essay on Stress.* Cambridge, MA: MIT press.

Halle, M. and J-R. Vergnaud (1987b). Stress and the cycle. *Linguistic Inquiry* 18: 45-84.

Halle, M. and Idsardi, W. (1996). 'General Properties of Stress and Metrical Structure'. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory (Pp. 403-443).* Oxford: Blackwell.

Halle, M. and Kenstovicz, M. (1991). 'The Free Element Condition and Cyclic versus noncyclic stress'. *Linguistic Inquiry 22*: 457 – 501.

Hannahs, S. J. (1995a). *Prosodic structure and French morphophonology.* Tübinger: Niemeye

Hannahs, S. J. (1995b). 'The phonological word in French'. *Linguistics* 33, 1125-1144.

Hannahs, S. J. (ms). 'Characterising French glides'. Draft, September 1998. University of Durham.

Harris, J. (1994). *English Sound Structure,* Oxford: Blackwell.

Hayes, B. (1980). *A metrical theory of stress rules.* Doctoral dissertation, MIT. Distributed by the IULC. New York: Garland, 1985.

Hayes, B. (1986). 'Inalterability in CV phonology'. *Language 62*: 321-351.

Hayes, B. (1989). 'Compensantory lengthening in moraic phonology'. *Linguistic Inquiry* 20: 253-306.

Hayes, B. (1990). 'Diphthongization and coindexing'. *Phonology 7*: 31 – 71.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies.* Chicago: University of Chicago Press.

Hock. H. H. (1991). *Principles of Historical Linguistics*, Berlin: Morton De Gruyter.

Hooper, J. (1972). 'The syllable in Phonological Theory'. *Language* 48: 525- 540.

Inkelas, S. (1996). 'The interaction of phrase and word rules in Turkish: An apparent paradox in the prosodic hierarchy'. *The Linguistic Review* 13: 193 - 217.

Ismail Dahaman. (1996). *Pedoman Ejaan dan Sebutan Bahasa Melayu.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Jakobson, R., Fant, G. and Halle, M. (1952). *Preliminaries to Speech Analysis.* Cambridge, MA: MIT Press.

Kager, R. (1989). *A Metrical Theory of Stress and Destressing in English and Dutch.* Dordrecht: Foris.

Kager, R. (1996). 'The Matrical Theory of Word Stress'. In Goldsmith, John A. (ed.), *The Handbook of Phonological Theory (Pp. 367-402).* Oxford: Blackwell.

Kaye, J. and Lowenstamm, J. (1984). 'De la syllabicité'. In F. Dell, D. Hirst and J. - R. Vergnaud (eds.), *Forme Sonore du Languge: Structure des représentations en phonologie* (Pp. 123-159). Paris: Hermann.

Kaye, J. and Lowenstamm, J. and Vergnaud, J. (1985). 'The internal structure of phonological elements : atheory of Charm and Government'. *Phonology Yearbook* 2 : 305-328.

Kenstowicz, M. (1994). *Phonology in Generative Grammar.* Oxford: Blackwell.

Kroeger, P. R. (1988). 'Sonority, strength, and the Malay-Javanic nasal prefixes'. Ms. University of Stanford.

Lass, R. (1976). *English Phonology and Phonological Theory.* Cambridge : Cambridge University Press.

Levin, J. (1984). 'Conditions on syllable structure and categories in Klamath phonology', In Mark Cobler, Susannah MacKaye, Michael T. Wescoat, eds., *Proceedings of the West Coast Conference on Formal Linguistic, Volume Three.* Stanford: The Stanford Linguistics Association.

Levelt, W. (1989). *Speaking: from intention to articulation.* Cambridge, MA: MIT Press.

Levelt, W. and and Wheeldon, L. R. (1994). 'Do speakers have access to a mental syllabary?' *Cognition 50*: 239-269.

Liberman, M. (1975). *The intonational system of English.* Doctorial dissertation, MIT. Distributed by the IULC.

Liberman, Mark and Alan Prince (1977). 'On stress and linguistic rhythm'. *Linguistic Inquiry 8*: 249-336.

Lléo, C. (1990). 'Homonymy and reduplication'. *Journal of Child Language 17*: 267-78.

Lufti Abas (1985). *Linguistik Deskriptif dan Nahu Bahasa Melayu.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

McCarthy, John and Alan Prince (1986). 'Prosodic Morphology'. MS, University of Massachusetts and Brandeis.

McCarthy, John and Alan Prince (1993). 'Generalized Alignment'. In Geert Booij and Jaap Van Marle, eds., *Yearbook of Morphology 1993*. Dordrecht: Kluwer Academic Publishers.Pp. 79-153.

McCarthy, John and Alan Prince (1994). 'The emergence of the unmarked: Optimality in prosodic morphology'. In Mercè Gonzàlez, ed., *Proceedings of the North East Linguistic Society 24.* Amherst, MA: Graduate Linguistic Student Association. Pp. 333-379.

McCarthy, John and Alan Prince (1995). 'Prosodic Morphology'. In John Goldsmith, ed., *The Handbook of Phonological Theory.* Oxford: Basil Blackwell.

McHugh, B. (1990). 'The phrasal cycle in Kirunjo Chaga tonology'. In S. Inkelas and D. Zec (eds.), *The Phonology-Syntax Connection* (pp-217-242). Chicago: University of Chicago Press.

Mohanan, K. P. (1986). *The Theory of Lexical Phonologi*, Netherlands: D.Reidel Publishing Company.

Nespor, M. and Vogel, I. (1986). *Prosodic Phonology*, Dordrecht: Foris.

Nik Safiah Karim. (1995). *Malay Grammar for Academics and Professionals.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Nik Safiah Karim, Farid M. Onn and Hashim Hj. Musa. (1989). *Tatabahasa Dewan, Jilid 2: Perkataan.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Pulleyblank, D. (1986). *Tone in Lexical Phonology*, Dordrecht: D. Reidel Publishing Company.

Prince, A. (1983). 'Relating to the grid'. *Linguistic Inquiry* 14: 19-100.

Prince, A. (1985). 'Improving tree theory'. In M. Niepokuj, M. VaClay, V. Nikiforidou, and D. Jeder (ds.), *Proceedings of BLS 11* (pp. 471-490). Berkeley: BLS.

Roca, I. (1994). *Generative Phonology*, London: Routledge.

Rose, Y. (2000). *Headedness and Prosodic Licensing in the L1 Acquisition of Phonology.* Ph.D. thesis, McGill University.

Rosenthall, S. (1989). *The Phonology of Nasal-Obstruent Sequences.* MA thesis, McGill University.

Sagey, E. (1986). *The Representation of Features and Relations in Non-linear Phonology*. PhD dissertation, MIT. Published in 1990. New York: Garland.

Sagey, E. (1988). 'On the ill-formedness of crossing association lines'. *Linguistic Inquiry* 19: 109-118.

Schane, S. S. (1973). *Generative Phonology*. Englewood Cliffs, NJ: Prentice-Hall.

Scullen, M. E. (1987). *French Prosodic Morphology: a unified account.* Bloomington: IULC.

Selkirk, E. (1980). 'The role of prosodic categories in English word stress'. *Linguistic Inquiry* 11: 563-605.

Selkirk, E. (1981). 'On prosodic structure and its relation to syntactic structure'. In *Nordic Prosody II*, T. Fretheim (ed.), 111-140. Trondheim: TAPIR.

Selkirt, E. (1984). *Phonology and Syntax: The Relation between sound and structure.* Cambridge, MA: MIT Press.

Spencer, A. (1996). *Phonology: Theory and Description*, Oxford: Blackwell.

Steriade, D. (1982). *Greek Prosodies and the Nature of Syllabification.* PhD dissertation, MIT.

Steriade, D. (1987). 'Vowel tiers and Geminate blockage'. Unpublished MS. MA: MIT.

Sauzet, P. (1996). 'Ordre des mots, ordre dans les mots' Language Française III, L'ordre des mots, 10-35. Paris: Larousse.

Tajul Aripin Kassin. (1996). *Assimilation Phenomena in Malay Dialects: Comparative Account*, Masters dissertation, Department of Language and Linguistics, University of Essex at Colchester.

Teoh Boon Seong. (1994). *The Sound System of Malay Revisited.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Trask, R. L. (1996). *A Dictionary of Phonetic and Phonology*, London and New York: Routledge

Trask, R. L. (1996). *Historical Linguistics*, London: Edward Arnold.

Vihman, M.M., Velleman, S.L., McCune, L. (1994). 'How abstract is child phonology? Towards an integration of linguistic and psychological approaches. 'In M. Yavas (ed.), *First and Second Language Phonology.* San Diego: Singular Publishing Group.

Vihman, M.M. (1996). *Phonological development and the origins of language in the child.* Oxford: Blackwell.

Vihman, M. M. (2001). 'Commentary on colloqium on bilingual acquisition of phonology.' 3rd Annual Symposium on Bilingualism, University of the West of England, April 2001.

Yunus Maris. (1980). *The Malay Sound System.* Kuala Lumpur: Fajar Bakti.

Wiese, R. (1996) *The Phonology of German.* Oxford: Claredon Press.

Wilbur, R. (1973). *The Phonology of Reduplication.* PhD. dissertation, University of Illinois, Urbana-Champaign. Distributed by Indiana University Linguistics Club.

Zaharani Ahmad. (1991). *The phonology and Morphology of the Perak Dialect,* Kuala Lumpur: Dewan Bahasa dan Pustaka.

Zaharani Ahmad (1992). *Fonologi Generatif.* (ed.) Kuala Lumpur: Dewan Bahasa dan Pustaka.

Zaharani Ahmad. (1996). *Teori Optimaliti dan Analisis Deretan Vokal Bahasa Melayu,* Jurnal Dewan Bahasa, Kuala Lumpur: Dewan Bahasa dan Pustaka.

Zaharani Ahmad (1998). *Phonology and Morphology Interface in Malay an Optimality Theoretic Account.* PhD. dissertation, Department of Language and Linguistics, University of Essex at Colchester.

Appendix I

Details of the Speakers

1. Name: **Norasniza Sailan**
   Address: Pt. Hj. Hassan,
   Batu Pahat,
   Johor
   Occupation: 3rd. Year student - Universiti Utara Malayasia

2. Name: **Dzulhaliezad Iskandar**
   Address: 86B, Kg. Tengah,
   Km. 5, Jalan Mersing,
   86000 Kluang,
   Johor.
   Occupation: 3rd. year student - Universiti Utara Malaysia

3. Name: **Maslita Abd. Aziz**
   Address: 15, Jln. Rotan Tekok,
   Taman Sri Jaya,
   Batu Pahat,
   Johor.
   Occupation: lecturer - Universiti Utara Malaysia

4. Name: **Tuan Zalizam Tuan Muda**
   Address: 74, Jln. Telipok,
   Taman Sri Mersing,
   86800 Mersing,
   Johor
   Occupation: lecturer - Universiti Utara Malaysia

5. Name: **Rahmat Ismail**

Address: 1001 Pontian Besar Kiri,

82000 Pontian,

Johor.

Occupation: Assistant Director of Labour and Manpower Department.

Appendix II

DATA SOURCES

(Word lists used in conversations)

A. Vowel sequences in SM

1. Morphological Simple

a.	/baek/	[bá.eʔ]	'good'
	/kaet/	[ká.etˀ]	'knit'
	/kaen/	[ká.en]	'cloth'
	/laen/	[lá.en]	'other'
	/maen/	[mā́.ẽn]	'play'
	/siang/	[sí.jaŋ]	'day'
b.	/belai/	[bə.lái]	'cherish'
	/buai/	[bú.wai]	'swing'
	/cuai/	[tʃú.wai]	'careless'
	/lambai/	[lám.bai]	'wave'
	/lalai/	[lá.lai]	'neglect'
	/nilai/	[nī́.lai]	'value'
c.	/bau/	[bá.u]	'smell'
	/baor/	[bá.o:]	'mixed'
	/laok/	[lá.oʔ]	'cooked fish'
	/maot/	[mā́.õtˀ]	'lethal'
	/saok/	[sá.oʔ]	'netted'
	/saoh/	[sá.oh]	'anchor'

d.	/halau/	[há.lau]	'eject'
	/ɾisau/	[ɾí.sau]	'worry'
	/pulau/	[pú.lau]	'island'
	/ɾandʒau/	[ɾán.dʒau]	'stake'
	/kilau/	[kí.lau]	'shine'
	/katʃau/	[ká.tʃau]	'disturb'
e.	/sətia/	[sə.tí.jə]	'loyal'
	/tʃəɾia/	[tʃə.ɾí.jə]	'happy/beam'
	/siapa/	[si.já.pə]	'who'
	/biak/	[bí.jaʔ]	'breed'
	/dia/	[dí.jə]	'he/she'
	/liaɾ/	[lí.ja:]	'wild'
f.	/biol/	[bí.jol]	'stupid'
	/siol/	[sí.jol]	'whistle'
	/tiop/	[tí.jopˀ]	'blow'
	/ɾioh/	[ɾí.joh]	'noisy'
	/piot/	[pí.jotˀ]	'great grand grandchild'
	/piutaŋ/	[pi.jú.taŋ]	'debtor'
g.	/kuala/	[ku.wá.lə]	'confluence'
	/buah/	[bú.wah]	'fruit'
	/buat/	[bú.watˀ]	'make'
	/muat/	[mṹ.wãtˀ]	'load/fill'
	/puaka/	[pu.wá.kə]	'haunt'
	/tua/	[tú.wə]	'old'

h.	/bueh/	[bú.weh]	'babble'
	/kueh/	[kú.weh]	'cake'
	/duet/	[dú.wet']	'money'
	/kues/	[kú.wes]	'push'
i.	/amboi/	[ám.boi]	'expression of exciting'
	/kaloi/	[ká.loi]	'a kind of fish'
	/tampoi/	[tám.poi]	'name of place'
	/dodoi/	[dó.doi]	'lullaby'
j.	/soal/	[só.ʔal]	'ask'
	/doʕa/	[dó.ʔã]	'invocation'

2. Loan Words

/mãʕãf/	[mã́.ʔãp]	'forgiveness'
/aulia/	[au.lí.ja]	'a good person'
/faʕẽdah/	[fa.ʔẽ́.dah]	'interest'
/geologi/	[geo.ló.gi]	'geology'
/saʕãt/	[sá.ʔãt]	'moment'
/ɣaeb/	[ɣá.ʔẽb]	'occult/supernatural'
/aidia/	[ai.dí.jə]	'idea'
/nasional/	[nã̀.si.jó.nãl]	'national'
/taʕãt/	[tá.ʔãt]	'loyal'
/mũʕãlaf/	[mũ.ʔã́.lap]	'convert'
/kreatif/	[krẽa.tif]	'creative'
/radio/	[ra.dí.jo]	'radio'
/traelər/	[tráe.lə:]	'trailer'
/saens/	[sá.ens]	'science'

/automatik/	[àu.to.mǎ.ti?	'automatic'
/auditorium/	[àu.di.to.ɾí.jum]	'auditorium'
/etis/	[é.tis]	'atheist'
/protin/	[pɾó.tin]	'protein'
/sterio/	[ste.ɾí.jo]	'stereo'
/tiori/	[ti.jó.ɾi]	'theory'
/dialek/	[di.já.le?]	'dialect'
/kriteria/	[kɾí.te.ɾí.ja]	'criterion'
/ertial/	[e.ɾí.jal]	'aerial'
/erograf/	[e.ɾó.graf]	'aerograph'
/disel/	[dí.sel]	'diesel'
/iodin/	[i.jó.din]	'iodine'
/primium/	[pɾi.mí.jum]	'premium'
/stadium/	[sta.dí.jum]	'stadium'
/nion/	[ní.jon]	'neon'
/liŋguistik/	[lìŋ.gu.wís.ti?]	'linguistics'
/gonoria/	[gò.nõ.ɾí.ja]	'gonorrhoea'
/alkaloid/	[al.ká.loid]	'alkaloid'
/paon/	[pá.on]	'pound'
/siɾios/	[si.ɾí.jos]	'serious'
/duet/	[dú.wet/	'duet'
/kualiti/	[kù.wa.lí.ti]	'quality'
/akuaɾium/	[à.ku.wa.ɾí.jum]	'aquarium'
/kuoɾum/	[kó.ɾum]	'quorum'
/kuota/	[kó.ta]	'quota'

3. **Morphological Complex (Prefixes: /di-/, /kə-/, /sə-/ + V...)**

a.	/di+ambel/	[di.ʔám.bel]	'taken'	(passive)
	/di+antara/	[dì.ʔan.tá.rə]	'in between'	(passive)
	/di+aer+i/	[dì.ʔa.é.ri]	'irrigated'	(passive)
	/di+ubah/	[di.ʔú.bah]	'moved'	(passive)
	/di+ulaŋ/	[di.ʔú.laŋ]	'repeated'	(passive)
	/di+ukor/	[di.ʔú.ko:]	'measured'	(passive)
	/di+ikat/	[di.ʔí.kat˺]	'tied'	(passive)
	/di+ikot+i/	[dì.ʔi.kót.ti]	'followed'	(passive)
	/di+iŋat+i/	[dì.ʔi.ŋát.ti]	'remembered'	(passive)
b.	/kə+ada+an/	[kə.ʔà.dá.ʔan]	'situation'	
	/kə+adel+an/	[kə.ʔà.dél.lan]	'justice'	
	/kə+aŋkoh+an/	[kə.ʔàŋ.kóh.han]	'arrogance'	
	/kə+enak+kan/	[kə.ʔè.nấʔ.kan]	'tastefulness'	
	/kə+esok+kan/	[kə.ʔè.sóʔ.kan]	'the next day'	
	/kə+əmas+an/	[kə.ʔə̀.mấs.san]	'golden'	
	/kə+utoh+an/	[kə.ʔù.tóh.han]	'strength'	
	/kə+utama+an/	[kə.ʔù.ta.mấ.ʔan]	'priority'	
	/kə+untoŋ+an/	[kə.ʔùn.tóŋ.ŋãn]	'profitable'	
c.	/sə+andai+ɲa/	[sə.ʔàn.dái.ɲã]	'supposing'	
	/sə+aŋkat+an/	[sə.ʔàŋ.kát.tan]	'troop/peer'	

/sə+ibu/	[sə.ʔí.bu]	'common mother/same mother'
/sə+ia/ - /sə+kata/	[sə.ʔì.jə.sə.ká.tə]	'unanimous'
/sə+ekor/	[sə.ʔé.ko:]	'numeral coefficient for animals'
/sə+əmak/	[sə.ʔə.mấʔ]	'common mother/same mother'
/sə+oraŋ/	[sə.ʔó.raŋ]	'a person'
/sə+olah/	[sə.ʔó.lah]	'as if'
/sə+utas/	[sə.ʔú.tas]	'numeral coefficient for string'
/sə+umpama/	[sə.ʔùm.pá.mõ]	'like'

4. Morphological complex (stem + suffix /-an/, /-i/).

a.	/buai+an/	[bu.wái.jan]	'cradle'
	/bəlai+an/	[bə.lái.jan]	'coaxing/caressing'
	/həlai+an/	[hə.lái.jan]	'piece'
	/kə+rai+an/	[kə.ra.ʔí.jan]	'entertainment'
b.	/bau+an/	[ba.ú.wan]	'perfume'
	/dʒaŋkau+an/	[dʒaŋ.káu.wan]	'snatching'
	/igau+an/	[i.gáu.wan]	'talk in one's sleep'
	/kilau+an/	[ki.láu.wan]	'glitter'
c.	/tʃatu+an/	[tʃa.tú.wan]	'ration'
	/lalu+an/	[la.lú.wan]	'passage'
	/buru+an/	[bu.rú.wan]	'hunting'
	/tiru+an/	[ti.rú.wan]	'imitation'

d.	/lalu+i/	[la.lú.wi]	'pass through'
	/kə+tahu+i/	[kə.ta.hú.wi]	'know'
	/dahulu+i/	[dà.hu.lú.wi]	'leading
	/tər+ɾadʒu+i/	[tə:.ɾa.dʒú.wi]	'head'
e.	/kə+sətia+an/	[kə.sə.tì.ja.ʔán]	'loyalty'
	/kə+tʃəɾia+an/	[kə.tʃə.ɾì.ja.ʔán]	'brightness'
	/kə+sədia+an/	[kə.sə.dì.ja.ʔán]	'readiness'
	/kə+bahagia+an/	[kə.bà.ha.gì.ja.ʔán]	'happiness'
f.	/guna+an/	[gù.nã.ʔán]	'using'
	/kə+suka+an/	[kə.sù.ka.ʔán]	'happiness'
	/kə+ɾadʒa+an/	[kə.ɾà.dʒa.ʔán]	'government'
	/dʒuta+an/	[dʒù.ta.ʔán]	'millions'
g.	/məŋ+luka+i/	[mə̃:.lù.ka.ʔí]	'hurting'
	/səɾta+i/	[səɾ.ta.ʔí]	'to participate'
	/mula+i/	[mṹ.la.ʔí]	'to start'
	/məŋ+lupa+i/	[mə̃:.lù.pa.ʔí]	'to forget'

B. Flap /ɾ/ in SM

1. Morphological Simple

/kəɾtas/	[kəɾ.tás]	'paper'
/səɾta/	[sə́ɾ.tə]	'together'
/cəɾmin/	[cəɾ.mĩn]	'mirror'
/bəɾnas/	[bəɾ.nã́s]	'brilliant'
/pəɾlu/	[pəɾ.lú]	'need'
/kəɾbau/	[kəɾ.báu]	'buffalo'

/pərnah/	[pər.nấh]	'ever'
/harta/	[hár.tə]	'property'
/warta/	[wár.tə]	'news'
/dərma/	[dər.mã]	'donation'
/korban/	[kór.ban]	'sacrifice'
/tərdʒon/	[tər.dʒón]	'dive'
/səpərti/	[sə.pər.tí]	'like'
/kəluarga/	[kə.lu.wár.gə]	'family'
/pərtiwi/	[pər.tí.wi]	'goddess of Earth'
/gərgasi/	[gər.gá.si]	'giant'
/skru/	[skrú]	'screw'
/struktor/	[strúk.to:]	'structure'
/spring/	[spríng]	'spring'
/traktor/	[trák.to:]	'tractor'
/drama/	[drá.mã]	'drama'
/plastər/	[plás.tə:]	'plaster'
/komplek/	[kóm.plek]	'complex'
/hipokrit/	[hi.pó.krit]	'hypocrite'

2. Morphological Complex

a) i.	/bər+ia/	[bə.rí.jə]	'to agree'
	/bər+ibu/	[bə.rí.bu]	'thousands'
	/bər+ulat/	[bə.rú.lat']	'worm-infested'
	/bər+ada/	[bə.rá.də]	'is'
	/bər+ekor/	[bə.ré.ko:]	'tailed'
	/bər+oleh/	[bə.ró.leh]	'get'
ii.	/pər+aku/	[pə.rá.ku]	'to admit'

/pər+ikat+an/	[pə.ɾi.kát.tan]	'federation'
/pər+elok/	[pə.ɾé.loʔ]	'to beautify'
/pər+untok+kan/	[pə.ɾun.tóʔ.kan]	'allocation'
iii. /tər+ambel/	[tə.ɾám.bel]	'taken'
/tər+uloŋ/	[tə.ɾú.loŋ]	'experienced'
/tər+ikat/	[tə.ɾí.kat̚]	'bound to'
/tər+aŋkat/	[tə.ɾáŋ.kat̚]	'lifted'
b). i. /bər+buru/	[bə:.bú.ru]	'hunting'
/bər+guna/	[bə:.gú.nə̃]	'useful'
/bər+dua/	[bə:.dú.wə]	'together'
/bər+puasa/	[bə:.pu.wá.sə]	'fasting'
ii. /pər+kata+an/	[pə:.kà.ta.ʔán]	'word'
/pər+pustaka+an/	[pə:.pùs.ta.ka.ʔán]	'library'
/pər+satu+an/	[pə:.sa.tú.wan]	'organisation'
/pər+waris/	[pə:.wá.ɾis]	'legatee'
iii. /tər+paku/	[tə:.pá.ku]	'nailed'
/tər+buka/	[tə:.bú.kə]	'opened'
/tər+lupa/	[tə:.lú.pə]	'forgot'
/tər+dʒumpa/	[tə:.dʒúm.pə]	'found'
/tər+nampak/	[tə:.nǎm.paʔ]	'seen'
c). i. /kibar+an/	[ki.bá.ɾan]	'flapping'
/hantar+an/	[han.tá.ɾan]	'sending'
/ekor+an/	[e.kó.ɾan]	'consequence'

/kə+bakar+an/	[kə.ba.ká.ran]	'fire'
ii. /aer+i/	[a.é.ri]	'to irrigate'
/məŋ+ekor+i/	[mə̃.ŋẽ.kó.ri]	'to follow'
/tabor+i/	[ta.bó.ri]	'to spread'
/məŋ+tʃəbor+i/	[mə̃ɲ.tʃə.bó.ri]	'to become involved in'
d). i. /gugor+kan	[gu.gó:.kan]	'drop'
/tabor+kan/	[ta.bó:.kan]	'scatter'
/hantar+kan/	[han.tá:.kan]	'send'
/tukar+kan/	[tu.ká:.kan]	'change'
ii. /subor+ɲa/	[su.bó:. ɲə̃]	'the flourishing'
/kubor+ɲa/	[ku.bó:. ɲə̃]	'his/her grave'
/sukar+ɲa/	[su.ká:. ɲə̃]	'the difficult'
/gambar+ɲa/	[gam.bá:. ɲə̃]	'the picture'

C. Vowel initial

/asah/	[á.sah]	'sharpen'
/aleh/	[á.leh]	'move'
/asoh/	[á.soh]	'take care of'
/əmas/	[ə.más]	'gold'
/elak/	[é.laʔ]	'avoid'
/esok/	[é.soʔ]	'tomorrow'
/ingat/	[í.ŋãtˀ]	'remember'
/intep/	[ín.tepˀ]	'spy'
/ikot/	[í.kotˀ]	'follow'
/otak/	[ó.taʔ]	'brain'

/oleh/	[ó.leh]	'by'
/oraŋ/	[ó.raŋ]	'human being'
/ubat/	[ú.bat˺]	'medicine'
/usek/	[ú.seʔ]	'to tease'
/untok/	[ún.toʔ]	'for'
/eksport/	[ék.spot˺]	'export'
/elektrik/	[i.léʔ.triʔ]	'electric'
/industri/	[in.dús.tri]	'industry'

Appendix III

Tapes of Word Lists and Conversations

Two tapes of the word lists and conversation recorded in 1998 in Malaysia are included (see Appendix II for the lists of words).