Prefixation and Reduplication in Malay: an Optimality-Theoretical Account

Sharifah Raihan Syed Jaafar

B.A (Linguistics), M.A (Linguistics), National University of Malaysia, Malaysia

A thesis submitted for the Degree of Doctor of Philosophy

Department of Linguistics and English Language

Lancaster University

United Kingdom

March 2011

ABSTRACT

This thesis investigates the morphology-phonology interface in Malay. The work is largely a corpus-based reanalysis of prefixation and reduplication. Based on two large different written corpora of Standard Malay (henceforth SM), the analysis permits us to make reliable and robust generalizations about how the language actually works. The data reveal that the language has a distinct co-existing phonological system. I will show that these co-existent grammars can be handled with Optimality Theory (henceforth OT), specifically in co-phonologies.

The reanalysis of prefixation places Malay in a wider context and examines, cross-linguistically, issues related to voicing and nasality. It is shown that nasal substitution, which is regularly used to eliminate nasal and voiceless obstruent clusters, fails to occur in some prefixed words. In the analysis, I propose that non-native words are not subject to the same phonological requirements as those imposed on native words. The constraint rankings must therefore be different from those found in native words which result in the blocking of nasal substitution at prefix-root junctures. The application of nasal substitution at prefix-prefix junctures is mainly determined by morphological factors rather than phonetic factors, due to a morphology-phonology interface constraint, i.e. EDGE-INTEGRITY.

The investigation of reduplication deals with total, partial and affixal reduplication. A recent theory of prosodic morphology – namely the Morpheme-Based Template or MBT (Downing 2006), motivated within OT (Prince & Smolensky 1993) – is applied to organize the morphological and prosodic factors that condition the size of prosodic morphemes (ibid.: 1). In the analysis, I propose that total and

affixal reduplication are best treated as compounding, rather than affixation, due to the disyllabic minimality condition. Considering the Perak dialect, light and heavy reduplication have been captured by associating each morphological construction with a different co-phonology.

This study also examines dialectal variation, comparing SM with three non-standard dialects with respect to prefixation. The analysis discovers some significant facts about the language. Since both voiceless/voiced obstruents undergo nasal substitution, the *N$\mathring{\text{Ç}}$ constraint has been replaced by a CRISP-EDGE[σ] constraint. Nasal deletion and nasalisation are also the strategies used to eliminate nasal and voiceless obstruent clusters. The different strategies applied can be satisfactorily explained in OT with its variable constraint rankings.

ACKNOWLEDGMENTS

I owe my deepest gratitude to a number of generous people who have largely given their support and contribution in completing this thesis. First and foremost I offer my greatest appreciation to my supervisor, Professor Francis X. Katamba, who has patiently supported and guided me throughout my thesis journey.

I would also like to thank *Dewan Bahasa dan Pustaka* (The Institute of Language and Literature) for allowing me to access their corpus data. I am also grateful to Associate Professor Ajid Che Kob, Associate Professor Dr. Mohd Fadzeli Jaafar and Mr. Zulkifli Ahmad for their contribution to my data supplement for the dialectal study. Without their cooperation I could not have obtained such pertinent data for my analysis. Special thanks go to Dr. Andrew Hardie, from Lancaster University, for his invaluable assistance with the corpus software, AntConc. His explanations of the use of the software have been much appreciated as this has made my analysis work much easier. Not forgotten are Professor Dr. Zaharani Ahmad, Deputy Dean of the Faculty of Social Science and Humanities from the National University of Malaysia (UKM), for his kind advice and encouragement throughout my study, and all the lecturers in the Linguistics Department at UKM.

I would like to dedicate my heartfelt thanks to my lovely parents, Syed Jaafar and Ramlah Saadon, and to my sisters, Sharifah Khairul Adibah, Sharifah Nabihah and Sharifah Nawal, for their enduring love and continuous support. Without them, the completion of this thesis would not have been possible.

I am indebted to the Ministry of Higher Education, Malaysia and the UKM for awarding me the SLAB scholarship and study leave to pursue my doctoral studies at Lancaster University, United Kingdom.

LIST OF TABLES

LIST OF FIGURES

LIST OF MAPS

ABBREVIATIONS

OT	Optimality theory
GTT	Generalised Template Theory
PBT	Prosodic-Based Template
MBT	Morpheme-Based Template
DBP	*Dewan Bahasa dan Pustaka* (The Institute of Language & Literature)
UKM	*University Kebangsaan Malaysia* (National University of Malaysia)
SM	Standard Malay
NS	Negeri Sembilan
PD	Perak dialect
ACT.	ACTIVE
PRF.	PREFIX
SUF.	SUFFIX
NOM	NOMINAL
VERBL	VERBALISER
CAUS.	CAUSATIVE
LOC.	LOCATIVE
ADJ.	ADJECTIVE
STEMEX	STEM EXTENDER
RED	REDUPLICANT
NAS ASS	NASAL ASSIMILATION
UG	UNIVERSAL GRAMMAR

Table of Contents

1. INTRODUCTION

1.1 Introduction

This study re-examines and re-analyses two morphological processes of word formation, prefixation and reduplication, in Malay. The analysis of prefixation includes both single and multiple prefixes which end with a nasal, as they exhibit interesting morpho-phonemic alternations. The discussion of reduplication covers the three types of reduplication that occur in Malay: total, partial and affixal. Previously, prefixation and reduplication have been widely discussed under conventional rule-based and non-linear autosegmental approaches. In this present study, however, these morphological processes will be given a new theoretical account according to Optimality Theory (henceforth OT). A number of issues and problems raised by prefixation and reduplication (see Section 1.3) will be captured more adequately under OT. In the theoretical framework of OT, grammars vary, one from another. This essential idea is one of the reasons for choosing OT to examine variation (i.e. different patterns of output) within the Malay language. The issue that will be taken up here is the variation yielded by the mean to avoid nasal and voiceless obstruent clusters. In this case, three dialects of Malay are observed: Perak, Kelantan and Negeri Sembilan (henceforth NS).

Two of the discussion topics above, i.e. prefixation and reduplication, will be analysed by using data obtained from the DBP-UKM (*Dewan Bahasa dan Pustaka – Universiti Kebangsaan Malaysia*) corpus. Since the corpus only provides data for standard Malay (henceforth SM), all the dialectal data used in this study, i.e. in Perak dialect partial reduplication and dialect variation which are discussed in Chapter 5 (see

Section 5.4) and Chapter 6, are based on previous scholars' work and the brief interviews[1] I conducted, respectively.

This study attempts to present a different way of analysing the three topics mentioned above, using an up-to-date theoretical framework developed within OT, namely Morpheme-Based Template (MBT), one of the versions of Generalised Template Theory (GTT) (Downing, 2006). This theory provides a complete explanation of the correlation between canonical morpheme form and prosodic morpheme shape, as well as word minimality condition. Prefixation and reduplication were chosen to be analysed in this study as MBT is seen as an adequate theory to account for the problems (see Section 1.3) raised in those topics. The results from the analysis of prefixation are essential for the next analysis i.e. reduplication. As the ideas proposed in MBT work well in accounting for prefixed words, I would also like to test how those ideas account for prefixed words that have been reduplicated. I shall discuss the theory of MBT further in the following chapter (see subsection 2.2.4).

This opening chapter offers an essential overview of the discussion developed in the subsequent chapters of this thesis. As the thesis is about Malay, I introduce some useful information about the language in Section 1.2. In Section 1.3, I present a statement of the problems which have arisen in both rule-based and autosegmental analyses. Section 1.4 highlights the research questions. Section 1.5 outlines the goals of the present work, while the last section, 1.6, presents the overall structure of this thesis.

[1] See Chapter 3 for details of what sort of data are used for the analyses of non-SM.

1.2 Introduction to Malay

Malay is a member of the Malayic sub-group of the Malayo-Polynesian branch of the Austronesian language family. It is widely used in a number of countries including Malaysia, Indonesian, Brunei, Singapore and surrounding areas. As stated in Act 152 of the Federal Constitution of Malaysia, Malay is the national and the official language of Malaysia. It is also the national language of Brunei and Indonesia and is one of the official languages of Singapore. The language has different names in those countries. Malay is called *Bahasa Malaysia* (Malaysia language) and *Bahasa Indonesia* (Indonesia Language) in Malaysia and Indonesia, respectively. In Malaysia, the term *Bahasa Malaysia* was used until the 1990s when it changed to *Bahasa Melayu* (Malay language). For certain political reasons, *Bahasa Melayu* had to revert to *Bahasa Malaysia*, the language which represents all races in Malaysia. Today, the Malay language is called *Bahasa Malaysia* and is mostly spoken in the Malay Peninsula and the coastlands of Sabah and Sarawak (Ahmad, 2005).

Like any other natural language, Malay has various regional and social dialects. The dialects of Malay coincide with the division of Malaysia into various states (cf.: Omar, 1993; Ahmad, 2005). Thus, the names of the dialects are the same as the names of the states in Malaysia, for example, Kedah dialect, Perak dialect, Terengganu dialect and so on (see Map 1). The standard version of *Bahasa Malaysia* originates from Johore-Riau Malay, which is mainly spoken in the southern part of the Malay Peninsula (Teoh, 1994; Ahmad, 2005). This dialect of Johore-Riau Malay has been chosen by a members community of the language to be the standard dialect of Malay as opposed to various other Malay dialects. This standard dialect is what we call *Bahasa Malaysia*. Speakers of Malay use *Bahasa Malaysia* to refer to standard

Malay (henceforth SM[2]). As the standard version of Malay, *Bahasa Malaysia* is used on formal occasions, for example in administration, education and the mass media. The reason why the Johore-Riau dialect was preferred over other dialects of Malay is because the morphological and syntactical of this dialect are closer to literary Malay (henceforth LM) (cf.: Omar, 1975; Onn, 1980; Ahmad, 2005). Between these two, Johore-Riau dialect is largely pronounced by Malay speakers as their daily communication form than LM. In what follows, let us see in what ways this dialect is close to LM.

Map 1: Malay dialects in the Malay Peninsula.

It has been observed that there are only four differences between the standard dialect and LM (Ahmad, 2005: 4). First, /a/ in word-final position is always retained as [a] in LM. In SM on the other hand, /a/ is realised as a schwa [ə]. I briefly present some relevant examples of this:

[2] In this study, 'Malay' is used as to refer to SM, including its dialects.

1. /a/ in word-final position.

	LM	SM
/saja/ 'I/me'	[saja]	[sajə]
/meʤa/ 'table'	[meʤa]	[meʤə]
/baʧa/ 'read/	[baʧa]	[baʧə]

Second, /r/ in word-initial and medial position is pronounced as a flap[ɾ] in LM, while in SM it is realised as a velar fricative [ɣ]. Third, /r/ in word-final position is always retained in LM, while in SM this segment is never pronounced. Let us see examples of each case, below:

2. a) /r/ in word-initial and medial position.

	LM	SM
/rumah/ 'house'	[ɾumah]	[ɣumah]
/kərtas/ 'paper'	[kəɾtas]	[kəɣtas]
/rasa/ 'taste'	[ɾasa]	[ɣasa]

b) /r/ in word-final position (Ahmad, 2005: 59):

	LM	SM
/kotor/ 'dirty'	[kotor]	[koto:]
/uker/ 'to carve'	[uker]	[uke:]
/pasar/ 'market'	[pasar]	[pasa:]

Fourth, the high front vowels, /i, u/ in the closed final syllable in LM, correspond to the input [i,u]. In SM however, they are realised as mid-vowels [e, o], as the following examples illustrate:

3. /i, u/ in closed-final syllable.

	LM	SM
/pilih/ 'choose'	[pi.lih]	[pi.leh]
/putih/ 'white'	[pu.tih]	[pu.teh]
/buluh/ 'bamboo'	[bu.luh]	[bu.loh]
/masuk/ 'to enter'	[ma.suk]	[ma.soʔ]

In Malaysia, SM which is largely used in most formal occasions as mentioned above has been taught to all students regardless of what their mother tongue is, in both primary and secondary school. This means that speakers who have already acquired their native mother tougue like any of Malay dialects before going to school, SM would be their second dialect. This group of speakers thus know more than one grammar, one is SM and the other one is the dialect they have acquired at home. Since the speakers know more than one dialect or grammar, they therefore have variation in their utterance.

As the national and official language of Malaysia, the Malay language or *Bahasa Malaysia* has undergone a long process of development towards its function as the national and official language. The development of Malay has changed the nature of the language with the infusion of Sanskrit, Portuguese, Chinese, Arabic, Javanese, English and many others.[3] In recent times, English words have been widely absorbed into the vocabulary of Malay. This happens when there are no appropriate Malay words that can be used, e.g. for specific words and scientific terms. Examples of such words are *concordance* (konkordan), *corpus* (korpus), *accelerometer*

[3] From http://ms.wikipedia.org/wiki/Bahasa_Melayu.

(akselerometer) and *probe* (prob).[4] This type of borrowing has become increasingly important to Malay, as demonstrated by the many new words and technology terms which now exist.[5] In such cases, borrowing involves using entire words from the source language. As well as borrowing words, Malay also borrows certain phonemes from other languages. All the borrowed phonemes are consonants, such as /f/, /ʃ/ and /x/ as in *factor* 'factor' (from English), *ʃarikat* 'company' (from Arabic) and *xamis* 'Thursday' (from Arabic), respectively. As stated in Ahmad (2005: 16), Malay has sixteen underlying consonants: /p, b, t, d, k, g, ʧ, ʤ, s, h, m, n, ɲ, ŋ, l and r/, and six vowels: /i, u, e, o, ə and a/. The consonants /f/, /ʃ/ and /x/ are not underlying consonants of Malay. Therefore, they are said to be borrowed phonemes in Malay.

It is worth knowing that the basic structure of Malay syllables is CV(C) (Teoh, 1994). This basic syllable structure signifies that the language disfavours consonant and vowel clusters in words. Vowel epenthesis and consonant deletion have always been used as ways to break up consonant and vowel clusters, respectively. Therefore, in all borrowed words, especially from English, which have consonant clusters, schwa is epenthesized between clusters, as in *class, glass, stem, sink,* to become [kəlas], [gəlas], [sətem] and [sinki], respectively. Meanwhile, words with vowel clusters which have been borrowed from Arabic, such as *saat* 'minute', *taat* 'obey', *jumaat* 'Friday' and *maaf* 'forgive', are pronounced as [saʔat], [taʔat], [jumaʔat] and [maʔaf], respectively, in Malay. The implication of the basic syllable structure CV(C) is that syllables in Malay can be CV or CVC, while syllable structures such as *V and *VC are not allowed (Ahmad, 1989). In the sound system of Malay, however, CV is favoured over CVC. This means that CV is an unmarked type of syllable in Malay.

[4] From Halimah Hj Ahmad at www.dbp.gov.my/mabbim/download.php?FilePoolID=139.

[5] With the awareness of how important science and technology is to the development of the country, most of scientific and technology terms have widely been used in Malay.

1.3 Statement of the problem

Prefixation and reduplication are seen as the most productive ways of forming new words in Malay.[6] Therefore, these morphological processes have received much formal attention from many Malay scholars. This can be seen in works such as: Hassan (1974, 1987); Othman (1981); Koh (1981); Yeoh (1988); Omar (1975, 1986, 1993); Karim et al. (1989, 1994); Karim (1995); Teoh (1994); Teoh and Ahmad (2006); Ahmad (1993, 1999, 2000a, 2000b, 2005). As productive ways to form new words, prefixation and reduplication raise a number of interesting phonological and morphological issues in Malay grammar. This section will list what the issues are, by showing how they have been treated unsatisfactorily in previous approaches, i.e, in ruled-based and autosegmental analyses. Each approach is discussed in subsections 1.3.1 and 1.3.2, respectively.

1.3.1 Rule-based analysis

We begin with prefixation. In most Austronesian languages including Malay, two types of clusters are disfavoured: (1) nasal and voiceless obstruents, and (2) nasal and sonorant clusters. In Malay, these clusters occur in three morphological environments; these are root-internal, prefix-root and prefix-prefix junctures. I shall mention here that the only type of clusters that occur at prefix-prefix junctures are nasal and voiceless obstruents. Nasal and sonorant clusters are not found at these junctures. Note that the occurrences of these two clusters in the two morphological environments are treated differently in the language.

[6] As well as prefixation and reduplication, compounding is another productive way of word formation in Malay (Karim et al., 1994). This type of word formation will not, however, be examined in this present study. For interested readers, they should consult work such as that of Ahmad (1999).

In Malay, nasal substitution is regularly applied to avoid nasal and voiceless obstruent clusters from emerging in surface representation at prefix junctures. Nasal substitution is the process of replacing a root-initial voiceless obstruent by a homorganic nasal (Pater 1999), as exemplified in (4a). If the root-initial consonant is voiced, then the nasal segment in the prefix assimilates to the place of articulation of the voiced obstruent. This results in a homorganic cluster in the surface representation, as in (4b).

4. (a) Nasal and voiceless obstruent clusters in Malay (from the DBP-UKM corpus).[7]

i) /məŋ-təmu-i/ [mə-nəmuwi]
ACT.PRF-meet-LOC.SUF 'to cause to meet'

ii) /məŋ-potoŋ/ [mə-motoŋ]
ACT.PRF-cut 'to cut'

(b) Nasal and voiced obstruent clusters.

i) /məŋ-dapat/ [məndapat]
ACT.PRF-get 'to get'

ii) /məŋ-bəri/ [məmbəɣi]
ACT.PRF-give 'to give'

In contrast, the occurrence of nasal plus voiceless obstruent clusters in root-internal situations in Malay is not resolved by nasal substitution. The cluster is resolved by a homorganic nasal, instead. The following examples in (5) show that root-internal occurrences of this cluster are permitted in the language. Yet, the examples in (5) prove that nasal and voiceless obstruent clusters are not entirely disallowed in Malay. This point will be discussed further in Chapter 6, subsection 6.3.1, particularly as to why this cluster emerges within roots.

---

[7] The examples are taken from the corpus used in this study i.e. the DBP-UKM corpus, see Chapter 3 for details.

5. Nasal and voiceless root-internal obstruent clusters (from the DBP-UKM corpus).

i)	/təmpatan/ 'local'	[təm.pa.tan]
ii)	/kumpulan/ 'group'	[kum.pu.lan]
iii)	/məntəri/ 'minister'	[mən.tə.ɣi]
vi)	/hamparan/ 'spread out'	[ham.pa.ɣan]
v)	/səməntara/ 'while'	[sə.mən.ta.ɣə]
iv)	/təntəram/ 'peaceful'	[tən.tə.ɣam]

The above examples, taken from the DBP-UKM corpus, clearly show that nasal and voiceless obstruent clusters emerge in the surface representation in root-internal position. It should be mentioned that the corpus I observe shows that the clusters also occur at prefix-root junctures in the language as well as within morphemes. Examples of the occurrence of nasal and voiceless obstruent clusters at prefix-root junctures are presented in (9).

The other clusters that the language does not permit are nasal and sonorant clusters. Nasal deletion is the strategy that the language uses to remove nasal and sonorant clusters, by deleting the nasal segments that precede sonorant consonants. I exemplify this with some relevant examples:

6. Nasal and sonorant clusters in SM (from the DBP-UKM corpus).

i)	/məŋ-lapor/ ACT.PRF-report 'to report'	[məlapo]
ii)	/məŋ-ŋuap/ ACT.PRF-yawn 'to yawn'	[məŋuwap]
iii)	/məŋ-ɲaɲi/ ACT.PRF-sing 'to sing'	[məɲaɲi]

In the above examples, we see that nasal and sonorant clusters occur at prefix junctures. Nasal segments that precede sonorant consonants are always deleted in order to get rid of such clusters. Now we see how the clusters behave within the roots. It has been claimed, in previous Malay studies (e.g.: Teoh, 1994; Ahmad, 1989), that root-internal nasal and sonorant clusters like those shown in (7) are found in Malay. However, the words in (7) are claimed to be non-native Malay words which have been borrowed from Indian and Chinese, as exemplified in 7(a) and 7(b), respectively.

7. Nasal and sonorant root-internal clusters in Malay (from Ahmad, 1989).

a) /maŋli/ 'Indian astrology' [maŋ.li]

b) /taŋloŋ/ 'tanglong' [taŋ.loŋ]

We now see how the two disfavoured clusters, (1) nasal and voiceless obstruents and (2) nasal and sonorant clusters, have been analysed in rule-based analysis. We examine first how nasal and voiceless obstruent clusters have been treated.

As illustrated in (5), since nasal and voiceless obstruent clusters in root-internal position are not banned in the language, no analysis concerning the clusters in this morphological environment is made in rule-based analyses. As postulated in these, in the process of prefixation to eliminate a sequence of nasal and voiceless obstruent, two rules are involved: (1) Nasal Assimilation, and (2) Voiceless Obstruent Deletion; these have to be applied in order. I show below how these two rules apply:

8. Nasal substitution in rule-based analysis.

Input	/məŋ-təmu-i/
1) Nasal assimilation	mən-təmu-wi
2) Voiceless obstruent deletion	mən-əmu-wi
Output	[mənəmuwi]

From the ordering of rules above, correct output is obtained whereby nasal and voiceless obstruent clusters do not emerge on the surface. In this study, I will argue that the analysis proposed by scholars for this group does not work for some prefixed words. As observed in the DBP-UKM corpus, there are counter-examples where the clusters emerge in the surface representations. The occurrence of nasal and voiceless obstruent clusters in some prefixed words, as in (9), poses a question, as the language does not allow clusters to emerge in the surface representation, yet there are counter-examples showing the presence of clusters on the surface. The examples in (9) are a second form of evidence showing that nasal and voiceless obstruent clusters are not entirely prohibited from the language (see subsection 4.2.1.2 for more data).

9. (i) /məŋ-tadbir/ [mən-tadbe]
ACT.PRF-administrative
'to administer'
(ii) /məŋ-protes/ [məm-pγotes]
ACT.PRF-protest
'to protest'

The voiceless obstruents [t] and [p] in the above examples remain undeleted after the assimilated nasal. This produces a sequence of nasal and voiceless obstruents in the word. I am certain that the situation described above is difficult to capture in rule-based analysis, where structural condition and structural change are linked by rules (Kager, 1999: 55). The rule, in this case, is that nasal assimilation and voiceless obstruent deletion then determine the structural change in words with response to the structural condition. In constraint-based theory, the structural change and structural condition are evaluated by constraint violations. In OT therefore, various types of structural changes may be triggered by a markedness constraint, depending on its interaction with the faithfulness constraint (ibid.). In accounting for nasal and voiceless obstruent clusters, the markedness constraint, *NÇ̥ (see subsection 2.4.1.1), is used to prevent the clusters from emerging in the surface representation. As nasal and voiceless root-internal obstruent clusters in Malay are not resolved by nasal substitution, as shown in (5), so *NÇ̥ must be ranked beneath the faithfulness constraint. To appreciate how those constraints interact in a constraint ranking, see Section 6.3.1, where a full discussion is offered.

Given the two pieces of evidence for the occurrence of nasal and voiceless obstruent clusters presented in (5) and (9), I would like to pose a question: why do nasal and voiceless obstruent clusters emerge in the surface representation although such clusters are banned in the language? The two cases where nasal and voiceless obstruent clusters emerge in the surface representation have been ignored in rule-based analysis. As far as I know, those cases have also not been discussed in autosegmental analyses, neither by Teoh (1994) nor by Ahmad (2000b). As I will show in 1.3.2, autosegmental analyses merely focus on nasal and voiceless obstruent clusters in monosyllabic words.

In explaining the process of prefixation, scholars using rule-based analyses have also put aside monosyllabic words in their analyses. Since there has been no discussion of monosyllabic words, scholars employing autosegmental analyses (e.g.:

Teoh, 1994; Ahmad, 2000b) have paid more attention to monosyllabic words in their analyses. These focus particularly on when nasal final prefixes, /məŋ+/ or /pəŋ+/, are attached to them, since schwa needs to epenthesize between them, as exemplified in (10), below:

10. (i) /məŋ-ə-**kod/** [mə.ŋə.kod]
   ACT.PRF-STEMEX-code 'to code'
   (ii) /məŋ-ə-**pam**/ [mə.ŋə.pam]
   ACT.PRF-STEMEX-pump 'to pump'
   (iii) /məŋ-ə-**bom/** [mə.ŋə.bom]
   ACT.PRF-STEMEX-bomb 'to bomb'

As we note from the rule-based analysis to account for nasal and voiceless obstruent clusters, nasal substitution seems to be the only way to eliminate these clusters. Perhaps, scholars using rule-based analyses have not taken into account monosyllabic roots. Therefore, they do not see other strategies, such as vowel epenthesis, which are used to avoid nasal and voiceless obstruent clusters in monosyllabic words. I argue that the analysis in (8), which has been used in rule-based analysis to avoid the clusters, is not applicable to monosyllabic roots. I now demonstrate the analysis in (8) for monosyllabic roots:

11.	Input	/məŋ-pam/
	1) Nasal assimilation	məm-pam
	2) Voiceless obstruent deletion	məm-am
	Output	*[mə.mam]

As I demonstrate in the above derivation, the solution offered by rule-based analysis fails to explain the process of the prefixation of monosyllabic roots. The output *[me-mam] derived from the analysis is incorrect; the correct output is [meŋ-ə-pam]. When prefixes are attached to monosyllabic roots, the nasal segments in the prefixes do not undergo nasal assimilation by the following voiceless obstruents because schwa is epenthesized between the prefix and the monosyllabic root. Therefore, the nasal segment becomes nasal velar [ŋ]. However, the voiceless obstruents of the roots remain undeleted. This means that both rules, nasal assimilation and voiceless obstruent deletion, which have to be applied in order, do not give a true explanation of how the process occurs.

As we notice, the nasal and voiceless obstruent clusters in (5) and (9) occur in root-internal position and at prefix-root junctures, respectively. It is worth mentioning that this cluster also occurs in another morphological environment; it is found at prefix-prefix boundaries in multiple prefixation, when two or more prefixes are attached to a root, as mentioned earlier in this subsection. The occurrence of nasal and voiceless obstruent clusters in multiple prefixation has also been discussed in rule-based analyses. By applying the same analysis and rules as in (8), the occurrence of nasal and voiceless obstruent clusters at prefix-prefix boundaries cannot be explained satisfactorily. This is because, in one case, i.e. in nominal prefixes, the clusters are resolved by nasal substitution, while in another case, i.e. in verbal prefixes, the clusters appear.

To the best of my knowledge, only two rule-based analyses concerning nasal and voiceless obstruent clusters in multiple prefixation have been performed by scholars. These are by Omar (1986) and Karim et al. (1989). How the clusters are analysed by these scholars will be further discussed in Chapter 2 (subsection 2.3.1). I shall now demonstrate how the analyses postulated by the scholars in this group pose a problem when accounting for nasal and voiceless obstruent clusters in multiple prefixation.

12. Nasal and voiceless obstruent clusters in multiple prefixation (from the DBP-UKM corpus).

    i) /**pəŋ-pər**-kaja-an/ [**pə.mər**.ka.ja.an]
    NOM.PRF-VERBL.PRF-rich-NOM.SUF
    'enrichment'

    ii) /**məŋ-pər**-luas-kan/ [**məm.pər**.lu.was.kan]
    VERBL.PRF-NOM.PRF-strength-CAUS.SUF
    'to cause to broaden'

When the two rules, nasal assimilation and voiceless obstruent deletion, are applied to the words in 12(i) and (ii), the outputs are:

Input	/**pəŋ-pər**-kaja-an/
1) Nasal assimilation	pəm-pər-kaja-an
2) Voiceless obstruent deletion	pəm-ər-kaja-an
Output	[**pə.mər**.ka.ja.an]

Input	/məŋ-pər-luwas-kan/
1) Nasal assimilation	məm-pər-luwas-kan
2) Voiceless obstruent deletion	məm-ər-luwas-kan
Output	*[**mə.mər**.lu.was.kan]

As we can see in the above examples, the rule ordering, nasal assimilation and voiceless obstruent deletion, as postulated in rule-based analysis to account for nasal and voiceless obstruent clusters, only works for the data in 12(i). These rules, however, fail to account for the data in 12(ii), as *[mə.mər.lu.as.kan] is not the right output though the cluster has been successfully eliminated. This clearly shows that the proposed solution to avoid nasal and voiceless obstruent clusters does not always work to explain the occurrence of the clusters in multiple prefixations. Thus in this study, I then propose that the application of nasal substitution in this morphological environment is mainly based on the morphological domain where the clusters occur. To put it differently, nasal substitution as a phonological process is associated with a particular morphological factor. In this case, the morphological factor which conditions the phonological process is the domain. To overcome this, we need a constraint which is able to define this issue of morphologically-conditioned phonology for the case that occurs in multiple prefixation. I suggest that the relevant constraint to account for this is EDGE-INTEGRITY (52). This morphology-phonology interface constraint, as we will see in Chapter 4 (Section 4.3), is able to account for the non-application of nasal substitution in 12(ii). Such a solution can only be applied in OT by formalizing the link between particular morphemes and particular phonological patterns in a non-derivational way.

We have just discussed how nasal and voiceless obstruent clusters are analysed in rule-based analyses. But what about the other clusters, i.e. nasal and sonorant? These clusters have not previously drawn as much attention among scholars who have studied Malay as nasal and voiceless obstruent clusters have. Perhaps, this cluster behaves grammatically, according to the phonological needs of the language where a nasal segment that precedes a sonorant is always deleted. The DBP-UKM corpus, used in this study, also proves that nasal deletion is always applied to break up nasal and sonorant clusters, as exemplified in (6) above. In short, this cluster does not pose any problem for the language as the cluster regularly undergoes nasal deletion. It also presents no challenge to the phonological system of the language. Nevertheless, this cluster will also be one of the topics under discussion in this study (see subsection 4.2.2 in Chapter 4).

All the problems mentioned above occur in prefixation. In reduplication, on the other hand, there are also problems that occur due to the inappropriate analyses on offer. Most of the problems arise from partial reduplication. This is because the phonological aspects involved in this type of reduplication are more interesting than other types of reduplication. Therefore, many previous studies have been concerned with partial reduplication (e.g.: Hassan, 1974, 1987; Othman, 1981; Omarm 1975, 1986, 1993). Although partial reduplication has received much attention among Malay scholars, the analyses pose a number of problems. I now show what these problems are.

First, the way that partial reduplication is described is unclear. Scholars in this group (e.g.: Hassan, 1974, 1987; Othman, 1981; Omar, 1975, 1986, 1993) describe partial reduplication as a process of copying the first syllable of the base. As we can see, in the following examples, the definition given does not in fact describe the real process of partial reduplication. The copying process is not of the first syllable of the base. Rather, only the initial onset of the root is copied into the reduplicative morpheme, while the vowel is always a schwa [ə], regardless of what the vowel in the initial syllable of the base is. The following examples offer a better illustration of this. For clarity, the reduplicative morphemes are underlined.

13. Partial reduplication in SM.

laki 'man'	<u>lə-</u>laki RED-man 'men'
kuda 'horse'	<u>kə-</u>kuda

	RED-horse 'trustle'
taŋgə 'stairs'	tə-taŋgə
	RED-stairs 'neighbours'
rambut 'hair'	rə-rambut
	RED-hair 'small hairs'

The above examples show that it is not the initial syllable of the base that is copied into the reduplicative morpheme. Thus the definition given by rule-based scholars cannot be used to explain how the copying process actually works.

As can also be seen in the above examples, partial reduplication in SM involves CV reduplicative morphemes. It is important to know that CVC reduplicative morphemes are found in Malay through one of its dialects, i.e. in the Perak dialect (henceforth PD), as is CV. In previous studies, both patterns of reduplicative morphemes, CV and CVC, have been discussed in separate analyses performed by Ahmad (2000b) and Aripin (2005), respectively. Both scholars have proposed CV and CVC templates for SM and PD, respectively. In our analysis, these two patterns of reduplicative morphemes will be discussed together, while CV and CVC templates will not be used. The CVC or heavy reduplicative morpheme in this study will be explained by the tendency for prosodic constituents to be of maximal size, while the CV or light reduplicative morpheme will be explained by the tendency for prosodic morphemes to have an unmarked syllable structure (Downing, 2006: 12). These opposing tendencies can be accounted for by proposing that heavy and light reduplicative morphemes are associated with distinct construction-specific constraint rankings or co-phonologies. How co-phonologies explain this case can be seen in Chapter 5 (Section 5.4).

Besides the problem of partial reduplication mentioned above, the process of copying prefixes into the reduplicant also needs explanation, since the prefixes are only copied in some reduplicated words, as shown in 14(a), while the copying process of prefixes into the reduplicant is not applicable in others, as seen in 14(b). This type of reduplication, where prefixed words are reduplicated, is called 'affixal reduplication' in this study.

14. (a) /məŋ-ə-lap/ [mə.ŋə.lap-ŋə.lap]
    ACT.PRF-STEMEX-wipe
    'to wipe repeatedly'
    (b) /məm-basuh/ [mə.ba.soh-ba.soh]
    ACT.PRF-wash
    'to wash'

Affixal reduplication has long been understood as a process of copying the prefix into the reduplicant. Given the examples shown in (14) above, I would like to highlight here that not all prefixes are copied into the reduplicant. As far as affixal reduplication is concerned, this point has not yet been raised by any scholars. It should be mentioned here that Onn (1980) discussed this type of reduplication in his analysis. However, that discussion does not deal with the process of copying prefixes into the reduplicant. It only describes how the copying process operates. Thus, I would like to bring this point into the discussion since there appears only to be imprecise explanation of how the reduplicant is copied (see Section 5.5 for the discussion).

1.3.2 Autosegmental analysis

When non-linear frameworks became prevalent in phonological theories, prefixation and reduplication in Malay were re-analysed by scholars who attempted to reassess any issues which had not been satisfactorily accounted for in rule-based analysis. Works like Teoh (1994) and Ahmad (1999 and 2000b) continue the discussion of Malay affixation, while Ahmad (2000b) and Aripin (2005) re-analyse the process of reduplication in SM and PD, respectively. Any issues left unexplained have been put forward for reanalysis via autosegmental analysis.

For example, Teoh (1994) and Ahmad (1999 and 2000b) attempt to explain how the issue arises in the process of monosyllabic bases. They demonstrate why voiceless obstruents remain while nasal segments in prefixes do not assimilate to the place of articulation of voiceless obstruents, where nasal segments always change to velar nasals. Since the nasal segments in prefixes always change to velar nasals, schwa is epenthesized between the nasal segments and the undeleted voiceless obstruents. Both scholars explain this issue by claiming that the lexical representation for monosyllabic roots is originally disyllabic, as Malay is a disyllabic language (see Chapter 4, subsection 4.2.1.1).

This claim leads however to another problem which relates to the disyllabicity minimality requirement of the language and the grammatical outputs produced from the analyses. The outputs derived lead to ungrammatical output of monosyllabic roots. Instead, what I found is that disyllabic lexical representations are not relevant for application to monosyllabic roots, as they do not emerge as disyllabic roots in the surface representation. In other words, monosyllabic roots emerge as the monosyllabic roots that they are. In this study, I will show that the solution offered by non-linear autosegmental analyses fails to account for this matter. As I will demonstrate in subsection 4.2.1.1, this matter can be explained in a more adequate manner by using the theory adopted in this present study, i.e. MBT.

As was mentioned in the previous subsection, monosyllabic bases have not been discussed in rule-based analyses. They have therefore been given attention by scholars in autosegmental analysis, in particular why schwa is epenthesized when nasal final prefixes are attached to monosyllabic roots. Since the focus is more towards monosyllabic words, the occurrence of nasal and voiceless obstruent clusters, such as in (9), has been ignored in scholars' analyses. As a result, this issue of nasal and voiceless obstruent clusters at prefix-root junctures has never been discussed in any previous studies of Malay, neither in rule-based nor in auto-segmental analyses. Hence, this issue has been selected for consideration in this study (see Chapter 4, subsection 4.2.1.2), as there are examples in the corpus that show the occurrence of the clusters.

Other than the abovementioned problems, other problems appear in non-linear autosegmental analyses, concerning reduplication. First, much work in autosegmental phonology, positing reduplication, resembles affixation (e.g.: Ahmad, 2000a; Aripin, 2005). What distinguishes them is that a reduplicant does have a phonologically-unspecified lexical entry where the phonological material is obtained from the base to which it is attached (McCarthy and Prince, 1993a, 1994, 1995b; McCarthy, 1995; cf. Ahmad, 2005). In this reanalysis, I argue that the claim presents a challenge to the grammar, particularly when disyllabic roots are totally reduplicated and the entire roots are copied. The size of the reduplicant then contains two syllabics which is larger than the regular size for an affix. In this case, can a reduplicant be considered an affix? This matter is highlighted in this study as no previous studies have covered it. In this study, this issue is discussed in Chapter 5, Section 5.3. The same question can also be asked of affixal reduplication since reduplicative morphemes in affixal reduplication are larger than monosyllabics, the same as in total reduplication (see Section 5.5).

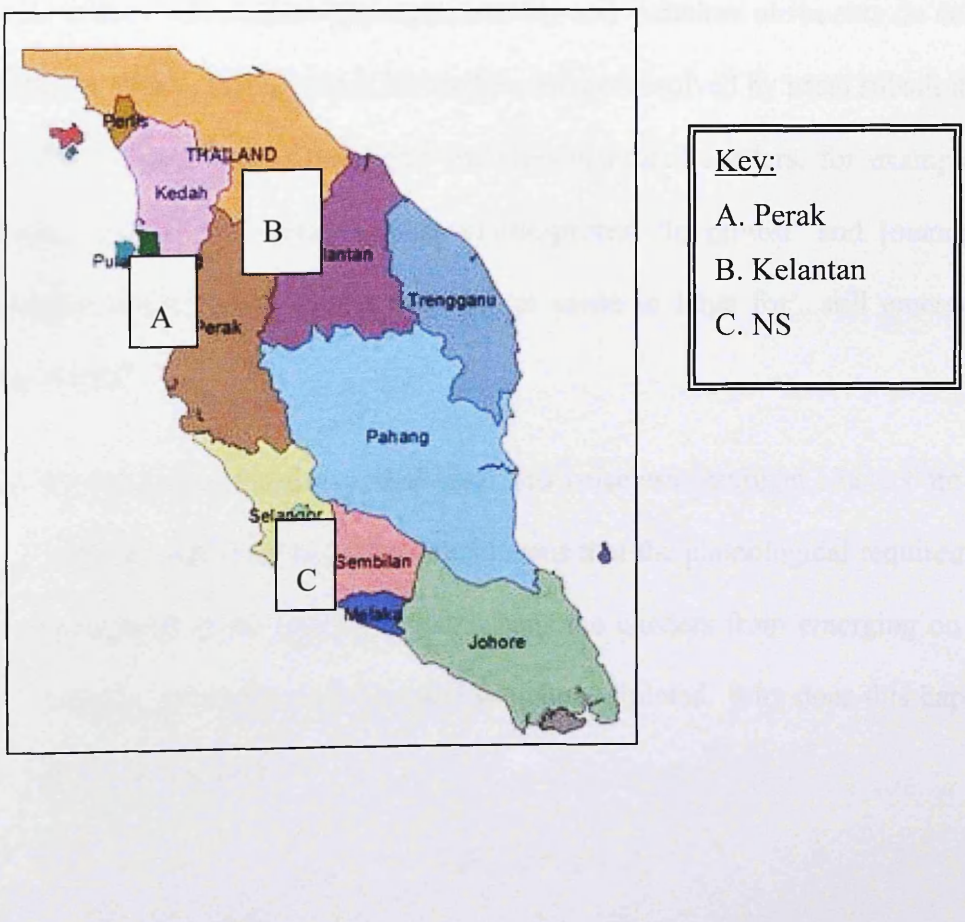
Second, CV and CVC templates are used by Ahmad (2000a) and Aripin (2005), in autosegmental analyses to account for partial reduplication in SM and PD, respectively. One thing I would like to point out here is that each template can only explain one process of partial reduplication. In other words, a template for one process cannot be used to account for another. This shows that only limited patterns of reduplicative morphemes can be analysed using a template analysis. As already noted, Malay has two patterns of partial reduplication – light and heavy. Does this mean the language needs more than one template to explain all the patterns? I believe that an appropriate phonological theory should be able to explain all patterns of reduplication without any specific template.

To explain the two patterns of reduplicative morphemes (i.e. light and heavy) in Malay partial reduplication, co-phonology analysis, as developed and motivated within OT (Orgun, 1996; Inkelas, 2008; Inkelas and Zoll, 2005), will be used. The concept of 'Markedness Reversals' in co-phonology enables us to account for light and heavy reduplication where reduplicative morphemes with identical categories show different patterns of markedness reduction (Downing, 2008) (see Chapter 5, Section 5.4, to see how markedness reversal applies in Malay). In other words, CV and CVC are both reduplicative affixes, though they have different patterns of markedness reduction. In a co-phonology analysis, this can be done by allowing the relevant markedness constraint to be re-ranked in different morphological constructions in the same language (ibid.). Chapter 5, Section 5.4, will demonstrate that the relevant markedness constraint is highly ranked in the constraint ranking of

CV reduplicative morphemes, while it is ranked lower for CVC reduplicative morphemes with the same constraints as CV reduplicative morphemes.

Other than reanalysing some issues in prefixation and reduplication, this thesis aims to show more clearly that the most important aspect in choosing OT is the study of variation. In an attempt to do that, as mentioned in Section 1.1, this thesis will take into account the issue of nasal and voiceless obstruent clusters in three Malay dialects: Perak, Kelantan and NS (see Map 2, below). By examining these three dialects, we will then be able to know how the clusters are resolved in non-standard Malay dialects. The areas labelled on the map of Malaysia, below, show where the dialects chosen for this study are spoken.

Map 2: The areas where the three dialects of Malay, Perak, Kelantan and NS, are spoken.



We note that most previous studies have only discussed voiceless-obstruent nasal substitution. Having referred to some relevant previous studies, I found some data which show that voiced obstruents also undergo nasal substitution. However, this situation occurs in only a few dialects of Malay, two of which are Perak (see subsection 6.3.1) and NS (see subsection 6.3.3). In my opinion, it is essential that this amazing phonological process be discussed further as it has never been discussed formally. Thus, this study will take this issue into account. We will see, in Chapter 6, how the idea of OT can satisfactorily explain the variation that occurs in the three dialects of Malay in terms of resolving voiceless/voiced obstruent nasal substitution.

1.4 Research Questions

This study seeks to answer the following research questions:

1) In most of the Austronesian languages, nasality and voiceless obstruents do not go together. In Malay, this sequence is therefore always resolved by nasal substitution. But why do sequences of nasal and voiceless obstruent clusters, for example in [sampan] 'small boat', [məm-protes] VERBL-protest 'to protest' and [məm-pər-bəsar-kan] VERBL-VERBL-large-CAUS.SUF 'to cause to large for', still emerge in Malay words?

    (a) We note, from the above, that nasal and voiceless obstruent clusters are not entirely prohibited in Malay. This means that the phonological requirement postulated in the language, which bans the clusters from emerging on the surface, is somehow obeyed and somehow violated. Why does this happen in Malay grammar?

(b) Is there any strategy applied in the language (in SM or its dialects) to eliminate nasal and voiceless obstruent clusters? If any phonological processes are applied, what are they?

(c) In eliminating the clusters, the markedness constraint *NÇ̥ plays a crucial role, ensuring that no nasal and voiceless obstruent clusters emerge in the surface representation. In two of the three dialects that this study examines, i.e. Perak and NS, voiced obstruents also undergo nasal substitution. In this case, does the *NÇ̥ constraint also play a role in the constraint hierarchy? Can *NÇ̥ be used to account for both cases, voiced/voiceless obstruent nasal substitution?

2) Reduplication has received much attention from previous Malay scholars (e.g.: Omar, 1986, 1993; Hassan, 1974, 1976; Othman, 1976, 1981, 1985; and many others) using various approaches. Nevertheless, there are some issues that need to be clarified. By focusing more on total, partial and affixal reduplication, this study intends to find out answers to the following research questions related to reduplication:

(a) Reduplication has been claimed as a process of affixation by previous Malay scholars (e.g.: Ahmad, 2001; Aripin, 2005). This claim will be revised to confirm whether or not reduplication is truly a process of affixation, by giving a theoretical justification based on the theory applied in this thesis.

(b) Why are the nasal segments in the prefixes are copied in some reduplicated words, as in [məŋəlap-**ŋ**əlap] and [mənəkan-**n**əkani] but not in [məmbandiŋ-bandiŋ], in Malay? What is the motivation for copying, or not, the nasal segment in the prefix into the reduplicant?

3) In Generative Grammar (Chomsky, 1965; Chomsky and Halle, 1968), phonological processes are presented by rewrite rules whereby one rule interacts with another rule through linear ordering. Rules apply one after another, with one rule's output being the following rule's input (Kager, 1999: 1). In contrast, in OT, constraints are the major focus. Constraints typically conflict with one another in order to satisfy a particular constraint. The form, or what is called the 'candidate' in OT, which incurs fewer constraint violations amongst various possible candidates will be selected as the most harmonic or optimal output. This selection of optimal output is related to the hierarchical ranking of constraints, whereby high-ranking constraints have more priority to be obeyed than low-ranking constraints (ibid.: 4). As constraints in OT are Universal, while rankings are not, so differences in ranking are the cause of variation (ibid.). In this study, the differences in ranking that cause variation will be applied to three selected dialects of Malay (i.e. Perak, Kelantan and NS) to account for the distinction between nasal and voiceless/voiced obstruent clusters root-internally, and prefix-root junctures in the dialects.

1.5 Research goals

This present study aims to fulfil two goals. The first one, a descriptive goal, is to provide a new analysis and explain the issues highlighted above in Malay grammar. No previous studies concerning prefixation and reduplication in Malay have attempted to analyse these issues by taking the canonical shape of prosodic morphemes into account. As we will see later in this thesis, the idea of the canonical shape of prosodic morphemes could solve all the problems in prefixation and reduplication. Thus I claim that this present study can be considered as the first formal analysis to adopt the ideas of GTT into Malay grammar.

Second, a theoretical goal of this study is to enhance the theory of GTT developed within OT as a formal comprehensive theory, defining the constant shape of prosodic morphemes by assigning them to the morphological category. In this theory, the size of prosodic morphemes is not determined by construction-specific constraints (like RED=Foot) (cited in Downing, 2006: 33). The final aim of this present study is to show the success of the theory in comparison to other previous approaches in accounting for the topics under examination. It is hoped that the analyses offered in this study can explain them more successfully than previous studies and that the results achieved in the analysis will be useful in the study of Malay grammar.

1.6 Organisation of this thesis

This thesis is divided into seven chapters. The organisation of the subsequent chapters is as follows. Chapter 2, Literature review: this chapter reviews what previous studies have done vis-à-vis the relevant issues in this thesis. Chapter 3, Data and Method: it describes the methodology that I have used for the research. This includes the data I have used to account for the phenomena concerned. Chapters 4, 5 and 6 present the analyses of prefixation, reduplication and dialectal variation, respectively. Chapter 7 concludes the thesis. The goal of this chapter is to summarise the important results of the three topics discussed: prefixation, dialectal variation and reduplication.

2. LITERATURE REVIEW

2.1 Introduction

There have been numerous studies concerned with prefixation in Malay, particularly the nasal final prefixes, /məN/[8] and /pəN/. A requirement of the language that a nasal occupying the coda position of a syllable has to be homorganic with the following onset consonant has attracted the attention of many Malay scholars (cf.: Hassan, 1974; Omar, 1986; Koh, 1981; Othman, 1983; Ahmad, 1993; Karim et al., 1989, 1994; Karim, 1995; Teoh, 1994; and many others). The process of combining final nasal prefixes and voiceless obstruent initial roots has been unsatisfactorily analysed by those scholars.

For example, nasal assimilation and voiceless obstruent deletion, as postulated in a rule-based approach, fail to account for the actual process of prefixation in Malay, whereby voiceless obstruents following nasal segments in some prefixes do not undergo the deletion process. As a result, nasal and voiceless obstruent clusters emerge in the surface representation. This disobeys absolutely the grammar of the language whereby nasal and voiceless obstruent clusters are not permitted to surface (see Chapter 4, subsection 4.2.1). I shall mention the analyses performed by rule-based and other scholars in non-linear autosegmental as I suggest that they pose a number of problems when accounting for Malay prefixation, particularly nasal final prefixes (see Chapter 1, Section 1.3). In this reanalysis study, all the problems can be

[8] Note that the archiphoneme [N] is widely used by previous Malay scholars (e.g.: Hassan, 1974; Omar, 1986; Karim et al., 1994; Karim, 1995) to represent the surface forms of underlying forms. Thus, I will retain this archiphoneme when discussing those previous works for ease of reference. However, this archiphoneme will not be used in this present study to represent the surface forms of the variants of Malay prefixes. Instead, /ŋ/ will be used to replace the archiphoneme.

accounted for in OT by categorising the lexicon of Malay into three different strata, i.e. monosyllabic foreign, undeleted voiceless plosives and native (see Figure 3 on p. 123).

It must be mentioned that those scholars have also analysed reduplication. Most of them approach reduplication from the perspective of morphology, and the effects of phonological processes on reduplication are only briefly discussed. This problem mainly arises in rule-based analyses (e.g.: Hassan, 1974; Omar, 1986; Koh, 1981; Othman, 1983; Karim et al., 1989, 1994; Karim, 1995). Although Aripin's (2005) and Ahmad's (2000a, 2005) analyses have discussed reduplication by incorporating both morphological and phonological aspects, the solutions they postulate are only applicable to particular aspects of reduplication. For example, partial reduplication has been discussed in Aripin (2005) and Ahmad (2000a, 2005) by focusing only on PD and SM, respectively. As a consequence, the solutions proposed by those scholars, e.g. the CV template, cannot be used to account for heavy reduplicative morphemes, as found in some dialects of Malay, such as PD. If a template analysis were used to analyse both types of reduplicative morpheme (i.e. light and heavy) in partial reduplication, would we need two different templates, one for each? This point will be put considered in the appropriate chapter (see Chapter 5, Section 5.4).

There are some substantial works on the morphology and phonology of Malay dialects. These include Musa (1974), Rufus (1966), Ahmad (1991), Che Kob (1985) and Abdul Malik (2004). Amongst these works, Musa (1974) and Abdul Malik (2004) are unpublished dissertations and theses. Note that those previous scholars focus more on dialectology studies, except for Abdul Malik (2004). The latter's study was concerned with vowel harmony in four dialects of Malay: Kelantan, Johor, Perak and Kedah. None of those studies is concerned with the phonological aspects of the dialects they examined. In my search through previous dialect studies, none of them has discussed the issue of nasal and voiceless obstruent clusters in Malay dialects. This is one of the reasons why this present study includes three Malay dialects, i.e. Perak, Kelantan and NS, for examination in terms of the issue under discussion. As well as looking at how nasal and voiceless obstruent clusters are resolved in those three dialects, this study will also pay attention to nasal and voiced obstruent clusters in which the combination of both segments results in some interesting phonological process in Malay grammar where voiced obstruents also undergo nasal substitution (see Chapter 6).

This chapter thus presents the background to literature relevant to the present study. In order to describe what has been done in previous research, this chapter is divided into two parts. These are: (1) the theoretical framework; and (2) Malay phonology scholarship, each of which will be discussed in Section 2.2 and 2.3, respectively. Section 2.4 presents important constraints to account for the problems that arise. Finally, Section 2.5 concludes the chapter.

2.2 Theoretical framework

This section will be discussing the theoretical frameworks that are relevant to the present study. There are four theoretical frameworks which will be discussed in turn: (1) rule-based approach, (2) non-linear autosegmental phonology, (3) OT and (4) MBT. The discussion of these theories is more to the principles and the central ideas proposed in those theories, as they are essential before the discussion of Malay phonology scholarship begins in the following section, Section 2.3.

2.2.1 Rule-based analysis

Phonological theory in the tradition of Chomsky and Halle (1968) as stated in *The Sound Pattern of English* (SPE) was based on rewrite rules (McCarthy, 2008: 1). A rewrite rule A → B / C____D is described as an input configuration and an operation that applies to it (Ibid.: 3). Basically, there are two ways that a set of phonological rules can be applied for a given Underlying Representation (UR). They are: (1) simultaneous rule application i.e. phonological rules are applied simultaneously to the UR, and (2) sequential rule application where phonological rules are applied one by one in sequence which is also known as rule ordering. Rule ordering means rule A must be applied before B. Rule B cannot precedes the application before rule A. In French for example, how /bɔn/ becomes [bɔ̃] 'good' is explained by applying two set of rules which have to be applied in order (Hyman 1975: 130):

15. (a) Vowel nasalisation

V → Ṽ / __N$ ($ indicates syllable boundary)

(b) Nasal deletion

N → Ø / Ṽ ___$

As claimed by Hyman, nasal deletion rule can only be applied when the nasal consonant is preceded by a nasalised vowel. In other words, vowel nasalisation must be applied first before the nasal deletion rule in order for nasal deletion to be effected. Otherwise, the output would be incorrect.

It is worth mentioning that this sequential rule application has also been applied in earlier scholars' works who have studied Malay. I shall point out in this study that there are problems occurring when rule application applies into Malay data, in particular prefixation and reduplication. I therefore argue that rule based analysis is not the best and adequate theory to account for Malay data. The further discussion of this can be seen in Section 2.3 (see subsection 2.3.1).

As the earlier theory of phonology, rule-based analysis has described a lot of phenomena of language. However, this theory has been criticised by phonologists who have realised that there were issues that this theory fail to explain. As claimed in McCarthy (2008: 1), rewrite rules can describe a lot of phenomena but they do a poor job in explaining how phonological systems fit together. Meanwhile, Kisseberth (1970) argues that rewrite rules are missing an important generalisation about the special role of surface structure constraints (cited in McCarthy 2008: 2). This point has been raised up when data from Yawelmani as illustrated below are considered:

16. Yawelmani final vowel deletion

a) /taxa:-kˀa/	[taxa:-kˀ] 'bring'
/taxa:-mi/	[taxa:-m] 'having brought'
b) /xat-kˀa/	[xat-kˀa] 'eat'
/xat-mi/	[xat-mi] 'having eaten'

The data above show that vowel at the end of a word is deleted, as in 16(a). In contrast to data 16(b), vowel at the end of a word is not deleted when it is preceded by consonant clusters. This is because the deletion of vowel would leave a consonant that cannot be syllabified - *[xatm].

17. Yawelmani vowel epenthesis

a) /ʔilk-hin/	[ʔi.lik.hin] 'sing (nonfuture)'
/lihm-hin/	[li.him.hin] 'run (nonfuture)'
b) /ʔilk-al /	[ʔil.kal] 'sing (dubitative)'
/lihm-al/	[lih.mal] 'run (dubitative)'

As we see in the above examples, a process of vowel epenthesis applies when there are triconsonantal clusters, as shown in (a). This process however, does not apply when the consonant clusters are smaller, as in (b). As argued by Kisseberth, it is possible for SPE style of rewrite rules to explain for example, when vowel at the end of a word can be deleted, as the process of vowel final deletion in Yawelmani does not apply to all vowel final words. He further says that vowel final deletion cannot create bad syllable while vowel epenthesis applies to eliminate the clusters (cited in McCarthy 2008: 2). This situation is called by Kisseberth as conspiracies and rewrite rules failed to explain. Since conspiracies cannot be explained in rewrite rules, I shall therefore argue that this approach is not appropriate to account for Malay data. As we will see later in the thesis, there are lot of conspiracies occur in Malay, particularly in prefixation and reduplication.

2.2.2 Non-linear autosegmental phonology

Morphophonological theory has undergone changes after the publication of SPE which was based on rewrite rules (Spencer, 1991). A new theory with no application of rules named Autosegmental Phonology was proposed by John Goldsmith (1976) in his doctoral dissertation. As stated in Spencer (1991: 133), this theory is essential to account for tone languages particularly African tone languages. This theory proposes that, there is a two-tiered representation in which tones are associated to tone-bearing segments (vowels or sometime sonorant consonants) according to certain universal conventions (Ibid. 133). Since we have more than one tier or line of phonological elements, for examples, a sequence of consonant and vowel phonemes in one tier, while a sequence of tones in another tier, therefore this representation is called non-liner or multilinear autosegmental phonology (Ibid.).

The different view of analysing data as offered by this theory has attracted many scholars in accounting for phonology and morphology processes. One of the morphological processes i.e. reduplication which appears to be fundamentally non-concatenative has been received much attention among scholars because it has important implications on phonology and morphology (Ibid.: 151)[9]. Before we see an instance of how reduplication is resolved by autosegmental phonology analysis, let us first observe the conditions stated in the theory. There are four conditions on linking the melody tier to the prosodic template which can be summarised as follows (Marantz, 1982 cited in Spencer, 1991: 152):

---

[9] How reduplication gives impact to phonology and morphology as said in the above sentence, see the reference cited in the text.

Table 1: Four conditions of melody linking in autosegmental phonology

Condition 1:	Melody consonants link to C slots and melody vowels link to V slots.
Condition 2:	Linking is strictly one-to-one. No multiple links are allowed.
Condition 3:	CV slots maybe pre-linked to specific phonemes. Pre-linking takes precedence over autosegmental linking from the root melody.
Condition 4:	Directionality of linking: either the leftmost melody phonemes link with the leftmost appropriate CV slot Or linking proceeds from left to right: or the rightmost melody phonemes link with the rightmost appropriate CV slot and linking proceeds right to left. Linking is 'melody driven' in the sense that the association algorithm starts with a melody phoneme and then tries to find an appropriate CV slot, not the other way around.

Agta reduplication is an instance which has always been associated to autosegmental phonology analysis. In this language, the reduplicative morpheme is not only attached to the left of the root, as a prefix, but it is also attached inside the root as an infix. The following examples give a better picture of both forms:

18. Agta reduplication (from Marantz, 1982)

a) bari 'body'	barbari-k kid-in 'my whole body'
b) mag-saddu 'leak' (verb)	mag-sadsaddu 'leak in many places'
c) ma-wakay 'lost'	ma-wakwakay 'many things lost'
d) takki 'leg'	taktakki 'legs'
e) ulu 'head'	ululu 'heads'

As exemplified above, the reduplication in Agta has variety of types. Reduplicative morpheme can be as a prefix which is attached before the root, as in 18(a), (b) and (d). It also can be as an infix where the reduplicative morpheme is copied inside the root, as exemplified in 18(c) and (e). As suggested by Marantz, these two types of reduplicative morpheme could be accounted for by proposing the following derivation:

19. Agta reduplication: /takki/ → [taktakki]

```
              t a k k i        t a k k i        t a k k i
              | | | | |   →    | | |            | | | | |
CVC     +     C V C C V        C V C      +     C V C C V

        t a k t a k k i
   →    | | | | | | | |
        C V C C V C C V   =   taktakki
```

The above derivation shows how [taktakki] is derived from /takki/. As proposed by Marantz, CVC is the size of reduplicative morphemes in Agta. This prosodic template is associated to the melody phonemes of *takki*. In accordance with autosegmental principles, melody consonants are linked to C slots, while melody vowels are linked to V slots. The linking process must however be one-to-one mapping. In the above derivation, we see that *ki* are not associated to any prosodic template of CVC. This is because the prosodic template CVC has already been linked to the melody phonemes i.e. *tak*. In autosegmental principle, any melody elements or prosodic template slots which left unassociated at the end of the derivation, they have to be deleted by the conventions (Spencer, 1991: 152). Therefore, *ki* which cannot be linked to the prosodic template of CVC do not appear in the reduplicative morpheme.

It ought to be mentioned that autosegmental phonology analysis has also been applied into Malay. A number of morphological processes like affixation and reduplication have been put forward for a new analysis. In the following section (see subsection 2.3.2), we will see how autosegmental phonology analysis is applied by Malay scholars to account for affixation and reduplication. In what follows, I discuss another theoretical framework i.e. OT which this present studies will be based on.

2.2.3 OT[10]

OT was put forward by Prince and Smolensky (1993) as a general approach to modelling human linguistic knowledge (Prince and Smolensky, 2004; McCarthy, 2006). Although OT has been largely associated with phonology and morphology, this theory has also been used in other subfields of linguistics, such as syntax and semantics (McCarthy, 2008).

The most important concept of the theory is that Universal Grammar (UG) consists of a set of violable constraints. The idea of the violability constraints in OT is different from that of classical rule-based theory in that UG is defined as inviolable principles and rule schemata (or 'parameters') (Kager, 1999). Although constraints can be violated, the violation is always minimal and happens only to satisfy a higher-ranking constraint (Archangeli, 1997). OT recognises two types of constraint: *faithfulness* and *markedness* (Kager, 1999). These constraints, faithfulness and markedness, are illustrated below.

[10] The discussion about OT here is very brief. Readers desiring a more detailed introduction to OT can consult the references cited in the text directly. Comprehensive references to OT can be found in: McCarthy and Prince (1993), Prince and Smolensky (2004), and some introductory books, such as Archangeli and Langendoen (1997) or Kager (1999). See McCarthy (2008) for good guidelines on implementing OT. Advanced readers can refer to Downing (2006), where the analysis of Malay in this study, particularly prefixation and reduplication, is based on the ideas of GTT, i.e. Generalised Template Theory.

Faithfulness constraints require that every segment in the input has a correspondent in the output. In other words, faithfulness constraints prohibit differences between input and output (McCarthy, 2008: 13). For example, in Malay, when the input /məŋ+pam/ maps to surface [mə.ŋə.pam], the faithfulness constraint against epenthesis, i.e. DEP-IO, is violated for this candidate.

Faithfulness constraints preserve lexical contrast (Kager 1999). They demand that linguistic forms be realised as closely as possible to their basic lexical forms. If there is any contrast in the output forms, it has to be a minimal amount, with not too much contrast to the basic lexical forms. Preservation of lexical contrasts is not only carried out by phonological elements (as this present study will focus on), it is also expressible by word structure (morphology) or phrase structure (syntax) (ibid.: 6).

Generally, there are three primary faithfulness-constraint families that represent the relational correspondence between S_1 (Input/Base) and S_2 (Output/ Reduplicant) (McCarthy and Prince, 1995). The three constraint families are defined below:

20. The faithfulness constraint families (McCarthy and Prince, 1995).

(i) MAX constraint family

General schema – every segment of S_1 has a correspondent in S_2.

MAX-IO – every segment of the input has a correspondent in the output (no deletion).

MAX-BR – every segment of the base has a correspondent in the reduplicant (total reduplication).

(ii) DEP constraint family

General schema – every segment of S_1 has a correspondent in S_2.

DEP-IO – every segment of the input has a correspondent in the output (no epenthesis)

DEP-BR – every segment of the reduplicant has a correspondent in the base (prohibits fixed default segmentism in the reduplicant).

(iii) IDENT(F) constraint family

General schema – Let α be a segment in S_1 and β be any correspondent of α in S_2. If α is [γF], then β is [γF] (correspondent segments are identical in feature F).

IDENT-IO – output correspondents of an input [γF] segments are also [γF].

IDENT-BR – reduplicant correspondents of a base [γF] segments are also [γF].

Markedness constraints require that output forms meet some criterion of structural well-formedness (cf. Kager, 1999). A markedness constraint assigns its violation without any reference to the input, as is required in faithfulness constraints. Markedness constraints play a key role as the grammatical factors that produce 'unmarked types of structure'. Unmarked structure is one of the two types of linguistic structure, the other one being a marked structure. Cross-linguistically, an unmarked structure is preferred and is basic in all grammars, while a marked structure is avoided and used by grammars only to create contrast (ibid.: 2).

What is unmarked structure? In the sound systems of languages, for certain types of structure, like segments, segment combinations or prosodic structures are universally favoured over others (Kager, 1999: 5). As exemplified by Kager (ibid.), front unrounded vowels are unmarked as compared to front rounded vowels, open syllables as compared to closed syllables, short vowels as compared to long vowels, and voiceless obstruents as compared to voiced obstruents. Kager (ibid.) adds that those marked structures are avoided in all languages, whilst in some languages they are banned altogether. To avoid all the marked structures, markedness constraints play a crucial role in ensuring only favoured structures emerge as the optimal output.

In the sound system of Malay, certain types of structures are favoured over others. For example, as was observed, schwa is favoured over any other vowels, and voiced obstruents following nasal segments are favoured over voiceless obstruents. These favoured structures form the unmarked structure in the language. The major focus of this present study is to examine how far voiceless obstruents after nasal segments are disfavoured in Malay. Since voiceless obstruents are disfavoured in the language, a markedness constraint which is relevant and crucial to account for this is *NC̥. The details of this constraint are discussed in 2.4.1.1, and to see how it plays its role in a constraint ranking, see Chapter 4. In this study, we will examine whether or not voiceless obstruents following nasal segments are completely banned in the language. To find this out, voiceless obstruents after nasal segments at prefix-root and prefix-prefix junctures (see Chapter 4 for the analysis), and root-internally (see Chapter 6 for the analysis), will be investigated. In Chapter 5 (Section 5.4), we will see that schwa is an unmarked vowel in Malay. Hence, it is used as a fixed vowel in the reduplicative morpheme in partial reduplication, regardless of what the vowel in the initial syllable of the base is.

There are five basic tenets of OT which can be summarized as follows (Table 2) (McCarthy and Prince, 1994: 335):

Table 2: Five basic tenets of OT

1)	Universality	UG provides a set of constraints that are universal and are universally present in all grammars.
2)	Violability	Constraints are violable, but only minimally so.
3)	Ranking	The constraint of Con are ranked on a language-particular basis. The notion of minimal violation is defined in terms of this ranking. A grammar is a ranking of the constraint set.
4)	Inclusiveness	The constraints hierarchy evaluates a set of candidates that are admitted by very general considerations of structural well-formedness.
5)	Parallelism	The best satisfaction of the constraints hierarchy is calculated over the whole hierarchy and the whole candidate set. There is no serial derivation.

In OT, the actual output of the underlying form is selected from a large set of potential surface forms called candidates. The selection of candidates is based on well-formedness constraint-system evaluation. As stated in McCarthy and Prince (1993, 1994), the representational structure of a grammar in OT can be summarised as follows:

21. GEN (in) = {cand 1, cand 2, … .}

EVAL ({cand 1, cand 2, … .})

The function of GEN (short for generator), will generate a large set of possible candidates (surface representation) from each input (underlying representation). The function of EVAL (short for evaluator) will evaluate the well-formedness of the possible candidates generated. The candidate which best satisfies, or minimally violates, the constraints in a hierarchy is termed as the optimal or most harmonic output, and constitutes the actual surface form in the language.

Disagreement between [cand.1] and [cand.2] in the constraint hierarchy is resolved by ranking the constraints in a strict dominance hierarchy (Prince and Smolensky, 1993). For example, [cand1] satisfies A and violates B, while [cand2] satisfies B and violates A. Since [cand1] is, by assumption, the actual output, then this suggests that constraint A must dominate B in the hierarchy of the language grammar. In OT, the constraint ranking is represented in the form of a constraint tableau, as illustrated in (22). There are some useful conventions to be found in the constraint tableau. Constraints in OT are represented from left to right, where the leftmost side is the highest-ranking constraint. Potential candidates are listed in vertical order. Violation of a constraint is marked by '*', while constraint satisfaction is unmarked. The violation of a constraint which is accompanied by an exclamation mark, '!', indicates a fatal violation. The optimal output is shown by a pointing finger '☞'.

22. /Input/ → [cand1], A >> B

/Input/	A	B
a.☞[cand1]		*
b. [cand2]	*!	

In the above tableau, the suboptimal candidate (b) is ruled out as it fatally violates A. Candidate (a), which violates the lower-ranking constraint B, is the optimal output. The violation of B, however, is not significant since the winner has already been determined. In OT, the remaining lower-ranking constraint becomes irrelevant once a winner emerges. The violation of constraints after that does not affect its grammaticality.

In (23), both candidates satisfy the highest-ranking constraint equally in the tableau. The satisfaction of constraint A cannot contribute to a decision to determine the winner between them. In this situation, the decision is made by consulting the next constraint, B. Since candidate (b) violates B, then candidate (a) emerges as the optimal output as this candidate does not violate B.

23.

/Input/	A	B
a. ☞ [cand1]		
b. [cand2]		*!

Besides the ways illustrated in (22) and (23), there are other ways in which candidates can interact, particularly when they are in a tie situation. In such a situation, where both candidates pass or fail the highest-ranked constraint equally, the next constraint in the hierarchy will facilitate continuance of the evaluation.

24.

/Input/	A	B
a. ☞ [cand1]	*	
b. [cand2]	*	*!

In (24), where the violation of A by both candidates cannot determine the optimal output, the evaluation passes on to the next constraint, B. Candidate 24(b) violates B while candidate 24(a) does not, thus candidate 24(a) is the optimal output.

The above tableaux show a single violation of a candidate. Next, I am going to show that, as well as a single violation, multiple violations of a candidate are likely to occur in a grammar and this situation must also be considered. The following tableau shows how multiple violations are represented:

25. Multiple violations: A >> B

/Input/	A	B
a.☞[cand1]		*
b. [cand2]		**!*

In OT, if a constraint is violated or satisfied equally by two candidates or more, the evaluation continues by consulting the next constraint in the hierarchy. As in the above tableau, both candidates, (a) and (b), satisfy A. The evaluation goes next to B, to determine the optimal output. In the above, it can be seen that both candidates violate B. However, the candidates violate B unequally, as candidate (a) violates B less than candidate (b). Therefore, candidate (a) is more harmonic than candidate (b), according to the evaluation of minimal violation, thus candidate (a) is the optimal output. As noted in McCarthy and Prince (1993: 88), constraint violations are not counted, but are merely a comparison of more versus less, thus it is a matter of ordering and not quantifying (cf. Prince and Smolensky, 1993).

We will see then, in the thesis, how constraint ranking and violation handle the Malay data. For example, *NC̥ as a markedness constraint plays a large role in ruling out candidates with nasal and voiceless obstruent clusters. This constraint must therefore be ranked above the faithfulness constraint UNIFORMITY, so that nasal and voiceless obstruent clusters cannot emerge as optimal output, as in (26):

26. (i) /məŋ-pukul/ [mə.mu.kul]
ACT.PRF-scold
'to scold'
(ii) /məŋ-kupas/ [mə.ŋu.pas]
ACT.PRF-peel
'to peel'

Due to the obedience of *NC̥, the examples in (26) violate UNIFORMITY (see 82) which bans nasal substitution. The violation of UNIFORMITY in those words satisfies the markedness constraint *NC̥. In this case, *NC̥ is satisfied by a process of nasal substitution. As we shall see in Section 4.2.1.1, in the group I name 'monosyllabic foreign', there is another method that the language uses to satisfy *NC̥. As well as nasal substitution, vowel epenthesis is also a way of avoiding nasal and voiceless obstruent clusters, as exemplified in (27). By epenthesizing a vowel between the prefix and the monosyllabic base, another faithfulness constraint named DEP-IO is violated, since the output segment does not have a correspondent input segment. Thus, DEP-IO should be ranked lower in the constraint ranking of the monosyllabic foreign group. In other words, due to satisfaction of *NC̥, the two faithfulness constraints, UNIFORMITY and DEP-IO, are violated in the constraint ranking in the SM grammar system.

27. a) /məŋ-ə-cam/ [mə.ɲə.cam]
    ACT.PRF-STEMEX-recognise
    'to recognise'
    b) /məŋ-ə-cap/ [mə.ɲə.cap]
    ACT.PRF-STEMEX-stamp
    'to stamp'

In phonology, the situations mentioned above are known as 'conspiracies', whereby one or more faithfulness constraints has to be violated in order to satisfy the markedness constraint(s) in the grammar. To put it differently, conspiracies are applied to get rid of the surface forms of a language that come from undesirable output or structure. This can be well explained in OT, a theory that allows and makes the most of conspiracies in phonology (McCarthy, 2002). In OT, since candidates are evaluated by output constraints, so conspiracy occurs in surface forms. As McCarthy (2008: 3) says, a conspiracy is referred to as an output configuration which involves several different operations, and those operations may participate in the conspiracy by being applied or failing to be applied, depending on the circumstances. We will see many more conspiracies that occur in the grammar of Malay when we come to the discussion of each topic.

In earlier OT, according to the theory, the grammar of a language is the total ordering of the constraints in CON (cf. McCarthy, 2002). There are, therefore, high-ranking and low-ranking constraints. However, in the recent revision of OT, the grammar of a language is not a total ordering but rather a partial ordering, which means that constraints that conflict may be unranked with respect to one another. This then leads to variation in the output form (ibid.). In addition to this, in recent work developed within OT, the Multiple Grammars theory (Anttila, 2007) of language or dialect variation is then used to account for variation occurring within an individual or dialect.[11]

Unlike other earlier phonological theories, OT formally expresses that the variation that occurs within a language is through differences in ranking. By analysing the data from three Malay dialects (i.e. Perak, Kelantan and NS), we will see, in Chapter 6 Section 6.3, how differences in ranking can explain the variation that occurs in the language. The central proposal of this theory is that a grammar is a hierarchical ranking of universal well-formedness constraints. Hence, in this analysis, the phonological diversity that occurs in the Malay dialects will be analysed by imposing the condition that each dialect has a particular constraint ranking, and that this results in the differences in output when saying the same word. This exciting concept of OT will thus be used in this thesis to account for the difference in the behaviour of nasal and voiceless/voiced obstruents, and of nasal and sonorant clusters, in those three dialects; this is discussed in Chapter 6. We will see in that chapter how exciting is the idea proposed in OT that language or dialect-particular ranking deals with more than one output produced from one input in those dialects.

We now move on to the next theory of prosodic morphology which is developed within OT named Generalised Template Theory (GTT). As this thesis adopts the ideas from GTT in general and MBT in particular, thus it is incredibly important to get some ideas of what the theory is about.

[11] Interested readers can consult the author's article to get more information about Multiple Grammar.

2.2.4 Generalized Template Theory – GTT

There are some important and useful ideas raised by GTT concerning certain morphological problems, such as affixation, reduplication and truncation, which occur in various languages. The ideas are: (1) canonical shape, (2) minimality requirements and (3) the minimum size of words.

The first idea is that this work aims to develop under the umbrella of the OT of canonical morphological shapes and its role in defining the constant shape of prosodic morphemes (Downing, 2006: 33). The central idea of GTT is that the size requirement for prosodic morphemes should fall out from their morphological category (affix, stem, root), rather than constructing a specific size, such as RED=σ (cf. Downing, 2006: 24). The size of the prosodic morpheme then follows from the canonical shape held by that morphological category (ibid.).

The second idea is that GTT provides a much better explanation of word minimality requirements. In GTT, these can be explained according to two versions of the theory developed within it: (1) Prosodic-Based Template (PBT); and (2) Morpheme-Based Template (MBT). These two versions offer different explanations to account for word minimality requirements. In PBT, words are subject to the same minimality condition. In PBT, BINARITY is the constraint which requires Feet to be minimally and maximally bimoraic or disyllabic.[12] Let us consider the following example from Choctaw:

[12] For details about the BINARITY constraint and the ideas of PBT, please consult Downing (2006).

28. VCV vs. CV verb stems (adapted from Downing, 2006: 53).

ani 'to fill' iʃ-ani 'for you to fill'

abi 'to kill' iʃ-bi 'for you to kill'

29. CV verb stem in Choctaw (from Downing, 2006: 54).

/bi/	HEADEDNESS	BINARITY	DEP-IO	*VV
a.☞(abi)			*	
b. (bi)		*!		
/iʃ-bi/				
c.☞(iʃbi)				
d. (iʃabi)			*!	

In the first candidate set, the optimal candidate (a), with vowel epenthesis, satisfies BINARITY, as it is minimally disyllabic. Candidate (b) however, with no epenthesis as the output form is the same as the underlying form, violates BINARITY. In the second candidate set, vowel epenthesis would not however apply when the prefix is attached to the monosyllabic root. This is because the prefix and root are disyllabic, and therefore epenthesis is not required. The solution that Choctaw produced to explain the disyllabicity minimality conditions using the idea of PBT seems plausible for the language.

In contrast to PBT, the idea of MBT is that words are not subject to the same minimality condition. As highlighted in MBT, different categories of words are subject to different minimality conditions (Downing, 2006: 100-1). For example, disyllabicity minimality conditions only hold for derived verb stems in Arabic and Modern Hebrew, while nouns can be monosyllabic (Downing, 2006: 101). Besides, for different categories of words, it is common for derived and underived words in the same category not to be subject to the same minimality conditions (Itô and Mester, 1990). For example, fusion is productive in nasal (N-) prefixes with a polysyllabic root in Javanese, while it is blocked in monosyllabic roots. As claimed by Uhrbach (1987, cited in Downing, 2006: 101), the lack of fusion in the monosyllabic roots of Javanese is because otherwise the output derived from nasal fusion would be subject to the disyllabic minimality constraint. Let us consider the following examples from Javanese nasal fusion:

30. Disyllabicity in Javanese nasal affixation (from Downing, 2006: 136).

/N-tulis/	Prosodic Stem	Nasal Fusion	Faith-IO
a. n-tulis		*!	
b.☞nulis			*
/N-bom/			
c.☞ŋəbom		*!	*
d. mbom	*!		

In discussing word minimality requirements, MBT also provides explanation of the minimum size that words must have, if the disyllabicity minimality requirements are not satisfied. The idea that words must be of a certain minimum size is another important advantage of MBT, in addition to the constant shape of prosodic morphemes; this was never considered by earlier approaches. This theory formally explains the idea of minimum size if the minimality requirements of a particular language cannot be met. In MBT, if words are not minimally disyllabic, then they have to satisfy the principle of binary minimality of branching words. The branching principle motivating the binary minimality is formalised by the markedness constraint HEAD BRANCH (31) and branching is defined by PROSODIC BRANCHING (32) (Downing, 2006: 122).

31. **HEAD BRANCH**

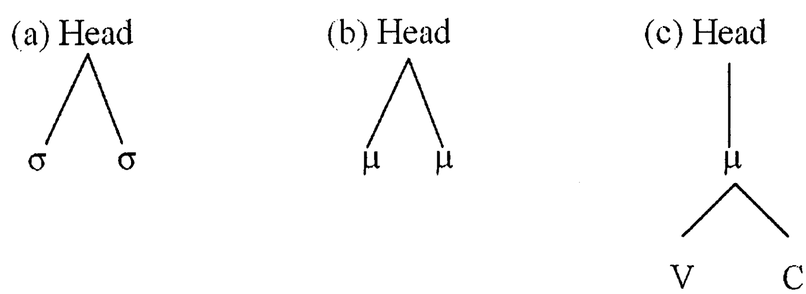
Lexical heads (roots) branch prosodically.

32. **PROSODIC BRANCHING**

A constituent branches *iff* it or its daughter contains more than one daughter.

Words that cannot meet the minimality requirements of a particular language have to satisfy one of the representations of Heads (roots), as illustrated in (33). All the representations as claimed by Downing satisfy HEAD BRANCH. The heads in 33(a) and (b) contain two syllables and moras as daughters, while in 33(c) the head dominates a mora with two daughters (Ibid.). The suitability between representations of Heads and a language differ from one language to another. By satisfying one of the representations of Heads, this could account straightforwardly for words that are not satisfied by the minimality requirements postulated by a particular language.

33. Representations of Heads (Roots) (Downing, 2006: 123).



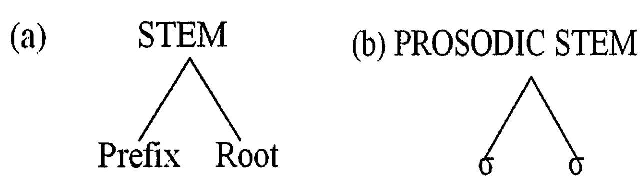
As this thesis concerns with prefixation and reduplication, thus it is worth discussing the analysis that has been done by Downing (2006) concerning those morphological problems. We begin with Downing's analysis of affixation. As claimed, stems that contain a prefix and root must satisfy the constraint called MORPHEME-SYLLABLE CORRELATION that can be defined formally as follows:

34. **MORPHEME-SYLLABLE CORRELATION** (adapted from Russell, 1997: 121, cited in Downing, 2006: 120)

Each morpheme prefix and root contains exactly one syllable.

The constraint in (34) means that a prefix must contain one syllable, and a root as well. Therefore, by MORPHEME-SYLLABLE CORRELATION, Stems must at least be disyllabic. The disyllabicity minimality required by MORPHEME-SYLLABLE CORRELATION (34) is a corollary to the constraint in (35), below, as it also formalizes the disyllabicity minimality requirement of Prosodic Stems (ibid.: 123).

35. **Prosodic Stem Minimality** (Downing, 2006: 124)



In the above diagram of Prosodic Stem minimality, the prefix contains one syllable. Cross-linguistically, the typical size of affixes is one syllable, or monomoraic. Since this theory predicts that all prosodic morphemes from the same morphological category should be subject to the same size restrictions (Downing, 2006: 26), then the size (i.e. one syllable or monomoraic) is claimed to be the canonical shape of that of prosodic morphemes (affixes). This means that affixes not containing their own syllable violate the MORPHEME-SYLLABLE CORRELATION (34), while a Stem that contains less than two syllables violates the Prosodic-Stem rule.

The next morphological problem that Downing has discussed is reduplication. Reduplication in Downing's analysis provides a new idea and analysis by claiming that reduplication is a process of self-compounding rather than of affixation. Since GTT predicts that the size of prosodic morphemes should fall out from their morphological category, then, in this regard, the size of reduplicative morphemes should also be monosyllabic or monomoraic as well, the same as for affixes. This idea has, however, been argued in other work on Prosodic Morphology, such as that of Eulenberg (1971), Niepokuj (1991), Inkelas (2005), Inkelas and Zoll (2005) and Downing (2006). Because the size of reduplicative morphemes exceeds one syllable, as required by affixes, then reduplication, particularly total reduplication, is best treated as a process of compounding rather than affixation.

Besides the above issue raised in MBT, concerning reduplication, there is another issue which is also important and relevant to Malay. This is about the difference sizes of reduplicative morphemes from the same type of reduplication. As stated in Downing (2006), it is a problem for MBT when several morphemes of the same category have different canonical shapes. For example, there are two types of reduplicative morphemes in Ilokano partial reduplication – light and heavy – as (36) illustrates:

36. Heavy and light reduplication in Ilokano (Hayes and Abad, 1989: 357, cited in Downing, 2006: 11).

(a) Light reduplication

bunéŋ 'kind of knife'	si-<u>bu</u>- bunéŋ 'carrying a *bunéŋ*'
sáŋit 'to cry'	ʔagin-<u>sa</u>- sáŋit 'pretend to cry'
pandilíŋ 'skirt'	si-<u>pa</u>-pandilíŋ 'wearing a skirt'

trabáho 'to work' ʔagin-<u>tra</u>-trabáho 'pretend to work'

(b) Heavy reduplication

kaldíŋ 'goat'	<u>kal</u>-kaldíŋ 'goats'
púsa 'cat'	<u>pus</u>-púsa 'cats'
sáŋit 'to cry'	ʔag-<u>saŋ</u>- sáŋit 'is crying'
trabáho 'to work'	ʔag-<u>trab</u>-trabáho 'is working'

It is difficult if not impossible for MBT to account for the two patterns of reduplicative morphemes in Ilokano. Therefore, an alternative solution has been suggested to resolve the case of Ilokano reduplication, where the reduplicative morphemes are different in canonical shape. The alternative solution applies co-phonologies, i.e. different constraint rankings. In co-phonology, languages which have two (or more) reduplication patterns can be neatly accounted for. So what are the advantages of applying co-phonology to account for cases like this? As claimed in Downing (2008), co-phonologies account well for the case where reduplicative morphemes of the same categories have different canonical shapes. The co-phonology analysis of the two patterns of Ilokano reduplication can be seen in the following tableaux. This discussion continues in Chapter 5, Section 5.4.

37. (a) Ilokano light reduplication (from Downing, 2006: 245).

MORPH-SYLL, NOCODA, *VV >> MAX-BR

/Heavy RED-trabaho/	MORPH-SYLL	NOCODA	*VV	MAX-BR
a. trab-trabaho	*	*!		***
b. tra:-trabaho	*		*!	****
c. trabaho-trabaho	**!			
d.☞tra-trabaho	*			****

(b) Ilokano heavy reduplication (from Downing, 2006: 246)

MORPH-SYLL >> MAX-BR >> NOCODA, *VV

Light RED- trabaho	MORPH-SYLL	MAX-BR	NO CODA	*VV
a.☞trab-trabaho	*	***	*	
b. tra:-trabaho	*	****!		*
c. trabaho-trabaho	**!			
d. tra-trabaho	*	****!		

Works such as Inkelas (1998), Inkelas and Orgun (1998), Inkelas and Zol (2000, 2005) and Orgun (1996, 1998) propose that different prosodic morphemes in the same language can have different optimal outputs. The different optimal outputs are accounted for by co-phonologies. In the case of Illokano presented in the above tableau, the markedness constraint, NOCODA, ranks highly for the light reduplicative morpheme, while it ranks low for the heavy reduplicative morpheme. In co-phonologies, the markedness constraint is allowed to have one ranking in one morphological construction, and the opposite ranking in the other. This is called "Markedness Reversals" in the theory where the relevant markedness constraint can be ranked in reverse in some morphological constructions in the same language (Downing, 2008). It is worth mentioning that the co-phonology analysis done by Downing on Ilokano reduplication is compatible with that of Itô and Mester (1999) where the markedness constraint, *N$\mathring{\text{Ç}}$, has been ranked differently in the lexicon of Japanese, as they show distinction in the degree of assimilation.

2.3 Malay Phonology Scholarship

In this section, I will discuss how rule-based, non-linear autosegmental phonology and OT have been applied in Malay scholars' works. In each subsection, we will see how prefixation and reduplication have been accounted for as well as the problems arisen in the analyses which afterwards need MBT as the solution to overcome them.

2.3.1 Rule-based analysis

Generally, most of earlier studies of Malay concerning prefixation and reduplication have applied rule-based analysis. The discussions were focused on the classification of forms, grammatical functions and the meanings of affixes (e.g.: Hassan, 1974, 1987; Omar, 1975, 1993; Karim, 1995; Onn, 1980). In Omar's (1986) description, for example, prefixation in Malay was described according to the meaning conveyed by the prefix, as the following examples show:

38. (a) məN- with the meaning of 'become like a... ' (from Omar, 1986: 106).

(i) /məN-abdi/ [məŋabdi]
ACT.PRF–slave
'become like a slave'

(ii) /məN-batu/ [məmbatu]
ACT.PRF- stone
'become like a stone'

(b) məN- with the meaning of 'producing a sound' (from Omar, 1986: 108)

(i) /məN-uak/ [məŋuwak]
ACT.PRF-croak
'to croak'

(ii) /məN-aduh/ [məŋaduh]
ACT.PRF-exclamation of pain
'to groan'

Prefixes which end with nasal consonants, such as /məN-/ and /pəN-/, have some allomorphs and their distribution depends on the phonological characteristic, namely the segment (Omar, 1986; Hassan, 1987; Karim et al., 1989). The nasal segments of those prefixes assimilate to the subsequent initial segments of the following syllable of a root. Each of these prefixes has six variants: [mə], [mən], [məm], [məŋ], [məɲ] and [məŋə]; and [pə], [pən], [pəm], [pəŋ], [pəɲ] and [pəŋə]. The occurrence of these variants depends on the first consonant of the root to which they are attached (Omar, 1986; Hassan, 1987; Karim et al., 1989). In the process of combining a root to a prefix, an alternation in some segment, at either prefix or root, will occur. As claimed by Malay scholars (e.g.: Hassan, 1974; Omar, 1986; Hassan, 1987; Karim et al., 1989, 1994; Karim, 1995; Ahmad, 1993), the process of combining a prefix to a root can be summarised as follows:

39.	mə- and pə-	before l, r, m, n, ɲ and ŋ.
	məm- and pəm-	before p, b, f, v
	mən- and pən	before t, d, c, ʤ
	məɲ- and pəɲ-	before c, s
	məŋ- and pəŋ	before k, g, h, x, ɣ, w and vowels
	məŋe- and pəŋə	before a monosyllabic base.

It has long been observed that the obstruent voiceless consonants, /p, t, k and s/, in Malay affixation are deleted when the consonants are concatenated with nasal final prefixes, /pəN-/ and /məN-/. Meanwhile, the phonological behaviour of the nasal segments in the prefixes is always homorganic to the following consonant of the root. Let us consider some relevant examples, below, as cited in Karim et al. (1994). Note that this phonological behaviour of nasal segments in prefixes is also found in the corpus observed in this study (see Chapter 4, subsection 4.2.1.2).

40. Nasal final prefixes in Malay (from Karim et al., 1994: 147).

a) /məŋ-pukul/ [məmukul]
ACT.PRF-scold 'to scold'
b) /məŋ-tari/ [mənari]
ACT.PRF-dance 'to dance'
c) /məŋ-karaŋ/ [məŋaraŋ]
ACT.PRF-compose 'to compose'
d) /məŋ-sinar/ [məɲinar]
ACT.PRF-ray 'to ray'

From the data above, it is clear that the nasal segment is homorganic with the following voiceless consonant of the root. Thus, the descriptive generalization that is presented by Karim et al. (1994), above, can be summarized as: the active prefix /meN-/ becomes [məm], [mən], [məŋ] and [məɲ] when the following consonant is /p/, /t/, /k/ or /s/, respectively. Although the examples given fulfil the descriptive rule, they may not be able to explain the real process of prefixation in Malay, since there is evidence that some voiceless obstruent consonants are not deleted when the combining process occurs. This phenomenon of undeleted voiceless obstruents, as claimed by scholars in many cases, has been retained. Most of them resort to the same solution, which is to treat the phenomenon as somehow exceptional.

The following data from the DBP-UKM corpus provide evidence of this phenomenon. The undeleted voiceless obstruents have resulted in nasal and voiceless obstruent clusters in surface representations, as in (41). I shall argue later, in Chapter 4, that the phenomenon treated as exceptional by Malay scholars (e.g.: Omar, 1986; Karim, 1995) can be accounted for by constraint-based analysis. The theoretical framework of OT offers a way of treating those problematic examples shown in (41) without claiming them to be exceptions. That is why OT is superior to other earlier approaches as it does not seek simplistic solutions. How OT deals with cases such as those in (41), and the same case as exemplified in (9), is discussed in Chapter 4, subsection 4.2.1.2.

41. Nasal final prefixes (from the DBP-UKM corpus).

(i)	/məŋ-paked͡ʒ-kan/	[mə**mp**aked͡ʒkan]
	ACT.PRF-package-CAUS.SUF	
	'to cause to pack for'	
(ii)	/məŋ-kelas-kan/	[mə**ŋk**əlaskan]
	ACT.PRF-class-CAUS.SUF	
	'to cause to categorise for'	
(iii)	/məŋ-kabul-kan/	[mə**ŋk**abulkan]
	ACT.PRF-grant-CAUS.SUF	
	'to cause to grant for'	

From the above review of Malay morphology scholarship, we see that the type of root that prefixes can be attached to is disyllabic roots. Apart from that, the aforementioned scholars have also analysed the other type of root that prefixes can be attached to, i.e. monosyllabic roots. Note that prefixes can also be attached to trisyllabic roots. Since most Malay words contain disyllabic roots (see Figure 4 on p. 133), the examples given above contain disyllabic roots. In what follows, I present how monosyllabic bases were analysed in previous studies.

In Malay, the process of prefixation to monosyllabic roots is somewhat different from the one for disyllabic roots. In the process of prefixation to monosyllabic roots, the initial voiceless consonants of roots are retained when nasal final prefixes are attached. This is in contrast to the process where nasal final prefixes are attached to disyllabic roots. As postulated by Malay scholars in rule-based analysis (e.g.: Karim et al., 1989: 110; Omar, 1986: 50, 121; Othman, 1981; Hassan, 1987; Abas, 1975), /məN-/ becomes [məŋə] if it is attached to a monosyllabic root. The nasal final prefixes /məN-/ and /pəN-/ each have only one allomorph, [məŋə] and [pəŋə], respectively.

Scholars using rule-based analysis provide no phonological explanation of why /məN-/ should always become [məŋə] for all monosyllabic roots, while this is not the case for disyllabic roots. This unexplained phonological matter will be discussed in detail in this study (see Chapter 4, subsection 4.2.1.1). As we shall see in that section, how /məN-/ becomes [məŋə] can be satisfactorily explained within the theoretical framework adopted in this study.

In the above literature review, we see that a single prefix is attached to roots. Now we shall review Malay scholars' work in situations where two prefixes are attached to a root, the so-called multiple prefixes. In previous studies, there are at least two analyses concerning multiple prefixation. These are, by way of example, Omar (1986) and Karim et al. (1989). There appears to be some disagreement between Omar's (1986) and Karim et al.'s (1989) views. The two contradicting views, regarding the deletion of voiceless obstruents when they concatenate to preceding prefixes, can be seen in the grammarians' analyses.

Omar (1986: 51) claims that a voiceless consonant which has already been attached to another prefix cannot be deleted when it is attached to a nasal final prefix, in order to maintain the form of the prefix, as can be seen in the examples below:

42. Multiple prefixes by Omar (1986).

(i) /məŋ-tər-tawa-kan/ [mən-tər-tawa-kan]
VERBL-NOM-laugh-CAUS.SUF
'unintentionally laugh'

(ii) /məŋ-pər-kata-kan/ [məm-pər-kata-kan]
VERBL-VERBL-say-CAUS.SUF
'to cause to discuss'

(iii) /pəŋ-kə-bumi-an/ [pəŋ-kə-bumi-jan]
NOM.PREP-earth-NOM
'burial'

As Omar (ibid.: 8, 51) claims, /t/ in 42(i) has been retained, whereas /t/ in /tulis/ [me-nulis] is deleted, since [ter-] in 42(i) is a prefix and the remaining [t] will maintain the position of /ter-/ as a prefix. The same reason is given for 42(ii) and (iii), for [per-kata-kan] and [ke-bumi-jan], respectively.

Karim's et al. (1989: 84) description, on the other hand, is slightly different from Omar's (1986) in the sense that, according to Karim (ibid.), a voiceless consonant must be deleted when it is concatenated to another prefix, as can be seen in the examples below:

43. Multiples prefixes by Karim et al., (1989).

(i) /pəŋ-kə-tua/ [pəŋətuə]
NOM-NOM-old
'headmaster/headmistress'

(ii) /məŋ-kə-təpi-kan/ [məŋətəpikan]
ACT.PRF-NOM-side-CAUS.SUF
'to cause to set aside for'

(iii) /pəŋ-kə-bumi-an/ [pəŋəbumijan]
NOM-NOM-earth-NOM.SUF
'burial'

From the above examples, /k/ in the second prefix in 43(i-iii) is deleted when concatenated to nasal final prefixes /məŋ-/ and /pəŋ-/, respectively, in the same way as in single prefixation. In contrast to Omar's (1986) analysis, /k/ in this case should not be deleted. Hence, the words in 43(i-iii) become [pəŋkətuə], [məŋkətəpikan] and [pəŋkəbumijan]. It is apparent that the difference in the analyses of the grammarians has resulted in different outputs.

Reflecting on Karim's et al. (1989) analysis, Karim maintains that voiceless consonants must be deleted when they are attached to another prefix, as can be seen in (43). However, Karim et al. (1989) express the same findings as Omar (1986) for the words /mən-tər-tawa-kan/ and /məŋ-pər-kata-kan/. The outputs for these words are not *[mə-nər-tawa-kan] and *[mə-mər-kata-kan], as postulated in their rule. On this matter, Karim et al. (1989: 109) claim that [mən-tər-tawa-kan] is the correct output, whereby /t/ is not deleted as it is considered an exception in the language. Meanwhile, /məmpər-/ in the word [məm-pər-kata-kan] is analysed as one morpheme, therefore it cannot be deleted and has to remain as [məmpər-]. Obviously, the grammarians are in agreement in terms of the spelled words but not in the explanation, in particular why [t] and [p] have not undergone the deletion process (Ahmad, 1993). Such analyses are weak observationally and lack explanatory adequacy as the voiceless consonants are either present, as in [mən-tər-tawa-kan], or absent, as in [pə-ŋə-bumi-jan].

Besides prefixation, reduplication has also been widely studied by Malay scholars, particularly partial reduplication. In rule-based analyses (e.g.: Hassan, 1974; Othman, 1981; Omar, 1993; Karim et al., 1994), partial reduplication in Malay is defined as a process of copying the initial syllable of the base. I exemplify some relevant examples for this. The initial syllable of the base is emboldened and the reduplicant underlined:

44.	(i) /**si**ku/	[sə-siku]
	elbow	RED-elbow 'square'
	(ii) /**ku**da/	[kə-kuda]
	horse	RED-horse 'trestle'
	(iii) /**bu**dak/	[bə-budaʔ]
	child	RED-child 'children'

The definition given by those scholars does not however really explain the actual process of partial reduplication, since the process of copying the first syllable of the base is not true for bases with heavy initial syllable. The following examples show this.

45.	(i) /**ram**but/	[rə-rambot]
	hair	RED-hair 'fine body hair'
	(ii) /**sən**di/	[sə-səndi]
	joint	RED-joint 'joints'
	(iii) /**taŋ**ga/	[tə-tanŋgə]
	stairs	RED-stairs 'neighbours'

The examples in (45) show that the reduplicant in partial reduplication does not derive simply from copying the initial syllable of the base, since the coda consonant of the base is not copied into the reduplicant. I therefore suggest that the definition needs to be revised in order to explain what exactly is copied from the base for both light and heavy initial syllables. In Chapter 5, Section 5.4, I provide a revised version of the copying process involving the reduplicant.

In the following subsection, we see how non-linear autosegmental phonology analysis applies in Malay scholars to account for both prefixation and reduplication.

2.3.2 Non-linear autosegmental phonology

In non-linear autosegmental phonology, Malay scholars i.e. Teoh (1994) and Ahmad (2000b) have focused on a particular issue in prefixation and reduplication. The process of prefixation to monosyllabic roots has been highlighted in both scholars' analysis. Teoh and Ahmad point out one important generalisation when monosyllabic roots are discussed. They point out that Malay is a disyllabic language. When scholars describe Malay as a disyllabic language, they mean that words in Malay minimally contain two syllables. Therefore, they postulate CV.CVC (Teoh, 1994) and V.CVC (Ahmad, 2000b) as templates for monosyllabic roots, since they are minimally disyllabic (see Chapter 4, subsection 4.2.1.1). Let us see how those scholars dealt with this matter in autosegmental analysis.

46. (a) Teoh (1994)

```
m ə ŋ     l a p
| | |     | | |
C V C C V C V C
```

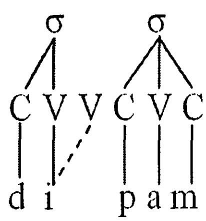
(b) Ahmad (2000b)

```
m ə ŋ   l a p
| | |   | | |
C V C V C V C
```

In these analyses, both Teoh and Ahmad postulate the above underlying representation when a monosyllabic root is attached to a nasal final prefix. From 46(a) and (b), we see that a surface monosyllabic root contains two syllables in the underlying representation. In my view, the templates proposed by both scholars seem to represent a problem in Malay grammar, because a monosyllabic root does not emerge as two syllables in the surface form. This claim obviously contradicts the present proposal. As I will argue further, in subsection 4.2.1.1, not all words in Malay must be disyllabic. That means they can also be monosyllabic. Given this fact about Malay, I will argue that a form such as *[ə.pam] is not grammatical, since schwa only appears when a nasal final prefix is attached to a monosyllabic root. To account for this, where words in Malay can also be monosyllabic in size, the idea of MBT regarding words' minimality is useful to explain the difference in the minimal size of words in Malay.

In addition to the proposed template for monosyllabic roots, Ahmad further claims that the VCVC template is able to account for prefix /di+/ with monosyllabic bases. He identifies a new output from the analysis. In order to fill the empty V slot, the final vowel in the prefix undergoes a process of vowel lengthening. I agree with Ahmad in this case where the vowel /i/ in the prefix gets lengthened. As I will claim later, in subsection 4.2.1.1, vowel lengthening occurs to satisfy the canonical shape for an affix. I briefly illustrate how the vowel in the prefix /di+/ gets lengthened:

47. Vowel lengthening in prefix /di+/ (from Ahmad, 2000b).



[di: pam]

On the other hand, in autosegmental analysis (e.g. Ahmad, 2000a), CV strings are used to account for the process of partial reduplication in SM. I argue that this CV template cannot account for both light and heavy reduplicative morphemes if Perak partial reduplication is considered. This is because reduplicative morphemes in PD do not only consist of the CV which is found in SM. In PD, CVC is one of the reduplicative morphemes, as well as CV. The size of the reduplicative morpheme is determined by the final syllable of the base. A base with a closed final syllable forms a CVC reduplicative morpheme, while a CV reduplicative morpheme is derived from a base with an open final syllable, as the following examples illustrate, cited in Aripin (2005: 171):

48. (a) Partial reduplication for bases with open final syllable

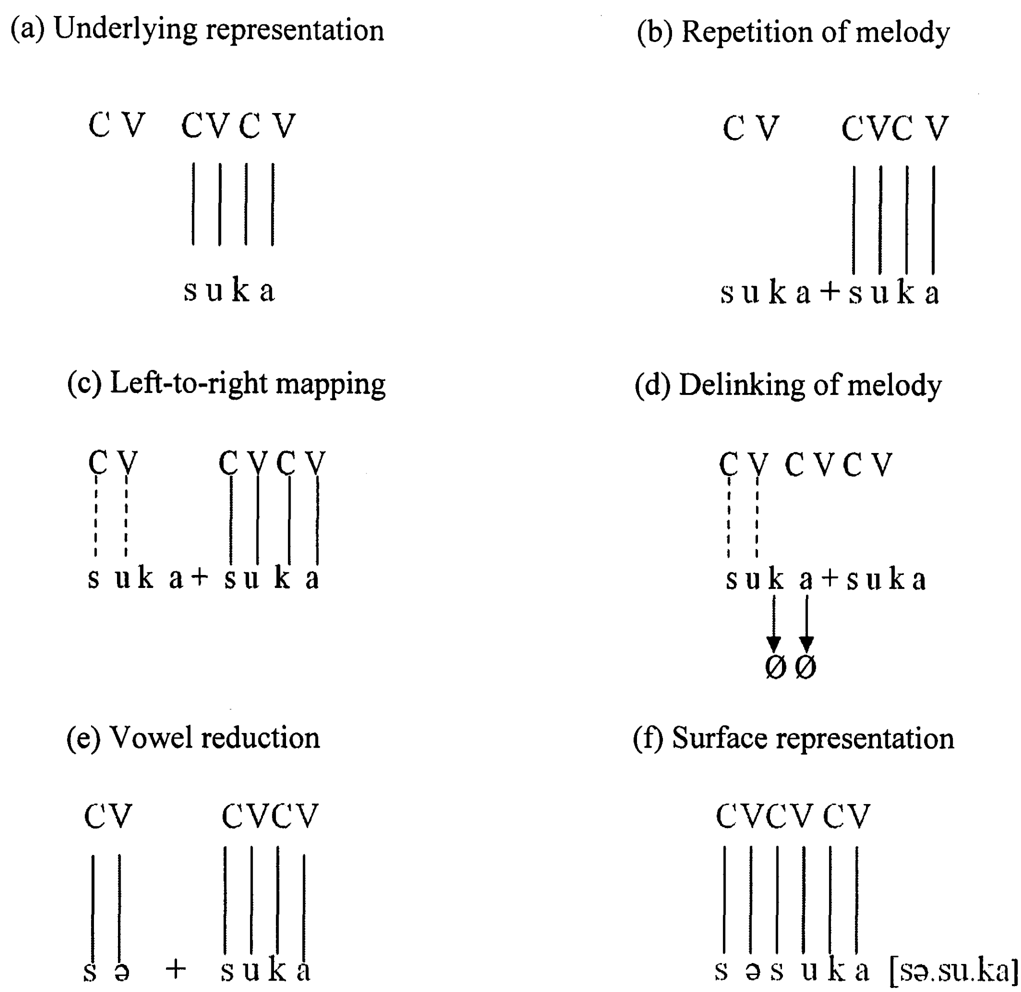
(i)	/taŋgə/ 'stairs'	[tə-taŋgə] RED-stairs 'neighbours'
(ii)	/kuda/ 'horse'	[kə-kudə] RED-horse 'trestle'
(iv)	/buku/ 'book'	[bə-buku] RED-book 'books'
(v)	/laki/ 'man/boy'	[lə-laki] RED-man/boy 'men/boys'

(b) Partial reduplication for bases with closed final syllable

(i)	/pətaŋ/ 'evening'	[pəm-pətaŋ] RED-evening 'evenings'
(ii)	/rambut/ 'shoot'	[rəʔ-rambuʔ] RED-hair 'shoots'
(iii)	/kawan/ 'friend'	[kəŋ-kawan] RED-friend 'friends'
(iv)	/təraŋ/ 'bright'	[tən-təraŋ] RED-bright 'brightest'

As we can see in (49), the CV template used in autosegmental phonology analysis, by Ahmad (2000a), to account for SM reduplication can satisfactorily explain CV reduplication. One of the reasons why the proposed CV template matches the data is that reduplication in SM only consists of CV elements as the reduplicative morpheme. This CV template, however, cannot be used to account for CVC reduplicative morphemes, as in 48(b). I will demonstrate in (50) how this CV template fails to account for such data in 48(b).

49. CV reduplication in SM (Ahmad, 2000a).



As can be seen, only consonant and vowel segments are linked into the CV template. Segments that cannot be linked into the template are deleted. If data from a Malay dialect, i.e. Perak, which consist of CVC reduplicative strings were to be considered in the analysis, I am convinced that the above CV template proposed by Ahmad (2000a) might have difficulty in capturing reduplication with CVC reduplicative morphemes, as in 25(b). I briefly demonstrate below how the CV template fails to capture CVC reduplicative morphemes.

50.

C V CVCVC   CV   CVCVC

pətaŋ   pə t a ŋ+pət aŋ   *[pə-pətaŋ]

It is apparent from the above representation that the CV template used in autosegmental analysis fails to account for CVC reduplicative morphemes. The output *[pə-pətaŋ] derived from the analysis is not the desired output. Like autosegmental analysis, other earlier analytical approaches that are ruled-based fail to account for the data in 48(b). Based on how reduplication is defined in ruled-based analysis (e.g.: Hassan, 1974; Othman, 1981, Omar, 1993; Karim et al., 1994), I too am convinced that the data in 48(b) are difficult if not impossible to capture. If the first syllable of /pətaŋ/ were copied, we would get the same output, *[pə-pətaŋ].

Thus, the present study will attempt to provide an analysis which can explain all the patterns of partial reduplication, both in SM and in non-standard dialects. As we will see in Chapter 5 (Section 5.4), I will argue that OT provides us with the best way of analysing and understanding the behaviour of partial reduplication in Malay. Given the insights afforded by OT, the two patterns for reduplicative morphemes – CV and CVC – can be accounted for in an adequate explanatory manner if the analysis of co-phonology – different constraint rankings – is adopted. This can be done by proposing that CV and CVC reduplication are associated with distinct, construction-specific constraint rankings, whereby the constraint ranking of CV reduplication is different from that for CVC reduplication. The analysis of this will be presented in further detail in Section 5.4.

2.3.3 OT

As we have just seen in the previous two subsections, rule-based and non-linear approaches have been applied in Malay scholars' works. In this subsection, I am going to discuss another approach in phonology i.e. which has also been applied in Malay. It is worth mentioning that, there is only work done by Ahmad (2005) who has applied OT in accounting for a number morphological and phonological processes i.e. syllable structure, suffixation, prefixation and reduplication. In what follows, I am going to discuss what has been done in Ahmad's work concerning prefixation and reduplication, as these topics are relevant to the present study.

In prefixation, two adjacent segments i.e. the final consonant of the prefix and the initial consonant of the root raise a number of issues in the morphological and phonological system of the language. The two adjacent segments can pose a problem if the combination of them results in the emergence of the prohibited clusters in the language i.e. (1) nasal and sonorant and (2) nasal and voiceless obstruent clusters.

Ahmad (2005) discusses the occurrence of nasal and sonorant clusters at prefix-root boundary. As was mentioned in Chapter 1, Malay does not permit nasal-sonorant clusters in a word. Teoh (1994: 97) says, "Nasal deletion in Malay as the phonotactic of the language in general does not allow sonorant clusters even across a morpheme boundary". Nevertheless, there are many instances where nasal and sonorant clusters occur in root-internally in the language as exemplified in (51). As was mentioned in Chapter 1, such examples in (51) however, are claimed by Ahmad (1989) and Teoh (1994) as non-native words of Malay.

51. Nasal and sonorant cluster root-internally (from Ahmad 2005: 120).

/loŋlai/	[lo*ŋl*ai] 'gracefully'
/paŋlima/	[pa*ŋl*ima] 'admiral'
/pərlu/	[pə*rl*u] 'necessary'
/pərnah/	[pə*rn*ah] 'ever'
/ʤumlah/	[ʤu*ml*ah] 'total'

In the OT analysis done by Ahmad (2005: 123), nasal and sonorant clusters in root-internally are captured by a constraint named, ROOT CONTIGUITY. By adding this constraint in the ranking, candidate with epenthesis or deletion segment cannot be chosen as the optimal output. In order to sonorant clusters candidate emerge as the winner, ROOT CONTIGUITY must be ranked higher. Therefore, sonorant clusters are preserved in root-internally. The following tableau illustrates the case we just discussed.

52. Preservation of sonorant clusters root-internally (from Ahmad 2005: 123)

/loŋlai/	ROOT CONTIG	OCP	MAX-IO$_{cons}$
a. ☞loŋ.laj		*	
b. lo.laj	*!		*
c. lo.ŋə.laj	*!		

Unlike in root-internally, nasal and sonorant clusters at prefix-root boundary are resolved by deleting one of the sonorant segments i.e. the first sonorant consonant is deleted. In order to account for the deletion one of the sonorant segments at the prefix-root boundary, an important constraint which is required to be added in the ranking is OCP. Work like Leben (1973), Goldsmith (1976), McCarthy (1986, 1988), Mester (1986), Itô and Mester (1986) and Yip (1988) suggest that the deletion of a sonorant consonant before another sonorant consonant must be driven by the OCP, which can formally be defined as follows:

53. **OCP ([+sonorant, +consonantal])**

Adjacent identical [+sonorant, +consonantal] features are prohibited.

As defined above, OCP requires the first sonorant is deleted before another sonorant at prefix-root boundary. We see in the following tableau how OCP plays its role in deleting the first of the two sonorant clusters at prefix-root boundary:

54. Nasal deletion at prefix-root boundary (from Ahmad 2005: 121)

/məŋ+lompat/	DEP-IO	OCP	ALIGN PREFIX	MAX-IO_{CONS}
a. məŋ.lom.pat		*!		
b. mə.ŋə.lom.pat	*!		*	
c. ☞mə.lom.pat			*	*

As well as nasal and sonorant clusters, there is another clusters discussed by Ahmad i.e. nasal and voiceless obstruent clusters. In discussing nasal and voiceless obstruent clusters in the language, Ahmad (2005) applies *NÇ̥ as the constraint which plays a crucial role to eliminate the clusters.

The clusters therefore undergo nasal substation. By satisfying this constraint, candidate with nasal substitution compels a violation of a faithfulness constraint named UNIFORMITY which prohibits two or more input segments from sharing an output correspondent (McCarthy & Prince 1995; McCarthy 1995; Lamontagne & Rice 1995 and Pater 1999, cited in Ahmad 2005: 127).

55. **UNIFORMITY** ('No Coalescence') (McCarthy and Prince, 1999: 296)

No element of S_2 has multiple correspondents in S_1.

For x, y ϵ S_1 and z ϵ S_2: if x R z and y R z, then x=y.

The interaction of *NÇ and UNIFORMITY can be seen in the following constraint ranking: DEP-IO_{VOW} >> *NÇ >> ALIGN PREFIX >> MAX-IO_{CONS} >> UNIFORMITY.

56.

/məŋ+pukol/	DEP-IO_{VOW}	*NÇ	ALIGN PREFIX	MAX-IO_{CONS}	UNIFORMITY
a. məm.pu.kol		*!			
b. mə.pu.kol			*	*!	
c. mə.ŋə.pu.kol	*!		*		
d.☞mə.mu.kol			*		*

It is clear from the above tableau, candidate (c) has been ruled out earlier as it fatally violates the highest constraint in the hierarchy, DEP-IO_{VOW}. Candidate (a) violates *NÇ, as it contains a sequence of nasal and voiceless obstruent. The remaining candidates (b) and (d) are in a tie position and subject to evaluation by the next constraint, ALIGN PREFIX. As we can see, both candidates violate ALIGN PREFIX because the right edges of the prefixes do not coincide with the right edges of both syllables. Since candidate (b) and (d) both violate ALIGN PREFIX, the evaluation is now proceed to the next constraint i.e. MAX-IO_{CONS}. Candidate (b) which deletes the nasal segment in the prefix incurs this violation. Thus the winner is candidate (d).

55. **UNIFORMITY** ('No Coalescence') (McCarthy and Prince, 1999: 296)

No element of S_2 has multiple correspondents in S_1.

For x, y ϵ S_1 and z ϵ S_2: if x R z and y R z, then x=y.

The interaction of *NÇ̥ and UNIFORMITY can be seen in the following constraint ranking: DEP-IO$_{VOW}$ >> *NÇ̥ >> ALIGN PREFIX >> MAX-IO$_{CONS}$ >> UNIFORMITY.

56.

/məŋ+pukol/	DEP-IO$_{VOW}$	*NÇ̥	ALIGN PREFIX	MAX-IO$_{CONS}$	UNIFORMITY
a. məm.pu.kol		*!			
b. mə.pu.kol			*	*!	
c. mə.ŋə.pu.kol	*!		*		
d.☞mə.mu.kol			*		*

It is clear from the above tableau, candidate (c) has been ruled out earlier as it fatally violates the highest constraint in the hierarchy, DEP-IO$_{VOW}$. Candidate (a) violates *NÇ̥, as it contains a sequence of nasal and voiceless obstruent. The remaining candidates (b) and (d) are in a tie position and subject to evaluation by the next constraint, ALIGN PREFIX. As we can see, both candidates violate ALIGN PREFIX because the right edges of the prefixes do not coincide with the right edges of both syllables. Since candidate (b) and (d) both violate ALIGN PREFIX, the evaluation is now proceed to the next constraint i.e. MAX-IO$_{CONS}$. Candidate (b) which deletes the nasal segment in the prefix incurs this violation. Thus the winner is candidate (d).

We have just discussed prefixation in OT work done by Ahmad. We now move on to reduplication. In analysing reduplication, Ahmad has given more attention among other issues to root reduplication and overapplication processes to reduplicative morphemes. I begin the review of Ahmad's work with root reduplication. There are two types of root have been analysed: (1) root ends with nasal consonant and (2) root begins with a vowel. In discussion these roots, Ahmad has highlighted a constraint that plays a crucial role i.e. ONSET which is driven by a process of glottal epenthesis, nasal assimilation and nasal coalescence.

The ONSET requirement is needed especially when roots begin with vowel are reduplicated. In Malay, glottal stop is a default consonant and is always used to break up vowel clusters in words. As we see in the tableau in (59), this default consonant is inserted to provide an onset to the syllable. Shown below are some of the examples of vowel initial roots given by Ahmad:

57. Vowel initial roots (Ahmad 2005: 144)

a) /api/ 'fire'	[api-ʔapi]
b) /alu/'welcome'	[alu-ʔalu]
c) /ikot/ 'follow'	[ikot-ʔikot]

In order to account for the root reduplication, R=ROOT has been proposed as an important constraint that must be added into the constraint ranking. This constraint ensures the reduplicant is identical to the root, as can be defined as follows:

58. **R=ROOT**

The reduplicant is identical to the base.

Since onset is highly required in the language, R=ROOT is violated in vowel initial roots that have been reduplicated. Since this type of root violates R=ROOT, this constraint therefore must be ranked beneath ONSET in order to ensure candidate with onset emerge as the optimal. To account for this, the following constraint ranking is proposed: ANCHOR RIGHT, ALIGN LEFT >> ONSET >> R=ROOT >> MAX-BR.

59.

/api+RED/	ANCHOR RIGHT	ALIGN LEFT	ONSET	R=ROOT	MAX-BR
a) a.pi-a.pi			**!		
b) ʔa.pi.-ʔa.pi		*!		*	
c) a.pi.ʔ-a.pi	*!				*
d)☞a.pi.-ʔa.pi			*	*	

In Ahmad's analysis, reduplication is claimed as suffix. Therefore ANCHOR RIGHT is the constraint that ensures the relation between the base and the reduplicative morpheme. Since candidate (c) is ill-anchored, it therefore has been ruled out. In candidate (b), glottal stop is inserted in both, the root and the reduplicative morpheme. The insertion of glottal stop in the left position of the root and the reduplicative morpheme results in the candidate violates ALIGN LEFT. Candidate (a) which does not have glottal stop in both, the root and the reduplicative morpheme violates ONSET twice. The later candidate i.e. candidate (d) which only violates ONSET once is chosen the optimal output.

The second type of root that Ahmad has analysed is roots end with nasal consonant. It is claimed that the final nasal consonant of the root does not undergo assimilation and coalescence processes when it is reduplicated. This is shown in the following tableau, as cited in Ahmad (2005: 145):

60. ALIGN RIGHT >> ALIGN NASAL >> *NÇ̊

/pandaŋ+RED/	ALIGN RIGHT	ALIGN NASAL	*NÇ̊
a)☞pandaŋ-pandaŋ		**	*
b) pandam-pandam	*!	*	*
c) panda-mandaŋ	*!	*	

It is shown in the above tableau that an assimilated and coalesced candidates i.e. (b) and (c), respectively violate ALIGN RIGHT. Therefore candidate (a) emerges as the optimal output as all the segments in the input are retained in the output.

From the discussion above concerning root reduplication, we note that Ahmad only accounts for disyllabic roots. Monosyllabic roots however have been missed in his analysis. Can monosyllabic roots be reduplicated as well as disyllabic roots? If so, how the process is. This is one of the points that this present study will be focus on when reduplication is analysed (see Chapter 5).

Besides roots reduplication, Ahmad has also discussed overapplication processes in both reduplicant to base and base to reduplicant. In the review of this work however, I will only discuss the latter as it is relevant to the present study. Overapplication process in base to reduplicant means that the PHONO-CONS applied in the base gives effect to the reduplicant. Before we go further, let us consider the examples below:

61. Overapplication process in base to reduplicant (from Ahmad, 2005: 150)

Roots	**Reduplicated forms**	
pukol 'hit'	mə̃mũkol-mũkol	*mə̃mũkol-pukol
pandai 'clever'	mə̃mãndai-mãndai	*mə̃mãndai-pandai

tari 'dance'	mə̃nãri-nãri	*mə̃nãri-tari
tahan 'stop'	mə̃nãhan-nãhan	*mə̃nãhan-tahan
kutep 'pick'	mə̃ŋũtep-ŋũtep	*mə̃ŋũtep-kutep
kumpol 'gather'	mə̃ŋũmpol-ŋũmpol	*mə̃ŋũmpol-kumpol
sikat 'comb'	mə̃ɲĩkat-ɲĩkat	*mə̃ɲĩkat-sikat

The above examples show that both, the bases and the reduplicative morphemes undergo nasal coalescence. The initial voiceless obstruents in the roots undergo nasal coalescence when prefixes ended with nasal are attached to them. The coalescence form then affects the reduplicant form. This situation as claimed by Ahmad as an overapplication process that occurs between base and the reduplicant which is caused by IDENT-BR [Nasal] constraint:

62. **IDENT-BR [Nasal]**

The correspondent segments in the base and the reduplicant must have identical values for the feature [nasal].

As we observed in the examples in (61), the reduplicant follows the base whereby the reduplicant also undergoes coalescence, as the base is. In other words, the reduplicant obeys the base than the root. In this case therefore, IDENT-BR [Nasal] should be ranked higher in the hierarchy than R=ROOT. I exemplify the following tableau:

63. Overapplication of nasal coalescence Ahmad (Ibid.: 152)

PHONO-CONS >> IO-FAITH >> IDENT-BR [Nasal] >> R=ROOT >> MAX-BR

/məŋ+tari+RED/	PHONO-CONS	IO-FAITH	IDENT-BR [Nasal]	R=ROOT	MAX-BR
a. məntari-tari	*!				***
b. mənari-tari		*	*!		**
c.☞mənari-nari		*		*	**

From the analysis of overapplication process, I agree with Ahmad that the copying of the coalesced nasal of the base into the reduplicant is an overapplication process. This analysis however, lack of some data which are also involved in reduplication. In analysing overapplication process occurs in reduplication, Ahmad has merely focused on disyllabic roots. It should be mentioned that monosyllabic roots also involve in overapplication process. This can be seen in the word like /pam/ 'pump' → [məŋəpam] → [məŋəpam-ŋəpam] 'to pump repeatedly', not *[məŋəpam-pam] whereby the nasal segment in the prefix is copied into the reduplicant.

As we have seen in the data in (61), the roots begin with voiceless stops. The data however have omitted voiced stops initial roots. In my opinion, voiced stop initial root is not included in the data is because this type of root is not relevant to the analysis as the reduplicant in voiced stop initial root is more faithful to the root rather than to the base. For example, the root /basoh/ 'wash' becomes [mə̃mbasoh], while [mə̃mbasoh] becomes [mə̃mbasoh-basoh] 'to wash repeatedly'. That is why this type of root has not been discussed in Ahmad. But, it is worth to explain in the analysis why the overapplication of nasal coalescence occurs when the root begins with voiceless stop, as shown in (61), while roots begins with voiced stop, the overapplication is not occurred. This is an important point that Ahmad has missed out.

The lacking of those data in Ahmad's analysis therefore will be accounted for in this study by offering an appropriate theory which is able to account for those data. An important point that this present study will be discussing is why the coalesced nasal in (61) is copied into the reduplicant, while the uncoalesced nasal as in [məmbasoh-basoh] which is not included in the analysis is not copied.

2.3.4 Morpheme-Based Template (MBT)

As we have seen in the above discussion of Malay phonology scholarship, all the problems that have been unsatisfactorily explained and have been left unexplained by previous scholars will be accounted for in this present study. In order to provide a satisfactory analysis, I therefore have chosen MBT which has been developed and motivated within OT. As a current theory of prosodic morphology, MBT has not yet been applied in any previous Malay studies. This study thus, intends to make use of the theory as it highlights a number of essential ideas that are useful to account for Malay data.

One of the ideas is, defining the constant shape of prosodic morphemes by their morphological category rather than constructing a specific size offers an advantage for the Malay data. For example, an earlier study of Malay (Ahmad, 2000a), which applies autosegmental phonology, proposed that the templatic size of reduplicative morphemes should only be CV elements. The proposed size of the templatic reduplicative morpheme could, however, only account for the data of SM reduplication. If the data of Malay dialects reduplication were to be considered, then the templatic size of CV would have difficulty in capturing the whole data for Malay reduplication since CVC reduplicative morphemes are found in Malay dialects, as illustrated in 48(b).

The idea of word minimality requirements as proposed in MBT is essential to account for Malay data. It is important to note that the theory proposes that there is a distinction between derived and underived words. These groups of words are not subject to the same minimality requirements. This idea is crucial to explain one of the points that this thesis will raise up i.e. not all words are subject to the same disyllabicity minimality requirements. This is because there seems to be a distinction between derived and underived words in Malay, whereby it is acceptable for underived words to surface as monosyllabic, while derived words are disyllabic. The difference in minimality requirements between derived and underived words, as proposed in MBT, will be used in this thesis as the supporting argument to argue against the claim made by previous studies (e.g.: Teoh, 1994; Ahmad, 2000b), where all words in Malay must at least be disyllabic. This idea proposed by MBT works for the Malay data. As can be seen in the following examples, in (64), underived monosyllabic words do not contain disyllabicity in their surface representation. For example, in monosyllabic underived words such as [pam] 'pump', [bom] 'bomb', [pos] 'post' and [mop] 'mope', the disyllabicity minimality requirements of the language cannot be satisfied unless, when the prefixes are attached, the disyllabicity minimality requirements of words can then be satisfied, as can be seen in (65):

64. Underived monosyllabic words (from the DBP-UKM corpus).

(i) /pam/ 'pump'	[pam]	*[ə.pam]
(ii) /bom/ 'bomb'	[bom]	*[ə.bom]
(iii) /mop/ 'mope'	[mop]	*[ə.mop]

(iv) /lap/ 'wipe' [lap] *[ə.lap]

65. Derived monosyllabic words (from the DBP-UKM corpus).

(i)	/məŋ-ə-kod/ ACT.PRF-STEMEX-code 'to code'	[mə.ŋə.kod]
(ii)	/məŋ-ə-pam/ ACT.PRF-STEMEX-pump 'to pump'	[mə.ŋə.pam]
(iii)	/məŋ-ə-bom/ ACT.PRF-STEMEX-bomb 'to bomb'	[mə.ŋə.bom]
(iv)	/məŋ-ə-pos/ ACT.PRF-STEMEX-post 'to post'	[mə.ŋə.pos]
(v)	/məŋ-ə-mop/ ACT.PRF-STEMEX-mope 'to mope'	[mə.ŋə.mop]

It is apparent from the above examples that the minimality requirements between derived and underived words in Malay need to be differentiated. To account for this, MBT was chosen to explain the differences that occur between derived and underived words, as it is a theory that provides a clear distinction between them.

In this reanalysis study, the new idea, where reduplication is treated as compounding, will be investigated by looking at total (66) and affixal (67) reduplication, since the reduplicative morphemes in both forms of reduplication tend to be more than monosyllabic. I briefly lay out some relevant examples of both forms of reduplication obtained from the DBP-UKM corpus to show that the size of the reduplicative morphemes is larger than the size required by an affix – a monosyllable:

66. Total reduplication (from the DBP-UKM corpus).

Roots		**Reduplicated words**
/kira/'count'	[kirə]	[kirə-kirə] 'calculation/approximately'
/kanak/ 'child'	[kanaʔ]	[kanaʔ-kanaʔ] 'children'
/mana/ 'where'	[manə]	[manə-manə] 'anything'
/undang/ 'rule'	[undaŋ]	[undaŋ-undaŋ] 'rules'
/cita/ 'ambition'	[citə]	[citə-citə] 'ambition'
/juara/ 'champion'	[juarə]	[juarə-juarə] 'champions'

67. Affixal reduplication (from the DBP-UKM corpus).

(i) /mə-mutus-i/ [məmutus-mutusi]
ACT.PRF-break-LOC.SUF
'to cause to break off'

(ii) /məŋ-pukul-i/ [məmukol-mukoli]
ACT.PRF-scold-LOC.SUF
'to cause to scold'

(iii) /məŋ-tekan-i/ [mənekan-nekani]
ACT.PRF-oppress-LOC.SUF
'to cause to oppress'

(iv) /məŋ-susah-i/ [məɲusah-ɲusahi]
ACT.PRF-difficult-LOC.SUF
'to cause to be difficult'

(v) /məŋ-puji/ [məmuji-muji]
ACT.PRF-praise-LOC.SUF
'to cause to praise'

It is apparent from the above examples that the reduplicative morphemes in total and affixal reduplication contain more than one syllable. In both types of reduplication, the size of the reduplicative morpheme is disyllabic. Because the size of reduplicative morphemes is larger than monosyllabic, I am therefore more inclined to agree with the claim that reduplication is a process of compounding. MBT provides a good explanation of this whereby reduplicative morphemes in both total and affixal are analysed as Prosodic Stems which are required to satisfy MORPHEME SYLLABLE-CORRELATION (34). I shall leave this point for further discussion in Chapter 5.

2.4 Which constraints are important?

This section discusses the constraints for Malay that are important to account for all the issues concerned in prefixation, reduplication and dialect variation. The constraints are discussed in two subsections. Subsection 2.4.1 discusses the relevant markedness constraints, i.e. *NC̥ and CRISP-EDGE [σ]; subsection 2.4.2 discusses faithfulness constraints which are treated in OT under Correspondence theory. This subsection will then have four subsections, as follows.

2.4.1 Markedness constraints

2.4.1.1 *NC̥

Nasal and voiceless obstruent clusters seem to be disfavoured in many languages (e.g. Indonesia, Javanese, Swahili, Toba Batak, Chamorro, Kaingang and many others, as cited in Pater, 1999). Therefore, the occurrence of this cluster in the underlying representation undergoes a number of phonological processes, such as post-nasal voicing, nasal deletion, denasalization and nasal substitution, as ways to get rid of the clusters in the surface representation (Pater, 1999). Therefore, as the formal driving force behind those phonological processes, *NÇ̥ is used as the output constraint for such clusters. This constraint is defined as follows:

68. ***NÇ̥**

No nasal/ voiceless obstruent sequences.

As was mentioned in Section 2.2.3, a voiceless obstruent following a nasal is less unmarked than a voiced obstruent is. This marked segment of voiceless obstruents is avoided in the language by considering a relevant markedness constraint, *NÇ̥, which ensures that no voiceless obstruent following a nasal occurs in the surface representation. This means that *NÇ̥ plays an important role as a markedness constraint in Malay grammar. Markedness constraints govern the form of linguistic structures (McCarthy et al., 1999). In OT constraint ranking, if a markedness constraint *NÇ̥ dominates a faithfulness constraint, for example MAX-IO, and no constraint dominates *NÇ̥, then no *NC-offending structure will appear as a surface representation, as in (26). In contrast, if *NÇ̥ is dominated by MAX-IO, then an *NÇ̥-offending structure emerges as a surface representation, as the following examples show:

69. (i) /məŋ-pukul/ *[məmpukul]
ACT.PRF-scold
'to scold'
(ii) /məŋ-kupas/ *[məŋkupas]
ACT.PRF-peel
'to peel'
(iii) /məŋ-kasih/ *[məŋkasih]
ACT.PRF-love
'to love'

Thus, it is crucial that *NÇ̥ is ranked higher in a constraint hierarchy of Malay grammar in order to avoid a marked structure, i.e. a nasal and voiceless obstruent cluster, from emerging as the optimal output. Since a markedness constraint such as *NÇ̥ only makes sense in OT, this theory seems to be able to explain markedness in the language, and particularly why [mə-mukul] is favoured over *[məm-pukul], by adopting a system of ranked constraints whereby MARKEDNESS >> FAITHFULNESS.

The situation in Malay is neither unique nor particularly unusual. It has long been observed that, phonetically, a voiceless obstruent following a nasal segment tends to cause difficulty for the articulatory system. As Huffman (1993: 310) observes, the velum rises very gradually during a voiced stop after a nasal segment because the nasal airflow will return to a value typical of plain obstruents during the release phase (cited in Pater, 1999). She further claims that the velum rises more leisurely for nasal and voiced obstruents than for nasal and voiceless obstruents where the velum needs to be raised quickly to produce a voiceless obstruent following a nasal segment.

In Pater's (1999) study of nasal substitution in Austronesian languages like Chamorro (Topping, 1973), Malagasy (Dziwirek, 1989), Indonesian including Javanese, and in several African languages (Rosenthall 1989: 50), it is shown how crucial *NÇ̥ is to disallowing nasal and voiceless obstruent clusters between a nasal final prefix and the initial consonant of a base. Therefore, in Indonesian for example, such an output yields [məmilih] from the input /məN+pilih/ because of the coalescence of segments (Pater, 1996: 6; Lapoliwa, 1981: 111). The following tableau, by Pater (1999), exemplifies the analysis of a sequence of nasal and voiceless obstruent in Indonesian for the word /məN+pilih/. In the tableau, we can see how the *NÇ̥ constraint prevents the sequence from emerging in the output.

70. NÇ̥ sequences in Indonesia (from Pater, 1999: 281, cited in McCarthy, 2004).

/məN+pilih/	DEP-IO	IDENT-IO[Nasal]	MAX-IO	ROOT-LIN	IDENT[ObsVce]	*NÇ̥	LINEARITY
a.☞ $məm_{12}ilih$							*
b. $məm_1p_2ilih$						*!	
c. $məp_1p_2ilih$		*!					
d. $məm_1b_2ilih$					*!		
e. $məp_2ilih$			*!				
f. $məŋ_1əp_2ilih$	*!						

In the above tableau, the failed candidate (f) is ruled out, as it incurs a fatal violation of DEP-IO, which is the highest constraint in the hierarchy. Candidate (c) violates IDENT-IO [Nasal] since the nasal segment of the prefix in the input is replaced by [p] in the output, whereas the constraint requires that the output segment must be identical to the segment in the input, which is nasal. Deleting one of the segments in the output results in candidate (e) violating MAX-IO. The IDENT [ObsVce] constraint requires that correspondent obstruents be identical in their specification for voice. Therefore, candidate (d) is non-optimal, since the voiceless obstruent /p/ in the input is replaced by the voiced obstruent [b] in the output. Candidate (b), with no nasal substitution, obeys LINEARITY. However, the obedience to this constraint is not significant because it is ranked beneath *NÇ̥. Therefore, candidate (a), with nasal substitution, is the optimal candidate because this candidate only violates the latter. Hence, the constraint ranking in Indonesian is as follows: DEP-IO, IDENT-IO [Nasal], MAX-IO >> ROOT LIN >> IDENT [ObsVce] >> *NC̥ >> LINEARITY.

Now I am going to show that Malay is another language that does not allow nasal and voiceless obstruent clusters in the surface representation. Nasal and voiceless obstruent clusters in the underlying representation in Malay are therefore regularly resolved by nasal substitution, as in Indonesian. Thus, the constraint *NC̥, which bans nasal and voiceless obstruent sequences, as used by Pater (1999) in analysing nasal substitution in Indonesian, will also be used to analyse nasal substitution in Malay. How *NC̥ plays its role in satisfying the phonological rules of the language will be demonstrated in Chapter 4.

In discussing nasal substitution in Austronesian languages, Pater (1996: 326) also points out that the fricative consonant /s/, following a nasal segment in Austronesian languages, such as Chamorro, Indonesian and Javanese, undergoes nasal substitution; /s/ as the initial consonant of a base alternates to nasal palatal /ɲ/, as the following examples show:

71. Nasal substitution in Indonesian, Chamorro and Javanese (from Pater, 1999).

/məN+sapu/	[məɲapu] 'to sweep' – Indonesian.
/man+saga/	[maɲaga] 'to stay' – Chamorro. (Topping, 1973: 50)
/N+sambuŋ/	[ɲambuŋ] 'to connect' – Javanese. (Poedjosoedarmo, 1982: 51)

The situation in Chamorro, Indonesian and Javanese shown above occurs in Malay as well. Observations from the corpus data show that for bases with initial fricative consonant /s/, nasals of the prefixes attached to them would alternate to [ɲ]. I lay out some of the examples from the corpus as evidence:

72. Nasal substitution in Malay: /s/ alternates to [ɲ] (from the DBP-UKM corpus).

(i)	/məŋ-sinar/	[məɲinar]
	ACT.PRF-ray 'to ray'	
(ii)	/məŋ-sayaŋ/	[məɲayaŋ]
	ACT.PRF-love 'to love'	
(iii)	/məŋ-sutʃi/	[məɲutʃi]
	ACT.PRF-chaste 'to chaste'	
(iv)	/məŋ-sikat/	[məɲikat]
	ACT.PRF-comb 'to comb'	

Despite the success of *NC̥ in eliminating nasal and voiceless obstruent clusters, in this thesis I will argue for a reanalysis of prefixation based on the evidence from the DBP-UKM corpus. As has been analysed by Malay scholars, in the process of prefixation to nasal final prefixes and voiceless obstruents, nasal substitution occurs. However, it is important to note that, in some Malay prefixed words, voiceless obstruents do not undergo nasal substitution, as shown in (41).

In optimality ranking, if voiceless-obstruent nasal substitution was to emerge as the optimal output, then *NC̥ must outrank any relevant faithfulness constraints. The converse state occurs if voiceless obstruents fail to undergo nasal substitution, as in (41). As was mentioned in Section 2.2.3, there are two families of constraints in OT: Faithfulness and Markedness. Because voiceless obstruents are retained, a relevant faithfulness constraint must be ranked higher than a markedness constraint, *NC̥. If this were to apply, a candidate without nasal substitution would be optimal. The interaction between these two constraints' rankings will generally determine the optimal output of a language, i.e. whether to choose to eliminate nasal and voiceless obstruent clusters or not. In this regard, faithfulness constraints should be ranked above *NÇ̥ since nasal substitution does not occur in (41), these are repeated in (73), below, for convenience. While other prefixed words undergo nasal substitution, *NÇ̥ is ranked above faithfulness constraints. This means we will have more than one constraint ranking to account for prefixation, since prefixed words in Malay have two patterns in the surface representation whereby one allows the occurrence of nasal and voiceless obstruent clusters and the other does not.

73. Nasal final prefixes (from the DBP-UKM corpus).

(i)	/məŋ-pakedʒ-kan/ ACT.PRF-package-CAUS.SUF 'to cause to pack for'	[mə**mp**akedʒkan]
(ii)	/məŋ-kelas-kan/ ACT.PRF-class-CAUS.SUF 'to cause to categorise for'	[mə**ŋk**əlaskan]
(iii)	/məŋ-kabul-kan/ ACT.PRF-grant-CAUS.SUF 'to cause to grant for'	[mə**ŋk**abulkan]

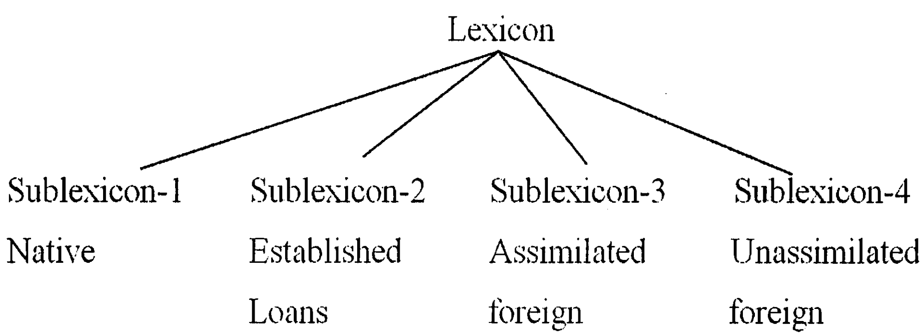
From the case presented in (73), I will claim in Chapter 4 (Section 4.2) that the occurrence of nasal-voiceless obstruent clusters at prefix-root junctures is due to non-native words. Since they are non-native, they cannot be analyzed as though following the same process as native words, and therefore nasal-voiceless obstruent clusters occur.

Although it has been claimed in previous studies (e.g. Hassan, 1974; Omar, 1986; Hassan, 1987; Karim et al., 1989, 1994; Karim, 1995; Ahmad, 1993) that nasal and voiceless obstruent clusters are not allowed to be present in the surface form, in some cases the clusters still emerge, as in (73), above. This thesis will thus argue that

Malay has a co-existent grammar which means that *NC̥ is closely obeyed in one of the sub-grammars, but not in the other.

It is important to note that such a situation also arises in Japanese vocabulary. Due to some reasons of assimilation, the lexicon in Japanese has been partitioned into four strata: native (Yamato), established loans, assimilated foreign and unassimilated foreign, as Figure 1 illustrates:

Figure 1: The Lexical Strata in Japanese Vocabulary



All the lexical strata categorised above are analysed according to the following constraints (see Table 3), which claim to operate in the phonological lexicon of Japanese (Itô and Mester, 1999b, cited in McCarthy, 2004: 554).

Table 3: Constraints operating in Japanese vocabulary

a)	SyllStruc	Syllable structure constraints.
b)	NoVoicedGem (NO-DD)	No voiced obstruent geminates (*bb, *dd, *gg, etc.)
c)	NoVoicelessLab (NO-P)	No singleton-p: a constraint against non-geminate [p].
d)	NoNas Voiceless (NO-NT)	Post-nasal obstruent must be voiced (*nt, *mp, ŋk).

In the Japanese phonological system, a syllable is controlled by a set of syllable structure constraints which include, among others, (1) *Complex where the language does not allow any complex onsets or codas in a syllable, and (2) Coda Condition, by limiting codas to place-linked consonants or segments without consonantal place (=nasal glide N) (cited in McCarthy, 2004: 555). In the phonological system of Japanese, voiceless obstruents can be geminate, for example *ok-kakeru, ot-tsuku* and *tsut-tatsu.* In contrast, if the consonant is a voiced obstruent, then the result is not a geminate but rather a cluster of homorganic nasal and voiced obstruents (ibid.). A constraint which plays an active role against voiced geminates is NoVoicedGem (NO-DD), as in (b). The third constraint, which is operative in the phonological lexicon of Japanese, is NoVoicelessLab (NO-P). This constraint states that voiceless labial /p/ must not appear as a singleton *[p]. Since there is a restriction preventing voiceless labials surfacing as singleton-p, this consonant appears as a voiceless labial geminate [pp], for example:

74. (a) ma-prefixation (from Itô and Mester, 1999b).

hiruma 'daytime'	map-piruma 'broad daylight'
hadaka 'naked'	map-padaka 'stark naked'

(b) Verbal root compounding (from Itô and Mester, 1999b):

hik- 'pull'	har- 'stretch'	hip-paru 'pull strongly'
tsuk- 'stab'	haʧir- 'run'	tsup-paʧiru 'dash, race'

In the analysis, Itô and Mester (1999b) claim that the native vocabulary of Japanese (Yamato) strongly obeys all the above constraints, while other lexical strata, Sino-Japanese, Assimilated Foreign and Unassimilated Foreign, violate one or more of them. In short, obedience to the four constraints in Yamato shows that those constraints are active in the phonological lexicon of Japanese, while the other three lexical strata, which do not fully obey the constraints, clearly show that they are non-native vocabulary. Therefore, they cannot obey all the constraints in the same way that native words do. For example, *NT is violated in the vocabulary of Japanese where the post-nasal obstruent is a voiceless obstruent, as in [šiNtai], because this incurs violation of NoNas Voiceless (NO-NT). Since Japanese loanwords behave differently from native words, Itô and Mester proposed different lexical strata to account for the different obedience of constraints by native and loanwords. The following table shows the degree of obedience of the constraints in every stratified lexicon, as cited in Itô and Mester (ibid.).

Table 4: The degree of obedience to active constraints for native and other lexicon strata in the vocabulary of Japanese

Lexical strata	SYLLSTRUC	NO-DD	NO-P	NO-NT
Yamato	√	√	√	√
Sino-Japanese	√	√	√	violated
Assimilated Foreign	√	√	violated	Violated
Unassimilated Foreign	√	violated	violated	violated

The situation above occurs in Malay prefixation as well. Observations from the corpus data show that this phonological restriction of disallowing nasal and voiceless obstruent clusters is obeyed, and thus *NC̥ can be satisfied. Although Malay disfavours the occurrence of nasal and voiceless obstruent clusters in surface representation, there are counter-examples which violate *NC̥. As shown in the examples in (41) and repeated in (73), the constraint cannot somehow be neatly satisfied in the language. Nasal and voiceless obstruent clusters in the examples below have not undergone nasal substitution, as is regularly applied in the language to eliminate nasal and voiceless obstruent clusters from emerging in the surface representation.

The phenomenon that occurs in Malay prefixation might possibly be accounted for by applying Itô and Mester's ideas on lexical strata in Japanese vocabulary. This is because their analysis clearly shows the distinction between native and loanwords in applying some of the phonological rules postulated in the language. As Itô and Mester (1999b, cited in McCarthy, 2004) demonstrate in their analysis, native words (Yamato) obey all the constraints, while loanwords violate some of them. The same goes for the Malay prefixation data. The surface representations without nasal substitution are loanwords, hence they violate the *NC̥ constraint, while others with nasal substitution are native. If Malay lexical items were ranked differently, according to the etymology of the words (native or loanwords) as in Japanese, I am sure we would get a better account of prefixation. Hence this idea will be demonstrated in the present analysis, based on evidence from the DBP-UKM corpus database. Therefore, it is suggested here that a reanalysis of prefixation should employ the OT analysis proposed by Itô and Mester (1993), in order to handle the problem that occurs in the language. We will then see that, by applying the idea of the analysis of lexical strata in the Japanese vocabulary to the Malay data, this would probably resolve the problem of why the *NC̥ constraint cannot be satisfied in some prefixed Malay words. The proposed analysis of different lexical strata to account for the issue of assimilation in Japanese vocabulary is interesting as the analysis could only be expressed within OT (Itô and Mester, cited in Roca, 1997). The differences in the phonological process that occur in prefixation can be clearly shown in the same set of constraints, though differing in ranking hierarchy. This analysis is discussed at length in the prefixation chapter, Chapter 4, of this thesis.

2.4.1.2 CRISP-EDGE [σ]

Another constraint, which is also crucial, is CRISP-EDGE. CRISP-EDGE [PrWD] requires that no element belonging to a Prosodic Word is linked to a prosodic category external to that prosodic word (Pater, 2001), as defined in (75), below:

75. **CRISP-EDGE [PrWD]** (adapted from Pater, 2001).

No element belonging to a Prosodic Word may be linked to a prosodic category external to that Prosodic Word.

In Pater's (2001) analysis of Austronesian nasal substitution revisited, he shows how CRISP-EDGE [PrWD] cannot prevent voiced obstruents from undergoing nasal substitution. A nasal preceding a voiced obstruent undergoes nasal assimilation instead of nasal substitution. Thus, CRISP-EDGE [PrWD] is crucial in ruling out the assimilated nasal from the following voiced obstruent. Pater claims that the interaction of CRISP-EDGE [PrWD] >> IDENT [PHAREXP],[13] is able to produce voiceless and voiced obstruent nasal substitution, as demonstrated by Pater (2001: 176) in the following tableaux:

[13] The IDENT [PHAREXP] constraint is discussed in subsection 2.4.2.4.

76. (a) Nasal substitution of voiced obstruents: CRISP-EDGE [PrWD] >> IDENT [PHAREXP]

$/N_1+B_2/$	CRISP-EDGE [PrWD]	IDENT [PHAREXP]
a.☞M_{12}		*
b. M_1B_2	*!	

(b) Nasal substitution to voiceless obstruents: CRISP-EDGE [PRWD] >> IDENT [PHAREXP]

$/N_1+P_2/$	CRISP-EDGE [PrWD]	IDENT [PHAREXP]
a.☞P_{12}		*
b. M_1P_2	*!	

By considering CRISP-EDGE [PrWD] in the above constraint ranking, it is seen that voiceless and voiced obstruents can undergo nasal substitution. This analysis seems crucial to the present study, particularly when nasal substitution in the three dialects of Malay (i.e. Perak, Kelantan and NS) is being discussed, since voiced obstruents following nasals can also undergo nasal substitution, for example in the PD, as exemplified in (77):

77. Nasal and voiced obstruent clusters in Perak (Ahmad, 1991).

(i)	/ŋ-bagi/	[magi]
	ACT.PRF-give 'to give'	
(ii)	/ŋ-dapat/	[napat]
	ACT.PRF-obtain 'to obtain'	
(iii)	/ŋ-gosok/	[ŋosoʔ]
	ACT.PRF-brush 'to brush'	
(iv)	/ŋ-basuh/	[masoh]
	ACT.PRF-wash 'to wash'	

In the above tableaux, as illustrated by Pater, CRISP-EDGE [PRWD] will most likely be able to account for the voiced obstruents nasal substitution that occurs in some dialects of Malay, since this constraint would permit both voiceless and voiced obstruents to undergo nasal substitution. To offer a better account for the Malay data, including SM and its dialects, presumably CRISP-EDGE [PrWD] would pose a problem if it were to be applied in the analysis. This constraint does though have its own advantage as it allows voiced obstruents to undergo nasal substitution, as illustrated in (76). However, if we refer to the requirement of the constraints, voiceless obstruents follow nasals root-internally, for example /sampan/, 'small boat', would have to undergo nasal substitution as well and the output would be *[saman] and not [sampan], because both segments are within the same prosodic word. There is no element in the word linked to the prosodic category external to that prosodic word.

The lack of nasal substitution root-internally prevents CRISP-EDGE [PrWD] from being satisfied. This will be discussed in greater detail in Chapter 6. From this, I will then suggest another constraint of CRISP-EDGE, i.e. CRISP-EDGE [σ] in (78), which I think might fit the whole dataset of Malay dialects. In the discussion of this, I will demonstrate how CRISP-EDGE [σ] is able to account for voiceless obstruents following nasals in both morphological domains: prefix juncture and root-internally.

78. **CRISP-EDGE [σ]**

No element belonging to a syllable may be linked to the adjacent syllable.

2.4.2 Correspondence theory

This section discusses some of the important faithfulness constraints that are used to analyse prefixation, reduplication and dialect variation. As stated in McCarthy (2008: 195), a faithfulness constraint assigns its violation based on disparities between the input and the output. Constraints of faithfulness require that output be as close as possible to input. Any input-output (henceforth I-O) disparity indicates a violation of the constraints. Besides the I-O relation, a faithfulness constraint also assigns its violation to any disparities between the base and the reduplicant (henceforth B-R) (McCarthy, 2008). The B-R relation occurs in reduplication where the constraints of reduplicative identity demand that the reduplicant be similar to the base. In OT, the relation between these two representations, I-O and B-R, is treated under Correspondence theory (McCarthy and Prince, 1995). Correspondence theory provides a general framework for defining faithfulness constraints (McCarthy and Prince, 1995, 1999, cited in McCarthy, 2008: 195). This theory proposes that each candidate supplied by GEN includes an output representation and a relation between the input and the output (ibid.). This is called the correspondence relation and is conventionally indicated by $\Re$. The formal definition of Correspondence theory is given below:

79. **Correspondence theory** (McCarthy and Prince, 1995: 262).

Given two strings S_1 and S_2, correspondence is the relation $\Re$ between the elements of S_1 to those of S_2. Elements[14] $\alpha \in S_1$ and $\beta \in S_2$, are referred to as correspondents of one another when $\alpha \Re \beta$.

In addition to the basic faithfulness constraints listed in (20) in subsection 2.2.3, there are other relevant constraints on correspondence elements, such as UNIFORMITY, INTEGRITY, ANCHOR and IDENT [PHAREXP]. The scope of these constraints is limited to a particular position such as the root, edge of a morpheme, word or morpheme internal position. In this study, there are four relevant faithfulness constraints which are important for analysing the issues concerned in prefixation, reduplication and dialect variation. The four constraints are: UNIFORMITY-ROOT, EDGE-INTEGRITY, ANCHOR RIGHT-BR and IDENT [PHAREXP]. The following subsections discuss all four faithfulness constraints briefly as they will become important to the exposition.

2.4.2.1 UNIFORMITY-ROOT

As mentioned in subsection 1.3.1, nasal and voiceless obstruent clusters are allowed root-internally. As exemplified in (5), nasal and voiceless obstruent clusters within roots are not resolved by nasal substitution, as regularly applied in the language. Harmonic nasals are applied instead.

[14] The 'elements' used here refers to segments, higher-order units of prosodic structure such as moras, syllables, feet, heads of feet, tones and distinctive features or feature nodes (McCarthy and Prince, 1995).

This blocking of nasal substitution also occurs in one of the Austronesian languages, i.e. in Indonesian. Indonesian has the same phonological requirements as Malay which bans nasal and voiceless obstruent clusters from occurring in the surface representation, however it does not ban the clusters root-internally. Let us observe the examples below, from Indonesian.

80. Root-internal occurrences of NÇ in Indonesian (Pater, 1993, cited in Kager, 1999: 75; Pater, 2001).

/əmpat/	[əm.pat] 'four'
/untuk/	[un.tuk] 'for'
/muŋkin/	[muŋ.kin] 'possible'

The blocking of nasal substitution within the roots in Indonesian has been accounted for by Pater (2001) who proposes a general ranking schema for the non-application of a phonological process within a root:

81. Non-application of a phonological process within a root.

Root-faithfulness >> Markedness >> Faithfulness

By applying the general ranking schema above, nasal substitution continues to apply at prefix junctures, but not within roots (Pater, 2001). Thus, the relevant faithfulness constraint to account for this is UNIFORMITY, and the root faithfulness constraint is UNIFORMITY-ROOT, as defined in (82) and (83), respectively:

82. **UNIFORMITY** ('No Coalescence') (McCarthy and Prince, 1999: 296).

No element of S_2 has multiple correspondents in S_1.

For x, y ϵ S_1 and z ϵ S_2: if x R z and y R z, then x=y.

83. **UNIFORMITY-ROOT**

The output reflects the precedence structure of the input segments of the roots, and vice versa.

84. UNIFORMITY-ROOT >> *NÇ̥ >> UNIFORMITY (from Pater, 2001: 162).

/əm$_1$p$_2$at/	UNIFORM ROOT	*NÇ̥	UNIFORM
a. əm$_{12}$at	*!		
b.☞əm$_1$p$_2$at		*!	

This analysis proposed by Pater, of the non-application of nasal substitution within a root, will also be applied to Malay, since the phonological process is also blocked within a root in Malay. How the blocking of nasal substitution within roots can be accounted for, as in Indonesian, will be presented in Chapter 6, subsection 6.3.1.

2.4.2.2 EDGE-INTEGRITY

As demonstrated in (12) (see Chapter 1, subsection 1.3.1), we note that nasal and voiceless obstruent clusters at prefix-prefix junctures in Malay are not entirely resolved by nasal substitution. This phonological process is applied only for nominal multiple prefixes /pəŋ+pər/, but not for verbal prefixes /məm+pər/. My assumption for why nasal substitution is blocked in /məm+pər/ might be due to non-phonological factors, since a nasal and voiceless obstruent sequence is there. I believe that there must be a reason why nasal substitution fails to apply in /məm+pər/. For this, see my discussion in section 4.3.

In this study, I propose that the occurrence of nasal and voiceless obstruent clusters at the surface representation in verbal multiple prefixes is due to a constraint requiring the integrity of the morphological constituent. In other words, the application of nasal substitution in multiple prefixation is mainly determined by the morphological input rather than the phonological input. This constraint, EDGE-INTEGRITY (McCarthy and Prince, 1995), ensures that morpheme boundaries coincide with syllable boundaries, as defined in (68):

85. **EDGE INTEGRITY** (McCarthy and Prince, 1995).

Edge segments in the input preserve their segments at the edge of the corresponding prosodic structure.

It is worth knowing how this morphological constraint is used to account for other issues or problems in other languages. Consonant copying in the Jeju dialect of Korean is one example showing how important the constraint, EDGE INTEGRITY, is in preserving the edge segments of a morphological unit in the underlying positions with their edge of a corresponding prosodic structure. Consonant copying in the Jeju dialect occurs at ...$C)_{PrWd}$ (V...., as in prefix-stems,[15] compound words and across words in phrases, as the following examples taken from Kang (to appear) illustrate:

86. (a) Consonant copying between prefix-stem.

(i)	/hot^h ipul/	hot.**t'**i.pul~hon.**ni**.pul	'thin comforter'
(ii)	/hot^h os/	hot.**t'**ot	'thin clothes'
(iii)	/c^{h}əs atəl/	c^{h}ət.**t'**a.təl	'the first son'

[15] In the Jeju dialect, prefixes and suffixes form their own prosodic words (see Kang: ibid.).

(b) Consonant copying in compounds

(i) /cicip ai/ ‘female child’ ci.cip.**p’**a.i ‘girl’

(ii) /kacuk os/ ‘leather clothes’ ka.cik.**k’**ot. ‘leather clothes’

(iii) /pitan os/ ‘silk clothes’ pi.tan.**n**ot ‘silk clothes’

(iv) /mul ankjəŋ/ ‘water glasses’ mul.**l**an.kjəŋ ‘goggles’

(c) Consonant copying across words in phrases

(i) /nun ətuk-ən/ nun **n**ə.tu.kən ‘the eyesight is not good’

eye dark-Inf.

(ii) /tap al-an/ tap.**p’**a.ran ‘knowing the answer’

answer know-Inf

As illustrated in the above examples, initial vowel morphemes satisfy their onset requirement by copying the preceding consonants. Kang proposes that ‘*consonant copy is insertion of a root node to the onset of the prosodic word initial vowel with additional featural copy from the preceding consonant at the prosodic word juncture*’. This is illustrated in the following diagram:

87.

$_{PrWd}$(... VC)(C̦V...)$_{PrWd}$ [in which C represents a copied root]

[F]

In the analysis, Kang claims that the consonant preceding the initial vowel of the following prosodic word is copied is due to EDGE INTEGRITY. Since a consonant is inserted before the initial vowel prosodic stem, so DEP-IO is ranked low in order to allow the copying of the consonant to occur. The following tableau is presented in Kang to explain the situation.

88.

/kacuk os/	ONSET	STEML	EDGE INTEG	STEMR	DEP-IO
a. PrWd PrWd σ σ σ [k a. c u k][o t]		*	*	*!	
b.☞ PrWd PrWd σ σ σ [k a c u k]k'[o t]		*	*		
c. PrWd PrWd σ σ σ [k a c u k][o t]	*!				

The failed candidate (c) violates the higher-ranked constraint, ONSET, since there is no consonant copy in the onset position. In order to obey the higher constraint, ONSET, candidates (a) and (b) both violate STEML and EDGE INTEGRITY since the left edge of /os/ is disrupted by another segment in the surface representation. The double linking of consonant [k] in the ambisyllabic representation in candidate (a) results in the candidate violating STEMR. As claimed by Kang, the evaluation above is important, particularly the role played by the crisp alignment constraints. He claims further that the right edge of a morphological unit is often disrupted due to ONSET satisfaction. When ONSET is satisfied by a copied segment, the crisp alignment requirement at the right edge emerges (Kang, to appear).

2.4.2.3 ANCHORING-BR

As pointed out above, reduplication is one of the topics that this thesis will be discussing. Like previous scholars (e.g.: Omar, 1993; Hassan, 1974; Othman, 1981, 1983; Abas, 1975; Che Kob, 1981; Karim et al., 1994; and many others), we will be focusing on partial reduplication as it is the most complex type of reduplication. Taking this into consideration, partial reduplication in the PD of Malay will be examined as well as SM.

As we saw in the preceding chapter, reduplicative morphemes in SM partial reduplication consist of CV elements. In PD, the reduplicative morpheme consists of both CV and CVC elements, as exemplified in the following examples, taken from Ahmad (1991):

89. (a) CV reduplicative morphemes in Perak.

(i)	/budaʔ/ 'child'	[budaʔ]	[bə-budaʔ] RED-child 'all kinds of children'
(ii)	/cərita/ 'story'	[cərita]	[cə-cərita] RED-story 'all kinds of stories'
(iii)	/kadʒi/ 'to study'	[kadʒi]	[kə-kadʒi] RED-to study 'to study repeatedly'
(iv)	/kira/ 'to estimate'	[kɛrɛ]	[kə-kɛrɛ] RED-estimate 'by my estimate'
(v)	/dulu/ 'long ago'	[dulu]	[də-dulu] RED-long ago 'very long ago'

(b) CVC reduplicative morphemes in Perak.

(i)	/pətaŋ/ 'evening'	[pətaŋ[	[pəm-pətaŋ] RED-evening 'every evening'
(ii)	/ʤaraŋ/ 'seldom'	[ʤaʁaŋ]	[ʤəɲ-ʤaʁaŋ] RED-seldom 'very seldom'
(iii)	/baraŋ/ 'thing'	[baʁaŋ]	[bəŋ-baʁaŋ] RED-thing 'all kinds of things'
(iv)	/sikit/ 'a little'	[siket]	[səʔ-siket] RED-a little 'very little'
(v)	/gelap/ 'dark'	[gəlap]	[gəʔ-gəlap] RED-dark 'very dark'
(vi)	/budaʔ/ 'child'	[budaʔ]	[bəʔ-budaʔ] RED-child 'all kinds of children'

Observe, in 89(b), that the PD has two different copying processes of CVC reduplicative morphemes, depending on what the final consonant of the base is. The difference in the copying process is on the second C of the reduplicative morpheme where it is based on the final consonant of the base. When the base ends with a glottal stop, the first C in the reduplicative morpheme is copied from the initial consonant of the base, while the second C copies the final consonant of the base. When the base ends with a nasal segment, this segment is copied into the final C in the reduplicative morpheme. As the copied nasal segment occupying the coda position of a syllable, the segment must be homorganic to the following onset consonant in the base. The nasal segment becomes [m], [n] or [ŋ] if the following consonant is [p, b], [t, d] or [k, g], respectively. The two groups of data below represent the two copying processes of CVC reduplicative morphemes in Perak just mentioned:

90. (a) Bases ending with a glottal stop.

(i)	/tembak/ 'shoot'	[temba?]	[tə?-temba?]
			RED-shoot 'shoots'
(ii)	/budak/ 'child'	[buda?]	[bə?-buda?]
			RED-child 'children'
(iii)	/kəʧe?/ 'small'	[kəʧe?]	[kə?-kəʧe?]
			RED-small 'very small'
(iv)	/siket/ 'a little'	[sike?]	[sə?-sike?]
			RED-a little 'very little'

(b) Bases ending with a nasal segment.

(i)	/pətaŋ/ 'evening'	[pəm-pətaŋ]
		RED-evening 'every evening'
(ii)	/bawaŋ/ 'onion'	[bəm-bawaŋ]
		RED-onion 'all kinds of onions'
(iii)	/kawan/ 'friend'	[kəŋ-kawan]
		RED-friend 'friends'
(iv)	/baʁaŋ/ 'thing'	[bəŋ-baʁaŋ]
		RED-thing 'all kinds of things'

The two faithfulness constraints discussed so far are important in accounting for nasal and voiceless obstruent clusters in the language. This subsection will discuss the next faithfulness constraint which is crucial in accounting for partial reduplication in Malay, i.e. ANCHORING-BR. Before showing why ANCHORING-BR is crucial to partial reduplication, I will explain what it is.

ANCHORING requires correspondence between two major landmarks in the base and reduplicant. These segments, base and reduplicant, must be at the left or right edge (Kager, 1999: 213). This requirement is stated in constraint format in (91):

91. **ANCHORING-BR** (Kager, 1999: 213)

Correspondence preserves alignment in the following sense: the left (right) peripheral element of R corresponds to the left (right) peripheral element of B, if R is to the left (right) of B.

ANCHORING-BR has been an important constraint for Diyari, Kinande and Samoan reduplication to account for reduplicative prefixes, suffixes and infixes, respectively. This constraint ensures the left edge of the reduplicant (in the case of a reduplicative prefix) corresponds to the left edge of base. Let us consider the following examples from each language:

92. ANCHORING-BR in Diyari, Kinande and Samoan reduplication (Kager, 1999: 212).

a. Diyari: t^{j}ilpa-t^{j}ilparku

b. Kinande: ku-gulu-gulu

c. Samoan: a-lo-lófa

The following correspondence diagrams demonstrate the obedience of ANCHORING–BR in the three languages, Diyari, Kinande and Samoan, as exemplified in Kager (1999: 212):

93.

```
(a)  tʲi l p a-       R     (b)   -g u l u   R    (c)  l o-    R
     | | | | |                     | | | |             | |
     tʲi l p a r k u  B         k u g u l u   B        l ó f a  B
```

As we see in 93(a) above, the leftmost segment in the reduplicative prefix [**tʲ**ilpa] in Diyari corresponds to the leftmost segment [**tʲ**ilparku] in the base. Meanwhile, in Kinande, the rightmost segment [gul**u**] in the reduplicative suffix corresponds to the rightmost segment [ku-gul**u**] in the base, as in 93(b). The leftmost segment [**lo**] in the reduplicant in the reduplicative infix in Samoan, in 93(c), corresponds to the leftmost segment in the base [**l**ófa].

This constraint seems useful to account for Malay partial reduplication which has two patterns of reduplicative morphemes: CV and CVC. The role of ANCHORING-BR becomes more important when CVC reduplicative morphemes are analysed, since the rightmost segment in the reduplicant must correspond to the rightmost segment in the base, as demonstrated in (94). Thus, the relevant constraint is ANCHOR RIGHT-BR. Therefore, I will make use of this constraint to analyse CV and CVC reduplicative morphemes in Malay partial reduplication, and these will be discussed further in Chapter 5, Section 5.4.

94. Correspondence diagram for ANCHOR RIGHT-BR.

```
   b ə ʔ-     R
       |
 b u d a ʔ    B
```

So far all the examples above, from Diyari, Kinande and Samoan, show the obedience of ANCHORING-BR. Now, I will show that this constraint can also be violated. As stated in Kager (1999: 213), ANCHORING-BR is a violable constraint. Perfective reduplication in Sanskrit is a good example of ANCHORING-BR violation. A reduplicant in Sanskrit simplifies the complex onset of the base by dropping the initial consonant of the cluster, /sk/ - [k] and /st/ - [t]. Here are the examples:

95. Sanskrit perfective reduplication (Kager, 1999: 213).

a) <u>ka</u>-skanda-a 'leap'

b) <u>ta</u>-stamba^{h}-a 'prop'

Due to onset simplification in perfective reduplication in Sanskrit, the leftmost segment in the base fails to correspond to the leftmost segment in the reduplicant. The following correspondence diagram gives a better picture of this violation.

96. Correspondence diagram for violation of ANCHORING-BR in Sanskrit (from Kager, 1999: 213).

```
  k a-        R
  | |
s k a n d a-a  B
```

We saw in (94) that ANCHORING-BR is obeyed in CVC reduplication in Malay. This constraint is however violated in CV reduplicative morphemes. The following correspondence diagram illustrates this violation:

97. Correspondence diagram for the violation of ANCHOR RIGHT-BR in Malay.

```
    b ə-    R
b u d a ʔ   B
```

2.4.2.4 IDENT [PHAREXP]

From Section 2.3 of the literature review in Chapter 2, we know that nasal and voiceless obstruent clusters have received much attention. However, because it is rare, it can be said that little attention has been paid in the literature to nasal substitution in voiced obstruents. In my search of the relevant literature, Pater's analysis (2001) of Austronesian nasal substitution revisited discusses the issue pointed out above. By giving examples from Muna, he posits a constraint, IDENT [PHAREXP], which specifically blocks voiced obstruents from undergoing nasal substitution.

98. **IDENT [PHAREXP]** (Pater, 2001).

Correspondent segments have identical values for the feature [PHAREXP]. If αRβ and α is [γ PHAREXP], then β is [γ PHAREXP].

In Muna, *-um-* appears as an infix with consonant initial roots, as exemplified in 99(a) (Pater, 2001: 163). It does not however appear as an infix if the initial consonant of the root is a labial consonant, as exemplified in 99(b), since the language has a phonotactic constraint that bans multiple labials in a root (Pater, 2001: 166). For a root beginning with initial labial consonants, the vowel and nasal segment in the infix are deleted, as can be seen in 99(b). If the labial consonant is voiceless, deletion of the vowel occurs and nasal substitution then follows (Pater, 2001: 164).

99. Affix *–um-* in Muna.

a)	/um+dadi/	[d*um*adi] 'live'
	/um+gaa/	[g*um*aa] 'marry'
	/um+rende/	[r*um*ende] 'alight'
	/um+solo/	[s*um*olo] 'flow'
b)	/um+baru/	[baru] 'happy'
	/um+pili/	[mili] 'choose'
	/um+ɓala/	[ɓala] 'big'
	/um+futaa/	[mutaa] 'laugh'

Now we see that –um- will only surface as infixation if the consonant-initial roots are not labial consonants. As shown in the examples above in 99(a), -um- would be infixed after the first consonant of the root, as in [dumadi], [gumaa], [rumende] and [sumolo], where the infix is italicised. In the examples in 99(b), the two types of consonant-initial roots, voiceless-labial and voiced-labial, undergo different processes. When the root-initial is a voiceless-labial, the vowel is deleted, while the two adjacent segments, nasal and voiceless labials, undergo nasal substitution; e.g., /um+pili/ 'choose' → [mili] and /um+futaa/ 'laugh' → [mutaa]. However, if the root-initial consonant is a voiced labial, only the vowel and nasal segment are deleted with no nasal substitution, such as in /um+baru/ 'happy' → [baru] and /um+ɓala/ 'big' → [ɓala].

According to Pater (2001), in order to satisfy the phonotactic constraint of the language, where multiple labials are not permitted, $*PL/LAB^2$ is crucial in banning such multiple labials from emerging on the surface. As –um- does not appear as an infix in one of the labial-initial roots that is a voiceless labial, then $*PL/LAB^2$ must be ranked above MAX-IO and UNIFORM, as shown below:

100. $*PL/LAB^2$ >> MAX-IO >> UNIFORM (adapted from Pater, 2001: 167).

/um_1+$p_2$ili/	$*PL/LAB^2$	MAX-IO	UNIFORM
a. $p_2um_1$ili	*!		
b. $p_2$ili		*!	
c.☞m_{12}ili			*

On the other hand, to account for the root-initial voiced labial consonant, IDENT [PHAREXP] is used to explain the blocking of any fusion between the nasal and the voiced labial obstruent (Pater, 2001: 169).

101. PL/LAB^2 >> IDENT [PHAREXP] >> MAX-IO >> UNIFORM

/um$_1$+b$_2$aru/	*PL/LAB^2	IDENT [PHAREXP]	MAX-IO	UNIFORM
a. b$_2$um$_1$aru	*!			
b.☞b$_2$aru			*	
c. m$_{12}$aru		*!		*

We noted that IDENT [PHAREXP] is a constraint that blocks a voiced obstruent from undergoing nasal substitution. This constraint is useful to account for nasal substitution in the dialects of Malay. As we will see in Chapter 6, nasal substitution is not limited to voiceless obstruents. For example, in the PD, both voiceless and voiced obstruents undergo nasal substitution, as shown in the following examples (Ahmad, 1991):

102. (i) /ŋ-puji/ [muji]
ACT.PRF-praise
'to praise'
(ii) /ŋ-bagi/ [magi]
ACT.PRF-give
'to give'

This constraint seems to be crucial for the analysis of Malay dialects, even though the constraint blocks voiced obstruents from undergoing nasal substitution. In Pater's (2001) analysis, IDENT [PHAREXP] emerges together with CRISP-EDGE [PRWD] (75). Pater demonstrates that in order for voiced-obstruent nasal-substitution candidates to be chosen as optimal outputs, CRISP-EDGE [PRWD] must rank above IDENT [PHAREXP]. As was mentioned in subsection 2.4.1.2, due to the blocking of nasal substitution in root-internal position in the PD for example, CRISP-EDGE [σ] is seen as more appropriate than CRISP-EDGE [PRWD]. Therefore, for the analysis of dialect variation in this thesis, CRISP-EDGE [σ] will be used. As we will see in Chapter 6, IDENT [PHAREXP] is ranked beneath CRISP-EDGE [σ] to account for voiced-obstruent nasal substitution. How these constraints interact with one another in choosing voiced-obstruent nasal substitution as harmonic candidates will be demonstrated in the discussion of Perak in subsection 6.3.1 and NS in subsection 6.3.3.

2.5 Conclusion

In my review of the literature, I have discovered that previous Malay morphological scholarship work concerning prefixation and reduplication had been described using inappropriate or inadequate approaches and theories. The analyses merely provided a general statement of the rules of phonological derivation. Moreover, no phonological motivation was provided, for example why a particular process or alternation occurred when nasal final prefixes are attached to certain initial consonants of roots. Thus, the problems and issues concerned have to date been accounted for unsatisfactorily. From the above discussion of previous OT work, MBT, which is one of the versions of GTT, is seen as being capable of accounting for the problems raised in this thesis more adequately and satisfactorily. The ideas proposed in MBT through GTT, such as the minimality condition, can account in an adequate manner for the minimal size of words in a language. This is the reason why MBT is chosen in this thesis to account for the Malay data, as it provides plausible phonological properties to explain words that do not satisfy the minimality requirements in particular languages.

We have also seen from the above discussion that OT seems to be able to handle some of the problems that occur in the language, particularly in prefixation and reduplication, including SM and its dialects. For example, the analysis of lexical strata in OT, proposed by Itô and Mester (1999), is predicted to solve the problem that occurs in Malay prefixation. As the lexicon is partitioned according to the etymological origins of words, so the lack of nasal substitution in prefixation can be plausibly resolved and explained. The discussion of this will be pursued in Chapter 4.

The analysis of co-phonology – different constraint rankings that can only be done in OT – was able to account for the two types of reduplicative morphemes in Ilokano reduplication. Like Ilokano, Malay reduplication will also be analysed by proposing a co-phonology analysis. See Chapter 5, Section 5.4, for the analysis.

3. DATA AND METHOD

3.1 Introduction

This chapter describes the research method which has been used to carry out the present study. My intention is to describe the methods of gathering data, of data categorisation prior to the analysis, and to comment on and discuss their advantages. In order to analyse the three topics concerned in this study, i.e. prefixation in SM, reduplication and dialect variation, three types of data are used: (1) corpus data, (2) secondary data from previous studies and (3) interviews. Corpus data from the DBP-UKM (The Institute of Language and Literature, National University of Malaysia) database were used to facilitate analysis of the issues raised in the process of prefixation in SM (Chapter 4) and also SM reduplication (see Chapter 5), while reduplication in the PD (also see Chapter 5) and dialectal variation (see Chapter 6) are based on secondary data from previous studies of Malay. An interview method was also used for dialectal variation analysis to supplement data from previous studies. Subsection 3.3.3 explains further why this additional method is used for dialectal variation.

This chapter is divided into three sections. In Section 3.2, I briefly explain the nature of the corpus database used in this present study, i.e. the DBP-UKM corpus. Section 3.3 describes the types of data used in this study and how they were analysed. Section 3.4 sketches out the advantages of choosing corpus data and Section 3.5 concludes the chapter.

3.2 DBP-UKM corpus

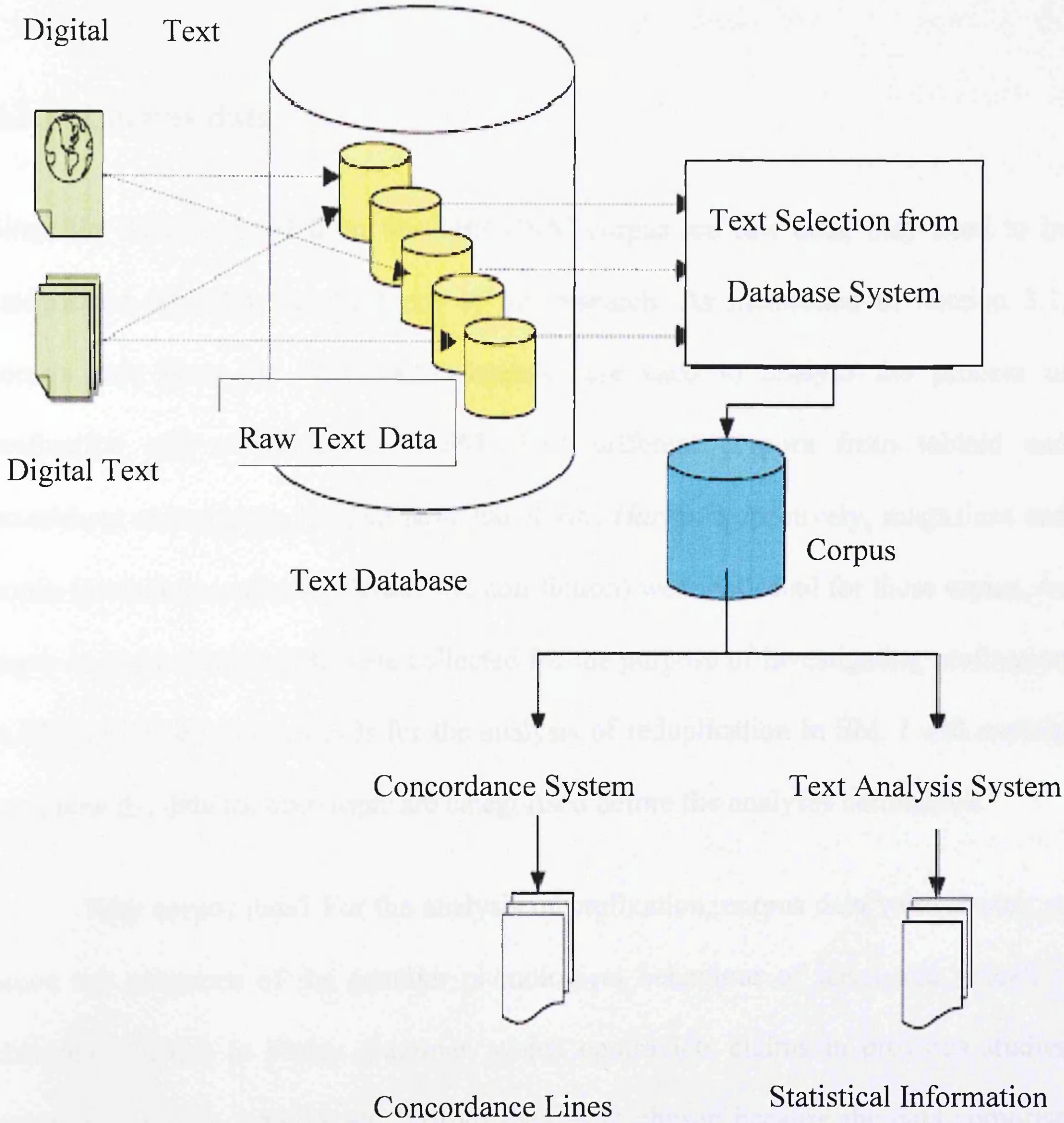
The DBP-UKM corpus database was constructed to provide an objective and authentic research facility to those interested in researching Malay. It was designed in the hope that corpus-based research might provide an impression of the typical behaviour of Malay words and phrases in real usage of the language in every aspect of life. The first language research effort was by DBP and began in 1983 with the Project for Computerized Text Analysis (Zaiton, 1987). As many as 2 million words were collected and used in the project through sampling techniques.

The DBP database is equipped with a system for processing the selection of texts. These can be processed in concordance form and analysed to obtain statistical data such as word frequency and word count. In the early stages, the collection of data into the database was somewhat opportunistic. Any material prepared in digital form or DBP books, magazines, papers, etc. were included in the database. The objective of the database construction, which was highlighted in 'DBP Aim Main Work 2001-2005', was to collect as many as 30 million words. By 2005, the total had reached 120 million words (Abdul Ghani et al.)[16]. The corpus consists of written classical texts (Malay histories and books) and modern texts, which are taken from books, newspapers and magazines. A verbal corpus is, however, still in the planning stage because the **marking up** of this type of corpus is more complicated than for written corpora. An entire view of the design of the DBP database is illustrated in Figure 2, below:

---

[16] The reference cited has not been published. Readers can obtain the article from the Institute of Language and Literature (DBP) or go to http://dbp.gov.my/korpus/korpus_DBP.pdf.

Figure 2: DBP Corpus Database System (adapted from Abdul Ghani et al.)



3.3 What data are used and how they are analysed?

In this section, I discuss what data are used to account for the three topics of concern in this thesis: prefixation, reduplication and dialectal variation. As mentioned in Section 3.1, three types of data are used for this study: corpus data, data from previous studies (secondary data) and interviews. I will start the discussion with the corpus data in Subsection 3.3.1. Subsection 3.3.2 presents the secondary data, followed by interview data in Section 3.3.3.

3.3.1 Corpus data

Since the data accessed from the DBP-UKM corpus are raw data, they need to be categorized according to the focus of my research. As mentioned in Section 3.1, corpus data from the DBP-UKM database are used to analyse the process of prefixation and reduplication in SM. Two different corpora from tabloid and broadsheet newspapers, i.e. *Harakah* and *Berita Harian* respectively, magazines and books (including academic, fiction and non-fiction) were collected for these topics. As many as one million words were collected for the purpose of investigating prefixation in SM, and five million words for the analysis of reduplication in SM. I will explain soon how the data for each topic are categorised before the analyses commence.

Why corpus data? For the analysis of prefixation, corpus data were chosen to prove the existence of the peculiar phonological behaviour of nasal and voiceless obstruent clusters in Malay grammar, which contradicts claims in previous studies regarding clusters. Additionally, corpus data were chosen because the data comprise examples of real usage of the language. As far as prefixation is concerned, the blocking of nasal substitution in a sequence of nasal and voiceless obstruents at prefix-root junctures, as presented in (9) (see Section 4.2.1.2), has never previously been discussed. Therefore, this present study attempts to show that this clustering is not always resolved by nasal substitution. This phonological process is somehow blocked from occurring between the clusters.

In SM reduplication, corpus data are used to reveal the use of the three types of reduplication discussed in this thesis: total, partial and affixal reduplication. To do this, five million words were collected. This begs the question of why so many words, some five million, were collected for this analysis. The answer is that reduplication, as one of the word formations in Malay, is not very productive compared to prefixation. Therefore, five million words were collected in the hope that we might obtain sufficient data to describe all three types of reduplication. There might be a problem if only one million words were collected. Most likely, we would not have enough data to explain adequately the use of those three types of reduplication in SM. The large corpus used in this thesis provides us with a large sample of tokens that allows us to make robust statistical generalizations for prefixation and reduplication. By this, I claim that this is the first study where a large corpus has been used to study the morphological and phonological system of Malay.

In what follows, I shall discuss how the categorisations of the collected corpus data were effected. Subsection 3.3.1.1 explains how the one million words were categorised before the analysis of SM prefixation started, while subsection 3.3.1.2 describes the categorisation of data for SM reduplication.

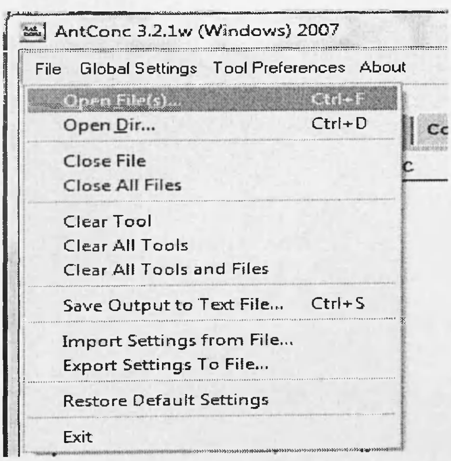
3.3.1.1 Data categorization for SM prefixation

Since this topic covers two types of prefixation, single and multiple prefixes, so the corpus data must be categorised according to the type of prefixes they belong to. It would be very difficult and time-consuming to categorise a million words without using corpus software. I have therefore used 'AntConc' corpus software to categorise the data.

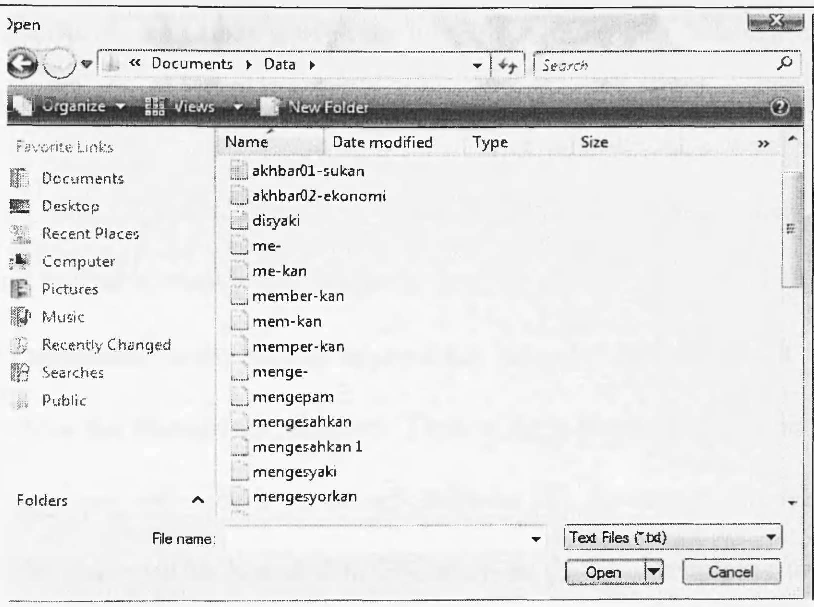
For single prefixation, I grouped the data according to the initial obstruent consonants of the bases: voiced and voiceless obstruents. These are two different sorts of data in which voiceless obstruents form the initial consonant of the root. They are voiceless obstruents with and without nasal substitution. The ones without nasal substitution are the type of data which violate the phonological requirements of the language since the voiceless obstruents remain undeleted. This type of data therefore violates *NÇ̥, the markedness constraint. I now explain how those groups, i.e. voiced and voiceless obstruents (with or without nasal substitution), are categorised using AntConc software.

(1) Voiced obstruents.

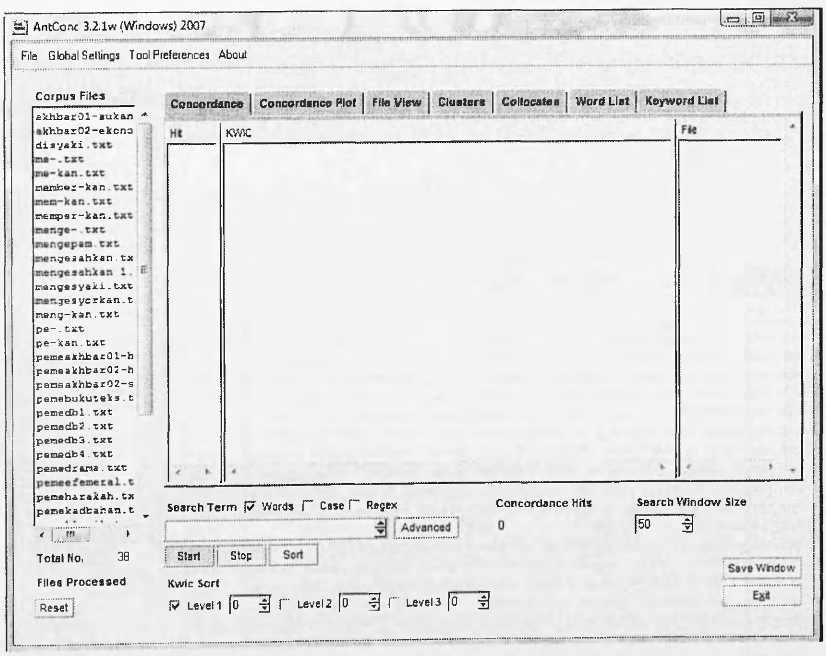
Before the relevant data for this group can be generated, we must choose a text file(s) where the data are stored by clicking on the **file** which is located at the top left of the software page and then select **open file(s)**. It looks like this:



A standard file-open will then appear. Double click on the text folder which contains the data, select the text files by clicking on them, and then click on the **open** button on the bottom – as the following screen shows:



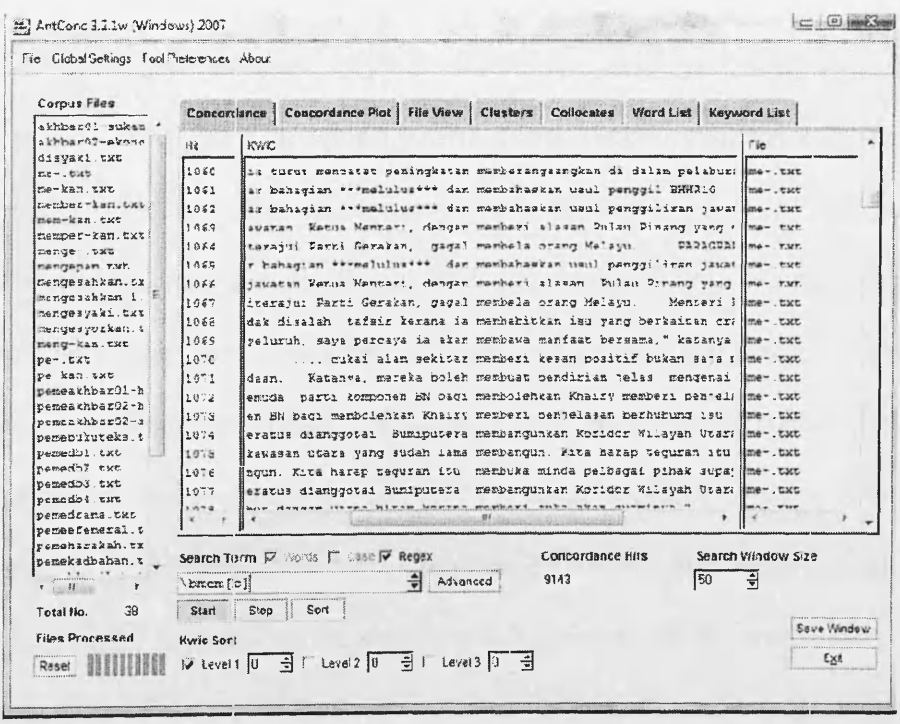
After all the text files have been selected, the data are now listed on the main page of AntConc, as shown below:



Now, we start searching for nasal and voiced obstruent clusters in the text files. To do that, we have to use some regular expressions (Regex) to search for the pattern we are looking for – nasal plus voiced obstruent. As the position of the nasal segment of the prefix is determined by the following initial consonant of the base, i.e.

voiced obstruent, we cannot search for whole allomorphs of /məŋ+/ i.e. [məm], [mən] and [məŋ] at one time. Searching for each of the allomorphs must be done one by one.

In what follows, I show how to search for the allomorph [məm]. Please bear in mind that the initial consonant following [məm] is a bilabial voiced obstruent, i.e. [b]. As just mentioned, some regular expressions must be used to search for the relevant words where the clusters are situated. Thus, a right character must be set for this. To search for [məm] followed by a voiced obstruent [b], the regular expression **\bmem[b]** is used. Observe that **\b** is added before **mem** in the regular expression **\bmem[b]** to indicate a word boundary. This regular expression, **\bmem[b]**, must be typed in the search term box on the main page of AntConc. Make sure to tick the **regex** box on the concordance screen. A concordance list will appear in the central area of the main page with all the occurrences of [mem+b]. All the steps mentioned above are shown in the screenshot below:



As we can see in this AntConc screenshot, the concordance list is for [mem+b]. To know how many examples of the word were found, just refer to the box of **Concordance Hits**. If we look at **Concordance Hits**, there are 9,143 words for [mem] + b initial base. Our search to find [mem] + b initial base examples of words in the corpus files is now done. To search for other voiced obstruents, such as [d] and [g], the steps discussed above are repeated.

(2) Voiceless obstruents.

To search for nasal plus voiceless obstruent clusters in the corpus is not as easy as searching for nasal plus voiced obstruent clusters. This is because nasal segments before voiceless obstruents undergone assimilation (for voiceless obstruents with nasal substitution). As we know, when a nasal combines with a voiceless obstruent, the nasal undergoes assimilation, while the voiceless obstruent is deleted. Therefore we get such outputs as /pəŋ-potoŋ/ → [pə-motoŋ] and /məŋ-tarik/ → [mə-nareʔ].

There are two possible ways to search for nasal and voiceless obstruent clusters in the corpus. First, if we set **\bpem** as the regular expression to search for /pəŋ+/ plus /p/ initial base, the concordance words that will appear can be: (1) the right words whereby the root actually begins with /p/ combined with a single nasal final prefix; (2) the wrong word where the root does not actually begin with /p/ combined with a single nasal final prefix but is a sonorant consonant instead, such as [pə-minat] where the underlying form is /pəŋ+minat/; (3) nominal multiple prefixes /pəŋ+pər/ → [pə-mər], as in /pəŋ+pər+badan+an] → [pə-mər-badan-an]. Second, if we only write **\bpe** as the regular expression, the software will generate all the words starting with [pə+]. Examples of the words that appear are as follows. For convenience, the words that start with [pə+] are underlined.

```
79 "Yang penting pemegang jawatan persatuan tidak bol BHES7
21 sendirian kerana pelumba di BHC34
424 Zahid yang juga pemenang pingat perak Kejohanan Lumb BHKS99
623 Daripada penelitian dan pemerhatian yang dibuat, didapati BHBC15
```

The examples of concordance words listed above are generated when the regular expression **\bpe** is used. None of the concordances listed above are words that we are looking for except concordance (79), which is the correct form of /pəŋ/ + /p/ initial base where the underlying form is /pəŋ+pəgaŋ/ → [pə-məgaŋ]. Concordance (21) is wrong since the initial consonant of the root is not a voiceless obstruent, i.e. [lumba]. The word [perak] in concordance (424) is not a prefixed word, it is a root word. Since the word starts with [pe] it also appears in the concordance list. The other form we get from the search is multiple prefixed words, such as in concordance (623), /pəŋ+pər+hati+an/ → [pə-mər-hati-jan].

To search for nasal and voiceless obstruent clusters in the corpus, I use a second way, i.e. **\bpe,** as the regular expression to find any initial voiceless roots that combine with prefix /pəŋ+/. Since the examples of words that appear in the concordance list contain more than one phonological character, the results can be categorised into five groups: (1) /pəŋ+/ combines with a voiceless obstruent initial root (with or without nasal substitution; (2) /pəŋ+/ with a sonorant initial root; (3) nominal prefixes /pəŋ+ mər/ → [pə-mər]; (4) /pəŋ+/ with a monosyllabic root; and (5) /pəŋ+/ with a voiced obstruent initial root. Thus the categorisation work has to be done manually whereby all the examples are categorised according to their phonological character. This means that we already have five types of data, all of which are useful for our analysis. I briefly lay out some examples from the concordance list to represent those groups.

(1) /pəŋ+/ with voiceless obstruent initial root.

(i) With nasal substitution

921 206 mengesan penipuan apabila pemeriksaan pengesahan BHA198

941 bahawa kemunculan tanda pemesongan bearis bukanlah alasan BHA176

980 Mengenai aduan ke atas pengilang atau pengimport yang disyaki

89177 Mengenai ekonomi pula, penubuhan Zon Pemprosesan Eksport adalah

(ii) Without nasal substitution.

2446 untuk memudah dan mempercepatkan pemprosesan permohonan

3748 selain pengalaman meluas dalam pentadbiran di kementerian

9643 berisi air dan memasukkan tiub pensterilan ke dalamnya untuk

89242 terus diberikan kepada kegiatan pengkomersialan keluaran

(2) /pəŋ+/ with sonorant initial root.

118 m Syarikat Harvard Heri 198 Malaysia 0 peratus dan dana BHBC78

108 204 kesan kepada pelanggan dan pemegang saham dari segi BHBC89

264 mereka untuk muncul selaku pemenang ketika lawan juga BHHS72

3025 hati, himpunan terbaru pewarna rambut Wella Decore menyediakan

(3) Nominal prefixes /pəŋ+mer/ → [pə-mər].

902 BSKL) membingungkan pemerhatian apabila terus mencatat BHDE81

903 Deutsche itu kerana pemerhati berpendapat ia mungkin BHDE26

937 teknologi pemerolehan minyak di tempat pengeluar BHFE61

(4) /pəŋ+/ with monosyllabic root

923 penipuan apabila pemeriksaan pengesahan permohonan BHIE78

992 ereka yang disyaki selain memperincikan rancangan pengeboman

7463 untuk mengatasi masalah kerosakan jentera pengepam air sejak

89266 Bluetooth(R) T305; kaca mata O ROKR; pengecaj mudah alih P790;

(5) /pəŋ+/ with voiced obstruent initial root.

947 angkan jadual supaya pemeteraian perjanjian penggabungan BHFE13

948 penggabungan an 208 a pemberita selepas majlis pemeteraian BHFE13

1018 keterangan daripada kira-kira 20 pendatang Indonesia kerana

1019 pendakwa ialah tempoh masa diambil oleh orang yang

The same situation occurs for the prefix /məŋ+/ plus initial voiceless obstruent base. All the groups mentioned above appear in the concordance list except for the third group. When **\bme** as the regular expression is entered into the **Search Term** box, we do not find any examples of words for the nominal prefixes /pəŋ+mər/ as we found before for the prefix /pəŋ+/. Verbal prefixes, i.e. /məŋ+per/ → [məm-pər], are found instead. Here are examples of words for those groups:

(1) /məŋ+/ with voiceless obstruent initial root

(i) With nasal substitution

20 iaitu membuat pemecahan secara mengejut. Dia yang BHLS54

33 kerana dikatakan tidak muntuk memikul tugas sebagai BHBC16

35 itu, cukup Itali itu, pernah menewaskan pemecut handalan BHj99

94 dwitahunan 570 ini boleh memisahkan antara pemenang dan BHES60

(ii) Without nasal substitution

1409 Menjadi harapannya, lirik yang dihasilkan tidak mengkhayalkan

4038 untuk menjadikannya lebih bijak dari segi memproses dan mengawal

6061 Penduduk Palestin sebelum ini pernah memfailkan saman terhadap

6691 di luar bangunan muzium. Muzium itu turut mempamerkan

(2) /məŋ+/ with sonorant initial root

12 419 rsama melolos dan melakukan pemecahan meninggalkan BHC135

113 hambar, dan melayakkan pemenang ke pertandingan Eropah BHES28

111 kuning. Ditanya meraih pengiktirafan selaku pemenangjersi BHJS18

5550 Umno sebagai parti pemerintah tidak cuba untuk mewujudkan sistem

(3) Verbal prefixes /məŋ+per/ → [məm-pər]

209 mereka ke Itali untuk berla 539 paksa mempercepatkan tarikh BHC43

185 berdiri di pentas pemenang, barulah McRae memperlihatkan BHES65

462 zi yang mengenal pasti mereka yang disyaki selain memperincikan

526 Majlis Usahawan di Peringkat Daerah (MPUD) dan memperkukuhkan

(4) /məŋ+/ with monosyllabic root

187 untuk keluar selaku pemenang, sekali gus mengesahkan BHES41

423 Polis mengesyaki pembunuhan Dr Joe Fernandez disyaki dilak

382 kred cuba mengebom sebuah pusat pemeriksaan tetapi BHIE78

5147 Cabaran yang dihadapi itu memaksa saya mengehadkan jumlah

(5) /məŋ+/ with voiced obstruent initial root

6 taktiknya melakukan pemecahan awal sebelum mendaki bukit BHJS96

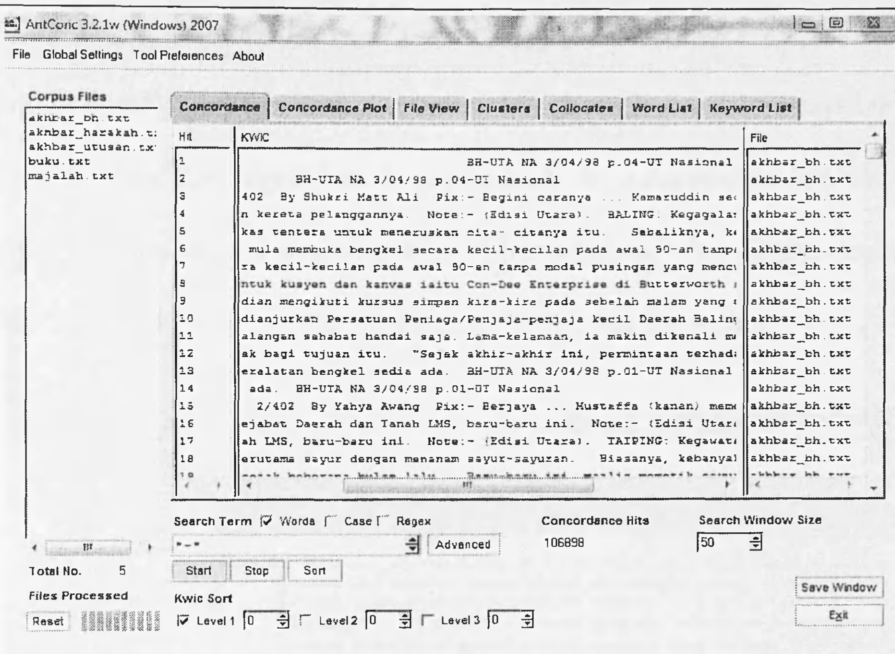
7 64 man Nieto Fernandez membuat pemecahan awal selepas BHJS96

104 meneruskann dirinya mampu menggantikan pemenang emas Brunei '99,

3.3.1.2 Data categorization for SM reduplication

As noted, reduplication is also one of the discussion topics, along with prefixation and dialect variation. The three types of reduplication, i.e. total, partial and affixal reduplication, discussed in this study are concerned with SM reduplication, with the exception of partial reduplication. As well as SM, partial reduplication in PD is also examined as partial reduplication in the dialect contains heavy reduplicative morphemes. As mentioned earlier in this thesis, all the data for SM examined in this study are based on the DBP-UKM corpus, while the data for the Malay dialects are based on previous scholars' work. Thus, the data used for the analysis of reduplication, which includes all three types of reduplication except for PD partial reduplication, are from the DBP-UKM corpus. As mentioned in subsection 3.3.1, in order to analyse SM reduplication, five million words of corpus data are examined.

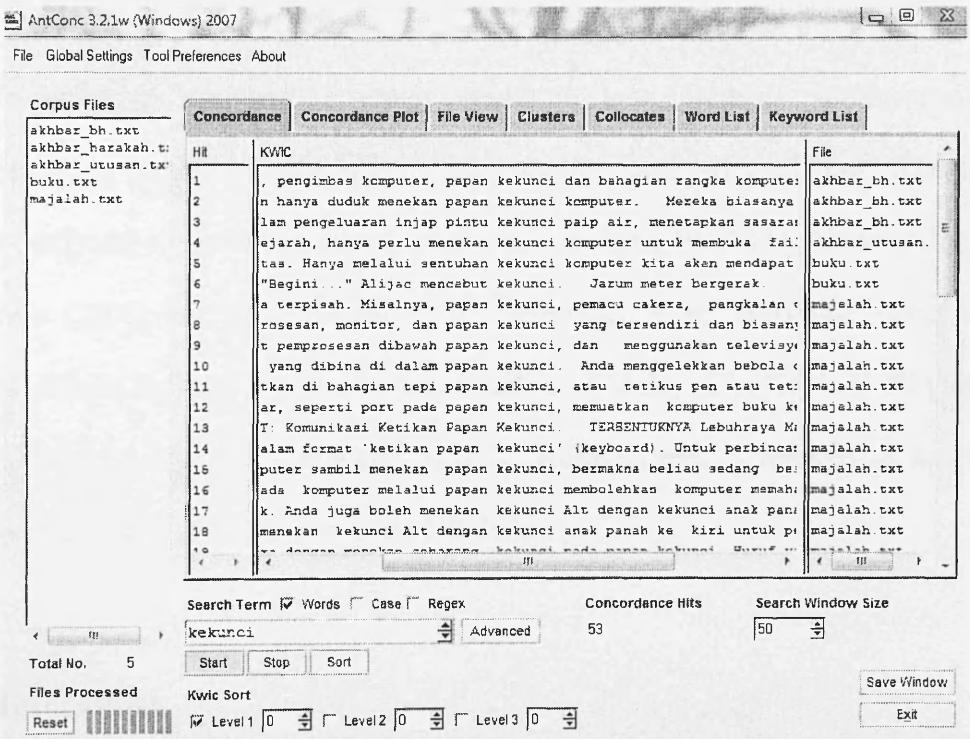
In what follows, I am going to explain how I worked with these five million words to analyse SM reduplication. As for prefixation, I also use AntConc to categorise the five million words into the three types of reduplication. I begin now to explain how the categorisation of data is effected. With the same steps applied to prefixation, we must first choose the text file(s) that are to be analysed. In Malay, reduplicated words are normally indicated by a hyphen '-' which is written between the reduplicative morphemes and bases. To search for reduplicated words in the data, this symbol is typed along with an asterisk (*) symbol before and after the hyphen in the **Search Term** box in the main page of AntConc. Different from the situation for prefixation, as '-' is not a regular expression, the **regex** box does not have to be ticked. By clicking on the **Start** button, all the reduplicated words appear on the screen.



It should be noted that a hyphen '-' is not used in partial reduplication as the reduplicative morpheme and base are combined into one, e.g. [ʤəʤari] 'radius' and [gəgendaŋ] 'eardrum'. Since a hyphen is not used in partial reduplication, the concordance list above contains four types of reduplication, i.e. total, rhyming, chiming and affixal reduplication. Therefore, they have to be separated according to their group.

As partial reduplication is not included in the above concordance list, a new search has to be performed. For this, no regular expressions or symbols are used to search for partially-reduplicated words. Another way of word search can be done using AntConc by simply typing the word we want in the box labelled **Search Term**. This way is fine for searching for partially-reduplicated words as the occurrences of such words are fewer than for total and affixal reduplication. Because partially-reduplicated words are few in number, speakers of Malay are able to remember all the words that belong to this group.

I now explain how the word [kəkuntʃi] 'password' is searched for. First, type the word [kəkuntʃi] in the **Search Term** box and then click on the **Start** button. Once again, make sure the **regex** box is unchecked. A concordance list for the word [kəkuntʃi] will appear in the central area of the window, with all the occurrences of *'kekunci'* in context. Observe that there are 53 examples altogether.



3.3.2 Secondary data

Secondary data from previous studies are used to analyse the issues raised in PD reduplication (only for partial reduplication) and dialectal variation. One might ask why secondary data are used to analyse these two topics and not corpus data. The reason is that the scope of investigation into those topics covers both SM and some of its dialects. Since the DBP-UKM corpus does not contain any data other than for SM, referring to relevant previous studies seemed the best way to obtain the data for Malay dialects. Secondary data from previous studies are used in the dialectal variation chapter, which is only concerned with Malay dialects. Likewise for the analysis of partial reduplication; since this topic seeks to explore all patterns occurring in Malay reduplication, some PD data were required for the analysis since SM only has light reduplication.[17] For this, the work of two scholars, Ahmad (1991) and Aripin (2005), is referred to.

In order to analyse the issues raised in both dialectal variation and reduplication, a number of relevant previous studies were selected. For dialectal variation, different works by various scholars are referred to. Works by Ahmad (1991) and Aripin (2005) are used to analyse the phenomenon of nasal and voiceless obstruent clusters in Perak. For the Kelantan dialect of Malay, Che Kob's (1985) and Teoh's (1994) studies are used, while work by Rufus (1966) is referred to as the source to analyse this phenomenon in the NS dialect.

3.3.3 Interviews

Interviews are another research method used in this present study as a data source. This method was adopted to obtain more data for the analysis of dialectal variation, particularly for the Kelantan and NS dialects. The data obtained from the previous studies (e.g.: Che Kob, 1985; Rufus, 1966) were not sufficient for the analysis. More data were needed in order to see how nasal and voiced/voiceless obstruent clusters really behave in the dialects. In order to get more data, two native speakers of NS (i.e. Dr. Mohd. Fadzeli Jaafar and Mr. Zulkifli Ahmad) and one for Kelantan (i.e. Assoc.

[17] As far as reduplication is concerned, reduplication in PD has been discussed much more in previous studies than for other dialects of Malay. The reduplication process in Perak is different from that of SM as heavy reduplicative morphemes in partial reduplication occur in this dialect.

Prof. Ajid Che Kob) were interviewed.[18] I prepared a short word list based on the data for PD obtained by Ahmad (1991) (see Appendix E for the word list). This was mainly to get the same list of words for the Kelantan and NS dialects as were found in Perak, so that the differences could be more easily seen. In the interview sessions, speakers were asked to pronounce all those words.

3.4 Advantages of choosing corpus data

I acknowledged earlier in this thesis that this is a reanalysis study concerning prefixation and reduplication in Malay. Numerous works have been done on prefixation and reduplication, with various approaches by scholars who have studied Malay (e.g.: Omar, 1975, 1986, 1993; Hassan, 1974, 1987; Karim et al., 1989; Karim, 1995; Onn, 1980, Ahmad, 1993, 2000a, 2000b, 2005; Teoh, 1994; and many others). However, little theoretical work has been done, particularly on prefixation. Work by Ahmad (2000b) and Teoh (1994) on prefixation was based on autosegmental analysis, while the only work on OT was done by Ahmad (2005). However, none of those studies used corpus data as the data source. The process of prefixation was analysed by scholars based on their intuition as native speakers. As a consequence, the results derived from the analyses could only account for some limited data. This might be the reason why previous analyses could not adequately explain the real process of prefixation. Moreover, earlier studies of Malay tended to treat counter-examples as being exceptional in the language since the rules they postulated do not work for their data. Thus their analyses were lacking in observational and explanatory thoroughness.

[18] I would like to thank Professor Ajid Che Kob, Dr. Mohd. Fadzeli Jaafar and Zulkifli Ahmad for their contributions to the data. As a native speaker of the NS dialect, Dr. Mohd. Fadzeli Jaafar admits that younger users of the dialect nowadays do not really speak the dialect, even at home. If they do speak it, they always mix it up with SM. Most probably, they have been influenced by SM which is taught at schools and universities in Malaysia.

In relation to this, this present study reexamines the process of prefixation in SM as well as reduplication, which has been questioned in the previous analyses. In my view, it is linguists' responsibility to think of all possible solutions to the problems that occur in the Malay language in a scientific way. This corpus-based study of Malay shows how corpus data can contribute to a comprehensive analysis of the issues arising. It is hoped that, by corpus analysis, prefixation and reduplication in SM can be explained in a way that corresponds to the way language users use the language in reality. McEnery and Wilson (2001: 1) state that: *'corpus linguistics is perhaps best described for the moment in simple terms as the study of language based on examples of 'real life' language use'*. They (ibid.) also state that *'Corpus linguists study real language, other linguists just sit at their coffee table and think of wild and impossible sentences'*. Therefore, there appears to be a problem in non-corpus studies as no statistical analysis of the data is involved. The absence of statistical data analysis prevents us from identifying the real usage of a particular pattern under investigation in a language.

3.5 Conclusion

I have now laid out and discussed all the research methods used in this present study. In the next chapter, we will first see how the problems that occur in Malay prefixation can be accounted for by the theory adopted for this study – MBT within OT.

4. PREFIXATION

4.1 Introduction

In this chapter, I shall examine the process of prefixation in Malay, which is one of the foci of this study. As is generally claimed (e.g.: Ahmad, 2000b, 2005; Teoh, 1994), Malay is one of the languages which do not permit nasal and voiceless obstruent clusters in their process of prefixation. Other languages with the same constraint include Indonesian (Pater, 1999), Toba Batak (Hayes, 1986), Kaingang (Henry, 1948; cf. Piggot, 1995), Chamorro (Topping, 1973: 50), Javanese (Poedjosoedarmo, 1982: 51), the Bantu languages such as Umbundu, Si-Luyana and OshiKwanyama (cited by Pater, 1999) and Mandar (Mills, 1975). Nasal substitution is one possible strategy for languages to rid themselves of nasal and voiceless obstruent clusters (Pater, 1999). In Malay, according to previous Malay scholars (e.g.: Hassan, 1974; Omar, 1986; Karim et al., 1994; Karim, 1995; and many others), this strategy is regularly applied to prevent the clusters emerging in the surface representation.

Observations from the DBP-UKM corpus show that the claim regarding nasal substitution postulated by those previous Malay scholars on prefixation does not hold for the whole dataset.[19] In an observation based on a sample of 1,141,816 words in SM, I note that 461 words with single prefixation contain nasal and voiceless obstruent clusters (see Appendix A), while 6,154 words have multiple prefixes, /məŋ+pər/ (see Appendix B). Moreover, there would be problems in accounting for the process of prefixation if the previous scholars' claims were applied to the data.

[19]The generalization postulated by previous studies can only explain some of the output derived from the process of prefixation. This is partly because the data they used were from their intuition as native speakers creating their own data.

This chapter will thus reanalyze the phenomenon of nasal and voiceless obstruent clusters in single prefixation, i.e. one prefix attached to a root, and multiple prefixation, i.e. more than one prefix attached to a root. These will be discussed in subsection 4.2.1 and section 4.3, respectively. In discussing the occurrence of nasal and voiceless obstruent clusters in those two types of prefixation, we will find out what causes the clusters emerge in the surface representation of single and multiple prefixation although they are supposed to be banned in the language. Later in this chapter, we will be able to obtain the answer to research question (1), i.e. why nasal and voiceless obstruent clusters occur in two of the three environments, prefix-root and prefix-prefix junctures (see Research Question 1(a) in Section 1.4), despite the fact that the language does not allow clusters to surface, as in [məm-protes] VERBL-protest 'to protest' and [məm-pər-bəsar-kan] VERBL-VERBL-large-CAUS.SUF 'to cause to enlarge'.

Besides nasal and voiceless obstruent clusters, this chapter will also discuss the other cluster that is involved in prefixation, that is nasal and sonorant (see subsection 4.2.2). As was mentioned in subsection 1.3.1, this cluster is also disfavoured in the language, along with nasal and voiceless obstruent clusters. Therefore, this cluster has been chosen for discussion in this study, in addition to nasal and voiceless obstruent clusters. As sonorants following nasals do not cause great problems when compared to voiceless obstruents, the analysis is more to demonstrate how nasal and sonorant clusters in the underlying representations are resolved in the language. We will investigate whether nasal substitution is also the way to avoid this cluster from emerging in the surface representation, or if there might be other phonological processes that are applied. The analysis of this will then answer part of Research Question 1(a).

In order to account for single prefixation, I am following Itô and Mester (1999) who analysed the structure of the phonological lexicon in Japanese. As mentioned in subsection 2.4.1.1, Malay has co-existent grammars, where one does not allow nasal and voiceless obstruent clusters, while the other one does. As we shall see in subsection 4.2.1, nasal substitution only applies when the roots are Malay native words. This means that *NC̥, a constraint which bans a sequence of nasal and voiceless obstruents on the surface, is obeyed for Malay native words but is violated for foreign words. Similar to Japanese, non-native words violate one or more constraints that are obeyed by the Yamato lexicon (Japanese native words) (see Table 4 in subsection 2.4.1.1). This analysis of single prefixation will thus proceed according to the etymology of the word: native or non-native. From this, I will then propose three categories for the lexicon in Malay as follows: monosyllabic foreign, undeleted voiceless plosive in loanwords, and native words, as illustrated in Figure 3, below:

Figure 3: Categories of the lexicon for Malay

MALAY LEXICON

Sub-lexicon 1	Sub-lexicon 2	Sub-lexicon 3
monosyllabic foreign	undeleted voiceless plosive in loanwords	native

The co-existent grammars also occur in multiple prefixation where one does allow the presence of nasal and voiceless obstruent clusters while the other does not. In the following examples, we see that the clusters occur in verbal multiple prefixes,

/məŋ+pər/. In nominal multiple prefixes, /pəŋ+pər/, on the other hand, the clusters undergo nasal substitution. Shown in 103(a) and (b) are some relevant examples of multiple prefixes, /məŋ+pər/ and /pəŋ+pər/, respectively.

103. a) Multiple prefixes with nasal and voiceless obstruent clusters - /məŋ+pər/.

(i) **məm-pər**-kuat-kan
VERBL.PRF-NOM.PRF-strong-CAUS.SUF
'to strengthen'

(ii) **məm-pər**-tadʒam-kan
VERBL.PRF-NOM.PRF-sharp-CAUS.SUF
'to sharpen'

b) Multiple prefixes without nasal and voiceless obstruent clusters - /pəŋ+pər/.

(i) **pə-mər**-badan-an
NOM.PRF-VERBL.PRF-body -NOM.SUF
'organisation'

(ii) **pə-mər**-kaja-an
NOM.PRF-VERBL.PRF-rich -NOM.SUF
'enrichment'

In this study, I will claim that nasal and voiceless obstruent clusters occurring in /məŋ+pər/ are due to the morphological boundary prefix-prefix where the clusters exist. I will argue, in Section 4.3, that nasal substitution does not necessarily apply whenever there is a nasal and voiceless obstruent cluster in which the phonetic environment is met. Although there is a nasal and voiceless obstruent cluster, nasal substitution cannot instantly be applied as the phonological process of nasal substitution is largely determined by the morphological domain of prefixes. The evidence strongly indicates that the morphological domain is the relevant conditioning factor that must also be taken into account to overcome the blocking of nasal substitution in multiple prefixation. It should be mentioned that this solution, where the application of nasal substitution is determined by looking at the morphological domain, has never been considered in previous Malay scholars' analyses. In OT, this matter can straightforwardly be accounted for by proposing a morphology-phonology interface constraint which is able to block nasal substitution from occurring at this morphological domain. To account for this, I introduce a constraint named EDGE-INTEGRITY (McCarthy and Prince, 1995), in (85), which is able to preserve the segments at the edges of both prefixes so that the segments cannot undergo nasal substitution. However, the above examples, where nasal and voiceless obstruent clusters undergo nasal substitution, will be analysed in this study as a case of analogy where the clusters have been accounted for in the same way as in single prefixation. See Section 4.3 for further discussion. These cases of inconsistent application of nasal substitution in Malay prefixation will be analysed by proposing a co-phonology analysis, i.e. of different constraint rankings. There are good reasons why co-phonology is preferred. This theory allows markedness and faithfulness constraints to be re-ranked in different morphological constructions, even in the same language (Downing, 2008). This power of co-phonology is essential to account for languages with complex morphologically-conditioned phonology (ibid.). From this, Malay prefixation will then have two constraint rankings: (1) Markedness >> Faithfulness in voiceless obstruents nasal substitution; and (2) Faithfulness >> Markedness in voiceless obstruents without nasal substitution.

In the following Section 4.2, I will be discussing single prefixation. This section is divided into two subsections to address: (1) nasal and voiceless obstruents and (2) nasal and sonorant clusters, which will be discussed in subsections 4.2.1 and 4.2.2, respectively. The discussion of nasal and voiceless obstruent clusters will be according to the type of root, i.e. monosyllabic in subsection 4.2.1.1, and disyllabic roots in 4.2.1.2.

4.2 Single prefixation

This section focuses on single prefixation. As already mentioned in Chapter 1 Section 1.3.1, two types of clusters are disfavoured in Malay: (1) nasal and voiceless obstruents and (2) nasal and sonorant clusters. Since these clusters are involved in single prefixation, this section will be divided into two subsections. Subsection 4.2.1 discusses nasal and voiceless obstruent clusters, subsection 4.2.2 discusses nasal and sonorant clusters.

It is worth mentioning that the language has different ways of eliminating nasal and voiceless obstruent clusters. The ways of eliminating the clusters are what we call 'conspiracies' (see Chapter 2, subsection 2.2.3). The way the language eliminates the clusters depends on the size of the root that the prefixes are attached to. Subsection 4.2.1 will discuss the clusters (i.e. nasal and voiceless obstruent clusters) in two different types of root: monosyllabic and disyllabic/polysyllabic roots, which are discussed in subsections 4.2.1.1 and 4.2.1.2, respectively.

4.2.1 Nasal and voiceless obstruent clusters

In this section, we are going to see how nasal and voiceless obstruent clusters behave for monosyllabic, disyllabic and polysyllabic roots.[20] The analysis in this section will

[20] It is important to note that original Malay words are mostly disyllabic (Ahmad 2007; Teoh 1994). One might easily predict that words with fewer or more syllables are probably non-native words. Readers can consult the cited references for further explanation.

thus be divided into two parts: (1) monosyllabic roots in subsection 4.2.1.1, and disyllabic / polysyllabic roots, which will be considered together, in subsection 4.2.1.2.

4.2.1.1 Monosyllabic Roots

As mentioned in Chapter 2, monosyllabic roots are also involved in the process of prefixation, as well as disyllabic/ polysyllabic roots. For monosyllabic roots, the process of prefixation is slightly different from that for disyllabic/ polysyllabic roots. The difference can be seen when schwa is epenthesized between monosyllabic roots and nasal final prefixes. As we will see in the following subsection, 4.2.1.2, there is no process like schwa epenthesis in the process of prefixation for disyllabic and polysyllabic roots to eliminate the clusters. In other words, epenthetic schwa is only applicable when nasal final prefixes are attached to monosyllabic roots.

Before I offer an OT account to explain the process of epenthetic schwa, let us observe the data in (104). The descriptive generalisations that can be summarised are as follows: (1) five prefixes can be attached to monosyllabic roots: /məŋ+/, /pəŋ+/, /di+/, /tər+/ and /bər+/; (2) schwa is epenthesized between monosyllabic roots and the nasal final prefixes, /məŋ+/ and /pəŋ+/ only; (3) the process of schwa epenthesis can not however be applied if the prefixes are /di+/, /tər+/ and /bər+/; and (4) the nasal segment in the prefix then alternates to a velar nasal [ŋ] after the process of epenthesis occurs.

104. Data for the Malay monosyllabic foreign lexicon (from the DBP-UKM corpus).[21]

a)	məŋ-ə-**cam** ACT.PRF-STEMEX-recognise 'to recognise'	(b)	məŋ-ə-**cap** ACT.PRF-STEMEX-stamp 'to stamp'
c)	məŋ-ə-**sah** ACT.PRF-STEMEX-validate 'to validate'	d)	məŋ-ə-**kod** ACT.PRF-STEMEX-code 'to code'
e)	məŋ-ə-**bom** ACT.PRF-STEMEX-bomb 'to bomb'	f)	pəŋ-ə-**bom**-an ACT.PRF-STEMEX-bomb-NOM.SUF 'bombing'
g)	məŋ-ə-**had** ACT.PRF-STEMEX-limit 'to limit'	h)	bər-**cap** REF.PRF-stamp 'to stamp'
i)	tər-**pam** PAS.PRF-pump 'pumped'	j)	tər-**had** PAS.PRF-limit 'limited'
k)	di-**had**-kan PAS.PRF-limit-CAU.SUF 'is limited'	l)	di-**lap** PAS.PRF-wipe 'is wiped'

An interesting question, which might be raised here is: what is the motivation for schwa epenthesis? This is an interesting question because schwa is not epenthesized in other types of roots, i.e. in disyllabic/ polysyllabic roots, as we will see later in subsection 4.2.1.2. Before I provide the answer to the question of what is the motivation of epenthetic schwa from the OT point of view, a first reason is that we can simply say that schwa might be epenthesized between a nasal final prefix and a monosyllabic root because the nasal segment in the prefix would not have a place of articulation to alternate to a velar nasal /ŋ/ if schwa was not added. If epenthetic schwa

[21] See Appendix C for the complete list of monosyllabic foreign lexicon data.

was not added to monosyllabic bases, then we could make a prediction for the possible output. Expected outputs are, by way of example:

105. Input	Output
a) /məŋ-pam/	*[məmam]
ACT.PRF-STEMEX-pump
'to pump'
a) /məŋ-bom/	*[məmbom]
ACT.PRF-STEMEX-bomb
'to bomb'

These outputs, *[məmam] and *[məmbom], are predictable if the regular rules for nasal assimilation and voiceless obstruent deletion, as postulated in ruled-based analyses to avoid nasal and voiceless obstruent clusters, are applied to monosyllabic bases. As noted in Chapter 1 Section 1.3.1, monosyllabic bases are not analysed in rule-based analysis. The outputs above are predicted simply on the basis of the two rules postulated in rule-based analyses to account for the clusters in the language. In the above examples in (105), since schwa is not epenthesized between the prefix and the monosyllabic root, so the regular phonological processes, nasal assimilation and voiceless obstruent deletion, are applied. Therefore we get outputs with nasal substitution, *[mə-mam] and *[məm-bom].

The essential point that I am trying to demonstrate here is that the process of prefixation for monosyllabic bases is not the same as that with other larger roots, such as disyllabic/ polysyllabic roots, as we will see in 4.2.1.2. The difference can be seen in (105) where the outputs *[mə-mam] and *[məm-bom] for the inputs /məŋ-pam/ and /məŋ-bom/ are not the desired outputs. In my opinion therefore, scholars implementing rule-based analyses should not have missed analysing monosyllabic bases in their studies. More importantly, such outputs as *[mə-mam] and *[məm-bom] are not found in the DBP-UKM corpus used in this study as initial voiceless plosives in monosyllabic bases do not undergo nasal substituion. I am convinced that this output is impossible to find, as it simply does not exist as any Malay word.

Second, schwa is inserted due to the disyllabicity minimality requirement of the language, as pointed out by Onn (1980), Maris (1980), Teoh (1994) and Ahmad (2000b), i.e. a word in Malay must have at least two syllables. Therefore schwa is by no means inserted to fulfil the word requirements of the language. Before we go further, let us consider how the disyllabicity minimality condition in Malay was accounted for by previous Malay scholars.

In linear studies for example, Onn (1980: 57) says that most of the 926 word stems[22] he observed are disyllabic, while trisyllabic stems number only 132. Based on his observations, he claims that most stems in Malay are disyllabic and therefore he claims further that Malay is a disyllabic language. Shown in (106), below, are some relevant examples:

106. Disyllabic and trisyllabic words (from Onn, 1980: 57).

(a) Disyllabic words

/ibu/	[i.bu] 'mother'
/bulan/	[bu.lan] 'moon, month'
/sumpah/	[sum.pah] 'to swear'
/ambil/	[am.bel] 'to take'
/inti/	[in.ti] 'the stuffing in... '
/ikat/	[i.kaʔ] 'to tie'
/təntu/	[tən.tu] 'sure'

[22] Note that 'stem' as used by Onn (1980) means root as used in this study. I will continue to use root in the whole of this thesis and will shortly argue for the significance of it, instead of stem, to refer to the same thing – a word with no prefix attached. See p. 137.

(b) Trisyllabic words[23]

/tuala/	[tu.wa.le] 'towel' (from Portugese)
/usia/	[u.si.ye] 'age' (from Sanskrit)
/binasa/	[bi.na.sə] 'to destroy' (from Sanskrit)
/səmua/	[sə.mu.wə] 'all' (from Sanskrit)
/ijazah/	[i.dʒa.zah] 'degree' (from Arabic)
/bəlandʒa/	[bə.lan.dʒə] 'expenses'
/binatang/	[bi.na.taŋ] 'animal'

I am in agreement with Onn on this point. The corpus data used for prefixation analysis shows that most roots[24] in Malay are disyllabic. Monosyllabic roots constitute only 395 words or 0.03% of all roots. Disyllabic roots comprise 68.4% (or 779,860 words). Trisyllabic/polysyllabic roots are also small in number. There are 31.6% or 360,814 words. I lay out some of the examples from the corpus in (107) and a graph (see Figure 4 below) showing the distinction between those roots:

107. Monosyllabic, disyllabic and trisyllabic roots (from DBP-UKM corpus).

(a) Monosyllabic roots

/pam/	[pam] 'pump'
/lap/	[lap] 'wipe'
/had/	[had] 'limit'
/bom/	[bom] 'bomb'

(b) Disyllabic roots

/kirim/	[ki.rem] 'send'
/kiri/	[ki.ri] 'left'
/dalam/	[da.lam] 'deep' / 'in'
/tidak/	[ti.daʔ] 'no'
/duda/	[du.də] 'widower'

23 Most of the trisyllabic words listed by Teoh here were originally borrowed words, and thus the number of words is much smaller than for disyllabic words.

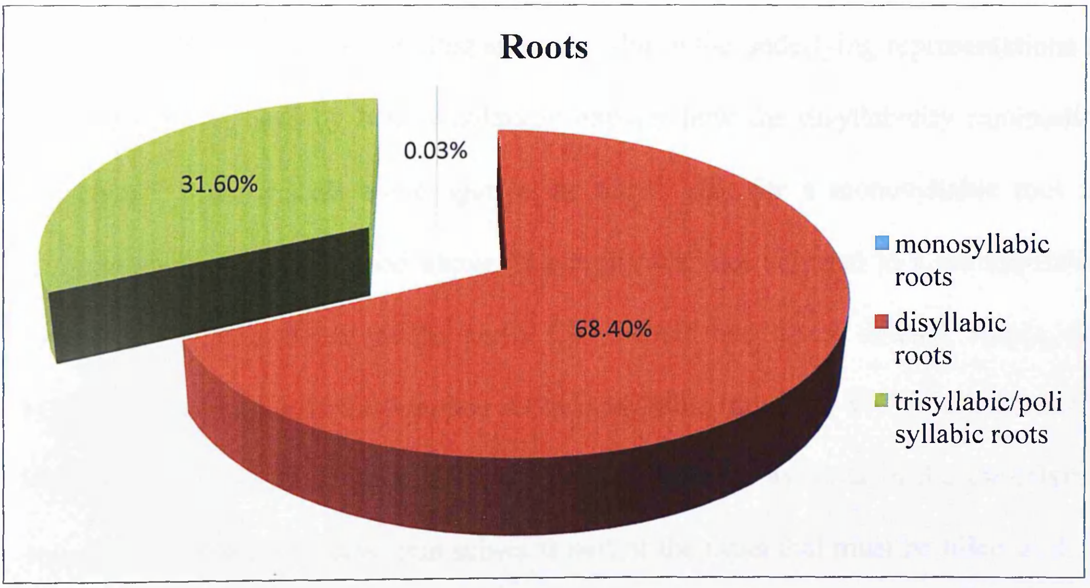
24 I use root as the replacement of word. See my argument for this on p. 139.

/luas/	[lu.was] 'large'
/petik/	[pe.teʔ] 'pluck'

(c) Trisyllabic/polysyllabic roots

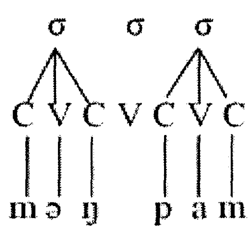
/teratai/	[te.ra.tai] 'lotus'
/perintah/	[pe.rin.tah] 'command'
/ansuran/	[an.su.ran] 'instalment'
/ibadat/	[i.ba.dat] 'worship'
/mereka/	[me.re.kə] 'they'
/dukacita/	[du.ka.ci.tə] 'sad'

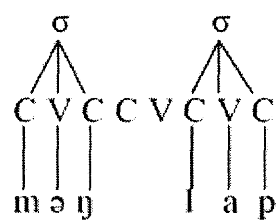
Figure 4: Monosyllabic, disyllabic, trisyllabic roots (from the DBP-UKM corpus).



Teoh (1994) and Ahmad (2000b), who agree with the disyllabicity minimality condition as postulated by earlier scholars, provide their own justifications for epenthetic schwa in monosyllabic bases from the viewpoint of non-linear analysis. Teoh (1994) claims that the template for a monosyllabic base is CV.CVC. He claims further that Malay is an inherently disyllabic language, and therefore this template is chosen as an underlying representation. Ahmad (2000b) claims the lexical representation for a monosyllabic word originally contains two syllables. In his analysis, he postulates that there is an empty V slot in the underlying representation. In my view, the empty slots, CV and V in Teoh's and Ahmad's analyses respectively, are mainly to ensure that the disyllabicity minimality condition of words is satisfied. The templates as claimed by those scholars are repeated below for convenience:

108. (a) INPUT (Ahmad 2000b) (b) INPUT (Teoh 1994)





As can be seen, we note that an empty slot in the underlying representations is the solution proposed by both scholars to explain how the disyllabicity minimality condition is met. I shall argue against the empty slot for a monosyllabic root as proposed by Teoh and Ahmad above. If a prefix was not attached to a monosyllabic root, then the skeletal tier for the prefix CVC would need to be deleted. Hence, the remaining skeletal tier is the template for monosyllabic roots, i.e. V.CVC (in Ahmad's UR) and CV.CVC (in Teoh's UR). In my view, the empty slots in the underlying representations seem to show that schwa is part of the input that must be filled in. It is important to note that schwa would not appear in monosyllabic roots. In other words, the presence of schwa can only be seen in the surface form when a prefix ends with a nasal attached to a monosyllabic root. As we saw earlier, in Chapter 2 subsection 2.2.4, the disyllabicity minimality condition in Choctaw is satisfied by epenthesizing /a/ in the monosyllabic root: /bi/ 'to kill' – [a-bi] 'for you to kill'. Although Malay is also a disyllabic language, this solution can not however be applied to account for monosyllabic roots, as analysed by Teoh (1994) and Ahmad (2000b). This is because schwa only appears when nasal final prefixes are attached to monosyllabic roots. Unlike Choctaw, with words like [a-bi] 'for you to kill', Malay does not allow V.CVC words to surface, e.g. *[ə.pam], *[ə.bom] and *[ə.cat]. The underlying representations proposed by Teoh (1994) and Ahmad (2000b) are thus unable to resolve the epenthetic schwa in monosyllabic roots.

In order to deal with this problem, I pose an alternative analysis that is couched in GTT terms. In this theory, as already mentioned in subection 2.2.4, the disyllabicity minimality requirement falls under: (1) the Prosodic Hierarchy-Based Template – PBT; and (2) the Morpheme-Based Template – MBT. In PBT, all words are subject to the same minimality condition (Downing, 2006: 127). Therefore, as we saw in Choctaw, /a/ is epenthesized in a monosyllabic root to satisfy the minimality constraint called BINARITY, but not if a prefix is attached. This idea of all words being subject to the same minimality condition leads however to the problem of accounting for canonical morpheme shape (ibid.).

In contrast to PBT, the idea of MBT is that not all words in a language have the same minimality condition (Downing, 2006: 100). For example, in some languages, different categories of words are subject to different minimality conditions. As Downing (2006) found in Arabic and Modern Hebrew, the disyllabicity requirement is only applicable to derived verb stems, while nouns can be monosyllabic (ibid.: 101).

It is instructive also to reflect on Féry (1991) who discusses the requirements of derived word disyllabicity in German infinitives. Féry claims that German infinitives are subject to a disyllabic minimality condition. As we can see in the following examples, the syllabic pronunciation of the infinitive suffix /-n/, in 109(a), is required in monosyllabic bases. In contrast, the infinitive suffix /-n/ is not required in bases with more than one syllable, as in 109(b).

109. Disyllabic minimum in German infinitives (Ussishkin, 2000: 38-9, cited in Downing, 2006).

(a)	sehen	[ze:.n] 'to see'
	bauen	[bau:.n] 'to build'
	fliehen	[fli:.n] 'to flee'
	woollen	[vɔln] 'to want'
(b)	fordern	[fɔrdərn] 'to demand'
	segeln	[ze:gəln] 'to sell'

There is further supportive evidence for this idea. Uhrbach's (1987) analysis of monosyllabic roots in Javanese shows that the output for nasal affixation must be minimally disyllabic, while monomorphemic roots can be monosyllabic (cited in Downing, 2006: 135). In other words, derived words in the language must contain at least two syllables, while underived words can be monosyllabic, as the following examples illustrate:

110. Nasal fusion in Javanese (Uhrbach, 1987: 233, cited in Downing, 2006: 102).

(a) Disyllabic roots

cukur	ɲukur 'shave someone'
bali	mbaleni 'return something'
tulis	nulis 'to write'
dudut	ndudut 'pull/interesting'
sapu	ɲapu 'broom/to sweep

(b) Monosyllabic roots

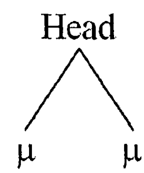
cet	ɲəcet 'to print'
bom	ɲəbom 'to bomb'
dol	ɲədol 'to sell'
tik	ɲətik 'to type/typewrite'
bis	ɲəbis '(to ride the) bus'

Hence, based on the idea of MBT, that not all words in a language have the same minimality condition, I offer a new analysis in order to account for monosyllabic roots in Malay. The disyllabicity minimality condition in Malay nasal substitution can be accounted for in a similar way as nasal fusion in Javanese. In MBT, there is a clear distinction in the minimality requirements between derived and underived words. As stated in Downing (2006: 128), underived words (monomorphemic) are optimally branching monosyllabic. For derived words, which consist of prefix and root (bimorphemic), they are minimally disyllabic by the MORPHEME-SYLLABLE CORRELATION seen in (34).

I will explain why monosyllabic roots in underived words are acceptable when surfacing as monosyllables. This situation can be formally explained by the principle of binary minimality branching words. The branching principle is formalised by the markedness constraint HEAD BRANCH, in (31), and branching constraint, PROSODIC BRANCHING in (32) (Downing, 2006: 122). As required by HEAD BRANCH (31), all lexical heads, which are roots, must branch prosodically. Roots are monomorphemic heads, while affixes are monomorphemic non-heads. Thus, they are not required to branch. Both of them (prefix and root) are predicted to be monosyllables as stated in MORPHEME-SYLLABLE CORRELATION (34). There are three representations where Heads can be performed, as shown in (33). Thus, I claim that Heads for Malay monosyllabic roots fall out from the representation in

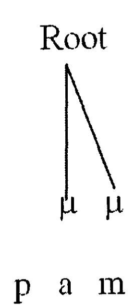
33(b) and are repeated in (111), below. This Head Branch can account straightforwardly for why underived words of monosyllabic roots in Malay do not violate the minimality word condition, as the Heads (roots) contain more than one daughter – two moras as daughters.

111. Representation of Head (root) for Malay.



From this representation of Heads, the construction for monosyllabic roots in Malay is that monomorphemic Heads satisfy PROSODIC BRANCHING (32) where a Head dominates two moras as its daughters.[25] This morphologically-based approach to word minimality can account straightforwardly for the problem of why words in Malay can be monosyllabic. Therefore, schwa does not need to be inserted in monosyllabic roots as the Heads (roots) already contain two moras, as the following diagram illustrates:

112. Representation of root /pam/.



[25] Interested readers can consult Dresher and van der Hulst (1998: 320) and Ussishkin (2000) for more details on Heads.

Given the insight from MBT, I have provided explanations of why monosyllabic roots in Malay surface as monosyllables whereas the language requires all words to be at least disyllabic. Thus, the templates CV.CVC and V.CVC, as proposed by Teoh (1994) and Ahmad (2000b), are deficient in explaining the disyllabicity minimality condition for monosyllabic roots.

There is another explanation that can be given for the case in question. It is also predicted that the disyllabic minimality requirement cannot be satisfied for monosyllabic roots in Malay because they are not Malay words. Teoh (1994: 102) claims that monosyllabic roots in Malay are borrowed words. This can be clearly seen in the statement that: *'I can off-hand think of only several monosyllabic words, and borrowing which manifest with /məŋ+/'*. As noted in Maris (1980: 10), '*there are very few simple words (stems) in Malay, probably about twenty in all, that are monosyllabic*'. This is consistent with what I found in the corpus, monosyllabic roots that have combined with a single prefix are not as numerous as disyllabic roots. Monosyllabic roots that have combined with a single prefix make up only 395 of the words found in the corpus (see Figure 4 in subsection 4.2.1.1).

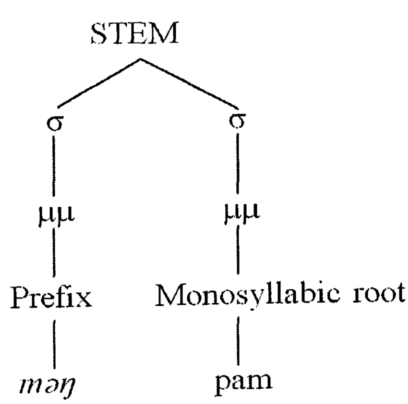
113. Monosyllabic roots are Malay borrowed words (from Teoh, 1994: 102).

cat	'to paint'	from Chinese
bom	'to bomb'	from English
pam	'to pump'	from English
had	'to limit'	from Arabic
sah	'to validate'	from Arabic'
syak	'to suspect'	from Arabic'

As stated in MORPHEME-SYLLABLE CORRELATION (34), Stems (referring to prefix and root) must minimally contain two syllables, one for each morpheme. Based on this, I will revise the claim made by previous scholars (e.g.: Teoh, 1994; Onn, 1980; Maris, 1980; Ahmad, 2000b) that ‘words’ in Malay must at least be minimally disyllabic. The word ‘word’, as used by those scholars, is unclear because it could refer to either root or stem. Contrary to this claim, I posit a new one, which is I think more accurate to account for Malay prefixed words. I posit that ‘stem’[26] in Malay must minimally contain two syllables. The disyllabic-minimality requirement for Stems is formalised as Prosodic Stem Minimality, as shown in (35). The constraint is a corollary of MORPHEME-SYLLABLE CORRELATION (34) (Downing, 2006).

From the data shown in (104), it is clear that nasal substitution in monosyllabic bases consist of bimorphemic stems (i.e. one morpheme for the prefix and another for the monosyllabic root). Thus, the construction of Prosodic Stem Minimality in Malay, for monosyllabic roots, can be illustrated as below:

114. Prosodic Stem for Malay Prefixation.



---

[26] I am following previous works such as Matthews (1991), Spencer (1991), Urbanczyk (1996), Stonhom (2004) and Downing (2006), where Stems are defined as polymorphemic Root-Affix complexes.

In order to account for Malay nasal substitution, a PROSODIC STEM will be high-ranked, as this constraint ensures stems (prefix and root) are minimally disyllabic. From the Prosodic Stem structure for Malay prefixation in (114), it is clearly seen that both prefixes and roots are minimally monomorphemic, as required by MORPHEME-SYLLABLE CORRELATION (34). As mentioned above, roots as monomorphemic Heads are predicted to be minimally monosyllabic and to contain two moras as daughters. It could be more – disyllabic or polysyllabic roots – as we will see in the following section.

Now we return to the question of what the motivation for schwa epenthesis is in monosyllabic bases, as raised in subsection 4.2.1.1. As we have just seen, schwa epenthesis is applied to monosyllabic bases; it was analysed by previous scholars (e.g.: Teoh, 1994; Onn, 1980; Maris, 1980; Ahmad, 2000b) and is a requirement of the word minimality condition in Malay. In contrast to those analyses, this study claims that schwa is epenthesized in order to satisfy the markedness constraint *NC̥, which requires that no nasal and voiceless obstruent clusters emerge in the surface form. By maintaining the analyses (i.e.: Teoh, 1994; Ahmad, 2000b) that the empty slot(s) are unable to resolve epenthetic schwa in a process for monosyllabic roots, I offer other possible solutions to resolve this problem. To give an OT account, this situation is captured by the constraint *NC̥ that has already been defined in (68).

As was mentioned in Chapter 2, subsection 2.4.1.1, nasal-voiceless obstruent clusters are disfavoured in a wide variety of languages (Pater, 1999). But why are voiceless obstruents following nasals disfavoured, while voiced obstruents are not? This is something that needs to be discussed by taking the phonetic aspect of voiceless and voiced obstruents into account. A nasal consonant is produced by lowering the velum in the mouth, allowing air to exit freely through the nose. The change from a nasal consonant to an obstruent causes the velum to be raised and this blocks the airflow from passing through the nose (Kager, 1999: 61). However, the process of raising the velum takes some time and is not complete at the time when the obstruent begins. At this point, there is a little air flowing out through the nose because the velum is not raised high enough. This is called 'nasal leak' (cf. Kager, 1999).

In Malay, therefore, nasal substitution is regularly applied when the input contains a nasal and voiceless obstruent to break up the cluster, i.e. nasal substitution occurs. The voiceless obstruent is deleted and leaves its place of articulation to the preceding nasal. It is one of the strategies for languages to get rid of the cluster, though there are other possible solutions, for example post-nasal voicing, nasal deletion and denasalisation (ibid.). In Correspondence Theory (McCarthy and Prince, 1995), the phonological process of nasal substitution is explained as a process of merging, between the nasal segment in the prefix and the voiceless obstruent, that can be interpreted as a two-to-one mapping from input to output. In other words, two segments in the input correspond to a single segment in the output. The correspondence relationship between the input and output of a sequence of /ŋ+p/ is illustrated in (115). As can be seen, the voiceless obstruent [m] in the output is obtained from two segments in the input, nasal segment /ŋ/ and place of articulation /p/.

98. Nasal substitution representation /məŋ$_1$+p$_2$ukul/ → [məm$_{12}$ukol]

Input	ŋ$_1$	p$_2$
Output	m$_{12}$	

For monosyllabic bases, nasal substitution does not however apply to break up the cluster, as presented in (104). The cluster is resolved by a strategy other than nasal substitution. This alternative strategy is epenthesis. Due to the *NC̥ constraint, I claim that the epenthetic schwa in monosyllabic bases is to satisfy the markedness constraint requirement. We have now seen that schwa epenthesis is also the strategy applied in the language to satisfy *NC̥. In OT, the markedness constraint can be satisfied in diverse ways, depending on the context (McCarthy, 2002: 95). In order to satisfy a particular markedness constraint, we have to rank the relevant markedness constraint above the faithfulness constraint in the constraint hierarchy: Markedness >> Faithfulness. As demonstrated in (115), nasal substitution is triggered by a conspiracy of the markedness constraint *NC̥. Now I am going to go into greater depth of how schwa epenthesis is triggered by a conspiracy to satisfy the same markedness constraint.

The process of schwa epenthesis has resulted in one segment being added to the prefix and the monosyllabic root. In OT, an additional segment, which is not part of the input, is analyzed under Correspondence Theory (McCarthy and Prince, 1995) (see Section 2.4.2). The general claim of this theory is that faithfulness to the input-output is a kind of requirement, whereby this pair of representations must be identical (Kager, 1999: 24). One example of the identity requirement, that is expressed by the faithfulness constraint, is DEPENDENCE. In this case, schwa, which is not found in the input, has been added to the output and is related to input-output faithfulness. Hence, the faithfulness constraint between input-output is DEP-IO, as defined in (116).

116. **DEP-IO**

Every segment in the input must have a correspondent in the output.

The violation of DEP-IO in the optimal output [məŋ$_1$əp$_2$am] is illustrated in the correspondence diagram below:

117. Correspondence diagram for epenthetic schwa

Input:	ŋ$_1$		p$_2$
	|		|
Output:	ŋ$_1$	ə	p$_2$

The discussion of monosyllabic bases so far has considered three constraints. These are Prosodic Stem, *NÇ̥ and DEP-IO. Because the optimal output violates DEP-IO by epenthesizing schwa, this constraint has to be ranked beneath *NÇ̥. Now, I shall demonstrate these constraints in the following tableau (where syllable boundaries are marked by a full stop ".").

118.

/məŋ$_1$əp$_2$am/	Prosodic Stem	*NÇ̥	DEP-IO
a. məm$_1$.p$_2$am		*!	
b. ☞mə.ŋ$_1$ə.p$_2$am			*

The presence of nasal-voiceless obstruent [mp] in candidate (a) has resulted in the violation of *NÇ̥. Although candidate (a) satisfies the faithfulness constraint, DEP-IO, by not epenthesizing schwa, it fatally violates the higher constraint *NÇ̥. In order to satisfy the *NÇ̥ constraint, schwa is epenthesized between the cluster, as in candidate (b). Since the cluster is separated by a schwa, candidate (b) satisfies *NÇ̥.

This is why I claimed above that schwa epenthesis is a conspiracy applied in the language to satisfy the markedness constraint, *NC̥.

Considering all the constraints in (118), another potential candidate that must be taken into consideration is *[mə.m$_{12}$am]. This candidate seems predictable if nasal substitution is applied. In OT, if a candidate satisfies a particular constraint, it may violate other constraints in the hierarchy. In this case, the new candidate *[mə.m$_{12}$am] satisfies *NC̥, as the cluster is resolved by nasal substitution. Nevertheless, it violates a constraint which bans nasal substitution, namely UNIFORMITY, that requires that no element of the output has multiple correspondents in the input, as has already been defined in (82).

The following diagram (119) shows how this constraint is violated. In the input, there are two segments, /ŋ+p/. In the output, these two segments become a single segment, [m], which has multiple correspondent segments when nasal substitution occurs. UNIFORMITY is a constraint which bans nasal substitution and is known as an anti-nasal substitution constraint. As we will see in this study, UNIFORMITY is violated by all candidates with nasal substitution. In order for a nasal substitution candidate to be chosen as the optimal output, this faithfulness constraint must be ranked beneath the markedness constraint, *NC̥. We see this ranking *NC̥ >> UNIFORMITY in the analysis tableau in (105).

119. Correspondence diagram for UNIFORMITY violation.

Input ŋ$_1$ p$_2$

Output m$_{12}$

Besides the phonological requirement mentioned above, Malay requires a nasal segment to occupy the coda position of a syllable; it must be homorganic to the following consonant. A nasal segment in the coda position assimilates to the place of articulation of the following onset consonant, e.g. /məŋ$_1$+t$_2$arik/ VERBL-pull → [men$_{12}$areʔ] 'to pull' and /məŋ$_1$+b$_2$aʧa/ VERBL-read → [məm$_1$b$_2$aʧə] 'to read'. I illustrate this in the following diagram of nasal assimilation:

120. Diagram of nasal assimilation.

/məŋ + baʧa/ [m ə m b a ʧ a]

{labial}

Due to this phonological requirement, another potential candidate that needs to be considered is *[məŋ$_1$p$_2$am]. The nasal segment [ŋ] in the prefix and the initial voiceless obstruent [p] in the surface representation give rise to a problem because the nasal segment [ŋ] is not homorganic to the initial consonant of the root, [p]. Note that in Indonesian, which also has the same requirement as Malay, the unassimilated nasal segment in the following consonant is resolved by a NASAL ASSIMILATION (henceforth NAS ASS) constraint (Pater, 2001), as defined in (121). This constraint will thus also be used for Malay in order to rule out any candidate without a homorganic nasal, such as *[məŋ$_1$p$_2$am].

121. **NAS ASS** (cf.: Jun, 1995; Padgett, 1995; Boersma, 1998; Pater, 2001).

A nasal must share place features with a following consonant.

Given the fact that NAS ASS is higher ranked in the hierarchy, it is plausible for nasal segment /ŋ/ in the prefix becomes [m] as in *[məm$_1$p$_2$am]. This possibility however can not be optimal since *NÇ̥ is ranked right after NAS ASS which causes *[məm$_1$p$_2$am] to be ruled out. In this situation, NAS ASS and *NÇ̥ do not conflict, and therefore they are not ranked with respect to each other. The interaction between the five constraints – Prosodic Stem, *NÇ̥, DEP-IO, UNIFORMITY and NAS ASS – is controlled by the following ranking: Prosodic Stem >> NAS ASS, *NÇ̥ >> UNIFORMITY >> DEP-IO. Bringing together all these constraints and all the candidates introduced thus far, I establish the following tableau:

122.

/məŋ$_1$+p$_2$am/	PrStem	NAS ASS	*NÇ̥	UNI	DEP-IO
a. məm$_{12}$am				*!	
b. məm$_1$p$_2$am			*!		
c. məŋ$_1$p$_2$am		*!			
d.☞mə.ŋ$_1$ə.p$_2$am					*

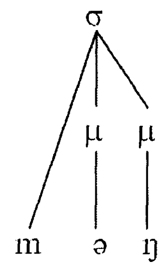
We have just seen how the nasal final prefix /məŋ+/, for example, behaves when it is attached to monosyllabic roots. Now, we examine how other prefixes, /di+/, /tər+/ and /bər+/, behave when they are combined with monosyllabic roots. We will then see, in the discussion of /di+/, /tər+/ and /bər+/, that another conspiracy applies in the process of single prefixation in Malay. These prefixes undergo vowel lengthening in order to satisfy the canonical shape of an affix. We note from the above discussion of monosyllabic roots that epenthetic schwa is a conspiracy that applies to monosyllabic bases to satisfy the markedness constraint, *NÇ̥. In prefixes /di+/, /tər+/ and /bər+/ on the other hand, vowel lengthening is applied in order to satisfy the canonical shape of an affix – monosyllabic bimoraic. I will first analyse the passive prefix /di+/. This is then followed by the other two prefixes, /tər+/ and /bər+/.

If we turn to the work done by Ahmad (2000b) in 108(a), he claims that the empty V slot in the template V.CVC, for monosyllabic roots, is mainly to ensure that schwa can be epenthesized in the derivation. By maintaining that the template can also be applied to prefix /di+/, he draws out a new output from the analysis. In order to fill the empty V slot, the final vowel in the prefix /di+/ undergoes a process of vowel lengthening. In the non-linear derivation, /i/ in the final prefix is associated with the empty V slot, hence it creates vowel lengthening in the passive prefix /di+/. As I mentioned in Chapter 2, I agree with Ahmad's (2000b) analysis. Thus, this analysis of vowel lengthening will be used in this study to account for the passive prefix /di+/.

In this study, however, the vowel lengthening which occurs in the prefix /di+/ is not due to an empty slot. The theoretical approach being applied here has nothing to do with a templatic constraint or segmental structures like skeletal C, V and X slots (e.g.: Marantz, 1982; Steriade, 1982; Archangeli, 1983; Clements and Keyser, 1983: Levin, 1983, 1985). In this study, GTT, developed within OT, proposes that the size requirement of prosodic morphemes should fall out from their morphological category (e.g. affixes, stems and roots), rather than constructing a specific size such as RED=σ (cf. Downing, 2006: 24). The size of the prosodic morpheme then follows from the canonical shape holding that morphological category (ibid.). This is how the theory avoids construction-specific constraints of size restrictions for prosodic morphemes.

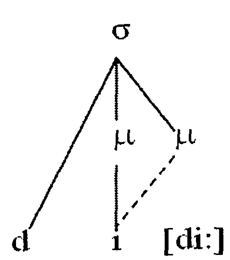
Based on the central idea of MBT, that the size of prosodic morphemes is assigned from their morphological category (e.g. affixes, roots or stems), I postulate that the canonical shape of affixes in Malay is monosyllabic bimoraic. In the following diagram, I illustrate the canonical shape for Malay prefixes – monosyllabic bimoraic prefixes.[27]

123. Canonical shape for Malay prefixes: /məŋ+/.



The above illustration is the prosodic structure of Malay prefixes which end with consonants such as /məŋ+/, /pəŋ+/, /bər+/ and /tər+/. The prosodic structure contains one syllable and is bimoraic, where each of the moras is in association with a schwa (nucleus) and a velar nasal /ŋ/ (coda). In the passive prefix /di+/, however, it does not contain a coda consonant as it ends with a vowel. This causes the second mora to be empty. In this case, how can the emptiness of the second mora be filled? By using OT, this question can be answered quite simply. I illustrate below how passive prefix /di+/, with no coda consonant, meets the canonical shape of Malay prefixes – monosyllabic bimoraic.

124. Canonical shape of Malay prefix /di+/ → [di:].

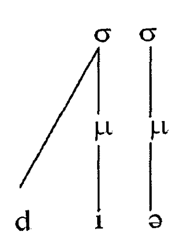


[27] This canonical shape for Malay prefixes may vary from other prefixes in other languages, similarly the size of roots. For example, in English and Ancient Greek, as claimed by Golston (1991), roots are minimally bimoraic, while affixes in the languages are minimally monomoraic.

As illustrated above, since the prefix does not contain a coda consonant, the second mora is filled by the preceding vowel, /i/. In Moraic Theory, a heavy syllable is said to have two moras, while a light syllable only has one (Durand and Katamba, 1995: 125). Besides having two moras, a long vowel (VV) is also counted as a heavy syllable (ibid.). Therefore, the above illustration is also considered a heavy syllable. Thus the canonical shape for Malay prefixes – monosyllabic bimoraic – contains CVC (123) or CVV (124) segments.

Other than the above solution proposed in (124), one might suggest other possible ways to satisfy the canonical shape of an affix. There are two possible ways that this prefix might be analysed. First, if schwa is epenthesized in prefix /di+/ to fill the second mora, then it creates two syllables. The MORPHEME-SYLLABLE CORRELATION (34) however demands that each morpheme (prefix and root) in stems consist of only one syllable. By inserting a schwa in the prefix /di+/, it will exceed the size requirement of affixes. Moreover, /di+/ is a single prefix and therefore it cannot have more than one syllable. The diagram of dissatisfaction for MORPHEME-SYLLABLE CORRELATION is illustrated below:

125. Violation of MORPHEME-SYLLABLE CORRELATION.



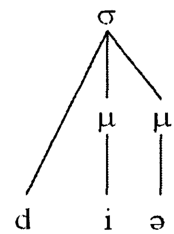
*[di.ə]

One might argue about the structure for the passive prefix /di+/ in (125). As I have said, schwa cannot be epenthesized to prefix /di+/ since the presence of schwa creates another syllable on the prefix. Therefore *[di.ə] violates the MORPHEME-

SYLLABLE CORRELATION (34). Moreover, as I claimed earlier in this chapter, the schwa epenthesis is essential to getting rid of nasal and voiceless obstruent sequences in monosyllabic bases. Therefore, schwa is not epenthesized in passive prefix /di+/ as there is no sequence of nasal and voiceless obstruents when /di+/ in (125) is attached to a monosyllabic root. Thus, this possible analysis of prefix /di+/ is not the best way to satisfy the canonical shape of an affix.

Second, another possible way to analyse the prefix /di+/ is to syllabify the schwa into the same syllable with /di+/ as *[die], as illustrated in (126). This might possibly be the optimal form for the prefix /di+/, by parsing the schwa tautosyllabically with the initial syllable. As has been claimed, not all vowel sequences in Malay can be syllabified tautosyllabically (Ahmad, 2005: 29). A sequence of vowels is claimed to be a diphthong if the second vowel is high, e.g. /pisau/ 'knife', /kurau/ 'a kind of fish', /gurau/ 'to joke' and /pantai/ 'beach' become [pi.saw], [ku.raw], [gu.raw] and [pan.taj], respectively. Thus, we cannot consider the output *[die] as a diphthong. It must be mentioned that Malay only has three diphthongs: [ai], [au] and [oi] (Ahmad, 1964; Hassan, 1974; Omar, 1975; Maris, 1980; Onn, 1980; Karim et al., 1989; Teoh, 1994; Ahmad, 2005), and *[iə] is not one of them. In short, parsing the schwa into the same syllable with [di] cannot be accepted, even though it satisfies MORPHEME-SYLLABLE CORRELATION (34). Thus, this analysis of prefix /di+/ cannot be used as a way to satisfy the canonical shape of an affix. The following diagram gives a better picture of this analysis:

126. Tautosyllabic schwa in the prefix *[diə].



*[diə]

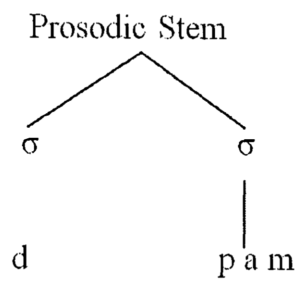
With the same constraints ranking that I have discussed so far, I establish a new tableau to account for the prefix /di+/.

127.

/di+pam/	PrStem	NAS ASS	*NC̥	UNI	DEP-IO
a. ☞di: .pam					*
b. di.pam	*!				
c. dpam	*!				

The failed candidates (b) and (c) fatally violate PROSODIC STEM (35). The prefix in candidate (c) does not have its own syllable, as illustrated in (128), below. Although candidate (b) contains two syllables, and therefore satisfies PROSODIC STEM (35), the prosodic structure of the prefix does not contain a monosyllabic bimoraic, as illustrated in (124). By contrast, the prosodic structure for Malay prefixes is satisfied in candidate (a) by lengthening the vowel in the prefix. Therefore it is chosen as the optimal candidate.

128. Violation of Prosodic Stem Minimality.



As illustrated in the above diagram, the PROSODIC STEM of candidate (c) *[dpam] only contains one syllable, which is from the root. The prefix does not contain any syllables, as there is no nucleus in the prefix to form a syllable. Thus, /d/ cannot be associated with its syllable. In this case, in order for /d/ to get its syllable, it must be syllabified into the same syllable with the root, /pam/. If this was applied, then the prosodic stem would contain only one syllable, which violates PROSODIC STEM (35).

Now we see how /bər+/ and /tər+/ satisfy the canonical shape of an affix – monosyllabic bimoraic. These two prefixes do not end with nasals. Therefore, we do not need a schwa here. These prefixes will be analysed here in the same way as I analysed the prefix /di+/. A question that might be posed here is how the lengthening process can be applied to these prefixes, since they end with a consonant that can be linked to the second mora of the canonical shape. The explanation of this is as follows.

It has long been observed that /r/ in Malay is never realised word finally (Maris, 1980; Omar, 1975; Onn, 1980, cited in Ahmad, 2005: 58). Some of the relevant examples are shown in (129). In a rule-based analysis, segment /r/ in stem final position is obligatorily deleted (Onn, 1980), while in non-linear analysis, /r/ is an optional delinking rule (Teoh, 1994: 43). Onn and Teoh's analysis can be explained in the following rules, i.e. in (130) and (131), respectively:

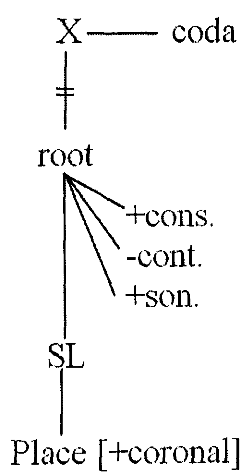
129. /r/ in stem finally (Ahmad, 2005: 59).

/kotor/ 'dirty'	[koto:]
/pasar/ 'market'	[pasa:]
/uker/ 'to carve'	[uke:]

130. /r/ deletion in rule-based analysis (Onn, 1980: 16).

$$r \rightarrow \emptyset \, / \, \_\_\_ \, \{\#, C\}$$

131. /r/ delinking rule (Teoh, 1994: 43).



The rules postulated by Onn (1980) and Teoh (1994) above have the same structural description where /r/segment in the coda position of a syllable is banned. Given this phonological rule, I propose that the final /r/ in the prefixes /bər+/ and /tər+/ must be deleted. Due to the /r/-deletion, these prefixes do not match the canonical shape of an affix as I proposed earlier. Hence, vowel lengthening is the best way to match the prefixes /bər+/ and /tər+/ to the second mora. In Malay, the CODA

CONDITION constraint plays a crucial role in banning /r/ in the coda position of a syllable. Adopting Itô and Mester's (1994) alignment formalism, the ALIGN RHOTIC (r, σ) (132) constraint, that requires the segment /r/ to be left-aligned with a syllable, is needed in our analysis. In order to satisfy this constraint, the vowel in the prefix will be lengthened.

132. **ALIGN-RHOTIC (r, σ)** (Itô and Mester, 1994).

Segment /r/ must be left-aligned with a syllable.

The tableau below gives a better picture of how ALIGN RHOTIC (r, σ) functions, and how the candidate with vowel lengthening surfaces as the optimal output. By satisfying this constraint, the candidate with vowel lengthening violates one faithfulness constraint, DEP-IO.

133.

/tər+pam/	PrStem	NAS ASS	*NÇ̥	UNI	ALIGN RHOTIC	DEP-IO
a. tər.pam					*!	
b.☞tə:.pam						*
c. tə.r.ə.pam	*!					*
/bər+cap/						
d. bər.cap					*!	
e.☞bə:.cap						*
f. bə.r.ə.cap	*!					*

Candidates (c) and (f) violate the highest constraint in the hierarchy, Prosodic Stem, where the prefixes do not consist of two moras. The epenthesis of schwa between the prefixes and roots cause the candidate to violate DEP-IO, since schwa has no correspondent in the input. As I claimed earlier in this chapter, the presence of schwa is only to avoid nasal and voiceless obstruent sequences. Since the prefixes /tər+/ and /bər+/ do not end with nasals, epenthetic schwa is not relevant. As ALIGN-RHOTIC (r, σ) is ranked higher than DEP-IO, candidates (b) and (e) are chosen as the optimal output.

4.2.1.2 Disyllabic and Polysyllabic Roots

In this subsection, I examine how disyllabic / polysyllabic roots surface when prefixes are attached. We will see later in this subsection that the other two groups in the Malay lexicon (see Figure 3 in Section 4.1) that I identified earlier, i.e. undeleted voiceless plosives in loanwords and native groups, are derived from these types of roots. Observation from the corpus shows that there are 461 out of 1,141,816 words where voiceless plosives following nasal segments that are not deleted (see Appendix A).

To begin with, I lay out some relevant examples illustrating how nasal and voiceless obstruent clusters surface in one of the two categories of the Malay lexicon: undeleted voiceless plosives in loanwords.

134. Group of undeleted voiceless plosives in the loanwords sub-lexicon (from the DBP-UKM corpus).

(a) Consonant clusters initial root

i) /məŋ-promosi/ [məm-promosi]
ACT.PRF-promotion 'to promote'

ii) /məŋ-kritik/ [məŋ-kritik]
ACT.PRF-critic 'to criticise'

iii) /məŋ-protes/ [məm-protes]
ACT.PRF-protest 'to protest'

iv) /məŋ-proses/ [məm-proses]
ACT.PRF-process 'to process'

v) /məŋ-transformasi/ [mən-transformasi]
ACT.PRF-transformation 'to transform'

vi) /pəŋ-struktur-an/ [pən-struktu-ran]
NOM.PRF-structure-NOM.SUF 'structure'

(b) Borrowed phonemes

i) /məŋ-ʃarat-kan/ [mən-ʃarat-kan]
ACT.PRF-condition-CAUS.SUF 'to cause to condition for'

ii) /məŋ-fokus/ [məm-fokus]
ACT.PRF-focus 'to focus'

iii) /məŋ-tadbir/ [mən-tadbir]
ACT.PRF-administrative 'administer'

iv) /məŋ-popular-kan/ [məm-popular-kan]
ACT.PRF-popular-CAUS.SUF 'to cause to be popular for'

v) /məŋ-fasakh/ [məm-fasakh]
ACT.PRF-divorce 'to annul a marriage'

vi) /məŋ-ʃukur-i/ [mən-ʃukur-i]
ACT.PRF-gratitude-LOC.SUF 'to cause to be grateful'

vii) /məŋ-xatan/ [məŋ-xatan]
ACT.PRF-circumcision 'to circumcise'

viii) /məŋ-tauhid/ [mən-tauhid]
ACT.PRF-monotheism 'oneness and unity of god'

ix) /məŋ-komersil/ [məŋ-komersil]
ACT.PRF-commercial 'to commercialise'

The descriptive generalisations that are observed in the above examples can be summarised as follows: (1) the voiceless obstruents remain undeleted even though nasal segments precede them; (2) the nasal segments in the prefixes are homorganic to the following initial consonants of the roots. I will first discuss the data in 134(a) and then the data in 134(b).

As already noted in Section 1.2, Malay disfavours consonant clusters in its surface representation. As we see in 134(a), the roots contain consonant clusters. The roots do not however undergo any phonological process, i.e. vowel epenthesis, to break up the clusters. I shall argue that the data in 134(a) are borrowed words because of the presence of consonant clusters *[CCV.] in the roots. In order to support this, I shall mention what previous Malay scholars have said about Malay syllable structure. In earlier studies of Malay phonology, such as Hassan (1974), Maris (1980) and Onn (1980), they discussed the basic syllable structure of Malay as being (C)V(C). Ahmad (2005) with his OT work states that Malay loan phonology offers good evidence that *COMPLEX is highly respected in the language. He claims that any consonant clusters in Malay are resolved by schwa epenthesis. This can be seen in borrowed English words such as *class, stem, post, glass* and *club*. These words are realised as [kelas], [setem], [pos], [gelas] and [kelab], respectively, in Malay (Ahmad, 2005: 18).

In fact, the data in 134(b) are also borrowed words. As we can see, the data in 134(b) are slightly different from the data in 134(a) in terms of the segments in the words. As we have already observed, in the data in 134(a), the roots contain more than one segment in the onset position. In contrast to the data in 134(a), if we take a closer look at the data in 134(b), the initial consonants of the roots are originally borrowed consonants. As mentioned in Chapter 1 Section 1.2, Malay has 16 underlying consonants: /p, b, t, d, k, g, ʧ, ʤ, s, h, m, n, ɲ, ŋ, l and r/, and six vowels: /i, u, e, o, ə and a/ (Ahmad, 2005: 16). As far as the underlying consonants are concerned, it is apparent that every word in 134(b) consists of non-underlying Malay consonants. For example, consonants /f/, /ʃ/ and /x/, in 134b (ii), (vi) and (vii) respectively, are not underlying Malay consonants. Thus they are all borrowed words. Observe that there are some examples which do not contain any borrowed consonant, as in 134(b) (iii), (iv), (viii) and (ix). These words are originally non-native words, see the examples in (135). Therefore, nasal substitution fails to apply.

135.	[məm-fokus]	from English
	[mən-ʃukuri]	from Arabic
	[məŋ-xatan]	from Arabic
	[mən-tadbir]	from Arabic
	[məm-popular-kan]	from English
	[mən-tauhid]	from Arabic
	[məŋ-komersil]	from English

The data in (134) present a different phonological pattern of non-native words in the language than the one discussed in subsection 4.2.1.1 – monosyllabic roots. The hierarchical ranking for this type of data in (134) is therefore different from the sublexicon of monosyllabic foreign. The new constraint ranking to account for the sublexicon of undeleted voiceless plossive is: PrStem >> NAS ASS >> DEP-IO >> UNIFORMITY>> *NÇ̥ >> ALIGN-RHOTIC, as demonstrated in the following tableau:

136. Constraint ranking for the group of undeleted voiceless plosive in loanwords.

/məŋ$_1$+p$_2$roses/	PrStem	NAS ASS	DEP-IO	UNI	*NC̥	ALIGN RHOTIC
a. məm$_{12}$ro.ses				*!		
b.☞məm$_1$p$_2$roses					*	
c. məŋ$_1$p$_2$roses		*!				
d. məŋ$_1$əp$_2$roses			*!			
/məŋ$_1$+t$_2$auhid/						
e. mən$_{12}$auhid				*!		
f.☞mən$_1$t$_2$auhid					*	
g. məŋ$_1$t$_2$auhid		*!				
h. məŋ$_1$ət$_2$auhid			*!			

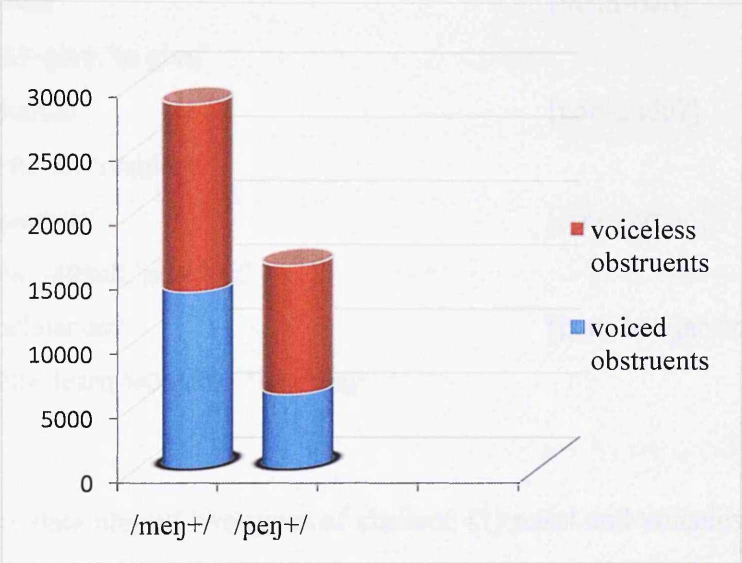
Candidates (c) and (g) are ruled out as they fatally violate the higher constraint NAS ASS, since the nasal segments in the prefixes are not homorganic to the initial consonants of the roots, [p] and [t], respectively. In the group of monosyllabic foreign (see subsection 4.2.1.1), vowel epenthesis is one of the strategies that can satisfy *NC̥. However this strategy cannot be satisfied in the group of 'undeleted voiceless plosive in loanwords', as DEP-IO is highly ranked in the constraint ranking. This shows that ə-epenthesis is not always a better option to break up nasal and voiceless obstruent clusters in the language. Therefore, candidates (d) and (h) cannot emerge as optimal output. In the above tableaux, as we can see, nasal substitution is a strategy to satisfy the constraint *NC̥, as candidates (a) and (e) show. However, these candidates fail to emerge as optimal output as the constraint which bans nasal substitution, UNIFORMITY, is ranked higher than *NC̥. Therefore candidates (b) and (f) are chosen as the optimal output.

Now, we come to examine the third category in the Malay lexicon, which is native. Here we will see whether the constraint ranking for the native Malay lexicon is the same as for the two groups we have discussed above. I am convinced that the constraint ranking of this group is different from those two groups as Malay does not permit nasal and voiceless obstruent clusters. This means the markedness constraint *NÇ̥ should be placed higher in the ranking compared to the constraint rankings of monosyllabic foreign (133) and undeleted voiceless plosives in loanwords (136).

As I shall argue in the next section, nasal and sonorant clusters are disallowed in the language as the two consonants share greater similarity in their features (see subsection 4.2.2). The similarity between the adjacent consonants leads to nasal deletion. Before we go to the point of similarity where consonants cause difficulties in an utterance, at this stage, I would like to argue that dissimilarity between consonants also causes the same problem. Nasals and voiceless obstruents share the same feature, i.e. [continuant]. Both consonants however have different features of [sonorant] and [voice]. The great dissimilarity in features between nasals and voiceless obstruents is the cause of why these consonants are not in harmony together.

As we will see in the following examples taken from the corpus, 24,657 voiceless obstruents following nasal segments undergo nasal substitution, where 14,623 words have been attached to prefix /məŋ+/ and 10,034 words to prefix /pəŋ+/. Meanwhile, 19,507 voiced obstruents after nasal segments undergo assimilation. From this last number, 13,731 and 5776 words are attached to prefixes /məŋ+/ and /pəŋ+/, respectively. This can be clearly seen in the following graph:

Figure 5: The number of voiced and voiceless obstruents following nasal segments in /məŋ+/ and /pəŋ+/.



Before I start the analysis, let us consider first relevant examples of the native group. More data from this group can be seen in Appendix D, attached to the end of this thesis.

137. Data for the native Malay sub-lexicon (from the DBP-UKM corpus).

(a) **Nasal and voiceless obstruent clusters**

(i) /məŋ-temu-i/ [mə-nemui]
ACT.PRF-meet-LOC.SUF 'to cause to meet'

(ii) /məŋ-potong/ [mə-motoŋ]
ACT.PRF-cut 'to cut'

(iii) /məŋ-kuat-kan/ [mə-ŋuwat-kan]
ACT.PRF-strong-CAUS.SUF 'to cause to strengthen for'

(iv) /pəŋ-pindah-an/ [pə-mindah-an]
NOM.PRF-migrate-NOM.SUF 'migration'

(v) /məŋ-kunjuŋ-i/ [mə-ŋundʒung-i]
ACT.PRF-visit-LOC.SUF 'to cause to visit'

(b)	**Nasal and voiced obstruent clusters**	
(i)	/məŋ-dapat/ ACT.PRF-get 'to get'	[mən-dapat]
(ii)	/məŋ-beri/ ACT.PRF-give 'to give'	[məm-bəri]
(iii)	/pəŋ-duduk/ NOM.PRF-sit 'resident'	[pən-duduʔ]
(iv)	/pəŋ-gəmpur/ NOM.PRF-attack 'attacker'	[pəŋ-gəmpo]
(v)	/pəŋ-belajar-an/ NOM.PRF-learn-NOM.SUF 'learning'	[pəm-belajar-an]

The above data present two types of clusters: (1) nasal and voiceless obstruent clusters and (2) nasal and voiced obstruent clusters. As the language disfavours nasal and voiceless obstruent clusters emerging in the surface representation, so the clusters have regularly been resolved by nasal substitution, as presented in the data in 137(a). In contrast, the voiced obstruents following nasals in 137(b) are preserved and do not have to undergo nasal substitution. The data in 137(b) are consistent with what Kager (1999: 61) says about the articulatory mechanisms of voiced obstruents, as mentioned in the previous subsection, on p. 141.

The data in (137) reveal that the case of nasal substitution in Malay only targets voiceless obstruents, i.e. not voiced obstruents. As has been discussed, *NÇ̥ is a constraint used to break up nasal and voiceless obstruent clusters from emerging in the surface representation. Nasal substitution therefore occurs. In contrast to nasal and voiced obstruent clusters, the clusters are permitted in the language therefore they do not have to undergo nasal substitution. A question which can be asked here is: What constraint in OT can be invoked to account for the blocking of voiced obstruent nasal substitution? Here I would like to propose a constraint which is able to prevent voiced obstruents from undergoing nasal substitution. As was mentioned in 2.4.2.4, the IDENT[PHAREXP] constraint (98), which specifically blocks voiced obstruents from undergoing nasal substitution, has been used to explain the blocking of voiced obstruent nasal substitution in Muna (Pater, 2001). Thus, in this study, I will make use of the constraint IDENT [PHAREXP] that served this purpose in the analysis of Muna, for the Malay data. The analysis of both voiceless and voiced obstruents follows.

In OT analysis, a candidate with nasal substitution can be obtained by ranking the constraint which bans the clusters, i.e. *NC̥, higher in the hierarchy. By satisfying *NC̥, a candidate with nasal substitution, for example [mən$_{12}$olaʔ] ‘to push’, violates a faithfulness constraint, i.e. UNIFORMITY. This is because both segments in the input /ŋ+t/ are not preserved as the input segments are not in a perfect one-to-one relation with the single output segment (Kager, 1999: 62).

I construct the following tableaux to account for the data in both 137(a) and (b). The tableaux demonstrate the satisfaction of *NC̥ and violation of UNIFORMITY by the optimal candidate, as just mentioned. It also shows how a voiced obstruent is blocked from undergoing nasal substitution by adding IDENT[PHAREXT] into the constraint ranking of the Malay native lexicon.

138. Constraint ranking of the Malay native lexicon.

PrSTEM >> NAS ASS >> *N$\c{C}$ >> IDENT [PHAREXP], ALIGN-RHOTIC >> DEP-IO >> UNIFORMITY

/məŋ$_1$+t$_2$olaʔ/	PrStem	NAS ASS	*N$\c{C}$	IDENT [PHAREXP]	ALIGN RHOTIC	DEP-IO	UNI
a.☞ mən$_{12}$olaʔ							*
b. mən$_1$t$_2$olaʔ			*!				
c. məŋ$_1$t$_2$olaʔ		*!					
d. məŋ$_1$ət$_2$olaʔ						*!	
/məŋ$_1$+d$_2$apat/							
e. mən$_{12}$apat				*!			
f.☞mən$_1$d$_2$apat							
g. məŋ$_1$d$_2$apat		*!					
h. məŋ$_1$əd$_2$apat						*!	

In the above tableaux, it is apparent that Malay disallows nasal and voiceless obstruent clusters in surface form, as *N$\c{C}$ is ranked higher in the hierarchy. Because of this, candidate (b) violates this constraint, since the cluster [nt] emerges in the surface. Candidate (c) violates NASS ASS because the nasal velar [ŋ] is not homorganic to the following obstruent [t]. Candidate (d) with schwa epenthesis violates DEP-IO as the segment does not have an input correspondent. Given the above constraint ranking, candidate (a) is the optimal output. It obeys *N$\c{C}$ as the cluster /nt/ is resolved by nasal substitution. This candidate however violates UNIFORMITY, as the two input segments are not preserved in the output. The violation of UNIFORMITY does not however play any significant role because the rest of the candidates have already been ruled out.

As we can see, IDENT [PHAREXP] can rule out candidate (e) with nasal substitution. Observe that *NC̥ does not play a role in this case. By adding IDENT[PHAREXP] to the constraint ranking, we now have an account of why nasal substitution is blocked between a nasal and a voiced obstruent cluster. However, it is crucial to note that voiced obstruents after nasals in Malay are not always preserved. Not surprisingly, these clusters also undergo nasal substitution, even though they are phonetically natural, as claimed by Kager (1999: 61). This situation occurs in some Malay dialects, for example Perak and NS. I shall discuss this later in this thesis, in the chapter on dialectal variation.

The tableau in (138) analyses disyllabic roots that combine with the nasal final prefix /məŋ+/. Now we see how /di+/ and /tər+/ behave when they are attached to disyllabic roots. The same hierarchical constraint ranking in (138), above, is applied to passive prefixes /di+/ and /tər+/ as well. Since there are no nasal segments in the prefixes, the NASS ASS and *NC̥ constraints do not play a crucial role. Let us now consider the tableau below:

139.

/di+bəlandʒa/	PrStem	NAS ASS	*NC̥	IDENT [PHAREXP]	ALIGN RHOTIC	DEP-IO	UNI
a. di. bəlandʒə	*!						
b. dbəlandʒə	*!						
c. ☞di: bəlandʒə						*	
/tər+tolaʔ/							
d. tər.tolaʔ					*!		
e. ☞tə:. tolaʔ						*	
f. tə.tolaʔ	*!						

We note from the above discussion that Malay has a co-existing distinct phonological system, or so-called co-phonology, in its single prefixation. This distinction could be captured adequately by a morphological construction with a distinct phonological process that has different constraint rankings. The same set of constraints is ranked differently according to a distinct phonological system where one allows nasal substitution, while the other one does not. The co-phonology developed within OT is completely different from earlier approaches such as rule-based analysis. In the latter, a phonological phenomenon is described as an input configuration and the operation that applies to it is rule ordering. If the data were to be analysed using rule-based analysis, optimal output with nasal and voiceless obstruent cluster would be impossible to obtain. In (140), I illustrate how rule-based analysis fails to handle undeleted voiceless obstruents after nasal segments in the process of single prefixation:

140.	Input	/məŋ-proses/
	1) Nasal assimilation	məm-proses
	2) Voiceless obstruent deletion	məm-roses
	Output	*[məm.roses]

4.2.2 Nasal and sonorant clusters

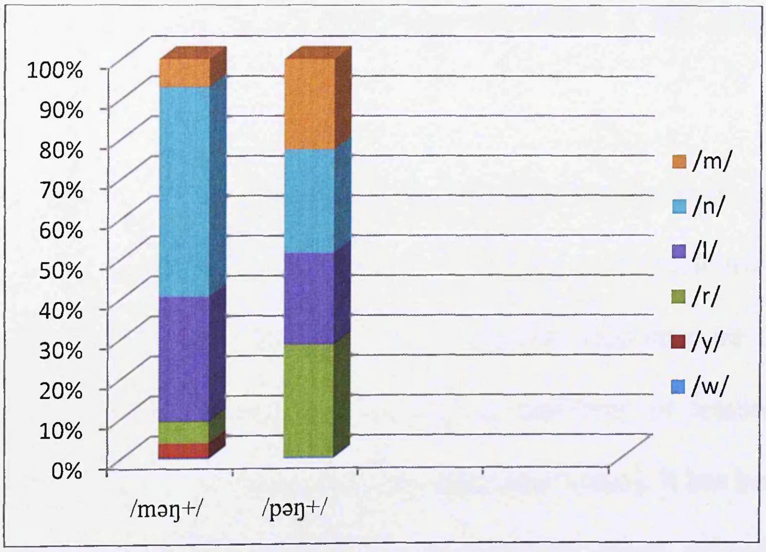
We have just seen, in the above discussion, that nasal and voiceless obstruent clusters are disfavoured in the language. I will also discuss here the fact that as well as simplifying NC̥ clusters, Malay also disfavours another type of cluster emerging in the surface representation, i.e. nasal and sonorant clusters. This is consistent with what I have observed in the corpus. Nasal segments followed by sonorant consonants are constantly deleted in the language. As was mentioned earlier in this thesis (see Chapter 1 subsection 1.3.1), nasal and sonorant clusters do not raise a particulary interesting issue when compared to nasal and voiceless obstruent clusters. There is not a single word in the data which shows the occurrence of this cluster in the surface representation.

In the previous section, we saw that two prefixes which end with nasal segments /məŋ+/ and /pəŋ+/ can be attached to obstruents, either voiced or voiceless. The same happens for sonorant initial roots. The two prefixes /məŋ+/ and /pəŋ+/ can also be attached to sonorant /m, n, l, r, y, w/ initial roots. I now provide the details of these two prefixes which have been attached to sonorant initial roots. The sonorant consonants which combine with the nasal final prefixes found in the corpus are as follows:

141.

	/məŋ+/	/pəŋ+/
/m/	1,306	4,769
/n/	9,711	5,435
/l/	5,785	4,832
/r/	1,003	5,904
/j/	681	1
/w/	43	73

Figure 6: The numbers of prefixes /məŋ+/ and /pəŋ+/ attached to sonorant initial roots (from the DBP-UKM corpus).



I shall now present the relevant data from the DBP-UKM corpus. More examples from this can be seen in Appendix E.

142. Data for nasal and sonorant clusters (from the DBP-UKM corpus).

a) /məŋ-lapor/ [mə-lapo]

ACT.PRF-report 'to report'

b) /məŋ-ɲuap/ [mə-ɲuwap]

ACT.PRF-yawn 'to yawn'

c) /məŋ-ɲaɲi/ [mə-ɲaɲi]
ACT.PRF-sing 'to sing'

d) /məŋ-raŋkul/ [mə-raŋkol]
ACT.PRF- embrace 'to embrace'

Observe that when the first initial consonant of the root is a sonorant, the preceding nasal is deleted. Thus nasal deletion is the way that Malay avoids nasal and sonorant clusters. In this case, nasal substitution is not applied to break up nasal and sonorant clusters. As Teoh (1994: 97) says, '*nasal deletion in Malay as the phonotactic of the language in general does not allow sonorant clusters even across a morpheme boundary*'. The statement made by Teoh is a description of the language. It does not explain the phonological motivation of why the clusters are disallowed, leading to nasal deletion. In the following discussion, I am going to offer an explanation of this.

Data such as those in (142) confirm that there are no words in the language containing a nasal followed by a sonorant. In Malay, a sonorant following a nasal segment is not allowed to surface as the two adjacent segments are not compatible phonetically. Nasals and sonorants share a great similarity of features where both segments are within the class feature of [sonorant] and [voice]. It has been observed in the literature on language production involving speech errors caused by similarity sound relations. Works such as Nooteboom (1967), Mackay (1970), Fromkin (1971) and Shattuck-Hufnagel and Klatt (1979), and many others, demonstrate that consonants sharing great similarity have a high potential to participate in a 'slip of the tongue' (cited in Rose and Walker, 2004). It has also been claimed in other work concerning language-production processing, done by Dell (1984, 1986), Stemberger (1985) and MacKay (1987), that the occurrence of similar but different consonants causes difficulties in an utterance that are mitigated by a shift towards identity (cited in Sharon and Walker, 2004). Because of the high similarity shared by consonants, misprononciation occurs in such phrases as '*subjects show*' which is pronounced as '*shubjects show*' (Shattuck-Hufnagel and Klatt, 1979). Therefore, near-identical sounds are preferred to identical forms. Thus, in Malay, two adjacent consonants, i.e. a nasal and a sonorant which have similarity in features, are neutralised by deleting the nasal before the sonorant to keep the similar segments distinct. There is another phonological reason why sonorant consonant that follows a nasal segment is banned. As stated in the Syllable Contact Law (Venneman 1998, Kaye 1990), an onset must be less sonorous than its preceding coda. As in Malay, the nasal coda in the prefix is less sonorous than /l/ or /r/ in the following onset. These two adjacent consonants therefore violate Syllable Contact Law. To overcome this, Malay has chosen to drop the nasal segment that precedes a sonorant consonant. I establish the following tableau with the same constraint ranking to account for nasal and sonorant clusters:

143.

/məŋ$_1$+l$_2$atih/	PrStem	NAS ASS	*NC̥	IDENT [PHAREXP]	ALIGN RHOTIC	DEP-IO	UNIFORMITY
a.☞ mə< >l$_2$ateh							
b. məŋ$_1$l$_2$ateh		*!					
c. məŋ$_1$əl$_2$ateh						*!	

It is clear from the above discussion on single prefixation that Malay does not allow two clusters: (1) nasal and voiceless obstruents and (2) nasal and sonorant clusters. To prevent nasal and voiceless obstruent clusters from appearing in the surface representation, vowel epenthesis, nasal asslimilation and nasal substitution are applied to monosyllabic foreign, undeleted voiceless plosives and native sublexical groups, respectively. For nasal and sonorant clusters, nasal deletion is applied instead.

The above discussion also shows that the UNIFORMITY constraint is fully obeyed in monosyllabic foreign and undeleted voiceless plosive groups compared to the native sub-lexical ones. In contrast, these two sub-lexicon groups disobey *NÇ̥, while the native sub-lexicon does not. These two differences in obedience and violation could be one piece of evidence that not all the data are Malay native words. The differences in obedience to the markedness constraint can be shown through differences in the constraint ranking, i.e. a co-phonology where the native sublexicon group has a different constraint ranking from non-native sublexical groups.

4.3 Multiple prefixation

We have just discussed the process of single prefixation in Malay. Now we turn our attention to multiple prefixation. It should be mentioned that, in prefixation, as many as three prefixes can be attached to a base to form multiple prefixation. As a source of evidence, the UKM-DBP corpus clearly proves that cases of three prefixes being combined to form multiple prefixation are very few. This is also mentioned in Onn (1980: 56), in that Malay affixes may take as many as four affixes, including a suffix. Here are some examples[28] (prefixes and suffixes are emboldened).

[28] Onn admits that such examples of multiple prefixes are very few in number.

144. Multiple prefixes (Onn, 1980: 56).

a) **bər-sə-kə**-tidur-**an**
REF-ADJ-NOM-sleep-NOM.SUF
'sleeping together without commitment'

b) **bər-kə-sə**-oraŋ-**an**
REF-NOM-ADJ-person-NOM.SUF
'a state of loneliness'

As this study is only concerned with nasal final prefixes, there is a maximum of two prefixes that can be attached to a root. Let us first consider the following examples in (145). I have classified the data into two groups according to the class derived from the multiple prefixes: (1) nominal and (2) verbal prefixes.

145. Multiple prefixes (from the DBP-UKM corpus).[29]

a) Nominal prefixes.

i) **pə-mər**-kaja-an
NOM.PRF-VERBL.PRF-rich-NOM.SUF
'enrichment'

ii) **pə-mər**-badan-an
NOM.PRF-VERBL.PRF-body-NOM.SUF
'organisation'

iii) **pə-məl**-bagai-an
NOM.PRF-VERBL.PRF-various-NOM.SUF
'variety'

[29] See Appendix B for the list of verbal and nominal prefix data.

b) Verbal prefixes

i) **məm.pər**.kuwat.kan

VERBL.PRF-NOM.PRF-strength-CAUS.SUF

'to cause to strengthen for'

ii) **məm-pər-**luas-kan

VERBL.PRF-NOM.PRF-strength-CAUS.SUF

'to cause to broaden for'

iii) **mən-tər-**tadʒam-kan

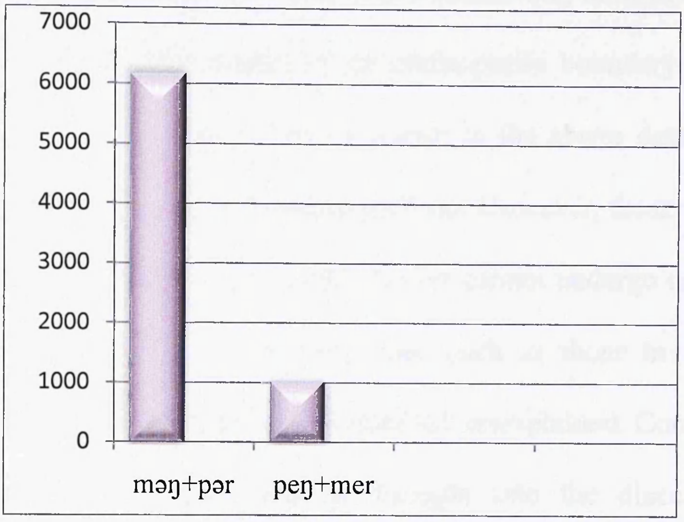
VERBL.PRF-VERBL.PRF-sharp- CAUS.SUF

'to cause to sharpen for'

A generalisation from the above examples can be summarised as: Nasal substitution occurs when the multiple prefixes produce a nominal prefixed word, as shown in 145(a). On the other hand, when the multiple prefixes form a verbal word, as in 145(b), nasal substitution is blocked.

Observations from the corpus show that 6,154 words are /məŋ+pər/, while only 1,020 words are /pəŋ+pər/. This shows that multiple prefixes with nasal and voiceless obstruent clusters occur more than the ones with nasal substitution. The difference in number between the prefixes can clearly be seen in the following graph:

Figure 7: The number of multiple prefixes /məŋ+pər/ and /pəŋ+pər/ (from the DBP-UKM corpus).



As already noted, the language does not allow nasal and voiceless obstruent clusters in the surface representation. Therefore, voiceless obstruents following nasals regularly undergo nasal substitution, as in 137(a). One question that can be asked here is: Is it obligatory for a sequence of nasal and voiceless obstruent to undergo nasal substitution? Or to put it in another way: Must nasal substitution be applied whenever there is a nasal and voiceless obstruent cluster since the phonetic requirements are already met? To answer this question in the context of multiple prefixation, I suggest that another factor, as well as the phonetic environment, i.e. the morphological environment, is worthy of consideration. By considering both factors, I assume that nasal substitution is blocked in multiple prefixation (i.e. at the prefix-prefix boundary). Why is it blocked in this morphological environment? To support this, I will then propose an EDGE-INTEGRITY (85) constraint which requires a morphological unit to preserve its edge segments in the underlying position by keeping them at the edge of a corresponding prosodic structure (Kang, forthcoming) in the constraint ranking.

Before the discussion begins, it is worth a brief overview of what has been done in previous studies. Works like Omar (1986), Hassan (1987), Kroeger (1988) and Pater (1999), cited in Ahmad (2005: 185), concerning multiple prefixation, claim that the morphological environment of the prefix-prefix boundary is impermeable to nasal substitution. This claim is true, as shown in the above data in 145(b), where nasal substitution does not occur between prefixes. However, those studies do not give any phonological explanation of why the clusters cannot undergo nasal substitution at this morphological boundary. Moreover, data such as those in 145(a), where the clusters undergo nasal substitution, have been left unexplained. Considering the above examples therefore, this issue will be brought into the discussion since nasal substitution is not fully active, as at the prefix-root juncture (i.e. single prefixation). Therefore, we need to revise the claims made in previous studies on nasal substitution at the prefix-prefix boundary. Before I offer an OT account, let us see first how the case under discussion has been resolved in ruled-based analyses.

In rule-based analyses, two rules, (1) nasal assimilation and (2) voiceless obstruent deletion, are applied to account for nasal and voiceless obstruent clusters. These two rules have to be applied in order, in that the nasal assimilation rule must precede the voiceless obstruent deletion rule. It is assumed that the same rules have also been applied to explain nasal and voiceless obstruent clusters in multiple prefixation, since there is a nasal and voiceless obstruent cluster. I illustrate how the rules apply:

146.

	Input:	/pəŋ+pər+badan+an/
(1)	Nasal Assimilation:	pəm+pər+badan+an
(2)	Voiceless Obstruent Deletion:	pəm+ər+badan+an
	Output:	[pə.mər.bada.nan]

From the above derivation, it shows that by applying the same rules, in order, to nominal multiple prefixes /pəŋ+pər/, the correct output is obtained. However, if this method of analysis were to be applied to another type of data, as we have in 145(b) for verbal prefixes /məŋ+pər/, we would instead get an incorrect output, as the following derivation shows:

147.

	Input:	/məŋ+pər+kuat+kan/
(1)	Nasal Assimilation:	məm+pər+kuwat+kan
(2)	Voiceless Obstruent Deletion:	məm+ər+kuwat+kan
	Output:	*[mə.mər.kuwat.kan]

The above derivation clearly shows that the two rules, taken in order, fail to account for /məŋ+pər/. The output derived from the above derivation is an output with nasal substitution. This is an incorrect output since /məŋ+pər/ does not undergo nasal substitution, as the data in 145(b) show. By assuming that nasal and voiceless obstruent clusters in multiple prefixation can be resolved in the same way as in single prefixation, some previous scholars (e.g.: Omar, 1993; Karim, 1995) have claimed that the data in 145(b) are to be considered as exceptional since the output does not obey the system of Malay grammar. Such a simple explanation has been given to account for the occurrence of nasal and voiceless obstruent clusters at the prefix-prefix juncture.

I assume that this claim was made by previous studies as they were only considering the phonetic factor in multiple prefixes. Since there is a nasal and voiceless obstruent cluster, hence the solution applied to a single prefix has also been applied to multiple prefixes. Therefore they simply claim that the non-application of nasal substitution in the verbal multiple prefixes /məŋ+pər/ is an exception. Here, I will put forward my argument about what I have claimed earlier, that nasal substitution does not actually occur in Malay multiple prefixation, and that it is also motivated by the morphological environment factor as well as the phonetic environment factor.

The data in 145(b) prove that nasal substitution is not active at the prefix-prefix juncture. The occurrence of nasal substitution, as in 145(a), is only a case of analogy, as I mentioned at the beginning of this chapter, where the solution applied in single prefixation has also been applied to multiple prefixes. To account for the blocking of nasal substitution between prefixes, we need to posit a new constraint, which ensures that no nasal substitution occurs in this morphological domain. OT is a theory which is able to account for a phonological process according to the morphological domain. To put it differently, unless the clusters occur in the right morphological domain, other than that, the phonological process of nasal substitution cannot be applied, even though all the phonetic requirements have been met. With the issue at hand, a morphology-phonology interface constraint is truly required to account for multiple prefixes. A faithfulness constraint proposed by McCarthy and Prince (1995), EDGE INTEGRITY, is a morphologically-conditioned phonology constraint which has been used to account for cases like this. The definition of EDGE INTEGRITY as proposed by McCarthy and Prince (ibid.) has been defined in (85) and is repeated below for convenience:

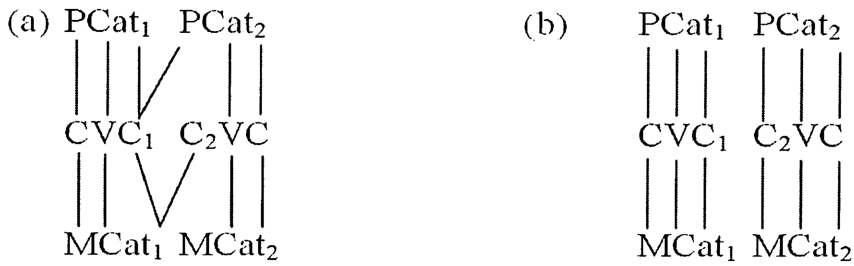
148. **EDGE INTEGRITY** (McCarthy and Prince, 1995).

A segment at the edge of a morphological constituent should be at the edge of a prosodic constituent, where the edges can be left, right or both.

As defined, EDGE INTEGRITY requires that the morphological unit preserve its edge segments in the input by keeping them at the edge of a corresponding prosodic structure. There is a strict faithfulness constraint on the segments at the edges so that every segment at the edge of a morphological unit is protected and is immune to phonological processes like epenthesis (Kang, forthcoming).

According to Kang (forthcoming), EDGE-INTEGRITY evaluates a segment and its affiliation. The segment is said to be violating the constraint if an initial or a final segment is not affiliated with the corresponding PCat's edge (ibid.: 7). The structure in 149(a) violates EDGE INTEGRITY since the final segment C_1 of $MCat_1$ is linked to $MCat_2$ and is not affiliated with $PCat_1$. Recall that the process of nasal substitution causes the two segments in the input to merge into a single segment in the output, due to the process of nasal substitution. Therefore, we see that the final segment C_1 of $MCat_1$ is also linked to the initial segment C_2 of $MCat_2$. The structure in 149(b) does not violate EDGE-INTEGRITY at all, since the two segments C_1 and C_2 are at the edges of their prosodic constituents.

149. /CVC + CVC/ ('+' stands for a morphological boundary) (from Kang, forthcoming).



The merging of the two segments C_1 and C_2 in the input into a single segment results in one less consonant in the output. The consonant then has to be syllabified into the second prefix, $MCat_2$. This results in the right edge of $MCat_1$ not being affiliated with the corresponding $PCat_1$. Thus I propose an EDGE-INTEGRITY (MWord, PrWord) constraint which focuses on the segments at the edges of morphological words and their prosodic words for Malay, as defined below:

150. **EDGE INTEGRITY (MWord, PrWord).**

Segments at the edges of a morphological word should be at the edge of a prosodic word at both edges.

Considering the constraint I have just introduced, I now establish the following constraint ranking for verbal multiple prefixes: PrStem >> EDGE INTEGRITY >> NAS ASS >> *NÇ̥ >> ALIGN RHOTIC >> DEP-IO >> UNI. Observe that EDGE INTEGRITY has been added into the constraint ranking.

151. Constraint ranking for multiple prefixes.

/məŋ$_1$+p$_2$ər+bəsar/	PrStem	EDGE INTEG	NAS ASS	*NÇ	ALIGN RHOTIC	DEP -IO	UNI
a. məm$_{12}$ərbəsar		*!			*		*
b.☞məm$_1$p$_2$ərbəsar				*	*		
c. məŋ$_1$p$_2$ərbəsar			*!		*		

The tableau above shows that the faithfulness constraint EDGE-INTEGRITY dominates the markedness constraint, *NÇ. Because of that, candidate (a), with nasal substitution, loses due to a fatal violation of the faithfulness constraint. In contrast, candidate (b) violates the markedness constraint *NÇ, as the candidate does not undergo nasal substitution. Since the markedness constraint *NÇ is ranked beneath the faithfulness constraint, EDGE INTEGRITY, the least unmarked output is preferable to the unmarked ones. Therefore [məm$_1$p$_2$ərbəsar] emerges as the winner, not *[məm$_{12}$ərbəsar]. This ranking, EDGE INTEGRITY >> *NÇ, can thus account straightforwardly for why nasal substitution does not occur in the environment of the prefix-prefix juncture.

The analysis of multiple prefixes above proves the fact that EDGE INTEGRITY provides us with a means to account for the problem of phonological opacity which arises from the interaction of phonological process (i.e. nasal substitution) and morphological domain (i.e. prefix-prefix juncture). This morphology-phonology interface constraint, which is found in OT and is used as a research tool to account for any morphologically-conditioned phonology issues, is absent in previous approaches, such as rule-based and non-linear autosegmental analyses. Thus, it is impossible to say that those previous approaches can perform the analysis correctly as OT does. Moreover, due to the morphological domain, nasal and voiceless obstruent clusters at the prefix-prefix juncture cannot undergo the regular process, nasal substitution. As a result, the constraint ranking for multiple prefixes is different from the one in the native group for single prefixation. As already demonstrated in the above analysis, the difference between multiple and single prefixes can be explained straightforwardly by applying co-phonology analysis. Co-phonology is the most morphologically sophisticated theory, as it assumes that morphological structure interacts with phonological constraints (Downing, 2006). This is the reason why OT is chosen as the theoretical framework in this study – it is a superior phonological theory that can account for a problem like this.

4.4 Conclusion

The above discussion on prefixation, which covers both single and multiple prefixation, has presented some important points about the grammar of Malay. Nasal and voiceless obstruent clusters are disfavoured in the language. This can be seen in the analysis of single prefixation, where words in the native group obey *NC̥ – the constraint which bans the clusters from occurring – while in the remaining groups, monosyllabic foreign and undeleted voiceless plosive in loanwords, they do not. Words in these two groups, as I have claimed, were originally non-native words, thus the regular process (i.e. nasal substitution) cannot be applied. It is impossible for borrowed words to be analysed in the same way as native words, since if the process for native words were to be applied to analyse borrowed words, then the surface output would become odd. The analysis above clearly shows that loanwords cannot be subject to a stricter set of constraints than those imposed on native words. By ranking the lexicon according to the etymology of words, this can limit the number of constraints used to analyse the process of prefixation in Malay. We probably need to consider more constraints to explain all the phonological matters regarding prefixed words in one hierarchy. As Inkelas, Orgun and Zoll (1997: 403) say, the model proposed by Itô and Mester 'would certainly limit the number of constraints available to define morpheme-specific co-phonologies to those true of the entire 'core' vocabulary'. One thing to note is that the phonotactics of loanwords cannot be totally changed to the phonotactics of native words, and therefore they cannot obey all the constraints obeyed by native words.

On the other hand, nasal and voiceless obstruent clusters at the prefix-prefix juncture are not resolved by nasal substitution. The clusters at this morphological boundary are permitted to emerge in surface representation as the edges of a morphological word are preserved by the faithfulness constraint called EDGE-INTEGRITY (150). However, as we saw, nasal substitution applies to nominal prefixes and is analysed in this study as a case of analogy. Since there is a nasal and voiceless obstruent cluster between the first and second prefix, it has been resolved by nasal substitution, as this is the way in which this cluster is resolved at the prefix-root boundary. To deal with this matter, I have argued that although there is a nasal and voiceless obstruent cluster, which means the phonetic environment is met and can thus be resolved by nasal substitution, the morphological environment must also be considered as a factor conditioning the clusters to undergo nasal substitution. By considering the phonetic and morphological environments, I have further claimed that nasal substitution is blocked between prefixes.

In the above analysis, it is clearly shown that OT offers a much better solution to handling all the problems in both single and multiple prefixes, rather than any other model. This theory could explain the application of nasal substitution and schwa epenthesis in disyllabic and monosyllabic bases, respectively. As we saw, stems with nasal and voiceless obstruent clusters in the underlying representation undergo some repair strategies in the surface representation, either nasal substitution or epenthesis schwa, depending on the size of the root. Those repair strategies apply in order to satisfy the markedness constraint in the hierarchy, i.e. *NÇ. In other words, epenthesis schwa and nasal substitution are the conspiracies applied in the language through prefixation to rid the surface representation of nasal and voiceless obstruent clusters. Another example of conspiracy applied in prefixation is vowel lengthening. We note that MBT developed within OT proposes that the size of prosodic morphemes should fall out from their morphological categories. By claiming the size of affixes in Malay is monosyllabic bimoraic, vowels in prefixes with no coda consonant have to be lengthened. Such a solution of vowel lengthening as offered in OT cannot be found in any other alternative.

For convenience, I briefly tabulate the rankings that outline the process of prefixation in Malay (native words):

152. Constraint rankings for single and multiple prefixation.

(1) Single prefixation

(a) Native:

PrStem >> NAS ASS >> *NÇ >> ALIGN-RHOTIC >> DEP-IO, UNIFORMITY

(b) Monosyllabic foreign:

PrStem >> NAS ASS, *NÇ >> UNIFORMITY >> ALIGN RHOTIC >> DEP-IO

(c) Undeleted voiceless plosive:

PrStem >> NAS ASS >> DEP-IO >> UNIFORMITY>> *N$\mathring{\text{Ç}}$ >> ALIGN-RHOTIC,

(2) Multiple prefixation:

PrStem >> EDGE-INTEGRITY >> NAS ASS >> *N$\mathring{\text{Ç}}$ >> DEP-IO >> UNIFORMITY >> ALIGN-RHOTIC

5. REDUPLICATION*

5.1 Introduction

In the previous chapter, I presented the processes of prefixation in Malay for single and multiple prefixes. In this chapter, I shall examine another morphological process: reduplication. As I mentioned in Chapter 4, the results from the analysis of prefixation are essential for the analysis of reduplication. It has been demonstrated in that chapter that Stems (i.e. prefix and root) must be disyllabic in size where each morpheme contains one syllable. In this chapter, we will investigate whether the disyllabic condition required on Stems also applies to prefixed reduplicants, which will be discussed in Section 5.5.

In Malay, reduplication has long received much formal attention, particularly from Malay scholars such as Omar (1986, 1993), Hassan (1974), Othman (1981), Onn (1980), Che Kob (1981), Karim et al. (1994), Karim (1995), Ahmad (2000a, 2005) and Aripin (2005).[30] It should be mentioned that, among those scholars, only Ahmad (2000a, 2005) and Aripin (2005) apply non-linear phonology analysis,[31] the other scholars apply rule-based analysis. Those previous analyses have discussed various types of reduplication such as total, partial, rhyming and chiming. Although Malay reduplication has been widely studied, there are some problems that have not been

*Acknowledgements: thanks to Moira Yip for her very valuable comments on this paper that was presented to the 4th 'Language at the University of Essex Postgraduate Conference' [LangUE].

[30] Reduplication has been analyzed by the previous scholars listed here, but is not specific to either SM or Malay dialects.

[31] Ahmad (2000a) and Aripin (2005) apply non-linear autosegmental analysis to account for partial reduplication in SM and PD, respectively. These two scholars' works will be reviewed in this chapter for the purpose of argument. Meanwhile, Ahmad (2005) uses OT to scrutinize Onn's work on Malay reduplication.

satisfactorily been explained by previous studies. To put it differently, previous analyses cannot neatly explain some processes of reduplicated words.

The discussion in this chapter will focus on the three types of reduplication, i.e. partial, total and and affixal. What issues in each of these types will be discussed? In partial reduplication, the inaccurate definitions given by previous scholars in rule-based analyses (e.g.: Hassan, 1974, 1987; Othman, 1980; Omar, 1975, 1986, 1993) will be revised. As was mentioned in subsection 1.3.1, partial reduplication was defined as a process of copying the first syllable of the base into the reduplicant. As I commented in the first chapter (see subsection 1.3.1), the definition of partial reduplication has only been simply described in rule-based analyses. Therefore, it will be revised in this present study, in section 5.4.

Besides that, two types of partial reduplication will be discussed, i.e. light and heavy reduplication. It must be noted that although partial reduplication in SM has been discussed in previous studies, such as Onn (1980) and Ahmad (2001), it has been discussed separately from the process of reduplication in any of the Malay dialects. This could be the reason why the two patterns of reduplicative morpheme (i.e. light and heavy) have never been discussed simultaneously in any previous study. Considering one of the Malay dialects, i.e. Perak, this chapter will be discussing heavy reduplicative morphemes in partial reduplication in PD,[32] as SM partial reduplication only contains light reduplicative morphemes. Why is PD chosen for analysis and not another dialect of Malay? This is because partial reduplication in PD has received more attention among Malay scholars (e.g.: Omar, 1985; Ahmad, 1991; Aripin, 2005)

[32] Partial reduplication in PD was selected for examination in this study because the reduplicative morphemes in partial reduplication in this dialect consist of heavy syllables which are not found in SM. For the analysis of heavy reduplicative morphemes in partial reduplication in PD, this study will be using the data from previous studies, such as Ahmad (1991) and Aripin (2005).

compared to other Malay dialects. As noted in Chapter 3, since Malay does not have a corpus for its dialects, so data from previous works are very much needed for this analysis. However, in order to analyze heavy reduplicative morphemes, which occur in PD, Ahmad's (1991) work has been referred to. All the data for SM reduplication on the other hand will be based on the five million words of data from the DBP-UKM corpus.

In affixal reduplication on the other hand, the issue of 'compounding' vs. 'affixation' will be highlighted in Section 5.5. This issue is worth discussing since previous Malay studies did not examine it seriously. It has been claimed in many previous studies (e.g.: Kim, 1996; Hendricks, 1999; Myers and Carleton, 1996; Keenan and Polinsky, 1998) that reduplication is a kind of affixation. In Malay, most previous studies (e.g.: Ahmad, 2001; Aripin, 2005) are more inclined to agree that reduplication resembles affixation. In my analysis, however, reduplication will not be claimed as a kind of affixation, as those previous studies claimed. Rather, reduplication in Malay will be analysed as being self-compounding. To account for this issue, total reduplication (discussed in Section 5.3) and affixal reduplication will be examined. Despite the fact that there is nothing much of interest in connection with the alternation segments in total reduplication, as the reduplicant copies the entire base, it is however crucial that this type of reduplication be discussed, since the reduplicative morpheme in total reduplication is essential to determining whether reduplication can be considered as a type of affixation or not. Total and affixal reduplication play a significant role in accounting for the canonical form in reduplicative morphemes. The size of the reduplicative morpheme in those types of reduplication determines whether reduplication is better treated as an affixation or self-compounding process.

Apart from the issue of 'compounding' vs. affixation', there are some other issues that have not been taken into account, even though they are crucial. When affixed words are reduplicated, the prefix is not carried along, as in 14(b). Rather, in some affixed words, the prefix is copied into the reduplicant, as in 14(a). I repeat the examples here for convenience. In the analysis, I will then propose a solution based on the ideas of MBT to overcome this problem.

153. (a) /məŋ-ə-lap/ [məŋ-ə-lap-**ŋə**-lap]
ACT.PRF-STEMEX-wipe
'to wipe repeatedly'

(b) /məm-basuh/ [mə.ba.soh-ba.soh]
ACT.PRF-wash
'to wash'

I begin the chapter with a brief note on Malay reduplication, which will be discussed in Section 5.2. In Section 5.3, I will focus particularly on total reduplication. In the following section, 5.4, partial reduplication in SM and PD will be discussed and, finally, in Section 5.5, I will demonstrate the process of affixal reduplication.

5.2 A brief note on Malay reduplication

Generally, reduplication in Malay has been divided into four groups. These are: (1) total reduplication (2) partial reduplication (3) rhyming reduplication and (4) chiming reduplication (Hassan, 1974; Ahmad, 2005: 143). Each of the groups will be described in turn.

First, total reduplication is a process of completely copying the base onto the reduplicative morpheme, like /makan/ 'to eat' – [makan-makan] 'to eat repeatedly'. It is also known as a process of doubling (Inkelas, 2005), where the reduplicant is always identical to the base from where it is copied. In this case, since the reduplicant copies the entire base, it is difficult to determine which is the reduplicant and which is the base or root. As stated in Wilbur (1973), since total reduplication copies the entire base, the reduplicant and base are identical. Because of this, there is no way to say the reduplicant is prefixed to the base, or the reduplicant is suffixed to the base. He also mentions that: 'There is no qualitative or quantitative difference between these two in a surface form, which is the original and which is the copy' (ibid.: 11). Wilbur's statement is upheld when we look at examples of Malay:

154. Total reduplication of roots (from the DBP-UKM corpus).

Roots		**Reduplicated words**
/marah/	[mãrah] 'scold'	[mãrah-mãrah] 'scold repeatedly'
/bulat/	[bulat] 'round'	[bulat-bulat] 'exactly'
/lagu/	[lagu] 'song'	[lagu-lagu] 'many songs
/təŋah/	[təŋah] 'middle'	[təŋah-təŋah] 'very in the middle
/baru/	[baru] 'new'	[baru-baru] 'very new

It is hard to determine the reduplicant from the reduplicated forms in the above examples, since the copying process involves copying the entire segment of the base. By contrast, in some languages, for example CiYao verbs reduplication, tone is not carried along in the reduplicant (Myers and Carleton, 1996). The reduplicated words are underlined:

155. CiYao Verb Reduplication (Myers and Carleton, 1996: 64, cited in Downing, 2006: 21).

Unreduplicated	**Reduplicated**
ku-téléka 'to cook'	ku-téléka-teleka 'to cook repeatedly
ku-wómbóka 'to save'	ku-wómbóka-womboka 'to save repeatedly'
ku-súlúmunda 'to sift (flour)'	ku-súlúmunda sulumunda 'to sift (flour) repeatedly'

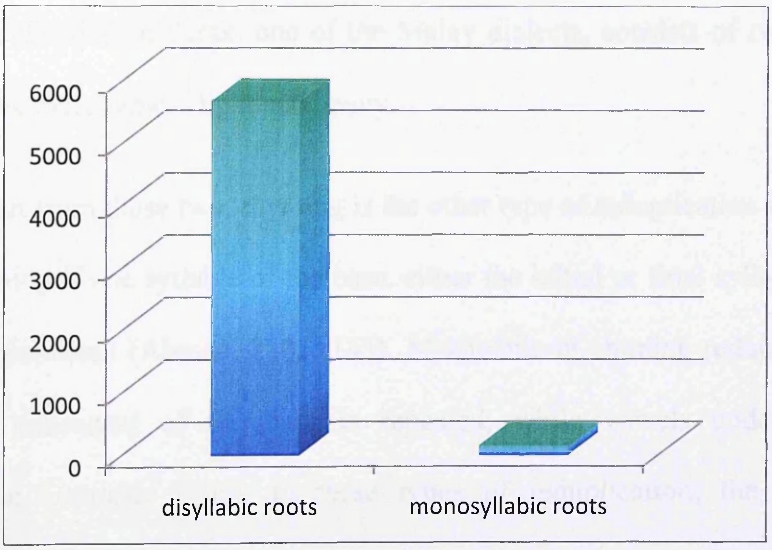
In terms of meaning, totally-reduplicated words denote plurality, repetition, continuity, intensity, extensiveness and reciprocity (cf.: Omar, 1993; Hassan, 1974; Ahmad, 2005: 141). I briefly lay out some relevant examples for each of these:

156. Meanings derived by totally reduplicating words (from Omar, 1993).

(a)	plurality	rumah-rumah 'houses'
(b)	repetition	muntah-muntah 'to vomit repeatedly'
(c)	continuity	lari-lari 'keep running'
(d)	intensity	benar-benar 'very true'
(e)	extensiveness	kawasan-kawasan 'spaces'
(f)	reciprocity	balas-balas 'revenge'

As claimed by Omar (1993), this type of reduplication is the most productive, compared to others. In my observations, based on five million words of corpus data, I firmly agree with Omar on this. The data contain a lot of examples of total reduplication in SM. There are 5,843 examples of total reduplication. This includes disyllabic roots (see Appendix F) and monosyllabic roots (see Appendix G). Most of the totally reduplicated words are disyllabic roots, 5,685 words. Monosyllabic roots comprise 158 words. The comparison of disyllabic and monosyllabic roots featuring total reduplication can be easily seen in Figure 8 below. In the corpus, this type of reduplication has the highest number among other types of reduplication, which are partial, rhyming, chiming and affixal. The details of both disyllabic and monosyllabic roots' total reduplication in SM will be discussed in the following Section 5.3.

Figure 8: Disyllabic and monosyllabic roots in total reduplication in SM (from the DBP-UKM corpus)



Second, partial reduplication is a process where only part of the base is copied. Unlike total reduplication, partial reduplication in SM is not a very productive form but it does exist as something productive in terms of phonological process. As I observed, there are only 4,180 words of partial reduplication in the corpus, as exemplied in (166) (also see Appendix H).

As stated in Omar (1975: 190), this type of reduplication became more productive after 1956 when the *Dewan Bahasa dan Pustaka – DBP* (The Institute of

Language and Literature) revived this morphological process as a procedure for coining new words. This procedure is designed particularly to convey scientific terms for concepts, which are mainly borrowed from English (cited in Ahmad 2005: 142). This includes both partial reduplication in SM and other dialects spoken in Malaysia. As claimed by Omar (1975), partial reduplication is widely used in colloquial Malay and in Malay dialects. In my search of the relevant literature (e.g.: Omar, 1985; Hassan, 1974; Ahmad, 1991; Aripin, 2005), I have found that partial reduplication in some Malay dialects consists of more than one pattern of reduplicative morpheme, i.e. light and heavy. In contrast, light reduplicative morphemes are the only pattern in SM partial reduplication, as exemplified in (166). As we shall see soon, in Section 5.4, partial reduplication in Perak, one of the Malay dialects, consists of two patterns of reduplicative morpheme – light and heavy.

Apart from those two, rhyming is the other type of reduplication in Malay. It is called rhyming if one syllable of the base, either the initial or final syllable, is copied onto the reduplicant (Ahmad, 2005: 143). Meanwhile in chiming reduplication, only the initial consonant of the base is repeated, while vowels undergo phonetic modification (Ahmad, 2005). In these types of reduplication, the reduplicative morpheme follows or precedes the base (Hassan 1974). As stated in Ahmad (2005), phonetic modifications in chiming reduplication are unpredictable, and so he claimed that it might be impossible to capture them with a rule. This might be the reason why rhyming and chiming in Malay have not received much attention from Malay scholars previously.

In the following are examples of each, i.e. rhyming and chiming reduplication.[33] Observe that in rhyming, the initial or final syllable of the base is copied, while in chiming only the initial consonant of the base is copied onto the reduplicative morpheme.

157. Rhyming reduplication (from the DBP-UKM corpus).

Roots	**Reduplicated words**
anaʔ 'child'	anaʔ-pinaʔ 'a large number of children'
saudarə 'relative'	saudarə-marə 'a large number of relatives'
sərtə 'along with'	sərtə-mərtə 'immediately'
bəŋkoʔ 'bent'	bəŋkaŋ-bəŋkoʔ 'to be crooked'

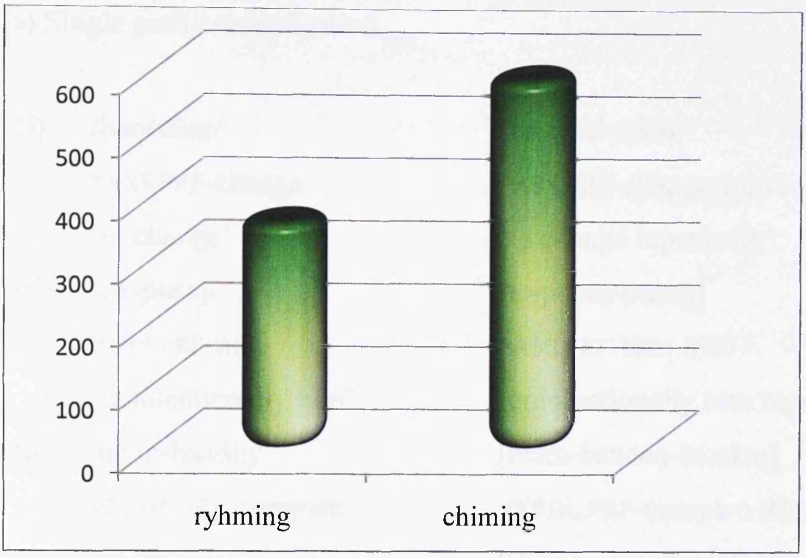
158. Chiming reduplication (from DBP-UKM corpus).

Roots	**Reduplicated words**
gunuŋ 'hill'	gunuŋ-ganaŋ 'hills'
tanah 'soil'	tanah-tanih 'various kind of soils'
asal 'origin'	asal-usol 'ancestor'

In the corpus, I observe that rhyming and chiming reduplication are not as productive as total and partial reduplication. There are 364 and 592 words for rhyming and chiming, respectively. The following graph presents the numbers for chiming (see Appendix I) and rhyming (see Appendix J) reduplication in SM.

[33] As has been claimed by Ahmad (2005), 'Rhyming and chiming reduplication are no longer productive in the language'. I will not therefore discuss these two reduplications; instead, I will refer the interested reader to sources such as Hassan (1974) and Onn (1980). For these types of reduplication, it has also been claimed by Omar (1982: 206) that there seems to be no rule to formulate the construction of them.

Figure 9: Rhyming and chiming reduplication in SM (from the DBP-UKM corpus)



Apart from the above-mentioned groups, there is another type of reduplication which I call 'affixal reduplication'. Affixal reduplication is a process whereby prefixed words are repeated. Most of the previous studies, for example Hassan (1974), Che Kob (1981) and Othman (1985), omit this type of reduplication from their lists. As far as I know, at least three Malay scholars have discussed affixal reduplication. These are Onn (1980), Omar (1982)[34] and Ahmad (2005). Taking this into account, it will therefore also be analysed in this thesis. Besides that, this type of reduplication is examined as it involves some interesting phonological processes, especially when the copying process occurs. The corpus shows that both single prefix and multiple prefixes can be reduplicated. However, the existence of these in the corpus is only occasional. Perhaps not all prefixed words in SM can be reduplicated. There are 612 words with single and multiple prefix reduplication, comprising 590 and 22 words, respectively. I briefly lay out some examples from the corpus. A full discussion of this topic can be found in Section 5.5.

---

[34] Note that affixal reduplication in Omar's (1982: 206) analysis is termed 'cross reduplication'.

159. Single and multiple prefix reduplication (from the DBP-UKM corpus).

(a) Single prefix reduplication

(i)	/bər-tukar/	[bə-tuka-tuka]
	PASS.PRF-change	PASS.PRF-change-RED
	'to change'	'to change repeatedly'
(ii)	/tər-pusiŋ/	[tə-puseŋ-puseŋ]
	PASS.PRF-turn	PASS.PRF-turn-RED
	'unintentionally turn'	'unintentionally turn repeatedly'
(iii)	/məm-bandiŋ/	[məm-bandeŋ-bandeŋ]
	VERBL.PRF-compare	VERBL.PRF-compare-RED
	'to compare'	'to compare with others'

(b) Multiple prefix reduplication

(i)	/məm-pər-bəsar-kan/	[məm-pər-bəsar-bəsar-kan]
	VERBL.PRF-NOM.PRF-large-CAUS.SUF	VERBL.PRF-NOM.PRF-large-RED-CAUS.SUF
	'to cause to enlarge for'	'to cause to enlarge for repeatedly'
(ii)	/di-pər-main-kan/	[di-pər-main-main-kan]
	PASS.PRF-NOM.PRF-play-CAUS.SUF	PASS.PRF-NOM.PRF-play-RED-CAUS.SUF
	'to be cheated for'	'to be cheated for repeatedly'

In the following sections, we are going to discuss the three types of reduplication chosen in this study i.e. total, partial and affixal. We begin the discussion with total reduplication, in Section 5.3.

5.3 Total Reduplication

As mentioned at the beginning of this chapter (Section 5.1), the reason why total reduplication is discussed in this present study is to support the argument I raised earlier that reduplication is a process of self-compounding, rather than affixation. In this section, I will provide my argument, based on the ideas of MBT, of why reduplication is best treated as self-compounding by examining the data for total reduplication. The discussion follows.

As far as reduplication is concerned, it has been treated as a kind of affixation. Reduplicative morphemes tend to be labelled as affixes, particularly concerning the aspects of their meaning and function (Downing, 2006). There are numerous examples of languages in the world where reduplication is considered as a form of affix, for example Korean (Kim, 1996) and Semai (Hendricks, 1999; Myers and Carleton, 1996; Keenan and Polinsky, 1998).

Malay scholars, for example Ahmad (2005) and Aripin (2005), have also proposed the same argument, with those scholars considering reduplication as a form of affix for SM and PD, respectively. Aripin (2005: 162) claims that reduplication is a process where the root receives an affixation. The reduplicative affix (i.e. reduplicative morpheme) in Aripin's analysis is represented as CVC elements. Likewise, Ahmad (2005: 141) claims that reduplicant R is the output form of the affix morpheme called /RED/. This affix morpheme has no phonological specification in its lexical entry. The reduplicant gets its phonological material from the base to which it is attached. In contrast, the affix morpheme in affixation is phonologically specified (Ahmad, 2005). In other words, the affix morpheme has its own phonological material, for example the prefixes /məŋ+/, /pəŋ+/, /di+/, /tər+/, /bər+/ and /sə+/, and the suffixes /-kan/, /-an/ and /-i/. The segmental form in affix morphemes is consistent, whereas the form in reduplicative morphemes changes for every word.

According to Moravcsik (1978) and McCarthy and Prince (1986), prosodic morphemes for reduplicative affixes typically contain one or two syllables, while other prosodic morphemes, like truncations, nicknames and roots, which can stand independently, tend to be disyllabic or polysyllabic (McCarthy and Prince, 1986; Hayes, 1995; Itô, 1990, cited in Downing, 2006). Given the insights from MBT, every morpheme has a canonical shape. It is common in many languages for affixes to be monosyllabic or monomoraic, while stems tend to be larger, typically foot-sized (Downing, 2006: 25). As reduplicative morphemes in total reduplication copy the entire base, so they cannot be said to be a kind of affixation, since the reduplicative morphemes would exceed the canonical shape of one affix. This poses a problem for the earlier claim that a reduplicant should be treated as an affix.

To deal with this, in recent prosodic morphology (e.g.: Eulenberg, 1971; Niepokuj, 1991; Downing, 2006), work concerning total reduplication has proposed that total reduplication is a process of compounding, rather than affixation. Such work claims that total reduplication is primarily a reduplication process and is a form of self-compounding (cited in Downing, 2006: 65). Those scholars also claim that, like compound words, total reduplication is made up of two lexical morphemes – base and reduplicant. Other scholars, such as Bauer (1988), claim that a reduplicated word is to be considered as compound if the base is entirely reduplicated. The claim that reduplication should be treated as forming compound words can clearly be seen in CiYao (155), as the reduplicative morphemes are disyllabic or polysyllabic (Downing, 2006). In this language, reduplicative morphemes contrast with ordinary affixes which typically contain one syllable. In CiYao, the reduplicative morphemes are polysyllabic.

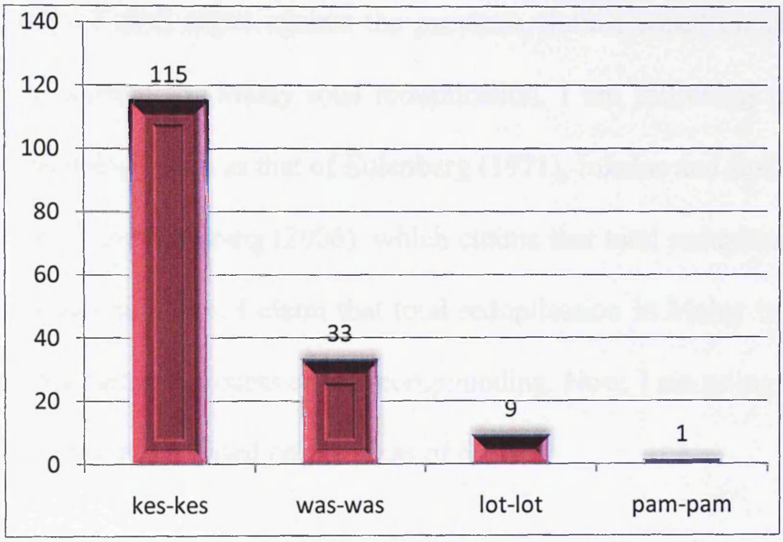
There is another instance, from other languages, for example Kambera, where the reduplicative morphemes contain more than one syllable. See the examples of total reduplication in Kambera below:

160. Total reduplication in Kambera (Klamer, 1998: 38, cited in Downing, 2006: 67).

ha-pungu = ha-pungu	'various <u>poles</u>'
ha-atu = ha-atu	'each and every <u>one</u> (people)'
pa-mula = pa-mula	'keep on <u>planting</u> (rice)'
ma-ramba = ma-ramba	'various (kinds of) <u>kings</u>'

The situations in CiYao and Kambera also occur in the process of doubling words/total reduplication in Malay where the reduplicative morphemes copy the entire base. Therefore, the size of the reduplicative morphemes is larger than one syllable, as in (154). The DBP-UKM corpus used in this study proves the fact that most of the reduplicative morphemes (5,685 words) in total reduplication that I observed are disyllabic. Because the roots are disyllabic, the reduplicative morphemes are also disyllabic, as reduplicative morphemes in total reduplication copy the entire root. However, there are totally-reduplicated words found in the corpus that are smaller than disyllabic, that is they are monosyllabic. In this case, the roots and reduplicative morphemes are monosyllabic. There are only 158 words of total reduplication with monosyllabic reduplicative morphemes in total (see Appendix G). The following graph presents the numbers for monosyllabic reduplicative morphemes in total reduplication:

Figure 10: The numbers for monosyllabic total reduplication in SM (from the DBP-UKM corpus)



Except for those four words with monosyllabic reduplicative morpheme presented in the graph, /kes/ 'case' → [kes-kes] RED-case 'cases', /was/ 'suspicion' → [was-was] RED–suspicion 'very suspicion', /lot/ 'group of' → [lot-lot] RED-group of 'more than one group of' and /pam/ 'pump' → [pam-pam] RED-to pump 'to pump repeatedly', the remainder of the total reduplication data found in the corpus contain disyllabic reduplicative morphemes. Now I draw out some of the examples of disyllabic total reduplication found in the corpus. For more examples, see Appendix F.

161. Disyllabic total reduplication (from the DBP-UKM corpus).

/pətani/ 'farmer'	[pətani]	[pətani-pətani] 'farmers'
/aŋgota/ 'member'	[aŋgotə]	[aŋgotə-aŋgotə] 'members'
/bila/ 'when'	[bilə]	[bilə-bilə] 'whenever'
/baru/ 'new'/ 'recent'	[baru]	[baru-baru] 'very new'/ 'recently'
/bənar/ 'true'	[bəna]	[bəna-bəna] 'very true'/ 'truly'

From the copious data for disyllabic total reduplication found in the corpus, we have concrete evidence to disagree with previous scholars' claims that reduplication is a form of affix. I shall argue against the previous claims, based on the data I have observed. To account for Malay total reduplication, I am following recent work in Prosodic Morphology such as that of Eulenberg (1971), Inkelas and Zoll (2000, 2005), Niepokuj (1991) and Downing (2006), which claims that total reduplication is a form of self-compounding. Thus, I claim that total reduplication in Malay is not a process of affixation but rather a process of self-compounding. Now, I am going to provide the justification of my claim based on the ideas of MBT.

Recall that MBT requires that all prosodic morphemes must be assigned a morphological category (e.g. root, stem, affix or RED). As we can see in (161), the reduplicative affix in Malay can be longer than one syllable. This is contrary to what MBT predicts for the typical size of an affix syllable – monosyllabic or monomoraic affixes. Because of that, it is plausible if the reduplicative affix, which contains more than one syllable, is analysed as either a Stem or Word. This is what those scholars have claimed for total reduplication, that it forms a compound word and can therefore no longer be said to be an affix. As a compound, a word with total reduplication, which is made up of two lexical morphemes, must be minimally bimorphemic, and therefore is required to be at least disyllabic. This requirement is formalised by a COMPOUND constraint which is analogous to PROSODIC STEM in (35) (Downing, 2006: 124).

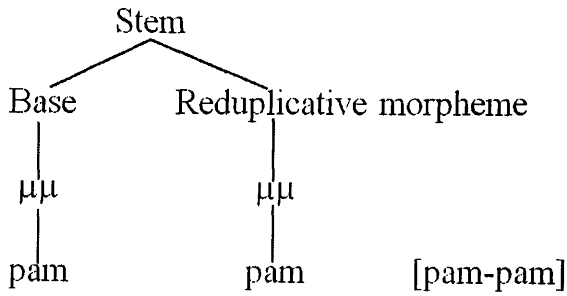
The data presented in Figure 10 show that monosyllabic roots can also be totally reduplicated. There is an important point that needs to be touched on when a monosyllabic root is the base for a reduplicative morpheme. As discussed at length in

Chapter 4 (subsection 4.2.1.1), even though the language requires a word to be at least disyllabic, this requirement however could not be satisfied in monosyllabic roots. Once again, the issue of disyllabicity minimality requirements will be addressed here.

Since reduplication in MBT is treated as self-compounding, the base and reduplicant are subject to similar minimality requirements because both of them are stems (Downing, 2006: 159). Therefore, they are defined as minimally disyllabic by PROSODIC STEM (35) (Downing 2006: 160). Based on the data in (161), the base and reduplicative are identical in size and, as such, both are minimally disyllabic. Therefore, the minimality requirements, as expected by MBT for base and reduplicative morpheme, do not encounter any problem with disyllabic roots.

For monosyllabic roots (monomorphemic roots) however, the bases do not contain two syllables and, therefore, the reduplicative morpheme would also be monosyllable as the bases are. To account for this, it is the same argument that I proposed earlier, in Chapter 4 (subsection 4.2.1.1), the disyllabicity minimality requirements for monosyllabic roots can be accounted for by the binary minimality principle, which is HEADS BRANCH (31). In this constraint, the base and reduplicative morpheme in monomorphemic roots reduplication are allowed to be monosyllabic bimoraic, the minimum size required for Heads (roots) (see the representation of Heads in (33). See the illustration below:

162. Representation of the minimum size of word in Malay:[35] /pam/ 'pump' → [pam].



The minimality requirements to account for Malay monomorphemic root reduplication can be clearly seen in the constraint ranking: HEADS BRANCH >> DEP–IO >> PROSODIC STEM. In this constraint ranking, as we see in the following tableau, HEADS BRANCH outranks DEP-IO where a mora is epenthesized to satisfy minimality requirements.

163.

/pam-STEM$_{RED}$/	HEADS BRANCH	DEP-IO, DEP-BR	PRSTEM
a.☞[pam]-[pam]			**
b. [pam]-[ə.pam]		*! (BR)	*
c. [ə.pam]-[pam]		*! (IO)	

As shown in the above tableau, when HEAD BRANCH is ranked high, it allows the base and reduplicative morpheme to be monosyllabic. Therefore, we see in the tableau that none of the candidate violates this constraint. Due to schwa epenthesis, candidates (b) and (c) violate DEP-BR and DEP-IO, respectively. Candidate (a) without schwa epenthesis, neither in the base nor in the reduplicative morpheme, emerges as the winner.

[35] As I said earlier, it is difficult to determine the base and reduplicative morpheme in total reduplication since the reduplicative morpheme copies the whole base. Because of that, the above representation is roughly sketched out simply to give a clearer view of the discussion.

Similar to Malay monosyllabic root reduplication, HEADS BRANCH has been used to account for monosyllabic (monomorphemic) roots in the process of total reduplication in Axininca Campa. Observe that HEADS BRANCH is high in the ranking to satisfy the bimoraic minimality requirement for monomorphemic bases (Downing, 2006: 160). By ranking HEADS BRANCH higher than other constraints in the ranking, the base and reduplicant, with less than two syllables respectively, can be chosen as the best output.

164. Axininca Campa monomorphemic base reduplication (from Downing, 2006: 161)[36]

/naa-Stem$_{RED}$ /	HEADS BRANCH	DEP-IO, DEP-BR	PRSTEM	MAX-BR
a.☞[naa]-[naa]			**	
b.[naa]-[naata]		*! (BR)	*	
c. [naata]-[naata]		*! (IO)		

We have just discussed total reduplication. The following section discusses partial reduplication.

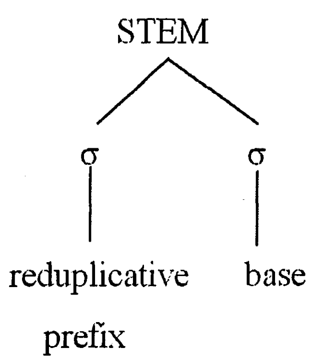
5.4 Partial Reduplication

Now we move on to partial reduplication. In this section, I am going to discuss the two patterns of reduplicative morpheme, light and heavy, that are found in SM and one Malay dialect, i.e. Perak.

[36] The tableau is from Downing (2006), though here it has been modified. The modification made in this tableau is only of the constraint. I do not include the ALIGN PREFIX constraint, since the constraint is not yet relevant to the discussion. The full tableau with the ALIGN PREFIX constraint can be seen later, in the section on affixal reduplication in Section 5.5, or by directly consulting Downing (2006: 161).

A key point to note here is that partial reduplication in Malay is not analysed as a process of self-compounding, as I claimed for total reduplication. As we will see later in this section, reduplicative morphemes in this type of reduplication are regarded as affixes, and therefore both reduplicative morphemes and bases are subject to the MORPHEME-SYLLABLE CORRELATION (34), which requires that each morpheme that is a prefix and a root contains exactly one syllable. Thus, stems must be at least disyllabic. In the case of reduplication, MORPHEME-SYLLABLE CORRELATION can be satisfied by ensuring each morpheme that is a reduplicative prefix and base contains exactly one syllable. I demonstrate in the following diagram how this constraint can be satisfied:

165. MORPHEME-SYLLABLE CORRELATION



As we will soon see, all partial reduplicated words satisfy the MORPHEME-SYLLABLE CORRELATION constraint, as reduplicative prefixes always contain one syllable while bases are always disyllabic or larger. Further discussion of the MORPHEME-SYLLABLE CORRELATION continues later, on p. 217. Now, we observe first what previous studies have said about partial reduplication.

Partial reduplication, as defined by McCarthy (1986), is a process where reduplicative morphemes copy the base to which they are attached. The copying process in partial reduplication is however incomplete (ibid.). Unlike total reduplication, Malay partial reduplication has received much formal attention by scholars working on Malay, including, for example: Omar (1986, 1993); Hassan (1974); Othman (1976, 1981, 1985); Abas (1988); Che Kob (1981); Karim et al. (1994); Karim (1995). Likewise, in many other languages in the world, such as Ilokano (McCarthy, 1986; Downing, 2006), Orokaiva (Sommer, 1981), Yoruba (Akinlabi, 1984; Pulleyblank, 1988), Kamrupi (Goswami, 1955) and Mokilese (Levin, 1983, 1985), much theoretical attention has been paid to partial reduplication. This is mainly because the reduplicative morphemes in this type of reduplication are quite different and interesting in their examination, and the patterns of reduplicative morphemes offer challenges to morphological and phonological theories. In Malay, this is the reason why partial reduplication has been widely discussed by previous scholars as it presents more interesting phonological and morphological processes, though it is not very productive in terms of total reduplication. In my observation of the corpus, there appear to be only 4,180 words of partial reduplication in SM out of five million words (see Appendix H). The corpus data really do prove that partial reduplication is present in a very small amount in the language.

Before we go further, let us consider first the following examples of partial reduplication in SM, taken from the DBP-UKM corpus. Observe that the reduplicative morphemes in this group consist of CV elements – light syllables.

166. Partial reduplication in SM (from the DBP-UKM corpus).

/bo.la/ 'ball' (round object)	[bə-bola] RED-ball 'ball' (foods e.g. fishball)
/taŋ.ga/ 'stairs'	[tə-taŋgə] RED-stairs 'neighbours'

/ka.ʧaŋ/ 'bean'	[kə-kaʧaŋ] RED-bean 'beans'
/po.hon/ 'tree'	[pə-pohon] RED-tree 'trees'
/kun.ʧi/ 'key'/ 'lock'	[kə-kunʧi] RED-key 'password'

In the above data we see that the C element in the reduplicative morpheme is copied from the initial consonant of the base, while the V element is a specific vowel, schwa. In fact, this is the correct process where the reduplicative morpheme in partial reduplication in SM is copied. With the case in hand, I would like to turn back to previous scholars' work on partial reduplication in SM. As noted earlier (see Chapter 2 Section 2.3), the way a reduplicative morpheme in this type of reduplication is copied was wrongly described by previous scholars, especially those who applied rule-based analysis (e.g.: Hassan, 1974, 1987; Omar, 1975, 1986, 1993; Karim et al., 1994, Karim, 1995; and many others). The reduplicative morpheme is not copied from the first syllable of the base, as described by those scholars. I shall argue that, if that description were to apply to those words in (166), then we would definitely get the wrong reduplicative morpheme for partial reduplication in SM. I demonstrate below how the copying process described by those previous scholars produces incorrect reduplicative morphemes:

167.

Input :	/bo.la/	/kun.ʧi/
Copy the first syllable:	bo-bola	kun-kunʧi
Output	*[bo-bola]	*[kun-kunʧi]

We have just seen in the above discussion how the copying process is described in rule-based analyses. Now, we shall see how the same case has been explained in segmental analysis. In works since Marantz (1982), segmental theories have been proposed as representing the template shape of reduplicative morphemes. Since reduplication is considered a type of prefixing, Marantz proposes that a reduplicative affix is obtained by copying the complete phoneme melody of the root and linking it to the affixal CV template (ibid., cited in Spencer, 1991: 151). Marantz's segmental theory was applied by Syed Ahmad (1998) to analyse partial reduplication in Malay. The concept of partial reduplication given by Syed Ahmad (ibid.) offers nothing different from the one given by previous scholars (e.g.: Omar, 1986, 1993; Hassan, 1974; Othman, 1981; Che Kob, 1981), where reduplication is defined as a process of copying the initial syllable of a base, while the vowel in the reduplicative morpheme is weakened to a central vowel 'schwa' (Syed Ahmad, 1998: 39). I briefly illustrate the derivation of the CV template as applied by Syed Ahmad:

168. CV Template for partial reduplication (adapted from Syed Ahmad, 1998).

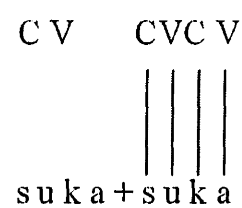
l a ŋ i t	l a ŋ i t	l a ŋ i t
C V C V C	CVCVC	CVCVC

From the derivation above, the whole word and segment of the base are copied into the template. Only the first two segments in the base, i.e. consonant and vowel, are associated to the CV templates. In accordance with autosegmental principles, any melody elements or prosodic template slots left unassociated at the end of the derivation will be deleted by convention (Marantz, 1980, cited in Spencer, 1991: 152). Therefore, the unlinked segments /ŋ/, /i/ and /t/ would have to be deleted. Then, the vowel in the CV template would be weakened to schwa and the output would be [lə.laŋit].

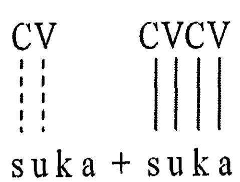
Another scholar, Ahmad (2000a), also applies template analysis to analyse partial reduplication in SM. The proposed template for reduplicative morphemes is CV as well. Differing from Syed Ahmad (1998), in Ahmad (2000a), the underlying form only contains the elements of CV, regardless of how long the base is. The melodies of the base are then repeated in the reduplicative morpheme and are associated to the skeletal tier through a process of left-to-right mapping. Any melodies left unassociated are deleted. The vowel in the reduplicative morpheme becomes a schwa through the process of vowel reduction. I briefly illustrate the process below:

169. (a) Underlying representation

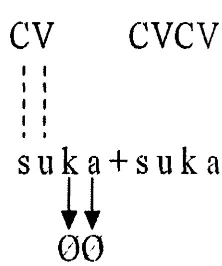
(b) Repetition of melody



(c) Left-to-right mapping



(d) Delinking of melody



(e) Vowel Reduction

CV CVCV

s ə + s u k a

(f) Surface representation

CVCV CV

s ə s u k a [sə.su.ka]

According to McCarthy and Prince (1986), 'CVC' as a reduplicative morpheme, for example, in the earlier stages seems to be plausible enough when templatic phonology is used to analyse other process of partial reduplication in other languages like Ilokano. It shows that this segmental approach finds it harder to account for a correct generalisation. I am in agreement with McCarthy and Prince on this matter. If we look at the CV template as proposed by Syed Ahmad (1998) and Ahmad (2000a) to account for SM partial reduplication, it poses a problem for the PD data in which the reduplicative morphemes also consist of both patterns, i.e. CV and CVC elements – light and heavy reduplicative morphemes. Certainly, the CV template can only be applied to SM reduplication, as CV is the only pattern of reduplicative morphemes in SM, and CV reduplicative morphemes in PD. This CV template, however, fails to apply to another pattern of reduplicative morphemes in PD, i.e. CVC. The two patterns of reduplicative morpheme in PD – CV and CVC – were exemplified in (89) (see Chapter 2 subsection 2.4.2.3), and are repeated here for convenience:

170. Partial reduplication in PD (from Ahmad, 1991).

(a) Light reduplicative morphemes

(i)	/budaʔ/ 'child'	[budaʔ]	[bə-budaʔ]
			RED-child 'all kinds of children'
(ii)	/cərita/ 'story'	[cərita]	[cə-cərita]

			RED-story 'all kinds of stories'
(iii)	/kaʤi/ 'to study'	[kaʤi]	[kə-kaʤi]
			RED-to study 'to study repeatedly'
(iv)	/kira/ 'to estimate'	[kɛrɛ]	[kə-kɛrɛ]
			RED-estimate 'by my estimate'
(v)	/dulu/ 'long ago'	[dulu]	[də-dulu]
			RED-long ago 'very long ago'

(b) Heavy reduplicative morphemes

(i)	/pətaŋ/ 'evening'	[pətaŋ]	[pəm-pətaŋ]
			RED-evening 'every evening'
(ii)	/ʤaraŋ/ 'seldom'	[ʤaʁaŋ]	[ʤəɲ-ʤaʁaŋ]
			RED-seldom 'very seldom'
(iii)	/baraŋ/ 'thing'	[baʁaŋ]	[bəŋ-baʁaŋ]
			RED-thing 'all kinds of things'
(iv)	/sikit/ 'a little'	[siket]	[səʔ-siket]
			RED-a little 'very little'
(v)	/gəlap/ 'dark'	[gəlap]	[gəʔ-gəlap]
			RED-dark 'very dark'
(vi)	/budaʔ/ 'child'	[budaʔ]	[bəʔ-budaʔ]
			RED-child 'all kinds of children'

Observe that there are two types of reduplicative morphemes in partial reduplication in PD: (1) light reduplicative morphemes, as in 170(a) and (2) heavy reduplicative morphemes, as in 170(b). For light reduplicative morphemes, the copying process is exactly the same as in SM, where the C element in the reduplicative morpheme is copied from the initial consonant of the base. The V element is schwa specific. Notice that in 170(b), the reduplicative morphemes are constructed of heavy syllables – CVC. The first CV has the same process as in SM, where the initial consonant in the reduplicative morpheme is copied from the initial consonant of the base, while the V segment is also schwa specific. From where does the final C in the reduplicative morpheme gets its phonological material? The final C in the reduplicative morpheme is copied from the final consonant of the base. Note that there are two types of final consonant of the base that are involved in the process of reduplication in PD. The consonants are nasal segments and consonant stops. The two final consonants then undergo certain phonological processes when the process of copying occurs. The phonological processes are debuccalisation and nasal assimilation. Debuccalisation converts consonant stops /t/, /p/ and /k/ into [ʔ] when the consonants are copied into the final C in the reduplicative morpheme in the syllable coda. The nasal assimilation process occurs when the nasal consonant is copied into the coda syllable in place of the final C of the CVC reduplicative morpheme. The nasal consonant assimilates to the following onset consonant in the second syllable.

Similarly, a reduplicative pattern, where the final consonant of the base is copied into the reduplicative morpheme, is also found in Semai reduplication. In this language, the reduplicative morpheme does not however contain a vowel. The reduplicative morpheme only contains CC elements whereby these consonants are copied from the initial and final consonant of the base, respectively. Let us first consider the Semai data, as illustrated below:

171. Semai reduplication (from Hendricks, 1999).

taʔəh	<u>th</u>-taʔəh	'appearance of large stomach constantly bulging out'
payaŋ	<u>pŋ</u>-payaŋ	'appearance of being dishevelled
sulɔŋ	<u>sŋ</u>-sulɔŋ	'the odd appearance of a snake's head'
cayɜm	<u>cm</u>-cayɜm	'contracted fingers of a human or animal, not moving'
ruhɔ:ɲ	<u>rɲ</u>-ruhɔ:ɲ	'the appearance of teeth attacked by decay'

Based on the Prosodic Morphology framework, where reduplicative morphemes are represented as prosodic units such as moras, syllables, feet, prosodic words and so on (McCarthy and Prince, 1990, 1993, 1994), those patterns of reduplicative morphemes, CV and CVC in SM and PD, can plausibly be accounted for. The success of prosodic units in describing reduplicative processes in many languages, such as Arabic (McCarthy and Prince, 1988), Sanskrit (see Steriade, 1988) and Kinande (see Mutaka and Hyman, 1990), has brought me to apply it to Malay reduplication. We will see later in this section how the prosodic template works on the two patterns of reduplicative morphemes – light and heavy.

It must be mentioned that Squamish and llokano are two languages that also have two reduplication patterns. Before we go to the Malay data, we first look at a case study of reduplication in Squamish and then continued by looking at Ilokano reduplication. As mentioned, co-phonology has been applied to Squamish which also has two reduplication patterns: (1) the reduplicative morpheme is CəC, regardless of what the base vowel is, as in 172(a); and (2) the reduplicative morpheme is CV whereby the vowel is copied exactly from the vowel in the base, as exemplified in 172(b). What I will demonstrate here is how elegant co-phonology is in accounting for the two patterns of reduplicative morphemes in the language. Shown below are examples of those two patterns. For convenience, the reduplicative morphemes are underlined:

172. Squamish's two reduplicative morphemes (from: Bar-el, 2000; Downing, 2006).

(a) CəC reduplicative morpheme

<u>p'əq'w</u> – p'éq'w	'yellow'
<u>təc</u>–téc	'skinny'

<u>k'wəs</u> – k'wás 'burn'

<u>təqw</u> – tóqw 'red codfish'

(b) CV reduplicative morpheme

<u>k'wá</u> – k'way? 'very hungry'

<u>sé</u> – siq 'fly'

<u>pó</u> – pum? 'swell'

The two patterns of reduplicative morphemes in Squamish has been analysed by applying co-indexed constraint analysis. This analysis has proposed the same ranking for CəC and CV as for reduplicative morphemes in Squamish. I exemplify the following tableau to show how co-indexed constraint analysis gives the wrong optimal output.

173. Squamish CəC and CV reduplicative morphemes – co-indexed constraint analysis (Bar-el, 2000, cited in Downing, 2006: 247).

/RED$_{AFX}$ -k'^{w}ay?/	*STRUC	*V-Place	MAX-BR ROOT	NO CODA	MAX-BR AFX
a.☹k'^{w}a-k'^{w}ay?	*	*!		*	**
b.💣k'^{w}ə-k'^{w}ay?	*			*	***
c. k'^{w}ay?-k'^{w}ay?	*	*!		**	
/RED$_{ROOT}$ -k'^{w}ás/					
d.☞k'^{w}əs-k'^{w}ás	*		*(ə)	**	
e. k'^{w}á-k'^{w}as	*	*!	**	*	*
f. k'^{w}ás-k'^{w}as	*	*!		**	

In the co-indexed constraint analysis shown above, the same constraint ranking optimizes a correct output for the CəC reduplicative morpheme, (d). With the same constraint used to account for the CV reduplicative morpheme, the analysis wrongly optimizes the optimal output, (b). Two suggestions have been proposed by Downing (2008) to solve the problem arising in the analysis. These suggestions are:

(1) Reversing the morphological labelling, i.e. from RED_{AFX} to RED_{ROOT} and vice versa and ranking the constraint *V-Place lower down in the hierarchy.

(2) Reversing the ranking of MAX-BR-ROOT and MAX-BR-AFX.

As claimed by Downing (ibid.: 7), the two solutions proposed to repair the co-indexed constraint analysis give the correct results. However, some problems occur when those solutions are applied. If solution (1) were to apply to the analysis, the result would conflict with the cross-Salish requirement that roots have minimal CVC, while affixes can violate this constraint. Although the correct results can be obtained when we reverse the ranking of MAX-BR-ROOT and MAX-BR-AFX, it is at the expense of violating what is claimed to be the universal ranking of these two constraints.

Since the two solutions proposed do not work well to account for CəC and CV reduplicative morphemes in Squamish, the best way to solve the problems is via co-phonology analysis. In co-phonology analysis, each of the reduplicative morphemes is labelled as a Root, that is $Root_1$ and $Root_2$. These Roots are analysed in a distinct constraint ranking where the same constraints apply for $Root_1$ and $Root_2$ but they are ranked differently in the hierarchy. The constraint rankings for $Root_1$ and $Root_2$ are as follows:

174. Co-phonology rankings for CəC and CV reduplicative morphemes in Squamish (Downing, 2006, 2008).

(a) $Root_1$ co-phonology: *V-PLACE >> MAX-BR >> NOCODA

(b) $Root_2$ co-phonology: NOCODA >> MAX-BR >> *V-PLACE

The following tableaux demonstrate how the above co-phonology constraint rankings are analysed:

175. Co-phonology analysis in Squamish reduplication.

(a) $Root_1$ – CəC reduplicative morpheme in Squamish.

/RED -k'ʷás/	MORPH -SYLL	*V-PLACE	MAX-BR (V-PLACE, SEG)	NOCODA
a.☞k'ʷəs - k'ʷás	*		*	**
b. k'ʷá- k'ʷas	*	*!	*	*
c. k'ʷás - k'ʷas	*	*!		**

(b) $Root_2$ – CV reduplicative morpheme in Squamish.

/RED - k'ʷayʔ/	MORPH-SYLL	NOCODA	MAX-BR (V-PLACE, SEG)	*V-PLACE
d.☞k'ʷa - k'ʷayʔ		*	**	*
e. k'ʷə - k'ʷayʔ		*	***!	
f. k'ʷayʔ- k'ʷayʔ		**!		*

As shown in 175(a), by ranking *V-PLACE higher in the hierarchy, this constraint optimizes candidate (a) with vowel reduction in the CəC reduplicative morpheme. Meanwhile, in 175(b), the markedness constraints NOCODA and MAX-BR ensures only the vowel in the base is copied exactly into the reduplicative morpheme. It is clear now that the co-phonology analysis as demonstrated above can account well for the two patterns of reduplication in Squamish.

According to Downing (2006: 248), the two reduplicative morphemes in Squamish are best treated by applying co-phonology analysis – a distinct constraint ranking. The reasons to prefer co-phonology over co-indexed analysis are, first, because the reduplicative morphemes CəC and CV with identical categories show different patterns in markedness reduction. As we saw in the above tableaux, a vowel in one of the reduplicative morphemes in Squamish must be a schwa, no matter what the vowel in the base is, as in the CəC reduplicative morpheme. In the other pattern however, i.e. CV, a vowel in this reduplicative morpheme can be any vowel, depending on what vowel in the base is. The difference in vowel reduction in the two reduplication patterns is analysed in co-phonology by reversing the ranking of NOCODA and *V-PLACE in the hierarchy, as co-phonologies allow markedness reversals. Meanwhile the second reason is that co-phonology is the more powerful theory when the morphological category and degree of markedness are not the ones expected in the language.

Besides Squamish, Ilokano is another language which also has two reduplication patterns and co-phonology is the best solution to explain those patterns too. As illustrated in Hayes and Abad (1989), Ilokano also has two types of reduplication. The two patterns of reduplicative morphemes in Ilokano could be adequately explained by proposing co-phonology. In co-phonology analysis, the heavy reduplicative morpheme in Ilokano is explained by the tendency for prosodic constituents to be of maximal size, while the light reduplicative morpheme is explained by the opposing tendency for some prosodic morphemes to have an unmarked structure (cf. Downing, 2006: 12). As claimed in Downing (ibid.), these opposing tendencies can be accounted for in co-phonologies analysis. Before we go to the analysis tableaux, let us consider examples of the two reduplication patterns:

176. (a) Heavy reduplication (Hayes and Abad, 1989: 357, cited in Downing, 2006: 11).

kaldíŋ 'goat'	kal-kaldíŋ 'goats'
púsa 'cat'	pus-púsa 'cats'
sáŋit 'to cry'	ʔag-saŋ- sáŋit 'is crying'
trabáho 'to work'	ʔag-trab-trabáho 'is working'

(b) Light reduplication:

bunéŋ 'kind of knife'	si-bu- bunéŋ 'carrying a *bunéŋ*'
sáŋit 'to cry'	ʔagin-sa- sáŋit 'pretend to cry'
pandilíŋ 'skirt'	si-pa-pandilíŋ 'wearing a skirt'
trabáho'to work'	ʔagin-tra-trabáho 'pretend to work'

As MBT has been applied to resolve the opposing cross-linguistic tendency, as in the two Ilokano reduplication patterns, it is proposed that the reduplicative morphemes – light and heavy – are affixes (Downing, 2006: 245). Reduplicative morphemes are treated as affixes, which means that both reduplicative morphemes in light and heavy reduplication must be monosyllabic in size, as required by the MORPHEME SYLLABLE CORRELATION constraint in (34). As we can see in (176), the two reduplicative morphemes can be light and heavy monosyllabic in size. Therefore, the difference in markedness of the two patterns of reduplicative morphemes, as claimed by Downing (2006: 245), is accounted for by proposing that each reduplication pattern introduces a distinct co-phonology. The following tableaux exemplify the analysis of Ilokano light and heavy reduplication in a version of MBT:

177. (a) Ilokano light reduplication (from Downing, 2006: 245).

MORPH-SYLL, NOCODA, *VV >> MAX-BR

/Heavy RED-trabaho/	MORPH-SYLL	NOCODA	*VV	MAX-BR
a. trab-trabaho	*	*!		***
b. tra:-trabaho	*		*!	****
c. trabaho-trabaho	**!			
d.☞tra-trabaho	*			****

(b) Ilokano heavy reduplication (from Downing, 2006: 246)

MORPH-SYLL >> MAX-BR >> NOCODA, *VV

Light RED- trabaho	MORPH-SYLL	MAX-BR	NO CODA	*VV
a.☞trab-trabaho	*	***	*	
b. tra:-trabaho	*	****!		*
c. trabaho-trabaho	**!			
d. tra-trabaho	*	****!		

As can be readily seen in the tableaux, the MORPH-SYLL constraint is highly ranked in both types of reduplication. Since the reduplicative morphemes are claimed to be affixes, the MORPH-SYLL constraint must be highly ranked in both types of Ilokano reduplication. However the other constraints, NO CODA, MAX-BR and *VV, need to change their positions in the ranking depending on the reduplicative shape. In order to allow light reduplication to be chosen as a target shape, the NOCODA and *VV constraints must be ranked higher in the ranking compared to the heavy reduplication ones. This is because the reduplicant syllable in light reduplication only consists of CV elements. Therefore, NOCODA must be ranked higher, so that any heavy reduplicant can be ruled out.

We have just seen two examples from Squamish and Ilokano reduplication. It has been shown that by using co-phonologies, the two patterns of reduplication in the two languages, i.e. CəC and CV in Squamish and light vs. heavy reduplication in Ilokano, can be accounted for straightforwardly. As discussed, co-phonologies are able to account for the differences in markedness of the two patterns of reduplicative morphemes in Squamish and Ilokano, as this theory allows markedness reversals. Since Malay also has two patterns of reduplicative morphemes, light and heavy, co-phonology analysis will be used to analyse them.

In co-phonology, a phonological function is associated with a morphological construction (Orgun, 1996, 1998; Inkelas, 1998, 2008; Inkelas and Zoll, 2005, 2007; Antilla, 2002; Downing, 2008; and many others). Scholars who have applied co-phonology analysis to analyse reduplication (e.g.: McCarthy and Prince. 1995; Myer and Carleton, 1996; and many others) claim that in morphological construction, specifically the Faithfulness constraint for example, Faith-BR can be interleaved into a fixed ranking of markedness constraints (cf. Downing, 2008). This allows some constructions to have a more (or less) marked structure in the output than others (ibid.).

By applying co-phonology to account for Malay reduplication (this includes SM and PD), the heavy reduplicative morpheme is explained by the tendency for prosodic constituents to be of maximal size, while the light reduplicative morpheme is explained by the opposing tendency for some prosodic morphemes to have unmarked structure and be distinctly ranked. These opposing tendencies can be accounted for by proposing that the heavy and light reduplicative morphemes are associated with distinct constraint rankings, i.e. co-phonology.

Considering both reduplicative patterns, including reduplication in SM and one of the Malay dialects, Perak, I shall now begin the analysis. Since Malay reduplication contains two types of reduplicative morphemes, as in Squamish and Ilokano, I will follow work such as that of Inkelas (1998), Inkelas and Orgun (1998), Inkelas and Zol (2000, 2005) and Orgun (1996, 1998) to propose that different prosodic morphemes in the same language can have different optimal output. They claim further that, although in the same language, the reason that different optimal output can be obtained is because each morphological construction is potentially associated with a different co-phonology: a construction-specific constraint ranking (Downing, 2006: 13). Therefore, light and heavy reduplication patterns in Malay must be accounted for by proposing co-phonology: a specific construction constraint ranking.

For ease of reference, since the light reduplicative morpheme in PD is exactly the same as the one in SM, I will combine the examples from SM and PD:

178. Two patterns of reduplicative morpheme in Malay.

(a) Light reduplicative morphemes.

(i)	/budaʔ/ 'child'	[budaʔ]	[bə-budaʔ] RED-child 'all kinds of children'
(ii)	/kun.ʧi/ 'key'/ 'lock'	[kun.ʧi]	[kə-kunʧi] RED-key 'password'
(iii)	/taŋ.ga/ 'stairs'	[taŋ.gə]	[tə-taŋgə] RED-stairs 'neighbours'
(iv)	/ka.ʧaŋ/ 'bean'	[ka.ʧaŋ]	[kə-kaʧaŋ] RED-bean 'beans'

(v)	/dulu/ 'long ago'	[dulu]	[də-dulu] RED-long ago 'very long ago'

(b) Heavy reduplicative morphemes.

(i)	/pətaŋ/ 'evening'	[pətaŋ[	[pəm-pətaŋ] RED-evening 'every evening'
(ii)	/ʤaraŋ/ 'seldom'	[ʤaʁaŋ]	[ʤəɲ-ʤaʁaŋ] RED-seldom 'very seldom'
(iii)	/baraŋ/ 'thing'	[baʁaŋ]	[bəm-baʁaŋ] RED-thing 'all kinds of things'
(iv)	/sikit/ 'a little'	[siket]	[səʔ-siket] RED-a little 'very little'
(v)	/gəlap/ 'dark'	[gəlap]	[gəʔ-gəlap] RED-dark 'very dark'
(vi)	/budaʔ/ 'child'	[budaʔ]	[bəʔ-budaʔ] RED-child 'all kinds of children'

Before we establish the constraint analysis, it is worth discussing first the relevant constraints to account for the two patterns of reduplicative morpheme partial reduplication. Following the analysis of Ilokano reduplication, reduplicative morphemes in both light and heavy reduplication in Malay are considered as affixes. They must therefore obey the MORPH-SYLLABLE CORRELATION as shown in (34) which requires affixes to be monosyllabic.

As proposed in co-phonology analysis, light reduplication is explained by the tendency for some prosodic morphemes to have an unmarked structure. Language typology and the widespread occurrence of processes which avoid codas suggest that the unmarked structure is for syllables not to have codas (cf. Kager, 1999). This unmarked structure is set in the following well-formedness constraint, named NOCODA:

179. **NOCODA** (syllables are open) (Kager, 1999).

$*C]_\sigma$

The above constraint requires that syllables are not allowed to have codas. As a coda is not present in the light reduplicative morpheme, then NOCODA must be highly ranked in the constraint ranking for light reduplicative morphemes. For heavy reduplicative morphemes, NOCODA should be ranked lower in the constraint ranking.

The second constraint which is crucial to accounting for the data is REDUCE. As we observed, a vowel in the two patterns of reduplicative morpheme is a fixed vowel, which is schwa, regardless of what the vowel in the base is.

180. **REDUCE**

Vowels lack quality.

The two constraints we just discussed are both markedness constraints which play important roles. As mentioned above, light reduplication is explained by the tendency for some prosodic morphemes to have unmarked structure. Here, I would like to point out work done by previous scholars regarding unmarked structure in reduplicative morphemes. In work such as that of Alderete et al. (1999), McCarthy and Prince (1994a, 1999) and Steriade (1988), it is said that reduplicative morphemes always illustrate unmarked structure, or what we call the Emergence of the Unmarked (TETU) (cited in Downing, 2006: 41). Those previous works claim that a marked structure is optimal in the base output, and is non-optimal in reduplicative morphemes. In OT, a marked structure in the base can be prevented from occurring in the reduplicative morpheme by ranking the relevant B-R faithfulness constraint beneath (some) markedness constraints (ibid.). I shall demonstrate shortly that this solution works well to produce unmarked structures in light reduplicative morphemes in Malay (see tableau 186).

We shall now discuss the relevant faithfulness constraints to account for the data. Recall that heavy reduplication is explained by the tendency for prosodic constituents to be of maximal size. The maximal size of prosodic constituents is stated in the MAXIMALITY CONDITION below:

181. **MAXIMALITY CONDITION** (Prince, 1985)

Units are of maximal size within the other constraints of their form.

As stated in Downing (2006: 12), this MAXIMALITY CONDITION optimizes the largest reduplicative syllable, which is formalised by the faithfulness constraint, MAX-BR.

182. **MAX-BR**

All segments of the base are contained in the reduplicative morpheme (no partial reduplication).

The faithfulness constraint, MAX-BR, requires that every element in the base has a correspondent in the reduplicative morpheme. This means that this constraint is violated more for light reduplicative morphemes as only a few segments of the base are copied into the reduplicative morpheme. As discussed earlier, the light reduplicative morpheme is a case of unmarked syllable structure. It is an open syllable, omitting any coda consonant. If [ŋ] in the first syllable from the base [taŋ.gə] were copied into the reduplicative morpheme, it would be syllabified as a coda syllable, from which the output would be *[təŋ-taŋge]. In this case, the faithfulness constraint, MAX-BR, must be ranked beneath the markedness constraint NOCODA in the constraint ranking of light reduplicative morphemes – NOCODA >> MAX-BR (see tableau in 186). But then when is a marked structure allowed in the reduplicative morpheme? A marked syllable structure occurs in heavy reduplicative morphemes. Thus, in order to get marked syllable structure in the reduplicative morpheme, MAX-BR must outrank NOCODA: MAX-BR >> NOCODA (see tableau in 187).

Observe that the heavy reduplicative morpheme which occurs in PD poses a fundamental challenge. As pointed out by Ahmad (2005: 182), the partial reduplication in PD and most of the Peninsular Malay dialects[37] does not meet the regularities of phonology and morphology in the language since the clusters of nasal and voiceless obstruents in the reduplicated words do not undergo the regular process, nasal substitution, e.g. [pəm-pətaŋ] – *[pəm-ətaŋ] and [kəŋ-kawan] – *[kəŋ-awan]. This process, which is invisibly active in the environment of the reduplicative morpheme, has posed a challenge to linguists writing about the grammar of Malay. It should be kept in mind that nasal and voiceless obstruent clusters are not permitted in the language. Therefore, this sequence undergoes nasal substitution whereby a voiceless obstruent is deleted, leaving its place of articulation to a nasal (Kager, 1999: 59). In this analysis, the undeleted voiceless obstruent after the nasal segment, as in 178(b), can be accounted for by proposing a constraint called ANCHORING. As the coda consonant in the reduplicative morpheme is determined by the final consonant of the base, the relevant ANCHORING constraint which plays a crucial role here is RIGHT ANCHOR-BR. It can be formally defined as follows:

[37] The author did not mention which Peninsular Malay dialects contain heavy reduplication.

183. **RIGHT ANCHOR-BR** (Kager, 1999: 251)

Any element at the designated periphery of S_1 has a correspondent at the designated periphery of S_2

This constraint requires that the right peripheral edge of the base must coincide with the right peripheral edge of the reduplicative morpheme. As a final consonant of the base will be copied into the coda position in the reduplicative morpheme, so RIGHT ANCHOR-BR is ranked higher in the hierarchy. This constraint is, however, ranked lower in light reduplication, because the right edge of the reduplicative morpheme does not coincide with the right edge of the base. I demonstrate the following correspondence diagram for obedience of RIGHT ANCHOR-BR. The violation of this constraint has already been demonstrated in (97).

184. Correspondence diagram for obedience of RIGHT ANCHOR-BR.

```
b u d a ʔ        base
        |
    b ə ʔ        reduplicative morpheme
```

As already mentioned, co-phonology is used in this present study to analyse the two reduplicative morphemes. I establish the following co-phonology rankings for Malay.

185. Co-phonology rankings for Malay.

(a) Light reduplicative morpheme co-phonology:

NOCODA >> REDUCE >> MAX-BR >> RIGHT ANCHOR-BR

(b) Heavy reduplicative morpheme co-phonology:

RIGHT ANCHOR-BR >> REDUCE >> MAX-BR >> NOCODA

Now I present the analysis of light and heavy reduplication in the following tableaux by considering all the constraints discussed above:

186. Light reduplication.

/Light RED-budaʔ/	MORPH -SYLL	NO CODA	REDUCE	MAX -BR	RIGHT ANCHOR -BR
a. bu-bu.daʔ	*		*!	***	*
b.☞bə-bu.daʔ	*			****	*
c. bud- bu.daʔ	*	*!	*	**	*
d. bu.daʔ-bu.daʔ	**!	*	*		
d. bəʔ-bu.daʔ	*	*!		***	

As discussed earlier, the unmarked syllable structure for light reduplicative morphemes is obtained by ranking the relevant markedness constraint, in this case NOCODA, above the faithfulness constraint, MAX-BR. Thus, candidates (c), (d) and (e) are ruled out. The remaining candidates in the hierarchy are now (a) and (b). Since REDUCE is high-ranked in the constraint ranking, candidate (a) is ruled out as the vowel in the reduplicative morpheme is a high vowel. Candidate (b) thus emerges as the optimal output.

187. Heavy reduplication.

/Heavy RED- buda?/	MORPH -SYLL	RIGHT ANCHOR -BR	REDUCE	MAX -BR	NO CODA
a. bu-bu.da?	*	*!	*	***	
b. bə-bu.da?	*	*!		****	
c. bud-bu.da?	*	*!	*	**	*
d. bu.da?-bu.da?	**!		**		*
e.☞bə?-bu.da?	*			***	*

As shown in (187), when RIGHT ANCHOR-BR ranks high in the hierarchy, it is optimal for the reduplicative morpheme to copy the rightmost segment of the base. As a glottal stop is the rightmost segment in the base, it must be present in the rightmost position of the reduplicative morpheme. Therefore candidates (a), (b) and (c) are ruled out. Next, candidate (d) with total reduplication is ruled out by the markedness constraint, REDUCE, because the vowels [u] and [a] are not reduced to a schwa. The remaining candidate, (e), which does not violate the highest constraint, RIGHT ANCHOR-BR, or REDUCE emerges as the optimal output. Since this candidate is not a total reduplication, it violates MAX-BR three times. These violations are however not significant as the optimal output has already been determined.

The data /buda?/ we analyse in the above tableaux, (186) and (187), end with a glottal stop. As we observed, two types of final consonant in the base are involved in heavy reduplication: (1) stop consonants and (2) nasal consonants (see 170(b) – heavy reduplication data in PD). Stop consonants have already been discussed above. Now, we turn to analyse other data for heavy reduplication where the base ends with a nasal consonant. With the same constraint ranking that I established in (187) above, for heavy reduplicative morphemes, I now demonstrate how a base ending with a nasal consonant copies the nasal consonant into the coda position in the reduplicative morpheme.

188.

/Heavy RED-pətaŋ/	MORPH -SYLL	RIGHT ANCHOR -BR	REDUCE	MAX -BR	NO CODA
a. pə-pə.taŋ	*	*!	*	***	
b. pət-pə.taŋ	*	*!		****	
c. pə.taŋ-pə.taŋ	**!		**		*
d.☞pəm-pə.taŋ	*			***	*

As shown in the above tableau, the final consonant of the base is a nasal consonant, /ŋ/. This consonant is copied into the heavy reduplicative morpheme and is syllabified as a coda consonant. When occupying the coda position of the reduplicative morpheme syllable, the velar nasal /ŋ/ assimilates to the following onset consonant, [p]. The velar nasal then becomes [m] after the assimilation process, while [p] remains in the onset syllable of the base. Nasal substitution as mentioned above is invisibly active in the environment of reduplicative morphemes. This situation is captured under the identity faithfulness constraint which is RIGHT ANCHOR-BR. Since the optimal output in (d) is not ill-anchored, it is chosen as the victor. As we can see in the above tableau (188), RIGHT ANCHOR-BR is also crucial to rule out other non-optimal candidates, hence it must be highly ranked.

The above analysis of Malay partial reduplication, which covers SM and PD, clearly shows that co-phonology can handle the two patterns of reduplicative morphemes – light and heavy. As can be seen, light and heavy reduplication patterns are analysed differently according to the specific-construction ranking or co-phonology. In co-phonology, the two different patterns of reduplicative morphemes are labelled as 'light' and 'heavy' and are ranked in a distinct constraint ranking, as I demonstrated in (186), (187) and (188). I believe that it is difficult if not impossible for the two patterns of reduplicative morpheme to be analysed by earlier derivational phonological approaches. Unlike co-phonology, which is developed and motivated in OT, it is a non-derivational theory. It formalizes the link between particular morphemes and particular phonological patterns in a non-derivational way (Downing, 2008). If a CV template was, for example, introduced to account for the two types of Malay partial reduplication, the heavy reduplicative morpheme would be impossible to resolve. As claimed by Moravcsik (1978), there are two problems with CV strings. The reduplicative syllable fails to match with the base because the reduplicative morpheme contains either more or fewer segments than the corresponding base. Hence, the CV reduplicative template, as proposed by Syed Ahmad (1998) and Ahmad (2000), fails to account for all the patterns of reduplicative morphemes in Malay partial reduplication.

As shown, the analysis of co-phonology can account well for the two reduplicative morphemes with identical categories though they show different patterns of markedness reduction. In co-phonology, the different patterns of markedness reduction are accounted for by reversing the markedness constraint in the ranking. This can be done by ranking the relevant markedness constraint above the faithfulness constraint in light reduplication, while in heavy reduplication the markedness constraint is ranked below the faithfulness constraint. This application is called 'Markedness Reversal', and is allowed in co-phonology. As demonstrated in the above tableaux, in (186), (187) and (188), the syllable markedness constraint

NOCODA can only be satisfied in light reduplication, not in heavy reduplication. Therefore, this syllable markedness constraint should be ranked higher in light reduplication to ensure the reduplicative morpheme contains no coda. In contrast to light reduplication, NOCODA is ranked lower in the heavy ones, since the reduplicative ends with a consonant. Therefore, by ranking NOCODA lower, a second consonant, which occupies the coda position of the reduplicative morpheme, is allowed to be present.

5.5 Affixal Reduplication

In this section I will discuss the process of reduplication in prefixed words. As claimed in previous work, for instance Onn (1980), Syed Ahmad (1998), Teoh (1994) and Ahmad (2005), prefixed words in Malay can also be reduplicated. This means that, as a productive morphological word formation, reduplication can be applied to both roots and affixed words. My observations from the corpus show that both single and multiple prefixed words in Malay can also be reduplicated. Although the existence of reduplicated words with multiple prefixes (22 words) does not turn up very frequently in the corpus compared to single prefixed words (590 words), it does however exist in Malay vocabulary as another form of reduplication in the language (see Appendix K). Therefore it is included for analysis in this study.

In what follows, I shall first discuss the process of affixation by presenting some relevant examples from previous work. Onn's (1980) work on affixation is chosen here as this work can be considered one of the comprehensive studies on affixation. Affixation in Onn has been broadly discussed more than other previous scholars' work. Before we are in a position to discuss affixal reduplicated words by referring to data obtained from the DBP-UKM corpus database, it is worth knowing what the process of affixation in Malay is. This will be demonstrated by reviewing Onn's work and the relevant phonological aspects found in affixed words, since surface forms produce certain effects in reduplicated words:

189. (a) V-final prefixes + V-initial stems (from Onn, 1980).

/di+ubah/	[di.ʔu.bah]	to alter (pass.)
/di+ambil/	[di.ʔam.bel]	to take (pass.)
/di+olah/	[di.ʔo.lah]	to describe (pass.)
/di+iris/	[di.ʔi.res]	to cut/to slice (pass.)
/di+aʤar+i/	[di.ʔaʤa.ri]	to teach (pass.)

(b) C-final prefixes + V-initial stems.

/bər+adu/	[bə.ra.du]	to compete
/pəŋ+usik/	[pə.ŋu.seʔ]	teaser
/pəŋ+asut/	[pə.ŋa.sot]	instigator
/məŋ+aʤar/	[məŋ.a.ʤa]	to teach
/məŋ+awas/	[mə.ŋa.was]	to supervise

The phonological facts displayed above can be summarised as follows: In the examples in 189(a), no matter what the root is, the prefix /di+/ will always be realised as [di]. When /di+/ is attached to a V-initial root, a glottal stop will be epenthesized to break up the derived vowel sequences (Onn, 1980). Meanwhile, in 189(b), the nasal segment in the prefix is maintained as a velar nasal and will be resyllabified to the following vowel, passing across the prefix boundary, likewise the /r/ segment in the prefix /bər+/.

As we can see in 189(a), the role of the glottal stop between the V-final prefix and the V-initial root is primarily to break up vowel sequences at the prefix-root juncture. Note that the glottal stop is a default consonant which is commonly used to break up any vowel sequences in Malay. Epenthesis of the glottal stop as a default consonant can be seen easily in Arabic loanwords, for example /maaf/ 'forgive', /ʤumaat/ 'Friday', /taat/ 'obey' and /saat/ 'second/minute'. These words are realised as [maʔaf], [ʤumaʔat], [taʔat] and [saʔat] (Ahmad, 1990). Additionally, glottal epenthesis is a result of the syllable structure in the language where a syllable must begin with an onset. In other words, a syllable without an onset is prohibited from emerging in the surface form.

Notice that the examples in 189(a) show that a glottal stop is epenthesized between the V-final prefix /di+/ and V-initial roots and is by no means to provide an onset to the syllable. In the examples in 189(b), all the roots begin with a vowel which means no onset, a glottal stop is however not inserted. Besides, by epenthesizing a glottal stop to provide an onset to a syllable, there is another way that a syllable can get its onset. The initial vowel in stems gets its onset from the preceding consonant in accordance with the Minimal Onset Satisfaction Principle (Roca, 1994). The final consonant of the prefix is syllabified to the initial vowel of the stem to provide an onset. This is in parallel to the above examples when C-final prefixes concatenate to V-initial stems, e.g. /bər+adu/ 'to compete' becomes [bə.ra.du]. The consequence of syllabifying the final consonant of the prefix to the initial vowel of the stem is that the prefix syllable produces a light syllable. The light syllable in the prefix then does not fit the size of the prosodic morpheme for affixes [σμμ] in Malay prefixation, as I maintained earlier. In order to ensure the prefix satisfies the size of that prosodic morpheme, as I have also argued, the best solution is that the vowel in the prefix needs to be lengthened.

It must be noted that Onn's (1980) analysis also observed that a vowel in the prefix would undergo compensatory lengthening. In his analysis, however, compensatory lengthening is due to: (i) the deletion of /r/ in a coda position. When [r] in a coda position is deleted, the additional mora to the coda consonant is linked to the preceding vowel. Therefore, the vowel gets lengthened. This process in which a coda consonant has an additional mora has been termed 'Weight by Position' by Hayes (1989, cited in Durand and Katamba, 1995: 127); (ii) when another consonant follows (ibid.: 16) and a vowel precedes a nasal and a consonant. I am not in agreement with this kind of lengthening and I shall argue against this shortly. The following are examples of each case:

190. (a) Lengthening due to r-deletion (from Onn, 1980).

/sabar/	[saba:]	patient
/lebar/	[leba:]	wide
/kərtas/	[kə:tas]	paper
/tər+duduk/	[tə:dudoʔ]	to sit unintentionally
/bər+kata/	[bə:katə]	to utter
/bər+dəbat/	[bə:dəbat]	to compete

(b) Lengthening of the vowel that precedes a nasal and a consonant (Onn, 1980: 38):

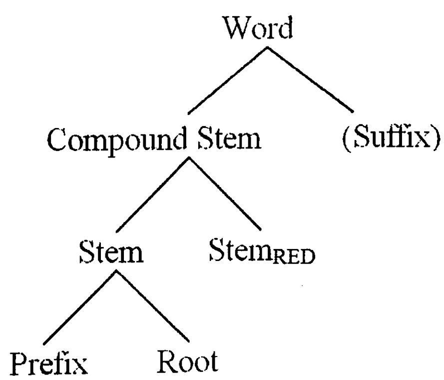
/baŋga/	[ba:ŋgə] 'to be proud'
/gurindam/	[guri:ndam] 'a kind of poetry'
/tumbaŋ/	[tu:mbaŋ] 'to fall'

/di+bantu/	[diba:ntu] 'to help (pass.)'
/di+ʔaŋgap/	[diʔa:ŋgap] 'to assume (pass.)'
/məŋ+daki/	[mə:ndaki] 'to climb (active)'
/məŋ+bayar/	[mə:mbayar] 'to pay (active)'

In this present study, I postulate another compensatory vowel-lengthening situation in Malay, in addition to the two examined by Onn (1980). This argument must, however, be differentiated from Onn's claims. According to Onn (ibid.), the vowel in the prefix will be lengthened although the final consonant is retained, as in [mə:ndaki] and [mə:mbayar]. In my view, vowel lengthening, as proposed by Onn, presents a problem in explaining why the vowel in the prefix should be lengthened whereas the final consonant is not deleted. There is no strong phonological motivation to account for this fact. Given the insight from Prosodic Morphology, where template morphemes are expressed in terms of prosodic units, I claim that vowel lengthening in the prefix is a kind of conspiracy in the language to satisfy PROSODIC STEM (35). Vowel lengthening is seen as a way to satisfy PROSODIC STEM because the other segments in the word cannot be linked to the empty mora, as it causes the association line to be crossed. In that sense, it is hard to agree with Onn (1980) that vowels are lengthened, as in [mə:ndaki] and [mə:mbayar]. I shall argue this based on the ideas of MBT where morphemes tend to have a 'canonical form' or a 'general phonemic shape' (Hockett, 1966a; Nida, 1949, cited in Downing, 2006). As I proposed earlier, the canonical shape for prefixes in Malay is monosyllabic bimoraic (123) – none is less or more. By lengthening the vowels in [mə:ndaki] and [mə:mbayar], this causes the prosodic size of the affix to exceed the canonical shape. Vowel lengthening, as I have demonstrated in this study can, however, only be applied in restricted environments. The environments where vowel lengthening occurs are: (1) when a root is concatenated with V-final prefix /di+/ to satisfy PROSODIC STEM, (35), as illustrated in (124); (2) when /r/ in the prefix /tər+/ and /bər+/ is deleted in order to satisfy ALIGN-RHOTIC, as in (133); (3) when the final consonant in the prefix is syllabified to the following syllable of the stem that begins with a vowel, as in 189(b).

Now we shall proceed to the analysis. We have seen that total reduplication in Malay resembles compounding, as has also been claimed in other languages. In Section 5.3, I have demonstrated data for total reduplication showing that reduplication in Malay is best considered as a type of compounding rather than affixation, as the reduplicative morphemes in total reduplication contain more than one syllable. As I have already mentioned in the discussion, work such as that of Eulenberg (1971), Niepokuj (1991), Inkelas (2005), Inkelas and Zoll (2000, 2005) and Downing (2006) claims that each half of the complex has the same morphological category, likewise in Malay. Thus, I illustrate the structure of reduplicative compounds for Malay reduplication:

191. Reduplicative compound structure for Malay.



I briefly present some of the relevant examples of affixal reduplicated words from the corpus. As mentioned earlier in this section, both single and multiple prefixes can be reduplicated in Malay. Thus, the data presented below are divided into these two types of prefixes:

192. (a) Data for reduplicated words with single prefix (from the DBP-UKM corpus).[38]

(i)	/tər-ʧari/ TRANS-find 'to find'	[tərʧari-ʧari] TRANS-find-RED 'able to find repeatedly'
(ii)	/di-alu-kan/ PASS.PRF-welcome-CAUS.SUF 'welcomed'	[di-ʔalu-ʔalu-kan] PASS.PRF-welcome-RED-CAUS.SUF 'very much welcomed'
(iii)	/mə-ruŋut/ VERBL.PRF-grumble 'to grumble'	[məruŋut-ruŋut] VERBL.PRF-grumble-RED 'to grumble repeatedly'
(iv)	/məm-bandiŋ/ VERBL.PRF-compare 'to compare'	[məm-bandiŋ-bandiŋ] VERBL.PRF-compare-RED 'to compare repeatedly'

(b) Data for reduplicated words with multiple prefixes (from the DBP-UKM corpus)[39]

(i)	/məm-pər-bəsar-kan/ VERBL-NOML-large-CAUS.SUF 'to cause to enlarge for'	[məm-pər-bəsar-bəsar-kan] VERBL-NOML-large-RED-CAUS.SUF 'to cause to enlarge for repeatedly'
(ii)	/məm-pər-main-kan/ VERBL-NOML-play-CAUS.SUF 'to cause to play for'	[məm-pər-main-main-kan] VERBL-NOML-play-RED-CAUS.SUF 'to cause to play for repeatedly'
(iii)	/di-pər-main-kan/ PASS.PRF-NOML-play-CAUS.SUF 'is mocked for'	[di-pər-main-main-kan] PASS.PRF-NOML-play-CAUS.SUF 'is mocked for repeatedly'

[38] See Appendix L for more data.
[39] See Appendix K for more data.

As we can see in the above examples, unlike in total and partial reduplication where roots are the input for the reduplicative morphemes, in affixal reduplication, in contrast, the affixed words are the input for the potential possible candidates. In MBT analysis, the input for the reduplicated words in this case is stems, which consist of prefix and root (191). Therefore, stems must at least be disyllabic, as required by the MORPHEME-SYLLABLE CORRELATION constraint (34). As explained in Chapter 2 subsection 2.2.4, this constraint is a corollary to the PROSODIC STEM constraint, as illustrated in (35). In MBT, the ALIGN PROSODIC STEM constraint defines the prosodic stem as the base for reduplicative morphemes (Downing, 2006: 159). The constraint, ALIGN PROSODIC STEM, can be formally defined as follows:

193. **ALIGN PROSODIC STEM** (Downing, 2006).

Align (L, RED; R, Prosodic Stem)

As formally defined by the ALIGN PROSODIC STEM constraint, a prosodic stem is the base for a reduplicative morpheme. Nevertheless, prefixes are not allowed to be copied into the reduplicative morpheme. As stated in Downing (2006: 159), prefixes are not copied except by reason of satisfying the minimality condition. It can be summarised that if the reduplicative morpheme has already satisfied the minimality condition, then copying the prefix is not plausible in this case. This is true for what I have observed in the corpus. Prefixes are not copied into reduplicative morphemes, only roots are copied. This can clearly be seen in the following examples taken from the corpus:

194.

(i)	/bər-pilih/	[bər-pilih-pilih]
	PASS.PRF-choose	PASS.PRF-choose-RED
	'to choose'	'to choose repeatedly'
(ii)	/məm-bəlek/	[məm-bəlek-bəlek]
	VERBL.PRF-scrutinise	VERBL.PRF-scrutinise-RED
	'to scrutinise'	'to scrutinise repeatedly'
(iii)	/bər-katə/	[bər-katə-katə]
	PASS.PRF-utter	PASS.PRF-utter-RED
	'to utter'	'to utter repeatedly'
(iv)	/tər-bawa/	[tər-bawa-bawa]
	TRANS-carry	TRANS-carry-RED
	'to carry unintentionally'	'to carry unintentionally repeatedly'
(v)	/məɲ-ʤadi/	[məɲ-ʤadi-ʤadi]
	VERBL.PRF-become	VERBL.PRF-become-RED
	'to become'	'to become seriously'

It is clear from the data that only roots are copied into reduplicative morphemes. Since the roots already consist of disyllabics and therefore the disyllabic minimality condition in the language is satisfied, the prefixes are thus not copied. It is crucial to demonstrate that reduplication with multiple prefixes also has the same process as in single prefix reduplication.

This idea of MBT, where prefixes are copied only to satisfy minimality, is important when we account for monosyllabic roots. As shown in the section on total reduplication (Section 5.3), base and reduplicative morphemes are also subject to the minimality condition in the language where a word must contain at least two syllables. In total reduplication however, this requirement cannot be met for monosyllabic roots.

Thus, this situation has been satisfactorily resolved in MBT by the HEADS BRANCH constraint (31), as I have shown in (162) and (163). In contrast to total reduplication, the requirement for disyllabicity minimality in affixal reduplication can be satisfied even in monosyllabic roots. This can be explained by the constraint I have just introduced above, i.e. ALIGN PROSODIC STEM (193). This constraint ensures that prefixes are copied into reduplicative morphemes to satisfy minimality. I will demonstrate the analysis of this shortly. Before that, let us consider the example below:

195. Data for reduplicated words with monosyllabic root.[40]

(i)	/mə-ŋə-pam/	[mə-ŋə-pam-**ŋə**-pam]
	VERBL.PRF-pump	VERBL.PRF-pump-PRF-RED
	'to pump'	'to pump repeatedly'

It is clear now, from the examples, that the minimality requirement in a language can be satisfied in monomorphemic roots in affixal reduplication. The minimality requirement is met in monomorphemic roots, as we can see where the prefix is also copied in the reduplicative morpheme. Since there is no prefix in total reduplication, so the reduplicative morpheme only contains one syllable. In affixal reduplication, copying the prefix into the reduplicative morpheme is a plausible way, as the prefix is also part of the base for the reduplicative morpheme. This situation can be accounted for by ALIGN PREFIX (196), ensuring that prefixes are copied to word-initial position only.

196. **ALIGN PREFIX**

AlignL (Prefix, ProsodicWord)

[40] Surprisingly, this is the only word found in the five million words of data that I examined.

Considering all the relevant constraints discussed above, I now establish the following tableau to account for reduplication with monosyllabic roots. The schematic ranking to account for this is: DEP-IO, DEP-BR >> ALIGN PRSTEM >> ALIGN PREFIX >> MAX-BR.

197.

/mə-ŋə-pam-STEM$_{RED}$/	DEP-IO, DEP-BR	ALIGN PRSTEM	ALIGN PREFIX	MAX-BR
a.☞[mə.ŋə.pam]-[ŋə.pam]			*	**
b. [mə.ŋə.pam]-[pam]		*!		****

Candidate (b) is ruled out because the reduplicative morpheme only contains one syllable. As reduplication is considered to be compounding, so it is must contain at least two syllables, one for each morpheme, as required by the PROSODIC STEM constraint. By copying the final syllable of the prefix into the reduplicative morpheme, candidate (a) does not violate PROSODIC STEM, as the reduplicative morpheme now contains two syllables. However, it does violate ALIGN PREFIX. As candidate (b) was ruled out earlier, then the violation of this constraint is not significant anymore and candidate (b) emerges as the winner.

Regarding the copying of prefixes into reduplicative morphemes, this is only allowed to satisfy minimality, as claimed in MBT, and I would like to add one important point which I found out from the corpus. Given the data for Malay reduplication, prefixes are not only copied to meet the disyllabicity minimality condition in the language, they are also copied into reduplicative morphemes to provide an onset to a syllable. As was discussed, the deletion of a voiceless obstruent in the process of prefixation to a nasal final prefix causes the lost of a voiceless segment on the surface. Although the segment has been deleted, it leaves its place of articulation to the nasal segment in the prefix (Kager, 1999: 59). Voiceless obstruent deletion has contributed that the stem root contains no onset and, therefore, as we have discussed earlier, the nasal segment in the prefix will be syllabified to the first syllable of the root stem in order to provide an onset to the syllable. As Onn (1980) also says, 'when the nasal assimilation deletes the original onset of the stem, the resulting nasal is treated as the new onset of the root'. Therefore, the reduplicated forms copy the nasal segment in the prefix into the reduplicative morphemes. If the nasal segment was not copied, then the initial syllable of the reduplicative morpheme would not have an onset. Below, I examplify the affixal reduplicated words found in the corpus. The bases are underlined.

198. Data for voiceless obstruent initial roots (from the DBP-UKM corpus).[41]

(i)	/mə-ŋoreʔ/ ACT.PRF-dig 'to dig'	[məŋoreʔ-ŋoreʔ] ACT.PRF-dig-RED 'to dig repeatedly'
(ii)	/mə-maŋgil/ ACT.PRF-call 'to call'	[məmaŋgil-maŋgil] ACT.PRF-call-RED 'to call repeatedly'
(iii)	/mə-nunʤuʔ/ ACT.PRF-show 'to show'	[mənunʤuʔ-nunʤuʔ] ACT.PRF-show-RED 'to show off'
(iv)	/mə-mukol/ ACT.PRF-beat 'to beat'	[məmukol-mukol] ACT.PRF-beat-RED 'to beat repeatedly'

Due to nasal substitution which is applied to break up nasal and voiceless obstruent clusters, reduplicative morphemes in affixal reduplication, as presented in

[41] For more data for reduplicated words with voiceless obstruents initial roots, see Appendix M.

the above examples, copy the nasal segments as the onset of syllables. In MBT, This phonological process that occurs in prefixed reduplicated words can be straightforwardly explained by the constraint called ALIGN ONSET, as defined below:

199. **ALIGN ONSET**

AlignL (Prosodic Stem, σ)

Considering all the constraints that I have used, and the new constraints I have just introduced here, ALIGN PROSODIC STEM, ALIGN PREFIX and ALIGN ONSET, I now establish a new tableau for affixal reduplication.

200.

/məmukol-STEM$_{RED}$/	DEP-IO, DEP-BR	ALIGN ONSET	ALIGN PRSTEM	MAX-BR
a.☞[məmukol]-[mukol]				
b. [məmukol]-[pukol]	*! (BR)			*
c. [məmukol]-[ukol]		*!		*

As shown in the tableau, the losing candidates, (b) and (c), are ruled out since they violate higher constraints. The presence of [p] in the reduplicative morpheme in candidate (b) causes the candidate to violate DEP-BR and IO, since [p] is not contained in both input and base. Candidate (c) violates ALIGN ONSET, as the reduplicative morpheme contains no onset. The solution proposed earlier, where the resulting nasal is treated as the new onset for the deletion of a voiceless obstruent, does not however apply to this candidate. Therefore, it violates ALIGN ONSET. Candidate (a), on the other hand, achieves this solution by syllabifying the nasal segment into the following syllable, obeying ALIGN-ONSET. This candidate is thus the winner.

Recalling the idea of MBT, where prefixes are only copied to satisfy the minimality condition in the language, I shall then make my argument based on the analysis made by Onn (1980), where prefixes (final nasals in prefixes) are copied into disyllabic bases, as shown in the data below:

201. Data for affixal reduplication (Onn, 1980).

Unreduplicated	**Reduplicated**
məmbalas	məmbalas-(m)balasi
məndakap	məndakap-(n)dakapi
məmbidas	məmbidas-(m)bidasi
mənjilat	mənjilat-(n)jilati

Observe too that, from the above examples, the nasal segment in the final prefix will also be repeated in the reduplicant. As I have already discussed, this repeated nasal occurs only if the roots are monosyllabic. In the above examples, however, the reduplicant is carried along with the nasal segment in the prefix although the roots are disyllabic. A question that can be posed is: Why do optional consonants, for example (m) as in [məmbidas-(m)bidasi], appear, even though the roots already have an onset and are disyllabic? I suggest that, in order to argue for the surface form derived from affixal reduplication where a nasal in the final prefix is also copied into the reduplicative morpheme, we should consider every segment in the word. As claimed in Teoh (1994: 28), based on Clement and Keyser (1983: 28), SM belongs to Type III, where CV or CVC are the core syllable structures in the language. According to this core syllable structure, the initial syllable [CCV.C] in the surface form of the reduplicated words above is not included in the core syllable structure. This matter can be plausibly accounted for by applying the analysis from MBT. A number of constraints must be considered in the tableau to account for the syllable structure of [CCV.CV], as proposed by Onn (1980). There are four relevant constraints that can be established here: NO CODA >> $\text{COMPLEX}_{\text{onst}}$ >> *VV. The effects of these constraints ranking are demonstrated in the following tableau.

202. Constraints accounting for the core syllable structure in Malay.

(a) NO CODA → syllables must not have a coda

(b) *COMPLEX → no more than one onset may associate with one syllable

(c) *VV → long vowels are marked (Rosenthall, 1994)

203.

	*NOCODA	*COMPLEX	*VV
a. CVC	*!		
b. ☞CV			
c. CCV		*!	
d. CV:			*

The analysis in the above tableau can account straightforwardly for the core syllable structure of a word in Malay. Candidate (c) cannot be chosen as the optimal output. The reason is that the syllable contains [CCV] elements where it is not the core syllable in Malay. Thus this candidate is ruled out by the constraint *$\text{COMPLEX}_{\text{onst}}$.

As we have already noted, Malay disfavours a syllable without an onset. This, however, does not mean that two onsets, or more than one onset, are allowed to be present in a surface form, as Malay also does not allow complex onset structures in a word. Candidate (b), which satisfies all the constraints in the hierarchy, is chosen as the victor. Candidate (a) violates the highest constraint in the hierarchy, NOCODA, since the syllable contains a coda. Although candidate (d) does not have a coda in the syllables, it does however violate *VV since the candidate does not contain at least two syllables. Therefore, the overcopying of the final nasal in the prefix onto the reduplicative morpheme, as in Onn's examples (201), is considered non-optimal output. Thus, it should not emerge as a surface form in reduplicated words.

The discussion on affixal reduplication above raises some important points. This type of reduplication is analysed as involving compound words where the reduplicative morphemes consist of disyllabics. After investigating affixal reduplication, we are now able to answer the Research Questions (2) relating to reduplication. Why for example is the nasal segment in the prefix copied in [məŋəlap-ŋəlap] and [mənekan-nekani], but not in [məmbandiŋ-bandiŋ]? The answer is that the nasal segment in the prefix, as in [məŋəlap-ŋəlap], is copied into the reduplicative morpheme to satisfy the word minimality condition in the language, i.e. stems must be at least disyllabic. This solution applies to all monosyllabic bases to ensure that the disyllabicity minimality of the language is satisfied. In the word [mənekan-nekani], for example, the nasal segment is copied mainly to provide an onset to the reduplicative morpheme as the base starts with an initial-voiceless obstruent, i.e. [t]. In short, there are two reasons why the nasal segment in the prefix is copied into the reduplicative morpheme in affixal reduplication: (1) to satisfy the disyllabicity minimality of the language; and (2) is to provide a new onset to the initial syllable of the reduplicative morpheme due to the deletion of a voiceless obstruent. From this, there is no reason why a nasal segment in the prefix, as in [məmbandiŋ-bandiŋ], should be copied since the reduplicative morpheme contains a disyllabic and has an onset.

Besides that, the analysis of affixal reduplication supports the claim I made earlier, i.e. that reduplication is compounding, not affixation. We noted that one of the reasons for copying nasal segments in the prefixes into reduplicative morphemes is to ensure that the reduplicative morphemes satisfy the disyllabic minimality requirements. This means that affixal reduplicated words contain more than one syllable. Therefore, they are not best treated as affixation because the size of the reduplicative morphemes is larger than the typical size of an affix.

5.6 Conclusion

The discussion presented above reveals some important points about the process of reduplication in Malay. As we have seen, MBT can satisfactorily account for the issues concerned in Malay reduplication. One of the issues is that reduplication resembles affixation. This has long been claimed by previous Malay scholars, so that the reduplicative morphemes in reduplication have been accounted for as affixes. I have argued in the analysis above, based on the ideas of MBT, that the previous claims pose a problem in accounting for the canonical shape of reduplicative morphemes, particularly in total and affixal reduplication. As discussed, the canonical shape of a prosodic morpheme is defined by the morphological category held by the prosodic morpheme. Thus, if the reduplicative morphemes in total and affixal reduplication were considered as affixes, then we would have a problem, because the sizes of the reduplicative morphemes are larger than monosyllabic – the typical size predicted for affixes. To account for this problem therefore, they (total and affixal reduplication) are no longer considered as affixation but rather as compounding.

Besides that, the issue of minimality requirements in reduplication has also been discussed. This issue has never been touched on in previous Malay studies where reduplication is concerned. In this study, how the word minimality requirement is satisfied in reduplicated words has been given formal attention by adopting MBT within OT analysis. By considering the disyllabicity minimality requirements in the language, we can therefore answer the question mentioned at the beginning of this chapter, concerning why prefixes are carried along when monosyllabic stems (prefixes and monosyllabic roots) are reduplicated, as in 153(a), but not carried along in disyllabic stems, as in 153(b). This is mainly to satisfy the disyllabicity minimality requirements in the language by obeying PROSODIC STEM (35), which demands that the prefix and root have at least disyllabicity in the constraint hierarchy. This is an exciting idea for word minimality when raised in a theory which then leads to a better account of this issue.

In addition to the idea of minimality, MBT can also explain why monosyllabic roots emerge as monosyllabic. In MBT, a constraint called PROSODIC STEM (35) requires that the base and reduplicative morpheme must at least be minimally disyllabic. It explains why monosyllabic roots violate this constraint, whereby the bases do not contain two syllables. Therefore, the reduplicative morpheme would also be monosyllabic, as the bases are. This matter has been well explained by MBT with the constraint named HEAD BRANCH (31) – the minimum size that roots must have (bimoraic).

Partial reduplication reveals that Malay does not only have one pattern of reduplicative morphemes. Considering SM and one of its dialects, i.e. PD, reduplication in Malay has two patterns for reduplicative morphemes – light and heavy reduplication. These two patterns for reduplicative morphemes could plausibly be accounted for by proposing co-phonologies – different constraint rankings. Different co-phonology constraint rankings are able to analyse both patterns adequately and satisfactorily, as this clearly discriminates between light and heavy reduplication. As discussed, heavy reduplication is explained by the tendency for prosodic constituents to be of a maximum size, while light reduplication is explained by the tendency for prosodic morphemes to have an unmarked syllable structure. These opposing tendencies, as shown in the analysis, can neatly be accounted for by distinct constraint rankings – a constraint ranking for each type of reduplication. As co-phonology allows markedness reversals, the relevant markedness constraint used to analyse light and heavy reduplication in SM and PD is NOCODA which can be re-ranked in the constraint rankings.

To sum up, it is clear that MBT developed within OT can explain some complex problems that previous studies failed to explain. Apart from that, some new issues have been raised here that were never been touched on in previous studies, for example that minimality requirements and a compounding process can also be explained by MBT. It is apparent from the above discussion that this theory works well in explaining those new issues.

6. DIALECTAL VARIATION

6.1 Introduction

In Chapter 4, we investigated how nasality and voicing are treated in the process of prefixation in SM. In this present chapter, we shall investigate how they are treated in three selected dialects of Malay i.e. Perak, Kelantan and NS (see Map 1). The investigation of nasality and voicing here includes: (1) nasal and voiceless obstruents, (2) nasal and voiced obstruents and (3) nasal and sonorant clusters. Why were only these dialects chosen to be analysed?

First, these dialects present different forms in their surface representation when treating nasal and voiceless obstruent clusters from SM, either within a root or at prefix-root junctures.[42] As we see in Table 5, a sequence of root-internal nasal and voiceless obstruents in some dialects of Malay, such as Perak, Johor and Kedah,[43] is not resolved by nasal substitution or by any other *NÇ effects, since it is allowed in the surface representation. In some dialects, for example Kelantan and NS on the other hand, a sequence of root-internal nasal and voiceless obstruents is treated differently, in that the sequence is not allowed to surface, even within a root. In these dialects,[44] the sequence would normally be resolved by deleting the nasal segment that the voiceless obstruent precedes.

---

[42] It is important to note that only nasal and voiceless obstruent clusters at prefix-root junctures are found in Malay literature. The same goes for nasal and voiced obstruent clusters. Perhaps the vocabularies of Malay dialects are rather simple. The words used in Malay dialects contain no complicated morphological forming words. This might be why there were no examples of these in previous studies.

[43] Note that three Malay dialects are listed here (i.e. Perak, Johor and Kedah) which permit nasal and voiceless obstruents cluster root-internally. However, in this present study, not all three dialects will be examined, due to time constraints, thus only Perak is chosen for examination.

[44] More examples of disallowing nasal and voiceless obstruent clusters root-internally in these two dialects will be presented in Section 6.3.1 where the issue is discussed.

Table 5:

The differences in root-internal nasal and voiceless obstruent clusters in SM and some of the dialects of Malay: Perak, Johor, Kedah, Kelantan and NS

SM	Perak	Johor	Kedah[45]	Kelantan	NS	
kampoŋ	kampoŋ	kampoŋ	kampoŋ	kapoŋ	kapoŋ	village
sampan	sampan	sampan	sampan	sapan	sapan	small boat
lompat	lɔmpat	lompat	lompat	lopaʔ	lopat	to jump
pintu	pintu	pintu	pintu	pitu	pitu	door
tempat	tempat	tempat	tempat	tepaʔ	tepat	place
taŋkap	taŋkap	taŋkap	taŋkap	takaʔ	takap	to catch
hantu	hantu	hantu	hantu	hatu	hatu	ghost

At a prefix-root juncture, a sequence of nasal and voiceless obstruents in Malay dialects is resolved in the same way as in SM, where this cluster undergoes nasal substitution, for example: /ŋ+puji/ → [muji] 'to praise' and /mŋ+pilih/ → [mmilih] 'to choose' in the Perak and NS dialects, respectively (see 6.3.2). It is worth noting that in Kelantan other strategies are applied to avoid nasal and voiceless obstruent sequences in the dialects, i.e. nasal assimilation and nasalisation.

Second, as we saw in Chapter 4 Section 4.2.1.2, voiced obstruent clusters in SM do not undergo nasal substitution. It ought to be mentioned that the reverse state occurs in some Malay dialects if the obstruent is voiced. In Perak, for example, voiced obstruents following nasals undergo nasal substitution, as in (102). Such a process of nasal substitution between nasal and voiced obstruents in Perak never occurs in SM.

[45] Note that Kedah has a number of sub-dialects, namely Perlis-Pulau Langkawi, Kedah Persisiran, Kedah Utara and Pulau Pinang (Omar, 1993: 190). Every sub-dialect shows variation from the others. The data from Kedah dialect listed here are from the sub-dialect Kedah Pesisiran, and may vary from other sub-dialects of Kedah.

This raises the question of how this phonological process can be resolved in OT analysis. To deal with this, it is necessary to posit a constraint that is able to rule out a nasal plus a voiced obstruent cluster, so that a candidate with nasal substitution can emerge as the optimal output. Therefore, I suggest that CRISP-EDGE[σ] (78) should be added to the constraint ranking of Malay dialects. However, obedience to CRISP-EDGE [σ] leads to a violation of another constraint that bans a voiced obstruent from undergoing nasal substitution, that is IDENT[PHAREXP] (98) (see Chapter 4 subsection 4.2.1.2). Thus, to account for voiced obstruent nasal substitution, as in Perak, these constraints should be ranked in the following order: CRISP-EDGE [σ] >> IDENT [PHAREXP] (see 6.3.3). The adoption of CRISP-EDGE [σ] seems more practical since it allows both voiceless and voiced obstruents to undergo nasal substitution rather than *NC̥. The advantage of using CRISP-EDGE[σ] can be seen in the discussion of Perak and NS dialects, where voiced obstruents can also undergo nasal substitution.

In order to do that, this analysis will be based on the data from Ahmad (1991) for Perak, Che Kob (1980) and Teoh (1994) for Kelantan, and Rufus (1969) for the NS dialect.[46] As mentioned earlier in this thesis (see Chapter 1 section 1.1), since the Institute of Language and Literature (DBP) only provides corpus data for SM, corpus data are not available for the dialects. For that reason, this chapter will draw on data gathered by the scholars mentioned above. Their studies on Malay dialects are essential for this analysis.

[46] It ought to be mentioned here that as well as the data from those scholars' books, some data were also obtained directly from native speakers of those dialects. Two native speakers of NS, Dr. Mohd Fadzeli Jaafar and Mr. Zulkifli Ahmad, and others, were interviewed, as well as there being a discussion with Assoc. Prof. Che Kob to obtain and reconfirm the data for Kelantan. The word lists for the three dialects which cover the three phonological phenomena – nasal and voiceless obstruents at root-internal and prefix-root junctures, and nasal and voiced obstruent clusters at prefix junctures – are provided in Appendix N.

In my search of Malay literature for Malay dialects, nasal and voiceless/voiced obstruent sequences behave distinctly for SM and among the dialects. The distinctions, or so-called 'variation' – different ways of saying the same word – have not however been discussed in great formal detail in previous works. The study of dialect variation by those previous scholars mentioned above is no more than just describing and listing the words used in the dialects they examined. Therefore, this analysis offers a different analysis from those previous studies in its analysing of different ways of saying the same word in the Perak, Kelantan and NS dialects. The variations occurring in the three dialects will be analysed with an adequate phonological theory adopted by this thesis, which is OT (see Section 6.2 for more detailed about variation in OT). By this, I claim that this analysis is the first OT account for those three clusters in Malay dialects. No generative grammar has discussed clusters, either within a dialect or between dialects, as this present study attempts to do.

The variation that occurs in Malay dialects, three of which were chosen for further discussion in this chapter, will be accounted for in terms of different constraint rankings. This analysis thus aims to show why dialects in Malay differ from one dialect to another and how OT as a theoretical framework can explain the differences that occur. Thereafter, we will see how OT can explain, in an adequate manner, the variations occurring among the dialects, as this is an important new contribution that OT has made (McCarthy, 2008: 260). It is crucial to note that this analysis does not claim to be the first study of dialect variation in Malay. It does however claim to be the first extensive and theoretical study of dialect variation concerning nasal plus voiceless andvoiced obstruent, and sonorant clusters in Malay.

In what follows, as this analysis of dialectal variation relies on an advantage of OT, dialect-specific ranking (see Chapter 2 Section 2.2.3), I will first discuss what variation means in the theory. This will then lead to the main intention of this present chapter, the data analysis, which will be discussed in Section 6.3. This section will be divided into three subsections whereby each subsection discusses one dialect. The Perak, Kelantan and NS dialects will be discussed in Sections 6.3.1, 6.3.2 and 6.3.3, respectively. In the concluding section, I will summarise the discussion of this chapter.

6.2 Variation in OT

Variation refers to the range of differences between languages or dialects within a language. It has become increasingly important to phonological theory over recent years (Coetzee and Pater, 2009). In the beginning, variation received much attention amongst sociolinguists. Therefore, it has always been explained from the viewpoint of sociolinguistics, i.e. it is due to external factors such as sex, age, style, register and social class (Anttila, 2002: 206). According to Anttila (2002), variation is also due to internal factors such as morphology, phonology, syntax and lexicon. With this in mind, this study is going to examine one of the internal factors, i.e. phonology, which is also a conditioning variation.

It is important to study the phonological facts of nasal and voiceless/voiced obstruents, and nasal and sonorant clusters, in three Malay dialects (Perak, Kelantan and NS). Why is a sequence of nasal and voiceless obstruent allowed within a root, but not at prefix-root junctures, as in PD? Moreover, the same sequence in other dialects of Malay, for example Kelantan and NS, is not allowed to emerge in either morphological environment. At this point, a rule-based analysis explanation does not seem to be helpful. To explain this, therefore, we need an appropriate phonological theory of language variation that can offer a good justification for the case. In recent work on phonological theory, OT has been used to understand the phenomenon of language or dialect variation (Antilla, 2002). The results yielded were illuminating, as variation could be captured more effectively and evidently within OT.

How can variation be accounted for in OT? The basic idea of OT is that UG consists largely of a set of well-formedness constraints, out of which individual grammars are constructed (cf. Prince and Smolensky, 2004) (see Section 2.2.3 in Chapter 2). This theory views UG as a set of violable constraints. By allowing constraints to be violable, we are able to see different patterns and variations between languages through such constraint violations. OT also suggests that each language has its own ranking as the grammar of a language uses a specific constraint ranking. Differences in ranking between language A and language B result in different patterns in the two languages, or so-called variation between languages. By differences in constraint ranking, this theory can formally expresses the distinction between each dialect or variation occuring in a language (McCarthy, 2008). Although constraints in UG in OT are allowed to be violated, violations do however occur in limited contexts (Archangeli and Langendoen, 1997). In the following section, we investigate how constraint ranking and violation handle the data for Perak, Kelantan and NS.

In the earlier stages of OT, known as classic OT, there was only one way that grammar could produce multiple outputs from a single input (McCarthy, 2008: 261). If there were two or more candidates with the same number of violation marks from every constraint in CON, then EVAL had to choose both as optimal outputs (ibid.). In addition to two or more candidates with the same number of violations, Anttila (1997) discussed variation in OT in a way that is similar to McCarthy's idea. Variation in a simple way, according to Anttila, occurs if the grammar is unable to distinguish between candidates. From this simple solution, Anttila exemplifies the following tableau whereby the candidates are both optimal, thus the grammar is unable to decide on the optimal output. Anttila argues that two optimal outputs derived from a set of constraints, as in (204), is a simple and truly poor way of dealing with variation. This, as claimed by Anttila, is not however the way to analyse variation in OT.

204.

	A	B	C
a. ☞ $cand_1$	*		
b. ☞ $cand_2$	*		

McCarthy (2008: 261) claims that candidates with the same number of violations probably never occur in terms of phonology and syntax, if the constraint sets are rich enough. To put it another way, a set of constraints should never produce more than one optimal output, as shown in the above tableau.

As stated in Guy (1997: 135), OT's hypothesis states that several categorical grammars are involved in the production of variability, with each grammar generating, discretely and invariably, a different one of variable outcomes. Another linguist's point of view concerning variation existing in languages can clearly be seen in Kiparsky's (1993, 1994) statement: 'Assumption: Variation comes from competition of grammatical systems (in the individual or in the community), not from a probabilistic component in the rules of the language'. As stated in Guy (1997: 135-6), OT is a theory which does not impose any limitations on sequential orders of the various constraints. Moreover, the theory explicitly attributes differences among languages to different orders of a universal constraint inventory (ibid.: 136). As has been suggested by the theory, the different orders of a universal constraint inventory in different languages might also be applied to differences within a language, such as regional dialects, social dialects and stylistic differences (ibid.). Several analyses concerning the variations within a language can be seen in work such as that of Iverson and Lee (1994), which concerns the differences between two regional dialects of Korean and between two stylistic varieties within one dialect. Anttila (1997) has analysed variable productions in Finnish, and Rose (1995) accounts for differences in the Gurage dialect (cited in Guy, 1997).

Recently, in works such as Anttila (1997, 2002, 2007), Boersma (1998), Kiparsky (1993a) and many others, variation has been analysed under the multiple grammars theory. This theory proposes that an individual grammar may have several different grammars (cited in Anttila 2007), which means that, in OT, a speaker of a language knows more than one constraint ranking. The existence of variation in an individual grammar occurs when the same individual speaks different forms at different times (ibid.: 519). The argument made by the theory for the occurrence of multiple grammars known by an individual is that it is as the result of multilingualism.

The most famous example from Kiparsky's analysis (1993a) of /t/ and /d/ deletion, in words like *'cost me', 'cost'* and *'cost again'* in American English, has always been referred to whenever variation is discussed. As proposed by Kiparsky (ibid.), t/d is deleted when it ends up being extra-syllabic (cf. Anttila, 2007: 522). The constraints used to analyse the phenomenon of t/d deletion are: *COMPLEX, ONSET, PARSE, ALIGN-LEFT-WORD and ALIGN-RIGHT-PHRASE, and these constraints are the possible grammars. The following tableau shows violations of the three candidates, as exemplified in Anttila (2007: 523):

205.

1. /kastəgɛn/	*COMPLEX	ONSET	ALIGN-L-W	ALIGN-R-P	PARSE
a. kast][əgɛn	*	*			
b. kas]t[əgɛn		*			*
c. kas][təgɛn			*		
2. /kast mi/					
a. kast][mi	*				
b. kas]t[mi					*
c. kas][tmi	*		*		
3. /kast/					
a. kast]]	*				
b. kas]]t				*	*

Under multiple grammars, those constraints are completely unranked with respect to one another (cited in McCarthy, 2008: 262). For example, there might be 120 possible ways to rank the constraints. The speakers of the language know the 120 grammars since there are 120 permutations of these constraints. Speakers, however, do not have to learn all 120 grammars that the language has. By allowing the constraint to be unranked, speakers have, in effect, internalised a system of 120 grammars (cited in McCarthy, 2008: 262).

6.3. Data Analysis

This section discusses how the three clusters, (1) nasal and voiceless obstruents, (2) nasal and voiced obstruents and (3) nasal and sonorant clusters, behave in the three dialects of Malay, Perak, Kelantan and NS, examined in this thesis. The behaviour of those clusters in three morphological domains, i.e. root-internally, prefix-root and prefix-prefix junctures, will be examined. We begin the discussion with PD in subsection 6.3.1. The other two dialects, Kelantan and NS, will be discussed in subsections 6.3.2 and 6.3.3, respectively.

6.3.1 Perak dialect

It is worth noting that one of the dialects that this study examines, PD, permits nasal and voiceless obstruent clusters in the surface representation root-internally (see Table 6, below). We begin the discussion of nasal and voiceless obstruent clusters within roots by first observing the data listed in the following table:

Table 6: Nasal and voiceless obstruent clusters in Perak (Ahmad, 1991).

SM	Perak	
teŋkat	teŋkat	'level'
tɔŋkat	tɔŋkat	'walking stick'
kampoŋ	kampoŋ	'village'
təmpat	təmpat	'place'
simpan	simpan	'to keep'
lambat	lambat	'late'
pintu	pintu	'door'

As illustrated in the above data, nasal and voiceless obstruent clusters within the roots appear in the surface representations of PD, the same as in SM. Observe that words such as [teŋ.kat], [kam.poŋ] and [təm.pat] in Perak do not undergo any phonological processes that serve to eliminate nasal and voiceless obstruent clusters, such as nasal substitution, nasal deletion or epenthesis. The above data obviously show that, despite the fact of nasal substitution, which is normally applied in SM to rid the language of nasal and voiceless obstruent clusters, it is however blocked within the roots in PD.

It is worth mentioning that the situation discussed above is also found in other languages that do not allow a sequence of nasal and voiceless obstruents, but that they emerge in the surface representation within the roots. One of those languages is Indonesian. The nasal substitution that is generally applied in Indonesian, to eliminate nasal and voiceless obstruent sequences at prefix-root junctures, is blocked at root-internal level as well. The consequence of blocking nasal substitution in Indonesian results in a sequence of homorganic nasal and voiceless obstruents in the surface representation, as exemplified below:

206. Root-internal occurrences of NÇ in Indonesian (from Pater 1999: 75; 2001).

/əmpat/	[əmpat]	'four'
/untuk/	[untuk]	'for'
/muŋkin/	[muŋkin]	'maybe'

Such a problem occurring in Indonesian has received much attention among theoretical linguists, particularly within OT (e.g. Pater 1999, 2001), as to why nasal substitution is blocked within roots. This poses a challenge to the theory when explaining the blocking of nasal substitution at root-internal position. In explaining the lack of nasal substitution at root-internal position, McCarthy and Prince (1994b, cited in Pater, 1996) claim that 'a large number of disparate phonological phenomena for instance, reduplicative and otherwise, result in a stricter Faithfulness requirement within the root than elsewhere in the word, that is the relative markedness of roots' (see also Urbanczyk, 1996). In OT, this situation is accounted for as faithfulness requirements are more strictly applied within the root than in non-root morphemes, such as affixes (McCarthy and Prince, 1995, cited in Kager, 1999: 75). To capture this situation, McCarthy and Prince (1994a) proposed a general ranking schema where root-specific versions of faithfulness constraints are ranked higher than the general version of these constraints: Root-Faithfulness >> Faithfulness.

Returning to our discussion of PD, since nasal substitution is also blocked within roots in the dialect, a root-specific constraint, which bans root-internal nasal substitution, is needed for the analysis of Perak. The relevance constraint of faithfulness that is able to capture the blocking of nasal substitution is UNIFORMITY, as defined in (82). We will see then, in this section, that UNIFORMITY, which is the general version of the faithfulness constraint, is ranked beneath the root-specific version of this constraint, i.e. UNIFORMITY-ROOT, in the constraint ranking of PD, as demonstrated in (209). The root specific constraint, UNIFORMITY-ROOT is defined in (83).

In the case where nasal substitution is blocked root-internally, UNIFORMITY-ROOT is ranked high so as to allow output containing a sequence of nasal and voiceless obstruents. There is a functional explanation of why nasal and voiceless obstruent sequences are allowed to be present root-internally. As asserted in Pater (1999), root-internal segments are more resistant to phonological processes than segments in other positions (cited in Kager, 1999: 75). In some of the morphological literature (e.g.: Mascaro, 1976; Kiparsky, 1982, 1993b; cf. Kager, 1999) it is said that 'there is a well-known class of processes that apply only across morphemes but fail to apply within the roots' (cf. Kager, 1999: 75). That is the reason why sequences of nasal and voiceless obstruents function root-internally in Perak and thus do not undergo the regular process, nasal substitution.

Before we proceed to the analysis of how UNIFORMITY-ROOT prevents root-internal segments from undergoing nasal substitution, we consider first what has been discussed in the relevant literature about this constraint. Pater (1999, 2001), for example, captured the case of the blocking of nasal substitution in Indonesian by applying the idea of McCarthy and Prince (1994b). Two root-specific constraints, LINEARITY-ROOT and UNIFORMITY-ROOT, as shown in 207(a) and 207(b), respectively, were used.[47] I exemplify in the following tableaux how Pater makes use of these two root-specific constraints to analyse the lack of nasal substitution in Indonesian:

207. (a) Root-internal NC̥ tolerance: ROOTLIN >>*NC̥ (adapted from Pater, 1999: 275).

/əm$_1$p$_2$at/	LINEARITY-ROOT	*NC̥	LINEARITY
a. əm$_{12}$at	*!		*
b.☞əm$_1$p$_2$at		*	

(b) Root-internal NC̥ tolerance: UNIFORMITY-ROOT >>*NC̥ >> UNIFORMITY (adapted from Pater, 2001: 162)

/əm$_1$p$_2$at/	UNIFORM ROOT	*NC̥	UNIFORM
a. əm$_{12}$at	*!		
b.☞əm$_1$p$_2$at		*	*

47 Pater's analysis concerning nasal substitution in Indonesian applied two constraints: LINEARITY is used in his earlier analysis, but UNIFORMITY in the revisited analysis on nasal substitution in Austronesian. According to McCarthy (1995), LINEARITY and UNIFORMITY are used to ban metathesis and coalescence, respectively.

The above tableaux clearly show that the root-specific constraints, LINEARITY-ROOT and UNIFORMITY-ROOT, play important roles in accounting for the blocking of nasal substitution root-internally in Indonesian. According to Pater (2001), by ranking the root-specific constraints UNIFORMITY-ROOT or LINEARITY-ROOT above *NC̥, nasal substitution can certainly be blocked from occurring within roots in Indonesian. To ensure nasal substitution is not blocked at prefix junctures, *NC̥ must outrank UNIFORMITY or LINEARITY, so that nasal substitution would continue to apply, as exemplified in the tableau below.

208. Nasal substitution is unblocked at prefix junctures in Indonesian: UNIFORMROOT >> *NC̥ >> UNIFORM (adapted from Pater, 2001: 161).

/məŋ$_{1}$p$_{2}$aksa/	UNIFORMROOT	*NC̥	UNIFORM
a.☞məm$_{12}$aksa			*
b. məm$_{1}$p$_{2}$aksa		*!	

As we can see in the above tableau, nasal substitution continues to apply at prefix junctures by employing the constraint ranking: UNIFORM-ROOT >> *NC̥ >> UNIFORM. Since the root-specific constraint is able to block nasal substitution only within a root, so this constraint is not violated by candidate (a), which undergoes nasal substitution. This candidate, however, violates UNIFORMITY, since the sequence of nasal and voiceless obstruents at the prefix juncture in the input is substituted into a single segment in the output. Considering the case under discussion, I am going to make use of Pater's analysis (2001) in order to capture the blocking of nasal substitution within roots in Perak, where a root-specific constraint, UNIFORMITY-ROOT, will be used in this analysis.

Now we shall begin the analysis. Since root-internal nasal and voiceless obstruent clusters in Perak are not resolved by nasal substitution, we need a root-specific faithfulness constraint, as discussed above, i.e. UNIFORMITY-ROOT, which is able to block the cluster from undergoing nasal substitution. By considering UNIFORMITY-ROOT in the ranking, a candidate without nasal substitution is preferred, i.e. [kam$_1$p$_2$oŋ] is preferred over *[kam$_{12}$oŋ], and thus emerges as the optimal output.

As we saw in Chapter 4 Section 4.2.1.1, schwa is epenthesized to break up the sequence of nasal and voiceless obstruents in prefixed monosyllabic words, since the voiceless obstruent is not deleted, even though the nasal final prefix is attached to it. This solution of epenthesizing a schwa can also be applied to eliminate a sequence of nasal and voiceless obstruents within a root. Thus, the potential candidate if schwa epenthesis were to apply is *[kaməpoŋ]. This candidate involves an additional correspondent in the output. Under Correspondence Theory, (see Chapter 2 subsection 2.4.2) the constraint that this candidate violates is DEP-IO (116). The diagram for DEP-IO violation is illustrated in (117). For this constraint, DEP-IO, and the other two faithfulness constraints, UNIFORMITY and UNIFORMITY-ROOT, discussed above, I establish the following constraint ranking for PD.

209. DEP-IO >> UNIFORMITY-ROOT >> UNIFORMITY

/kam$_1$p$_2$oŋ/	DEP-IO	UNIFORMITY-ROOT	UNIFORMITY
a.☞kam$_1$p$_2$oŋ			
b. kam$_{12}$oŋ		*!	
c. kam$_1$əp$_2$oŋ	*!		

Here we see that epenthesis is not a better way to break up the cluster as DEP-IO is highly ranked, as in candidate (c). The root-specific constraint, UNIFORM-ROOT, blocks candidate (a) from undergoing nasal substitution. Therefore it is chosen as the winner. However, candidate (b) with nasal substitution violates UNIFORMITY-ROOT. I illustrate below the difference between [kam$_1$p$_2$oŋ] and [kam$_{12}$oŋ], as in candidates (a) and (b), respectively. Candidate (b) with nasal substitution violates the root-faithfulness constraint, UNIFORMITY-ROOT. The subscripted number is used to indicate the correspondence relationship.

210. Correspondent diagram of UNIFORMITY-ROOT violation.

Candidate (a)	Candidate (b)
X_1 X_2	X_{12}
| |	|
k a m$_1$ p$_2$ o ŋ	k a m$_2$ o ŋ

As well as the above candidates in (209), we shall consider other potential candidates that might be generated. Another potential candidate that must be taken into consideration is *[kapoŋ]. In *[kapoŋ], the nasal segment is deleted and this has resulted in one segment in the input having no correspondent in the output. The relation between input and output is called Faithfulness–IO, where faithfulness to the input and output identity is a type of requirement in which a pair of representations must be identical, as stated in Correspondence Theory (79) (McCarthy and Prince, 1995, cited in Kager, 1999: 24). In this case, the Faithfulness-IO constraint that is crucial to account for nasal deletion is MAX-IO:

211. **MAX-IO** (Kager, 1999: 24)

Every segment in the input must have a correspondent in the output.

The constraint in (211) requires every element in the input to have a correspondent in the output. The violation of MAX-IO in the suboptimal candidate *[kapoŋ] is illustrated in the correspondence diagram, below:

212. Correspondence diagram for nasal deletion: [kampoŋ] → *[kapoŋ].

Input	m	p
		|
Output		p

Considering the potential candidate *[kapoŋ] in the tableau, I establish the following part of the constraint ranking: DEP-IO >> MAX-IO >> UNIFORM-ROOT >> UNIFORM. Now we have two relevant faithfulness constraints, DEP-IO and MAX-IO, in the ranking for further evaluation.

213. DEP-IO >> MAX-IO >> UNIFORM-ROOT >> UNIFORMITY

/kam$_1$p$_2$oŋ/	DEP-IO	MAX-IO	UNIFORMITY-ROOT	UNIFORMITY
a.☞kam$_1$p$_2$oŋ				
b. kam$_{12}$oŋ			*!	
c. kam$_1$əp$_2$oŋ	*!			
d. kap$_2$oŋ		*!		

As can be seen, candidates (c) and (d), which undergo schwa epenthesis and nasal deletion respectively, to eliminate nasal and voiceless obstruent clusters, are ruled out because the candidates incur violation of DEP-IO (116) and MAX-IO (211), respectively. Candidate (b), which undergoes nasal substitution, violates UNIFORMITY-ROOT. Although nasal substitution is a process of merging two segments in the input into a single segment in the output, which results in one segment less in the output, it does not however violate the faithfulness constraint, MAX-IO, as with candidate (b). Indeed, MAX-IO requires every segment in the input to have a correspondent in the output; in nasal substitution however, the two [mp] segments share a single output correspondent (Pater, 2001: 167). Therefore, candidate (a) with lack of nasal substitution is chosen as the optimal output as this candidate violates none of the constraints.

We saw from the above discussion that in order for a candidate without nasal substitution to emerge as the optimal output, the root-specific faithfulness constraint, UNIFORMITY-ROOT, is high in the ranking. This constraint is able to rule out any candidate with nasal substitution as it requires the output segments to be as faithful as possible to the input segments. As /kam$_1$p$_2$oŋ/ is a root word, the output must be faithful to the input which is a root.

We now should move on to the next discussion of nasal and voiceless obstruent clusters at prefix-root junctures in Perak. The same as in SM, PD also disfavours the same two clusters in the surface representation at prefix-root junctures: (1) nasal and voiceless obstruents and (2) nasal and sonorant clusters. We start the discussion with nasal and voiceless obstruent clusters in this dialect. As we shall see in the following discussion, in order to eliminate nasal and voiceless obstruent clusters, a markedness constraint, CRISP-EDGE, is added into the constraint ranking of PD. The discussion is as follows.

As stated in Ahmad (1991: 54), there are two prefixes ending with a final nasal in Perak: /ŋ+/, a verb forming prefix, and a noun forming a prefix, /pəŋ+/. The verb forming prefixes, i.e. the velar nasal /ŋ+/, has two allomorphs, /ŋ+/ and /məŋ+/. These two allomorphs can be used interchangeably in the dialect. Thus, he claims that words like /ŋ-tɔʁɛh/ ACT.PRF-tap → [nɔʁɛh] 'to tap' and /ŋ-ɲadi/ ACT.PRF-be → [ɲadi] 'to be' can also be pronounced as [mənoʁɛh] and [məɲadi], respectively (ibid.: 55). However, the form with velar nasal /ŋ+/ is used more prominently in Perak compared to the other one which has a sound more like SM.

Those prefixes, /ŋ+/, /məŋ+/ and /pəŋ+/, exhibit the same phonological behaviour where the nasal segments in the prefixes undergo phonological alternation, as can be seen in the following data:

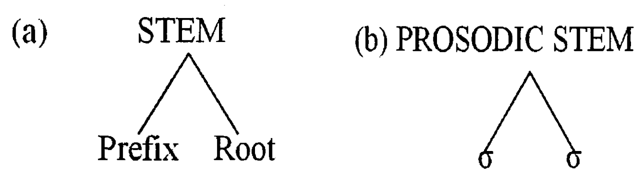
214. Nasal and voiceless obstruent clusters in Perak (Ahmad, 1991: 54-5).

(i)	/ŋ-pudʒi/ ACT.PRF-praise 'to praise'	[muji]
(ii)	/ŋ-taŋkap/ ACT.PRF-catch 'to catch'	[naŋkap]
(iii)	/ŋ-kutɛt/ ACT.PRF-pick up 'to pick up'	[ŋutɛt]
(iv)	/pəŋ-taŋkap/ NOM-catch 'catcher'	[pənaŋkap]
(v)	/pəŋ-kajɔh/ NOM-ride 'rider'	[pəŋajoh]
(vi)	/ŋ-tɔʁɛh/ ACT.PRF-tap 'to tap'	[nɔʁɛh]

It is clear, from the above list, that a sequence of nasal and voiceless obstruents undergoes nasal substitution in Perak. Since nasal substitution at a prefix-root juncture is used to break up the cluster, UNIFORMITY is violated, as the two segments in the input are mapped to a single segment in the output. Hence, UNIFORMITY must be ranked beneath the other faithfulness constraints in the hierarchy. From this, the output undergoes nasal substitution and can emerge as a winner.

As discussed in Chapter 2 subsection 2.2.4, stems that contain a prefix and root must satisfy MORPHEME SYLLABLE CORRELATION (34). This constraint requires a prefix to contain one syllable, and the root must also contain one syllable. This constraint therefore requires that stems (i.e. prefix and root) must be at least disyllabic. The disyllabicity minimality required by MORPHEME-SYLLABLE CORRELATION (34) is a corollary to the constraint named PROSODIC STEM, as illustrated in (35), and is repeated below for convenience. As we are examining nasal and voiceless obstruent clusters at prefix-root junctures, PROSODIC STEM is crucial to ensure that the prefix and the root are at least disyllabic. This prosodic stem minimality constraint will then be added into the constraint ranking of PD.

215. Prosodic Stem Minimality (Downing, 2006: 124).



Observe that in (208), *NÇ in Pater's analysis is ranked beneath the root-specific constraint, UNIFORMITY-ROOT, mainly to ensure that a sequence of root-internal nasal and voiceless obstruents does not undergo nasal substitution. An important point that must be addressed here is that, in this analysis, CRISP-EDGE, which bans any element linked to a prosodic word and may be linked to a prosodic category external to that prosodic word, will be used instead of *NÇ to avoid nasal and voiceless obstruent clusters in the surface representation. Besides that, CRISP-EDGE is also crucial to rule out nasal and voiced obstruent clusters in the surface representation. In order to account for voiced obstruent nasal substitution in the Perak and NS dialects, which allow voiced obstruents to undergo nasal substitution, I will make use of the constraint CRISP-EDGE [σ], as in (78), which has been used by Pater (2001) to analyse the same sequence in Muna when undergoing nasal substitution (see Chapter 2, subsection 2.4.1.2). As we shall see later in this section, CRISP-EDGE [σ] becomes more crucial in accounting for nasal and voiced obstruent clusters at prefix-root junctures in PD. As I will demonstrate in (229), this constraint is able to rule out candidates with a sequence of nasal and voiced obstruents in the surface representation at prefix junctures.

Besides that, CRISP-EDGE [σ] is also crucial when we come to examine nasal and voiceless obstruent clusters root-internally in Kelantan and NS, whereby [gat_2oŋ] is preferred over *[gan_1t_2oŋ]. In the following subsection, 6.3.2, I will explain CRISP-EDGE [σ] in great detail when we discuss nasal and voiceless obstruent clusters root-internally in Kelantan. As the case in Kelantan provides us with more understanding about the importance of CRISP-EDGE [σ] in root-internal situations, I will explain why CRISP-EDGE [σ] is chosen for this analysis of Malay dialects and not CRISP-EDGE [PRWRD], as used by Pater (2001) to account for Indonesian nasal substitution. Below is how CRISP-EDGE [σ] is defined. This constraint was defined in (78) and is repeated here for convenience:

216. **CRISP-EDGE [σ]**

No element belonging to a syllable may be linked to an adjacent syllable.

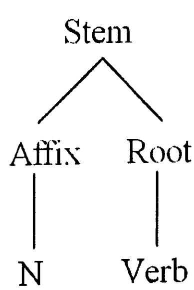
As just mentioned above, CRISP-EDGE [σ] rules out candidates with both nasal and voiceless/voiced obstruent clusters in the surface representation. This means this constraint will prevent *[$m_1p_2$udʒi] from emerging as a winner. The only candidate that this constraint prefers is a candidate with nasal substitution, i.e. a candidate without a nasal and voiceless obstruent cluster. The potential candidate is [m_{12}uʤi]. Although this candidate obeys CRISP-EDGE [σ], it does violate another constraint which requires edge segments in the input to preserve their segments at the edge of the corresponding prosodic structure, as stated in a constraint named EDGE-INTEGRITY, which is defined below. This constraint has already been introduced in Chapter 2 subsection 2.4.2.2, in (85).

217. **EDGE INTEGRITY** (McCarthy and Prince, 1995).

Edge segments in the input preserve their segments at the edge of the corresponding prosodic structure.

Before I establish a new tableau to evaluate a sequence of nasal and voiceless obstruents at prefix-root junctures in Perak, I would like first to show a case of nasal substitution in Javanese. It is worth knowing that the situation regarding nasal and voiceless obstruent clusters at prefix-root junctures in Perak is the same as nasal fusion in Javanese, as demonstrated in (30) (see Chapter 2, subsection 2.2.4), where the prefix also shares the same syllable with the initial syllable of the root. Since stems in Javanese must be minimally disyllabic, so nasal fusion is accounted for as follows:

218. Javanese nasal fusion structure (from Downing, 2006: 135).



The above construction clearly shows that nasal fusion in Javanese contains bimorphemic stems (affix and root). Therefore, they are subject to PROSODIC STEM

(35) (Downing, 2006: 135). I exemplify again, below, the tableau in (30) for Javanese nasal fusion:

219. Disyllabicity in Javanese nasal affixation (Downing, 2006: 136).

/N-tulis/	PRSTEM	Nasal Fusion	Faith-IO
a. n-tulis		*!	
b.☞nulis			*
/N-bom/			
c.☞ŋəbom		*	*
d. mbom	*!		

The important point I would like to address here is that, for optimal output, candidate (b) in the above tableau does not violate PROSODIC STEM, as it is disyllabic. As we will see in the following tableau, this optimal output is the same as the optimal output in Perak, /ŋ_1+$\text{p}_2$udʒi/ → [m_{12}udʒi], because the prefix is a nasal segment and therefore the voiceless obstruent following the prefix is deleted.

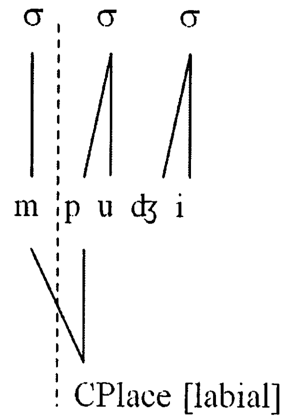
Putting together all the constraints we have discussed thus far, I establish the following part of the constraint hierarchy for PD: PROSODIC STEM >> DEP-IO >> MAX-IO >> UNIFORM-ROOT >> CRISP-EDGE [σ] >> EDGE-INTEG >> UNIFORM.

220. Nasal and voiceless obstruent sequences at prefix-root junctures in Perak.

/ŋ₁+p₂udʒi/	PRSTEM	DEP-IO	MAX-IO	UNI-ROOT	CRISP-EDGE[σ]	EDGE-INTEG	UNI
a. ☞ m_{12}udʒi						*	*
b. $m_1p_2$udʒi					*!		
c. $m_1əp_2$udʒi		*!					
d. $p_2$udʒi	*!		*				

All candidates in the above tableau satisfy PROSODIC STEM, as the stem contains two syllables. In order to eliminate nasal and voiceless obstruent sequences at prefix junctures, schwa is epenthesized between the two segments, as in candidate (c), and the nasal prefix is deleted, as in candidate (d). This evaluation reveals that the epenthesis of schwa and nasal deletion in candidates (c) and (d), respectively, can never be optimal, because these candidates disobey DEP-IO and MAX-IO, respectively, which are ranked higher in the hierarchy. The competing candidates, (a) and (b), do not violate these constraints. Without undergoing nasal substitution, candidate (b) obeys UNIFORM, because this constraint works against segmental fusion or coalescence (McCarthy and Prince, 1999). This candidate, however, violates CRISP-EDGE [σ], as the two segments are doubled-linked. The diagram for CRISP-EDGE [σ] violation of this candidate is illustrated in (221). Since CRISP-EDGE [σ] is ranked above UNIFORMITY, so candidate (b) is ruled out. Therefore, candidate (a), with nasal substitution, is the winner, as it only violates the latter.

221. Diagram of CRISP-EDGE [σ] violation: [m$_1$p$_2$udʒi] with multiple linking.



Another potential candidate that must be taken into consideration is *[ŋ$_1$p$_2$udʒi]. The nasal segment in this potential candidate is not homorganic to the following voiceless obstruent. Thus it violates NAS ASS, as defined in (121). This constraint is also used by Pater (2001) to account for voiceless obstruent nasal substitution in Indonesian. In Indonesian, NAS ASS outranks CRISP-EDGE [PrWD] and UNIFORM in Indonesian nasal substitution. I exemplify the analysis by Pater in the following tableau with the constraint hierarchy: NAS ASS >> CRISP-EDGE [PrWD] >> UNIFORM.

222. Indonesian nasal substitution: NAS ASS >> CRISP-EDGE [PrWD] >> UNIFORMITY (Pater, 2001: 175).

/məŋ$_1$+p$_2$ilih/	NAS ASS	CRISP-EDGE[PrWD]	UNIFORM
a.☞məm$_{12}$ilih			*
b. məm$_1$p$_2$ilih		*!	
c. məŋ$_1$p$_2$ilih	*!		

As we can see in the tableau, the unassimilated nasal to the following consonant, as in candidate (c), violates NAS ASS. According to Pater (2001: 175), in order to obtain a candidate with nasal substitution as the optimal output, NAS ASS and CRISP-EDGE [PrWD] must dominate UNIFORMITY.

Following the above analysis, the potential candidate *[ŋ₁p₂udʒi] and the constraint NAS ASS will be added into the constraint ranking of Perak for further evaluation. The relevant constraint ranking is now as follows: PROSODIC STEM, DEP-IO >> NAS ASS >> MAX-IO >> UNIFORM-ROOT >> CRISP-EDGE [σ] >> EDGE-INTEG >> UNIFORM.

223. Constraint ranking for Perak.

/$ŋ_1$+$p_2$udʒi/	PRSTEM	DEP-IO	NAS ASS	MAX-IO	UNI -ROOT	CRISP-EDGE[σ]	EDGE-INTEG	UNIFORMITY
a. ☞m_{12}udʒi							*	*
b. $m_1p_2$udʒi						*!		
c. m_1 ə$p_2$udʒi		*!						
d. $p_2$udʒi				*!				
e. $ŋ_1p_2$udʒi			*!					

From the above discussion, we note that nasal and voiceless obstruent clusters are disfavoured at prefix-root junctures in the PD. But what about the occurrence of the same cluster at prefix-prefix root junctures in the dialect? Are nasal and voiceless obstruent clusters also banned? This merits a question since the cluster can also occur at prefix-prefix root junctures, as in SM. As Ahmad (1991: 80) says, most derived words in Perak have only one layer of prefixation or suffixation. Roots can only be attached to one prefix or suffix, or a combination of both, but not a sequence of the two. From this, it means that PD cannot receive more than one prefix in prefixed words as in SM. Thus, multiple prefixes can not possibly be found in the vocabulary of Perak.

In what follows, I am going to discuss the other cluster which is no less crucial to be explored, i.e. nasal and sonorant clusters. As mentioned, this cluster is also not permitted to emerge in the surface representation at prefix-root junctures in Perak. As discussed in Chapter 4, previous Malay scholars (e.g.: Teoh, 1994; Ahmad, 2000) claimed that root-internal nasal and sonorant clusters in SM, as presented in (7), are all non-native Malay words which have been borrowed from Chinese. In my search of the previous literature for Perak, primarily Ahmad (1991), as this dialect is based on his work, the examples provided show no root words with nasal and sonorant clusters. In my assumption, perhaps the situation that occurs in SM regarding nasal and sonorant clusters root-internally also occurs in Perak. However, there is not a single example of this cluster within a root to be found in Ahmad's work.

On the other hand, there is a short discussion on nasal and sonorant clusters at prefix-root junctures in this dialect that is presented in Ahmad (1991). In PD, the final nasal of the prefix is claimed to be deleted when it occurs with roots beginning with nasals and liquids /l/ and /ʁ/ (ibid.: 60). I now lay out all the examples given by Ahmad (ibid.):

224. Nasal and sonorant clusters at prefix-root junctures in Perak (Ahmad 1991: 60).

/pəŋ+maɛn/	[pə-maɛn]	'something which is used in play'
/pəŋ+naoŋ/	[pə-naoŋ]	'something which shades'
/pəŋ+ɲaɲi/	[pə-ɲaɲi]	'singer'

/pəŋ+ʁompaʔ/	[pə-ʁompaʔ]	'robber'
/pəŋ+lakon/	[pə-lakon]	'actor'/'actress'
/məŋ+latɛh/	[mə-latɛh]	'to train'
/məŋ+lətop/	[mə-lətop]	'explode'
/məŋ+ʁokɔk/	[mə-ʁokɔk]	'to smoke'
/məŋ+ʁompaʔ/	[mə-ʁompaʔ]	'to rob'

The hierarchical ranking that was established before is: PRSTEM >> DEP-IO >> NAS ASS >> MAX-IO >> UNIFORMITY-ROOT >> CRISP-EDGE[σ] >> EDGE INTEGRITY >> UNIFORMITY. The effects of this ranking on nasal and sonorant clusters are illustrated in the following tableau:

225.

/məŋ$_1$+ʁ$_2$okɔk/	PRSTEM	DEP IO	NAS ASS	MAX -IO	UNI-ROOT	CRISP--EDGE[σ]	EDGE-INTEG	UNIFORMITY
a.☞ məʁ$_2$okɔk				*				
b. məŋ$_1$ʁ$_2$okɔk			*!					
c. məŋ$_1$əʁ$_2$okɔk		*!						

The above tableau shows that a sequence of nasal and sonorant is also disallowed in the dialect of Perak, along with nasal and voiceless obstruent clusters as discussed earlier. The occurrence of a nasal and sonorant cluster in the underlying representation in Perak, as we can see in the above tableau, is resolved by nasal deletion. This phonological effect of nasal deletion can clearly be explained by the above constraint ranking in (225). In order for [mə$_1$ʁ$_2$okɔk] to be chosen as the optimal output, the IO-faithfulness constraint, MAX-IO, is ranked below DEP-IO as

C-deletion is a better option than V-epenthesis to avoid nasal and sonorant sequences in the dialect.

The analysis thus far has discussed two consonant clusters that are banned at prefix-root junctures in Perak: nasal and voiceless obstruent clusters, and nasal and sonorant clusters. Now we see how voiced obstruents in the dialect undergo nasal substitution. Why does a sequence of nasal and voiced obstruent not need to undergo nasal substitution? This is because the sequence allows a more leisurely rising of the velum than nasal and voiceless obstruents (Huffman, 1993: 310, cited in Pater, 1999). Thus, there is no need for a voiced obstruent following a nasal segment to undergo nasal substitution. Some dialects of Malay, however, prove the fact that a voiced obstruent following a nasal segment will also undergo nasal substitution. This interesting process that occurs in Malay dialects cannot be found in SM. As discussed in Chapter 4, nasal substitution is only applied to break up a sequence of nasal and voiceless obstruents due to the high obedience to *NÇ̊. Therefore, a word like /məŋ+paksa/, 'to force', surfaces as [məm$_{12}$aksa], where the nasal and voiceless obstruent cluster undergoes nasal substitution. Based on evidence from the data for Perak, I claim that nasal and voiced obstruent clusters also undergo nasal substitution. Thus this section will be discussing this issue further, using apparatus available within OT. Before we proceed to the analysis, let us first observe the Perak data.

226. Nasal and voiced obstruent clusters in Perak (from Ahmad, 1991).

(i)	/ŋ-bagi/	[magi]
	ACT.PRF-give 'to give'	
(ii)	/ŋ-dapat/	[napat]
	ACT.PRF-obtain 'to obtain'	
(iii)	/ŋ-gosok/	[ŋosoʔ]

ACT.PRF-brush 'to brush'

(iv) /ŋ-basuh/ [masoh]

ACT.PRF-wash 'to wash'

The above data reveal that voiced obstruents after nasals also undergo substitution. The question to be addressed in this subsection is why the initial voiced obstruent of the root needs to be substituted since the language does licence nasal-voiced obstruent clusters in a word. What argument in OT can be offered to explain the phenomenon of substituting a voiced obstruent following a nasal segment?

As discussed in the literature review, a voiced obstruent in Muna can be blocked from undergoing nasal substitution by adding a constraint named IDENT [PHAREXP]. This can be seen in the tableau presented in (101), in Chapter 2 subsection 2.4.2.4. As presented in Pater (2001), a voiced obstruent can be blocked from undergoing nasal substitution by ranking IDENT [PHAREXP] above CRISP-EDGE [PrWD] (Pater, 2001: 176). The following tableau demonstrates how this ranking blocks a voiced obstruent from undergoing nasal substitution, but not a voiceless obstruent.

227. IDENT [PHAREXP] >> CRISP-EDGE [PrWD] (Pater, 2001: 176).

/məŋ$_1$+b$_2$əli/	IDENT[PHAREXP]	CRISP-EDGE[PrWD]
a.☞məm$_1$b$_2$əli		*
b. məm$_{12}$əli	*!	
/məŋ$_1$+p$_2$ilih/		
c.☞məm$_{12}$ilih	*	
d. məm$_1$p$_2$ilih		*!

We can see in the above tableaux that a voiced obstruent following a nasal segment can be blocked from undergoing nasal substitution by the constraint ranking: IDENT [PHAREXP] >> CRISP-EDGE [PrWD]. However, with the ranking reversed, CRISP-EDGE [PrWD] >> IDENT [PHAREXP], both voiceless and voiced obstruents are subject to fusion (Pater, 2001: 176), as the following tableau demonstrates:

228. CRISP-EDGE [σ] >> IDENT [PHAREXP] (Pater, 2001).

/məŋ$_1$+b$_2$əli/	CRISP-EDGE[PrWD]	IDENT[PHAREXP]
a. məm$_1$b$_2$əli	*!	
b.☞məm$_{12}$əli		*
/məŋ$_1$+p$_2$ilih/		
c.☞məm$_{12}$ilih		*
d. məm$_1$p$_2$ilih	*!	

Considering the ranking CRISP-EDGE[PrWD] >> IDENT[PHAREXP], as in the above tableau, as well as a voiceless obstruent, a voiced obstruent following a nasal segment at a prefix-root juncture can also undergo nasal substitution. To account for the case in Perak, nasal substitution with voiced obstruents can be attributed to the ranking of IDENT [PHAREXP] beneath CRISP-EDGE [σ][48], as demonstrated in the following tableau:

[48] Observe that although I make use of Pater's (2001) work regarding both voiceless and voiced obstruent being able to undergo nasal substitution by reversing the ranking IDENT [PHAREXP] >> CRISP-EDGE[PrWD] to CRISP-EDGE[PrWD] >> IDENT [PHAREXP], I will continue to use CRISP-EDGE [σ] in this analysis as I have already applied it in the analysis of nasal and voiceless obstruent clusters in PD.

229. Nasal substitution with voiced obstruents in Perak: CRISP-EDGE [σ] >> IDENT [PHAREXP].

/ŋ$_1$+b$_2$agi/	CRISP-EDGE[σ]	IDENT [PHAREXP]
a.☞m$_{12}$agi		*
b. m$_1$b$_2$agi	*!	

The above analysis clearly shows that a sequence of nasal and voiced obstruents can also undergo nasal substitution. Most previous studies of nasal substitution in Malay (e.g.: Hassan, 1974, 1987; Omar, 1975, 1993; Karim, 1995; Onn, 1980) only relate this phonological process to nasal and voiceless obstruent clusters. Therefore, *NÇ̥ plays an important role in their analysis. This present analysis of Malay dialects reveals that the *NÇ̥ constraint, which prohibits nasal and voiceless obstruent sequences, cannot satisfactorily explain the phenomenon of voiced obstruent nasal substitution. Hence, to provide a better explanation, *NÇ̥ is replaced by another constraint, CRISP-EDGE [σ], which is able to capture nasal substitution in both voiceless and voiced obstruents. In Chapter 4, I demonstrated the importance of the *NÇ̥ constraint in eliminating the occurrence of nasal and voiceless obstruent clusters in SM. In order to satisfy this markedness constraint, the clusters undergo nasal substitution. It is however difficult to explain the process of nasal substitution with voiced obstruents since it is beyond the limits of *NÇ̥ to eliminate a sequence of nasal and voiced obstruents. I establish the following tableau to see how difficult it is for the constraint to capture voiced obstruent nasal substitution. If *NÇ̥ were to apply, then the optimal output would fall to candidate (b) which is not the correct output in the dialect.

230. *NC̥ >> IDENT [PHAREXP]: *[mbagi]

/ŋ$_1$+b$_2$agi/	*NC̥	IDENT[PHAREXP]
a. m$_{12}$agi		*!
b.☞m$_1$b$_2$agi		

As we can see, candidate (a) is ruled out because it violates IDENT [PHAREXP]. On the other hand, candidate (b), which does not violate any constraint in the hierarchy, is chosen as the winner. This is an incorrect output since a voiced obstruent after a nasal segment in the dialect must undergo nasal substitution. It is clear that *NC̥ in the above constraint hierarchy does not play a significant role in ruling out a sequence of nasal and voiced obstruents, as in candidate (b), since the constraint is specifically for nasal and voiceless obstruent clusters. This shows that *NC̥ cannot be used to account for voiced obstruent nasal substitution as occurred in PD above. From this, we have now answered the Research Question in 1(c) (see Section 1.3). Thus, *NC̥ must be replaced by CRISP-EDGE [σ] which is able to rule out both voiceless and voiced obstruents following nasal segments. By considering the constraint ranking CRISP-EDGE [σ] >> IDENT [PHAREXP] above, I establish a new hierarchy of constraint ranking for PD to account for voiced obstruent nasal substitution.

231. Voiced obstruent nasal substitution in PD.

/ŋ$_1$+b$_2$agi/	PRSTEM	DEP IO	NAS ASS	MAX -IO	UNI-ROOT	CRISP--EDGE[σ]	EDGE-INTEG	IDENT [PHAREXP]	UNIFORMITY
a.☞ m$_{12}$agi								*	*
b. m$_1$b$_2$agi						*!			
c. ŋ$_1$b$_2$agi			*!						

The failed candidate, (c), violates NAS ASS as the nasal segment in the prefix does not assimilate to the place of articulation to the following onset consonant. Since the NAS ASS constraint is ranked higher in the hierarchy, candidate (c) is ruled out. The competing candidates now are (a) and (b). As we see, by ranking CRISP-EDGE [σ] above IDENT [PHAREXP], the candidate with a nasal and voiced obstruent cluster, candidate (b), is ruled out. Thus candidate (a) emerges as the winner as it only violates the latter constraint, IDENT [PHAREXP], which bans voiced obstruents from undergoing nasal substitution.

It is now apparent from the above tableau that the CRISP-EDGE [σ] constraint cannot limit voiced obstruents from undergoing nasal substitution in Perak, which does occur in Muna, as presented in Chapter 2 subsection 2.4.1.2. As we saw in (231), CRISP-EDGE [σ] also permits voiced obstruent nasal substitution. The effect of CRISP-EDGE [σ] is to permit both voiced and voiceless obstruents to undergo nasal substitution, thus this constraint is preferred for the analysis of Malay dialects, rather than *NÇ, which limits nasal substitution to voiceless obstruents only. Since the role of *NÇ is limited, it cannot handle the case being discussed here.

6.3.2 Kelantan dialect

In the above discussion, we have seen that Perak is one of the Malay dialects that lacks nasal substitution within a root. However, as we shall see in this section, a reverse state of affairs occurs in Kelantan. Different from Perak, root-internal nasal and voiceless obstruent clusters in Kelantan undergo nasal deletion. To begin with, let us first consider the following examples:

232. Nasal and voiceless obstruent clusters in Kelantan (from Che Kob, 1985).

SM	KD	
[bantal]	[bata:]	'pillow'
[sampul]	[sapo:]	'envelope'
[tampar]	[tapa:]	'to slap'
[pantun]	[patoŋ]	'poem'
[santan]	[sate]	'coconut milk'
[kelantan]	[kelatɛ]	'name of state'
[untuŋ]	[ʔutoŋ]	'profit'
[gantuŋ]	[gatoŋ]	'to hang'
[paɲʧuŋ]	[paʧoŋ]	'to lop off'

From the above examples, nasal and voiceless obstruent clusters are not manifested phonetically in Kelantan root-internally. As these clusters are disfavoured in the dialect, a nasal segment preceding a voiceless obstruent would be deleted. This can clearly be seen in, for example, the word /gantoŋ/ which becomes [gatoŋ], where the nasal segment undergoes a process of deletion as the way to get rid of nasal and voiceless obstruent clusters in the surface representation.

Because a nasal segment before a voiceless obstruent in Kelantan is deleted, the optimal output violates the faithfulness constraint, MAX-IO, that requires input segments to have output correspondents. Therefore, MAX-IO must be ranked beneath constraints disfavouring other outcomes. The following tableau demonstrates the evaluation:

233.

/gan$_1$t$_2$oŋ/	DEP-IO	UNI-ROOT	UNI	NAS ASS	EDGE-INTEG	MAX-IO
a.☞gat$_2$oŋ						*
b. gan$_{12}$oŋ		*!				
c. gan$_1$ət$_2$oŋ	*!					

Candidate (b) violates UNIFORM-ROOT because of the sequence of nasal segment and voiceless obstruents within the word which undergo nasal substitution. Candidate (c) violates DEP-IO since epenthetic schwa is applied between the nasal segment and the voiceless obstruent. Candidate (a) is optimal as it satisfies the highest-ranked constraints defining nasal deletion as a way of avoiding a sequence of nasal and voiceless obstruents. As the result of nasal deletion, candidate (a) violates MAX-IO. The violation of MAX-IO is, however, not significant as the optimal output has already been determined.

When evaluating the grammar of Kelantan nasal deletion, another potential candidate that must be taken into consideration is *[gan$_1$t$_2$oŋ]. This potential candidate might possibly be included in the constraint ranking of Kelantan, since nasal and voiceless obstruent sequences within a root can also have opportrunity to surface, as occurs in Perak. To complete our analysis, we must explain why this candidate cannot be chosen as the optimal candidate in Kelantan. Thus, there must be another constraint, which bans a nasal and voiceless obstruent sequence within a root, and this constraint must rank above MAX-IO in order to rule out [gan$_1$t$_2$oŋ] and [gat$_2$oŋ] being able to emerge as the optimal candidate. A possible solution to account for why such output as [gan$_1$t$_2$oŋ] cannot be chosen as the optimal output in the Kelantan dialect can be found by appealing to the constraint CRISP-EDGE [σ].

Returning to the point I mentioned in the preceding subsection, i.e. why CRISP-EDGE [σ] is chosen in this analysis of Malay dialects rather than CRISP-EDGE [PrWD], I now provide the justification for choosing this constraint. According to Itô and Mester (1999a), three types of CRISP-EDGE have been used in the literature to solve certain phonological matters. There are: (1) CRISP-EDGE [PrWD], as defined in (75), which was used by Itô and Mester (1999) in their analysis of the prosodic morphology of Sino-Japanese. Pater (2001) has also used this constraint when accounting for nasal substitution in Indonesian. (2) CRISP-EDGE [Ft] was used to explain the ambisyllabicity in most of dialects of English which is claimed only to be possible in non-foot-initial position (Kiparsky, 1979).[49] The other one is (3) CRISP-EDGE [σ] which is used to rule out gemination and similar cases such as double-linking (Itô and Mester, 1999a).

As I suggested in subsection 6.3.1, the phenomenon of nasal and voiceless obstruent clusters in Malay dialects is best analysed by positing CRISP-EDGE [σ] (78). CRISP-EDGE [PrWD] in (75), which was used by Pater (2001) to account for nasal substitution in Indonesian, was only used to discuss the clusters at prefix-root junctures. The analysis did not however consider the clusters occurring root-internally. In that case, the use of CRISP-EDGE [PrWD], as in Pater's (2001) analysis

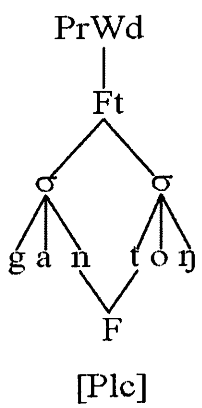
[49] Refer to the cited reference for more detail.

concerning nasal substitution in Indonesian, is plausible as nasals and voiceless obstruents are not in the same prosodic word, and therefore the clusters undergo nasal substitution.

If CRISP-EDGE[PrWD] were to apply when accounting for the clusters of root-internal nasal and voiceless obstruents in Indonesian, as applied at prefix-root junctures, it could not explain why the clusters occur root-internally, since they belong to the same prosodic word. The same problem applies to some dialects of Malay, such as Perak, where the clusters are permitted within roots. It would be difficult to capture the case if CRISP-EDGE [PrWD] was applied. Candidates without nasal substitution or nasal deletion in Kelantan could not be ruled out by CRISP-EDGE [PrWD], since the clusters are within the same Prosodic Word. I now illustrate the following prosodic word structures of root-internal nasal and voiceless obstruent clusters in Kelantan:

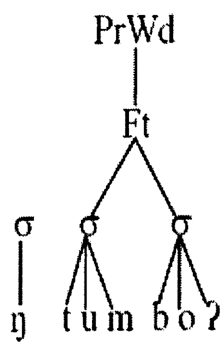
234. Prosodic word structures in Kelantan.

(a) Root-internal



[gantoŋ]

(b) Prefix juncture



[ŋnumboʔ]

As required by CRISP-EDGE [PrWD] in (75), no element belonging to a Prosodic Word may be linked to a prosodic category external to that of the Prosodic Word. This means that the sequence of nasal and voiceless obstruents in 234(a) does not violate CRISP-EDGE [PrWD], as the nasal segment and the voiceless obstruent belong to the same Prosodic Word and are not linked to a prosodic category external to that of the Prosodic Word. CRISP-EDGE [PrWD] is, however, violated in 234(b), as the voiceless obstruent [t] is an element that belongs to the Prosodic Word which is linked to the nasal segment in the prefix, which is the element that belongs to the prosodic category external to that Prosodic Word.

Returning to our discussion of the Kelantan dialect, if CRISP-EDGE[PrWD] were to apply in the analysis, the candidate *[gan$_1$t$_2$oŋ] from the input /gan$_1$t$_2$oŋ/ could not be ruled out, since it does not violate CRISP-EDGE[PrWD] and, therefore, the desired optimal output [gat$_2$oŋ], could not emerge as the winner.

The following tableau sums up the argument I have just made, showing that suboptimal candidate *[gan$_1$t$_2$oŋ] cannot be ruled out by CRISP-EDGE[PrWD] with the same schematic ranking that is in (233): DEP-IO >> UNIFORM-ROOT >> UNIFORM >> NAS ASS >> EDGE-INTEGRITY >> CRISP-EDGE[PrWD] >> MAX-IO, as demonstrated in (235). This ranking fails to account for the phenomenon under discussion, since it yields an incorrect output:

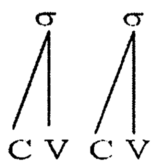
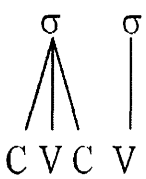
235.

/gan₁t₂oŋ/	DEP-IO	UNI-ROOT	UNIFORMITY	NAS ASS	EDGE-INTEG	CRISP-EDGE [PrWD]	MAX-IO
a.☹ gat₂oŋ							*
b. gan₁₂oŋ		*!					
c. gan₁ət₂oŋ	*!						
d.☞gan₁t₂oŋ							

The above tableau shows that CRISP-EDGE [PrWD] cannot successfully determine which candidate should emerge as the optimal output. As we also see in the above tableau, candidate (d), which emerges as the winner, is incorrect output as a nasal preceding a voiceless obstruent root-internally in the Kelantan dialect would have to be deleted. Candidate (a), indicated with an unhappy face, should be the correct output; however, it does not emerge as the winner in the above constraint ranking. As an alternative solution therefore, I have suggested that CRISP-EDGE [σ] must be considered in the constraint ranking of Malay dialects.

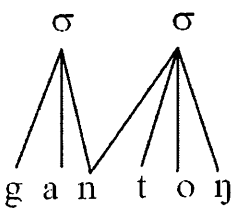
By replacing CRISP-EDGE [PrWD] with CRISP-EDGE[σ] in the constraint ranking, *[gan₁t₂oŋ] can be ruled out and, therefore, the desired optimal output, [gat₂oŋ], can be obtained. CRISP-EDGE [σ] is able to rule out *[gan₁t₂oŋ] because the homorganic nasal segment has the same place of articulation as the following voiceless obstruent. By having the same place of articulation, these two adjacent segments are linked together in the same syllable. According to the obedience and violation of CRISP-EDGE [σ], as exemplified by Itô and Mester (1999a) and shown in Table 7, I claim that the violation of CRISP-EDGE [σ] by candidate *[gan$_1$t$_2$oŋ] falls out from the representation in (d).

Table 7: Obedience and violation of CRISP-EDGE [σ] (from Itô and Mester, 1999a: 209).

Obedience of CRISP-EDGE [σ]	Violation of CRISP-EDGE [σ]
a 	c. 
b. 	d. 

For a clearer view, I illustrate in the following diagram how *[gan$_1$t$_2$oŋ] violates CRISP-EDGE [σ], as shown in (d) in the above representation.

236. Representation of CRISP-EDGE [σ] violation for *[gan$_1$t$_2$oŋ].



As established in the earlier tableau (235), CRISP-EDGE [PrWD] is not violated by this candidate and, therefore, it surfaces as optimal output. Considering CRISP-EDGE [σ] and the potential candidate *[gan$_1$t$_2$oŋ], I re-establish the following tableau for the Kelantan dialect, which gives a better picture of how the candidate *[gan$_1$t$_2$oŋ] can be ruled out and the candidate with nasal deletion, [gat$_2$oŋ], surfaces as optimal output:

237. Nasal and voiceless root-internal obstruent clusters in Kelantan.

DEP-IO >> UNIFORMITY-ROOT >> UNIFORMITY >> NAS ASS >> EDGE-INTEG >> CRISP-EDGE [σ] >> MAX-IO

/gan$_1$t$_2$oŋ/	DEP-IO	UNI ROOT	UNIFORMITY	NAS ASS	EDGE-INTEG	CRISP-EDGE[σ]	MAX IO
a.☞gat$_2$oŋ							*
b. gan$_{12}$oŋ		*!					
c. gan$_1$ət$_2$oŋ	*!						
d. gan$_1$t$_2$oŋ						*!	

It is clear, from the above tableau, that CRISP-EDGE [σ] is the better constraint, rather than CRISP-EDGE [PrWD]. This constraint can prevent candidate (d), with a nasal and voiceless obstruent sequence, from emerging as the winner. Since *[gan$_1$t$_2$oŋ] violates CRISP-EDGE [σ], candidate (a), which only violates the latter, i.e. MAX-IO, emerges as the correct output.

The deletion of nasals preceding voiceless obstruents root-internally clearly shows that Kelantan is quite direct in banning nasal and voiceless obstruent clusters in the dialect when compared to Perak. Now we examine the occurrence of this cluster at prefix-root junctures in Kelantan. The process of prefixation to a nasal final prefix in Kelantan dialect is, however, somewhat different from Perak. Before we proceed to the analysis, let us observe the data for nasal and voiceless obstruent sequences at prefix-root junctures in Kelantan.

238. (a) Nasal and voiceless obstruent clusters in Kelantan (from Teoh, 1994).

(i)	/ŋ-toloŋ/	[nnũloŋ]
	ACT.PRF-help 'to help'	
(ii)	/ŋ-piker/	[mmĩke:]
	ACT.PRF-think 'to think'	
(iii)	/ŋ-tukar/	[nnũka:]
	ACT.PRF-change 'to change'	
(iv)	/ŋ-katʃau/	[ŋŋãtʃa]
	ACT.PRF-disturb 'to disturb'	
(v)	/ŋ-kipas/	[ŋŋĩpah]
	ACT.PRF-fan 'to fan'	
(vi)	/ŋ-tumbuk/	[nnũmboʔ]
	ACT.PRF-pound 'to pound'	

(b) Nasal and voiceless obstruent clusters in Kelantan (from Che Kob, 1985)[50]

(i)	/ŋ-kanduŋ/	[ŋŋãnduŋ]
	ACT.PRF- contain 'to contain'	
(ii)	/ŋ-toreh/	[nnɔ̃ɣɛh]
	ACT.PRF-tap 'to tap'	
(iii)	/ŋ-susup/	[ɲɲũɣuʔ]
	ACT.PRF-infiltrate 'to infiltrate'	
(iv)	/pəŋ-tolak/	[nnɔ̃lɔʔ]
	NOM.PRF-push 'pusher'	
(v)	/pəŋ-takoq/	[nnãkoq]
	NOM.PRF-scared 'scared person'	
(vi)	/pəŋ-tari/	[nnãri]
	NOM.PRF-dance 'dancer'	

[50] Note, some of the examples are taken from Che Kob (1985) and some are obtained from the interview session with him in person. More examples of nasal and voiceless obstruent clusters at prefix-root junctures in Kelantan can be seen in the word list in Appendix N.

The phonology of the Kelantan dialect in (238) can be summarised as follows: (i) a nasal segment assimilates to the place of articulation of a following voiceless obstruent; (ii) a voiceless obstruent becomes a nasal segment, which has the same place of articulation; and (iii) the first vowel following a nasal segment is nasalised. From the above examples, we see that voiceless obstruents undergo alternation to nasals when preceded by a nasal segment in the prefix. The alternation segment of voiceless obstruents to nasals creates consonants clusters at the prefix juncture, as illustrated in (238).

Before we proceed any further, it is worth discussing the single velar nasal /ŋ+/ as an active prefix in the data shown above. Based on my discussion with the author of the book titled *'Dialek Geografi Pasir Mas'*, Che Kob, /ŋ+/ can also be used as /məŋ+/ in the underlying form. These two active prefixes can be in free variation or in complementary distribution in the dialect. The single nasal prefix /ŋ+/ can be used for any consonant initial base, while /məŋ+/ can only be attached to bases that start with /m, n, ɲ, ŋ, l, ɣ/. As mentioned in subsection 6.3.1, /ŋ+/ in PD has, as claimed by Ahmad (1991), two allomorphs, /ŋ+/ and /məŋ+/. The velar nasal /ŋ+/ is a variant of /məŋ+/. This prefix undergoes an initial syllable reduction (ibid.: 55). Syllable reduction has been claimed by previous scholars to be a common case in most Malay dialects, such as in Kedah (Collins, 1986: 10; Omar, 1986: 99), Kelantan (Che Kob, 1985: 270), Terengganu (Omar, 1985: 287), Pahang (Collins, 1983: 104) and Perak (Ahmad, 1991: 55). Thus, in our analysis, /ŋ+/ is used to account for any initial obstruent consonants bases in the dialect (these include voiceless and voiced obstruents), while /məŋ+/ will be used for initial sonorant bases.

Turning back to the data presented in (238) above, a voiceless obstruent following a nasal segment alternates to a nasal consonant which takes the place of articulation of the voiceless obstruent that remains in the nasal consonant. For example, if the voiceless obstruent following the nasal segment in the prefix is /p/, then it alternates to [m] on the surface. This process of alternation from /p/ to [m] I here call nasalisation. The representation of nasalisation is illustrated in the following diagram:

239. The representation of nasalisation.

Input:	ŋ	p
	|	|
Output:	m	m

The alternation of a voiceless obstruent to a nasal consonant in Kelantan at prefix-root junctures is quite similar to a case in the Bantu language, Kikongo (Meinhof, 1932; Dereau, 1955; Webb, 1965; Ao, 1991; Odden, 1994; Piggott, 1996, cited in Rose and Walker, 2004), as exemplified in (240). In this language, a voiced obstruent in the suffix /-idi/ alternates to a nasal consonant [-ini] when preceded by a nasal consonant at any distance in the stem (Rose and Walker, 2004). Given the case in Kikongo, we could say that the initial consonant in the root in Kelantan alternates to a homorganic nasal when preceded by a nasal final prefix, as we can see in (238). See the examples of Kikongo data from Rose and Walker (ibid.) below:

240. Kikongo suffix /-idi/ becomes [-ini].

1(a)	m-[bud-idi]$_{stem}$ 'I hit'	2(a)	tu-[kun-ini]$_{stem}$ 'we planted'
(b)	n-[suk-idi]$_{stem}$ 'I washed'	(b)	tu-[nik-ini]$_{stem}$ 'we ground'

Note that the phonological process that occurs in the Kelantan dialect of Malay has also been examined by Pater (1999). He states that, as well as in Indonesian, Kelantan Malay also undergoes nasal substitution that includes nasal deletion. As we have just seen, nasal and voiceless root-internal obstruent clusters in Kelantan undergo nasal deletion, and this is consistent with what Pater claimed. However, based on the above data for Kelantan prefixation, I would like to revise the claim made by Pater (1999).

If we observe the process of prefixation to a nasal final prefix in the dialect, we can see that the clusters are not resolved by nasal deletion, as applied root-internally. I claim that there are two processes that occur at the prefix-root juncture in Kelantan dialect to eliminate nasal and voiceless obstruent clusters, which I refer to here as nasal assimilation and nasalisation. This means that there is more than one method that the dialect uses to get rid of the clusters, as well as the nasal deletion which is applied root-internally (see the illustration in 237). Probably, the claim made by Pater (1999) was only concerned with nasal and voiceless root-internal obstruent clusters, without considering the occurrence of clusters at prefix-root junctures. Thus, the claim is not correct in its explanation of the occurrence of clusters at any morphological domain in the dialect.

For a clearer view, I illustrate the representations of nasal deletion, nasal assimilation and nasalisation in Kelantan Malay, as follows:

241. The representations of nasal deletion, assimilation and nasalisation in Kelantan.

	Kelantan Malay (Root-internally)	Kelantan Malay (Prefix-juncture)	
Input:	n t	ŋ t	t
Output:	t	n	n
Process:	nasal deletion	assimilation	nasalisation

The phonological processes that occur at prefix-root junctures in Kelantan, as we saw, are quite distinct in that the voiceless obstruents do not undergo the process of deletion, as is regularly applied in the language. As I have just claimed, the voiceless obstruents at the prefix-root junctures in Kelantan undergo nasalisation instead of deletion. In an OT account, the irregularity process that occurs at the prefix juncture in Kelantan is an explainable phenomenon. This irregular behaviour has never been discussed in previous Malay studies. Pater's analysis (1999) of nasal substitution in Austronesian languages only discussed nasal and voiceless obstruent clusters within the roots in Kelantan. In OT, the reason why voiceless obstruent deletion does not take place in the optimal output for this dialect can be explained by the candidate to hand not being the candidate best satisfying the constraint hierarchy. By employing the same ranking hierarchy that we have in the analysis of nasal and voiceless root-internal obstruents in Kelantan, as demonstrated in (237), together with the constraint PROSODIC STEM (35) which is required to account for stems (i.e. prefix and root), I establish the following tableau:

242. PROSODIC STEM >> DEP-IO >> UNIFORMITY-ROOT >> UNIFORMITY >> NAS ASS >> EDGE-INTEG >> CRISP-EDGE [σ] >> MAX-IO.

/ŋ₁+t₂umboʔ/	PRSTEM	DEP-IO	UNI-ROOT	UNIFORMITY	NAS ASS	EDGE-INTEG	CRISP-EDGE[σ]	MAX-IO
a. $ŋ_1$ $t_2$umboʔ					*!			
b. n_{12}umboʔ				*!		*		
c.☞$n_1n_2$umboʔ								
d. n_1ə$t_2$umboʔ		*!						
e. $n_1t_2$umboʔ							*!	

As can be seen, the failed candidate (b), which undergoes nasal substitution, is ruled out because it involves a violation of UNIFORMITY. Because UNIFROMITY is ranked higher in the hierarchy, candidate (b) cannot emerge as the optimal output. By violating UNIFORMITY, it explains why nasal and voiceless obstruent clusters at the prefix juncture in Kelantan do not undergo the regular process of nasal substitution. In order to break up [nt] clusters, schwa is epenthesized in candidate (d). This possible way that nasal and voiceless obstruent clusters can be eliminated is, however, not the best way, as schwa epenthesis incurs a fatal violation of DEP-IO. As required by NAS ASS, the nasal segment in a coda position must be homorganic to the following consonant. The unassimilated nasal before the voiceless obstruent in candidate (a) thus violates this constraint, as the nasal segment [ŋ] in the prefix is not in the same place of articulation as the following consonant [t]. Unlike candidate (a), the nasal segment in the prefix in candidate (e) assimilates to the following consonant [t]. However, the voiceless obstruent does not undergo the process of deletion, hence it violates CRISP-EDGE [σ], which prohibits multiple linking between syllables. Therefore, candidate (e) is non-optimal. Candidate (c) is the optimal output, since this candidate violates none of the constraints in the hierarchy.

We have seen in the above discussion that Kelantan disfavours nasal and voiceless obstruent clusters at prefix-root junctures. Are any clusters banned in the dialect apart from nasal and voiceless obstruents? In what follows, I am going to discuss another cluster that occurs at prefix-root junctures that is both nasal and sonorant. Before we see how this cluster is treated in the dialect, it would be helpful to consider some relevant examples from it:

243. Nasal and sonorant clusters at prefix-root junctures (from Che Kob, 1981).

(i)	/mŋ+ɣəseʔ/ ACT.PRF-hot 'to express the very hot weather'	[m-ɣəseʔ]
(ii)	/mŋ+masôʔ/ ACT.PRF-cook 'to cook'	[m-masôʔ]
(iii)	/mŋ+rapeq/ ACT.PRF-nonsense 'to talk nonsense'	[m-rapeq]
(iv)	/mŋ+minuŋ/ ACT.PRF-drink 'to drink'	[m-minuŋ]
(v)	/məŋ+lətuʔ/ ACT.PRF-explode 'to explode'	[m-lətuʔ]
(vi)	/ŋ+ɣəkɔh/ ACT.PRF-crack 'to crack'	[m-ɣəkɔh]
(vii)	/mŋ+ɣutuh/ ACT.PRF-collapse 'to collapse'	[m-ɣutuh]

Recall, as mentioned above (see p.292), /məŋ+/ can be attached only to certain initial consonants bases, i.e. /m, n, ɲ, ŋ, ɣ, l/. Therefore, this prefix is used to account for nasal and sonorant clusters at prefix-root junctures. The examples above show that sonorant consonants following nasal segments at prefix-root junctures are not deleted. When initial sonorant bases are combined with /məŋ+/, the final nasal segment of the prefix, i.e. /ŋ/, is deleted. From this, the output form must be, for example, [mə-masôʔ] or [mə-lətuʔ]. From the data shown above, why is schwa absent from the prefix with only [m] left? I am inclined to claim that the reason why only [m] has shown up in the data is because the prefix has undergone so-called syllable reduction. Speakers of Kelantan tend to minimize a syllable in a word. When they speak, one or more syllables of the words will be silent. Thus, outputs like [m-masôʔ] and [m-lətuʔ] are produced instead of [mə-masôʔ] and [mə-lətuʔ], respectively. Therefore, only the [m] segment in the prefix is realised in output forms. This situation of syllable minimisation also occurs within roots in the dialect, as the following examples illustrate. In the examples in (244), the syllable minimisation is more like schwa deletion as only the schwa in the first syllable is deleted.

244. Syllable minimization in roots (from Che Kob, 1985).

	SM	KD
(i)	[kəliliŋ] 'around'	[kliliŋ]
(ii)	[səlimut] 'duvet'	[slimũʔ]
(iii)	[kəramat] 'supernatural power'	[kɣamãʔ]
(iv)	[təmbaga] 'copper'	[tmãgɔ]
(v)	[səluar] 'pants'	[sluwa]
(vi)	[bəɣuwaŋ] 'bear'	[bɣuwɛ̃]

I now establish a new tableau with the same constraint ranking as in (242) to account for sequences of nasals and sonorants in the Kelantan dialect.

245.

/ŋ₁+r₂apeq/	PRSTEM	DEP-IO	UNI-ROOT	UNIFORMITY	NAS ASS	EDGE-INTEG	CRISP-EDGE[σ]	MAX-IO
a. $r_2$apeq	*!							*
b.☞$m_1r_2$apeq								
c. m_1ə$r_2$apeq		*!						

As shown in the tableau, candidate (a) violates the highest constraint in the ranking as the candidate has no prefix. Although the failed candidate (c) avoids the cluster via schwa epenthesis, it fatally disobeys DEP-IO. Thus, schwa epenthesis is not the way to get rid of nasal and sonorant clusters. The remaining candidate (b) is thus the optimal candidate.

Unlike Perak, Kelantan does not allow voiced obstruents to undergo nasal substitution. Nasal and voiced obstruent clusters in this dialect only undergo nasal assimilation, with no deletion of voiced obstruents, the same as in SM. Nasal assimilation and undeleted voiced obstruents are illustrated in the following examples:

246. Nasal and voiced obstruent clusters in Kelantan.

(i) /ŋ+guntiŋ/ [ŋ-guntiŋ]
ACT.PRF-scissor 'to cut'

(ii) /ŋ+gatal/ [ŋ-gata:]
ACT.PRF-itchy 'itching'

(iii) /ŋ+batal/ [m-bata:]
ACT.PRF-cancel 'to cancel'

(iv) /pəŋ+mabuk/ [p-mãbuʔ]
NOML.PRF-drunk 'drunker'

(v) /ŋ+gosok/ [ŋ-gɔsɔʔ]

NOML.PRF-brush 'to brush'

(vi) /ŋ+bedil/ [m-bəde:]

ACT.PRF-gun 'to shoot'

To complete the analysis, the full hierarchy of constraints as established for the Kelantan dialect is: PROSODIC STEM >> DEP-IO >> UNIFORMITY-ROOT >> UNIFORMITY >> NAS ASS >> EDGE-INTEG >> CRISP-EDGE [σ] >> MAX-IO

247.

/ŋ+guntiŋ/	PRSTEM	DEP-IO	UNI-ROOT	UNIFORMITY	NAS ASS	EDGE-INTEG	CRISP-EDGE[σ]	MAX-IO
a.☞ $ŋ_1g_2$untiŋ							*!	
b. $g_2$untiŋ	*!							*
c. $ŋ_{12}$untiŋ				*!				

The above tableau shows that candidate (a), with nasal assimilation and an undeleted voiced obstruent, emerges as the winner. A candidate with nasal deletion has been ruled out early as it fatally violates the PRSTEM constraint. As the PRSTEM constraint requires each prefix and root to contain one syllabic, candidate (b) violates this constraint as the candidate only contains a root. Meanwhile, candidate (c) with nasal substitution is not chosen as the winner as it incurs violation of UNIMORMITY that bans nasal substitution. Therefore, candidate (a) is the optimal candidate. Since the voiced obstruent is not deleted, this candidate violates CRISP-EDGE[σ] because the nasal segment in the prefix and the initial voiced obstruent base are linked together.

6.3.3 Negeri Sembilan dialect

We now shall discuss how the situations discussed in the Perak and Kelantan dialects occur in another dialect of Malay concerned in this study; this is NS. We begin the discussion by presenting some examples of NS dialect nasal and voiceless root-internal obstruent clusters. Observe that this cluster is resolved in the same way as in Kelantan, whereby nasal segments preceding voiceless obstruents undergo deletion. Let us consider the following data:

248. Nasal and voiceless obstruent clusters in the NS dialect.[51]

SM	**NS**	
[ʤempot]	[ʤaput]	'to fetch'
[dentom]	[datum]	'an explosive sound'
[tampa:]	[tapo]	'to slap'
[kampoŋ]	[kapoŋ]	'village'
[sampan]	[sapan]	'small boat'
[teŋkaʔ]	[tekaʔ]	'level'
[simpan]	[sipan]	'to keep'

The data in (248), above, show that nasal and voiceless obstruent clusters within roots in NS undergo nasal deletion. Because of this, the root-specific faithfulness constraint, UNIFORMITY-ROOT, must be low in the constraint ranking. Now I establish the following tableau to account for root-internal nasal and voiceless obstruent clusters in NS.

[51]The data presented here are from Rufus (1966) and also the interview conducted with two native speakers of NS: Dr. Mohd Fadzeli Jaafar and Mr. Zulkifli Ahmad, as mentioned in Chapter 3 subsection 3.3.3. A list of additional data of NS is enclosed at the end of the thesis and was obtained from those two native speakers of NS. I would like to thank them for providing me with more data from NS for my analysis.

249.

/dan₁t₂um/	DEP-IO	CRISP-EDGE[σ]	NAS ASS	UNI-ROOT	MAX-IO	UNIFORMITY	EDGE-INTEG
a. dan_1t_2um		*!					
b.☞dat_2um					*		
c. $dan_{12}um$				*!			
d. $dan_1ət_2um$	*!						

As shown in the above tableau, when CRISP-EDGE [σ] is high-ranked, candidate (a) with a nasal and voiceless obstruent cluster is ruled out. Candidate (c) with nasal substitution is ruled out by the constraint on anti-nasal substitution, UNIFORMITY-ROOT, while candidate (d) with schwa epenthesis is ruled out by DEP-IO. Thus the latter candidate, (b), is more harmonic than the former, even though the nasal segment is deleted.

The same as in the two dialects discussed above, NS also does not permit two types of clusters at prefix-root junctures in the surface representation. The two clusters are: (1) nasal and voiceless obstruents and (2) nasal and sonorant clusters. Now I will first examine how nasal and voiceless obstruent clusters at prefix-root junctures in NS are treated. Nasal and sonorant clusters will follow. Before we go to the data analysis, it is worth knowing the allomorph structure of nasal final prefixes /mŏN-/ and /pŏN-/ in the NS dialect as they give effect to the phonological processes applied in the dialect, especially in the process of alternation segments. As stated in Rufus (1966: 48), the prefix /mŏN-/ has two sets of basic allomorphs which are /mŏN-/ and /$Ø^{n-}$ /.

According to Rufus, /mŏN-/ occurs regularly with any initial-consonant of the root except for /ʧ/. The use of /mŏN-/ as a prefix is shown in the following examples:

250. Prefix /mŏN-/ in Kelantan (from Rufus, 1966: 52).

	Roots	**Prefixed words**
(i)	/hisap/ 'suck'	[mo-hisap] ACT.PRF-suck 'to suck'
(ii)	/kɑreh/ 'hard'	[mo-ŋɑreh-kan] ACT.PRF-hard-CAUS.SUF 'to cause to make hard'
(iii)	/pukɷ/ 'scold'	[m-mukɷ] ACT.PRF-scold 'to scold'
(iv)	/lawan/ 'fight'	[m-lawan] ACT.PRF-fight 'to fight'
(v)	/ganti/ 'replace'	[m-ŋanti] ACT.PRF-replace 'to replace'

The allomorph /$Ø^{n-}$ / is claimed by Rufus (ibid.: 52) to be a free alternant of /mŏN-/ that can be attached with a consonant stop initial-root or /s/. For example:

251. Prefix /$Ø^{n-}$ /in NS (from Rufus, 1966).

	Roots	**Prefixed words**
(i)	/caŋkɷ/ 'hoe'	[ɲaŋkɷ]=[m-ɲaŋkɷ] ACT.PRF-hoe 'to hoe
(ii)	/guntiŋ/ 'scissors'	[ŋuntiŋ]=[m-ŋuntiŋ] ACT.PRF-scissors 'to cut'
(iii)	/ʤoŋõq/ 'see'	[mñoŋõq]=[ñoŋõq] ACT.PRF-see 'to see'
(iv)	/sɑbɷt/ 'say'	[ñɑbɷt]=[mñɑbɷt] ACT.PRF-say 'to say something'
(v)	/pilih/ 'choose'	[mmilih]=[milih]

ACT.PRF-choose 'to choose'

(vi) /kaʤa/ 'chase' [ɲaʤa]=[maʤa]

ACT.PRF-chase 'to chase'

The examples show that the two prefixes in the NS dialect have no constant form when they are attached to roots. It is said that the prefix /mŏN-/ can be attached to those listed initial roots but, in the surface representation, the prefix can be [mo-] or [m-], as shown in (250). Likewise, for /Ø$^{n-}$ /, somehow this prefix surfaces as an assimilated nasal, [N-] or [m-], as exemplified in (251). It can be said that the two prefixes do not seem to show any difference in their surface forms. My justification of this is that if we refer to the examples in 250(v) and 251(ii), even though the /g/ initial-roots are combined with two different prefixes, the same outputs are obtained. The two different prefixes, /mŏN-/ and /Ø$^{n-}$/, surface as [m-]. Other examples are also found in Rufus where two different roots begin with /b/, /basωh/ 'wash' and /baʧa/ 'read', surface as [mo-masωh] ACT.PRF–wash 'to wash' and [m-maʧo] ACT.PRF–read 'to read', respectively.

Based on the above examples, it seems that the case under discussion is almost the same as the one in Perak. Recall that /ŋ-tɔʁɛh/ ACT.PRF-tap → [nɔʁɛh] 'to tap' and /ŋ-ɟadi/ ACT.PRF-be → [ɲadi] 'to be' in Perak can also be pronounced as [mə-nɔʁɛh] and [mə-ɲadi], respectively (Ahmad, 1991: 55). Thus, [mo-masωh] and [m-maʧo] in NS may also be realised as [m-masωh] and [mo-maʧo], respectively. In my analysis, /Ø$^{n-}$ /, which surfaces as an assimilated nasal, as in /ʤoŋõq/ 'see' → [ñoŋõq] ACT.PRF-see 'to see' and /pilih/ 'choose' → [**m**ilih], has undergone syllable reduction.

As we will see, nasal and voiceless obstruent clusters in NS at the prefix-root juncture are not very different from those in Perak, as the nasal segment of the prefix and the initial voiceless obstruent undergo the regular process of nasal substitution. Let us first observe the data for the NS dialect:

252. Nasal and voiceless obstruent clusters in NS (from Rufus, 1966).

(i)	/mŏN-pilih/	[mmilih]
	ACT.PRF-choose 'to choose'	
(ii)	/mŏN-pukul/	[mmukol]
	ACT.PRF-scold 'to scold'	
(iii)	/mŏN-tarik/	[mnariʔ]
	ACT.PRF-pull 'to pull'	
(iv)	/mŏN-kətuk/	[mŋətoʔ]
	ACT.PRF-knock 'to knock'	

As can clearly be seen, voiceless obstruents following nasals are deleted, while the nasals in the prefix assimilate to the same place of articulation with the following voiceless obstruent. Putting together all the constraints from (249) and the additional constraint which is important to account for STEMS (i.e. prefix and root), PRSTEM yields the following set of rankings to account for nasal and voiceless obstruent clusters at prefix junctures in NS: PROSODIC STEM >> DEP-IO >> CRISP-EDGE[σ] >> NAS ASS >> UNIFORM-ROOT >> MAX-IO >> UNIFORM >> EDGE-INTEG.

253.

/mŏN_1+$p_2$ilih/	PRSTEM	DEP-IO	CRISP-EDGE[σ]	NAS ASS	UNI-ROOT	MAX-IO	UNIFORMITY	EDGE-INTEG
a. mŏN_1p_2ilih				*!				
b.☞mŏm_{12}ilih							*	
c. mŏ$m_1p_2$ilih			*!					
d. mŏm_1ə$p_2$ilih		*!						
e. mŏ$p_2$ilih						*!		

The failed candidate (d) with a schwa epenthesized between the prefix and the root is ruled out as a consequence of violating DEP-IO. By having the same place of articulation, candidate (c) violates CRISP-EDGE [σ] as the two adjacent segments [m] and [p] are linked together in the same syllable. Since the nasal segment in candidate (a) does not assimilate to the following consonant, this candidate violates the NAS ASS constraint. Deleting the second nasal in the prefix causes candidate (e) to violate MAX-IO because the one of the input segments does not have an output correspondent. Candidate (b) with nasal substitution violates UNIFORMITY. The violation is however in significant as all other candidates have already been ruled out.

Rufus (1966) mentioned in her work that there are three types of affix in the NS dialect. These are prefix, circumfix and suffix. In the discussion of prefixes in the dialect, all the examples of data given are only concerned with single prefixes. There is no discussion of multiple prefixes being used in the dialect. Is it probable that the dialect has no multiple prefixes? To confirm this, two previous scholars' works have been referred to, Omar (1985, 2008) and Hussein (1973). Both scholars discuss the NS dialect in their analysis, but the discussions do not aim to provide a comprehensive study of it as they focus on other dialects of Malay. No examples of multiple prefixed words are found in either scholar's work.

Now we come to examine the other cluster which is nasal and sonorant. As we have seen in the two dialects discussed above, nasal deletion is the way to prevent this cluster from occurring in the surface representation. Do nasal and sonorant clusters have the same solution that the Perak and Kelantan dialects have? We begin the discussion by presenting some relevant examples from Rufus (1966):

254.

(i)	/mŏN-raso/	[mŏ-raso]
	ACT.PRF-taste 'to taste'	
(ii)	/mŏN-ratap/	[mŏ-ratap]
	ACT.PRF-wail 'to wail'	
(iii)	/mŏN-latus/	[mŏ-latus]
	ACT.PRF-erupt 'to erupt'	
(iv)	/mŏN-lawan/	[mŏ-lawan]
	ACT.PRF-fight 'to fight'	
(v)	/mŏN-repot/	[mŏ-repot]
	ACT.PRF-complain 'to complain'	
(vi)	/mŏN-ɲaɲi/	[mŏ-ɲaɲi]
	ACT.PRF-sing 'to sing'	
(vii)	/mŏN-rumpωt/	[mŏ-rumpωt]
	ACT.PRF-grass 'to do gardening'	

It is clear from the above examples that nasal and sonorant clusters are disfavoured in the NS dialect. The occurrence of this cluster in the surface representation is resolved by nasal deletion, in exactly the same way as it is in the Perak and Kelantan dialects. I establish a new tableau with the same constraint ranking as in (253) that was established above to account for a sequence of nasal and sonorants in the dialect.

255.

/mŏN_1+$r_2$aso/	PRSTEM	DEP-IO	CRISP-EDGE[σ]	NAS ASS	UNI-ROOT	MAX-IO	UNIFORMITY	EDGE-INTEG
a. mŏN_1r_2aso				*!				
b.☞mŏ$r_2$aso						*		
c. mŏN_1ə$r_2$aso		*!		*				

With the constraint ranking established in the above tableau, a sequence of nasal and sonorants is not selected as the winner, as seen in (a). Candidate (b) with nasal deletion emerges as the winner, instead.

As mentioned earlier in this chapter, voiced obstruents in two of the three dialects of Malay discussed in this chapter, Perak and NS, undergo nasal substitution. We have already discussed the case of voiced obstruent nasal substitution in Perak in subsection 6.3.1. Now we shall discuss the other dialect in which voiced obstruents also undergo nasal substitution, i.e. NS. We first begin the discussion by presenting the data from NS:

256. Nasal and voiced obstruent clusters in NS (from Rufus, 1966).

(i) /mŏN-baca/ [mmaco]
ACT.PRF-read 'to read'

(ii) /mŏN-bunuh/ [mmunoh]
ACT.PRF-kill 'to kill'

(iii) /mŏN-dapat/ [mənapat]

ACT.PRF-get 'to get'

(iv) /mŏN-ʤəmput/ [mŋaput]

ACT.PRF-invite 'to invite'

(v) /mŏN-ʤawap/ [mɲawap]

ACT.PRF-answer 'to answer'

(vi) /mŏN-bawa/ [pəmawaʔ]

ACT.PRF-carry 'carrier'

The data presented above show that nasal and voiced obstruent clusters in NS undergo the same process as in Perak, whereby this cluster also involves nasal substitution. The process is exactly the same as in Perak. Nasal segment in the prefix assimilates to the place of articulation of the following voiced obstruent. The voiced obstruent then is deleted. Now, we discuss how this situation is accounted for in a tableau analysis.

As I have demonstrated in Section 6.3.1, CRISP-EDGE [σ] is ranked above IDENT [PHAREXP] in order to rule out candidates with a sequence of nasal and voiced obstruents (see tableau in 229). By ranking CRISP-EDGE [σ] >> IDENT [PHAREXP], *[$m_1b_2$agi] does not emerge as the optimal output in Perak. The following tableau demonstrates how this ranking, CRISP-EDGE [σ] >> IDENT [PHAREXP], can rule out candidates with a sequence of nasal and voiced obstruents.

257. Nasal substitution with voiced obstruents in NS: CRISP-EDGE [σ] >> IDENT [PHAREXP].

/mŏN$_1$+b$_2$aco/	CRISP-EDGE[σ]	IDENT [PHAREXP]
a.☞ mŏm$_{12}$aco		*
b. mŏm$_1$b$_2$aco	*!	

As demonstrated in the above tableau, when CRISP-EDGE [σ] is ranked high, it is optimal for a sequence of nasal and voiced obstruents to undergo nasal substitution, as in candidate (a). Competing candidate (b) violates this constraint. The two segments [m] and [b] are linked together into one syllable as they share the same place of articulation.

Putting those two constraints together, along with the other constraints used for NS earlier, I now establish a new tableau to account for voiced obstruent nasal substitution in the dialect. Observe that the constraint ranking is still the same as for voiceless obstruent nasal substitution as demonstrated in (253), and for the nasal and sonorant clusters in (255).

258. Voiced obstruent nasal substitution in NS dialect.

/mŏN$_1$+b$_2$aco/	PRSTEM	DEP-IO	CRISP-EDGE[σ]	NAS ASS	UNI-ROOT	MAX-IO	UNIFORMITY	IDENT [PHAREXP]	EDGE-INTEG
a. mŏN$_1$b$_2$aco				*!					
b. mŏm$_1$b$_2$aco			*!						
c.☞mŏm$_{12}$aco								*	*

As shown in the above tableau, CRISP-EDGE [σ] dominates IDENT [PHAREXP] in the ranking. This directly explains why a voiced obstruent following a nasal segment undergoes nasal substitution.

The analysis presented above shows that two clusters, i.e. nasal and voiceless obstruents and nasal and sonorant clusters, are also disfavoured in the three Malay dialects of Perak, Kelantan and NS. Analysis of Perak and NS dialects reveals however that nasal and voiced obstruent clusters are also disfavoured in the dialects, as well as the two clusters.

6.4 Conclusion

The analysis of dialectal variation in Malay shows that the three clusters, (1) nasal and voiceless obstruents, (2) nasals and sonorants and (3) nasal and voiced obstruents, behave differently with respect to the phonological processes of the dialects that they belong to. The differences between the phonological processes are analysed in this study as the result of each dialect imposing a different particular dialect ranking. It is apparent from the analysis that the differences in ranking proposed in the theory can account straightforwardly for the multiple outputs from a single input which occur within a language or dialect variation.

As we have seen in the analysis, nasal and voiceless obstruent clusters are prohibited from emerging in the surface representation of all three of the dialects discussed above. In two of the three dialects, Kelantan and NS, the clusters are more strictly prohibited since they are not allowed to emerge even within roots. The root-internal clusters in these two dialects have been resolved by nasal deletion. In order for candidates with nasal deletion occurring root-internally in Kelantan and NS dialects to be chosen as optimal output, CRISP-EDGE[σ] is the constraint which is able to rule out candidates that spare UNIFORMITY-ROOT, which bans nasal substitution within a root. CRISP-EDGE[σ] is therefore ranked above MAX-IO in the ranking for the Kelantan and NS dialects, since the deletion of nasal segment results in the candidate violating MAX-IO.

Unlike Perak, clusters are allowed root-internally. In the process of prefixation to a nasal final prefix, nasal substitution violates UNIFORMITY as the constraint works against segmental fusion or coalescence (McCarthy and Prince, 1999). This constraint leads to a straightforward account of the lack of nasal substitution root-internally in Perak by employing UNIFORMITY-ROOT. As mentioned, such cases are captured, by McCarthy and Prince (1995b), by a general ranking schema in which a root-specific version of the faithfulness constraints must be ranked higher than the general version of these constraints, (cited in Kager, 1999: 76). By ranking UNIFORMITY-ROOT above UNIFORMITY, a sequence of nasal and voiceless obstruents can be prevented from undergoing nasal substitution within a root. The preservation of this cluster within roots, as in PD, is the consequence of a candidate's output best satisfying the root-specific constraint UNIFORMITY-ROOT in the hierarchy. Yet, this could also be another piece of evidence that a root is more faithful than other morphological affixes, as claimed by McCarthy and Prince (1995b).

The occurrence of nasal and voiceless obstruent clusters at prefix-root junctures is, on the other hand, completely banned in all three dialects. As discussed, nasal substitution is also the way to break up clusters at prefix-root junctures. However, in Kelantan, nasal substitution is not the way that is used to break up clusters. In Kelantan, nasalisation and assimilation are applied instead. The different ways of resolving nasal and voiceless obstruent clusters in the dialects of Malay discussed in this study show that nasal substitution is not only the strategy for eliminating nasal and voiceless obstruent clusters. From the discussion, we have seen that nasal deletion, nasalisation and assimilation, along with nasal substitution, are the strategies applied in Malay, including both SM and a number of its non-standard dialects. This provides us with the answer to the Research Question 1(b) (see Section 1.3 in Chapter 1).

The second cluster that this chapter has discussed is nasals and sonorants. We note from the above discussion that the occurrence of nasal and sonorant clusters is the same as in SM, where the clusters only occur at prefix-root junctures in non-standard dialects of Malay, i.e. in the Perak, Kelantan and NS dialects. As we saw, nasal and sonorant clusters at prefix-root junctures are always resolved by deleting the nasal segments in the prefixes. In this way, nasal deletion, as shown above, is applied in the three dialects concerned in this study. Thus, it can be said that nasal deletion is the only way that Malay, including SM and its dialects, uses to avoid nasal and sonorant clusters.

The other cluster discussed in this chapter is nasal and voiced obstruents. As presented in Chapter 4, nasal and voiced obstruent clusters are allowed in the surface representation in SM. Therefore, the clusters do not have to undergo nasal substitution, as nasal and voiceless obstruent do. In contrast, in non-standard dialects of Malay, nasal and voiced obstruent clusters also undergo nasal substitution as occurs in the Perak and NS dialects. By employing the constraint of faithfulness to obstruent voicing IDENT[PHAREXP], the case where voiced obstruents following nasals at prefix junctures in Perak and NS undergo nasal substitution is an explainable phenomenon. IDENT[PHAREXP] is ranked above CRISP-EDGE[σ] to stop a voiced obstruent from undergoing nasal substitution in Indonesian (Pater, 2001). However, in this study, IDENT[PHAREXP] is ranked beneath CRISP-EDGE[σ], so that nasal substitution is not limited to voiceless obstruents. In other words, the ranking CRISP-EDGE[σ] >> IDENT[PHAREXP] allows both voiced and voiceless obstruents following a nasal segment to undergo nasal substitution.

Since voiced obstruents can also undergo nasal substitution, then *NÇ has been replaced by CRISP-EDGE[σ], as *NÇ only allows voiceless obstruents to undergo nasal substitution. As demonstrated in (230), this constraint fails to rule out candidates with a nasal and voiced obstruent cluster and thus it cannot be used to account for both voiceless and voiced obstruent nasal substitution. By employing CRISP-EDGE[σ] in the analysis of non-standard dialects of Malay, nasal and voiced clusters can be ruled out because nasal and voiced obstruents are two adjacent segments that are linked together in the same syllable. In order to satisfy CRISP-EDGE[σ],the clusters undergo nasal substitution. Voiced obstruent nasal substitution is one of the important points this study makes. Nasal substitution does not only apply between a cluster of nasal and voiceless obstruents. It also occurs for a voiced obstruent following a nasal consonant. The data from the the NS and Perak dialects are evidence that a sequence of nasal and voiced obstruents can also undergo nasal substitution.

For convenience of reference, I briefly tabulate the constraint rankings in the three dialects of Malay, Perak, Kelantan and NS, which have been discussed in this chapter:

259. Constraint rankings for the three Malay dialects: Perak, Kelantan and NS.

(a) Constraint ranking in Perak:

PROSODIC STEM >> DEP-IO >> NAS ASS >> MAX-IO >> UNIFORMITY-ROOT >> CRISP-EDGE[σ] >> EDGE-INTEGRITY >> IDENT[PHAREXP] >> UNIFORMITY

(b) Constraint ranking in Kelantan:

PROSODIC STEM >> DEP-IO >> UNIFORMITY-ROOT >> UNIFORMITY >> NAS ASS >> EDGE-INTEGRITY >> CRISP-EDGE [σ] >> MAX-IO

(c) Constraint ranking in NS:

PROSODIC STEM >> DEP-IO >> CRISP-EDGE [σ] >> NAS ASS >> UNIFORMITY-ROOT >> MAX-IO >> UNIFORMITY >> IDENT[PHAREXP] >> EDGE-INTEGRITY.

7. CONCLUSION

7.1 Conclusion

In this study, I have examined two types of morphological process, prefixation and reduplication, which play decisive roles in forming new words in Malay. Besides that, dialect variation in three selected Malay dialects, Perak, Kelantan and NS, has also been investigated. This chapter concludes the discussion and analysis of those topics. By applying current prosodic morphological theory developed within OT, i.e. MBT, all the problems mentioned in Chapter 1 Section 1.3 can be captured more adequately. I briefly summarise each of the topics discussed in this thesis.

For prefixation, the discussion has been based on one million words gathered from the DBP-UKM database. The large dataset used for prefixation provides strong evidence to account for how nasal final prefixes really behave when they are attached to voiceless, voiced and sonorant initial-roots in SM. Observations of the one million words from that corpus reveal that nasal and voiceless obstruent clusters at prefix-root junctures are not entirely prohibited in SM as there are counter-examples which show that the clusters exist somehow in the language. As I have claimed in the analysis, the occurrence of nasal and voiceless obstruent clusters is due to the etymology of the words, as observed by partitioning the data into three lexical strata, i.e. monosyllabic foreign, undeleted plosive in loanwords, and native. The analysis of lexical strata shows that non-native words (i.e. monosyllabic foreign and undeleted plosive loanwords) are not subject to the same phonological requirements as those imposed on native words. Therefore, each lexical stratum has been analysed with distinct construction-specific constraint rankings or so-called co-phonologies.

Based on the central idea of MBT, that the typical size of prosodic morphemes should fall out from their morphological categories (Downing 2006), I proposed that the canonical shape of affixes in Malay is monosyllabic bimoraic. Thereofore, affixes in Malay must be monosyllabic bimoraic. However this requirement, postulated for affixes, gives rise to a problem with the prefix /di+/ as it does not fit the typical size for an affix. In order to satisfy the canonical shape, monosyllabic bimoraic, compensatory lengthening has to be applied to the prefix /di+/, so the vowel in the prefix has to be lengthened, likewise for prefixes /bər+/ and /tər+/. Due to the phonological requirement of the language that [r] in the coda position of a syllable has to be deleted, the vowels in the prefixes also undergo vowel lenghthening. In this case, to satisfy the constraint called ALIGN-RHOTIC that requires segment /r/ to be left aligned with a syllable, vowel lengthening applies. Such a solution is applied to satisfy one or more constraints in the hierarchy that can be well explained by OT with its idea of conspiracies. Compensatory lengthening is applied to the prefixes /di+/, /bər+/ and /tər+/ so as to satisfy the canonical shape for an affix, i.e. monosyllabic bimoraic.

In the analysis, I also claimed that monosyllabic words cannot be subject to the disyllabicity minimality of the language. Regarding this claim, I have argued on the basis of two previous studies, i.e. Teoh (1994) and Ahmad (2000b), who claimed that monosyllabic words in Malay actually contain two syllables. Those scholars propose CV.CVC (Teoh) and V.CVC (Ahmad) as the lexical representations for monosyllabic words. I argued that this idea is not plausible as schwa only appears when nasal final prefixes are attached to monosyllabic roots. There is no such output as *[ə.pam] or *[ə.bom] to be found in the corpus. Therefore, the idea proposed by those scholars cannot be accepted, as *[ə.pam] or *[ə.bom] do not exist in the language.

Based on the idea of MBT, that not all words can be subject to the same minimilaity condition, I argued that monosyllabic words in Malay emerge as monosyllabic, which they are. The disyllabicity minimality requirements postulated in the language are only applicable when monosyllabic words are combined with nasal final prefixes. As also proposed in MBT, underived words are not subject to the same words minimality requirements as derived words. Therefore, based on this, I claim that the disyllabic minimality postulated in Malay can only be met in derived words. Underived words can be monosyllabic. This matter is explained by HEAD BRANCH, the minimum size of roots. Although underived words can contain less than two syllables, these words must however meet the minimum size of words proposed by MBT. The minimum size can be one of the Heads, as presented in (43).

As I have claimed, the size of monosyllabic words falls out from the representation in (b). Therefore, underived words with monosyllabic roots in Malay do not violate the minimality of words condition. This idea of the minimum size of roots, as proposed in the theory, is essential when analysing Malay data as monosyllabic words do not surface as disyllabic, as required by word minimality in the language. Such an explanation was omitted from the earlier model, however it can be well explained in MBT.

This analysis of prefixation is not limited to nasal and voiceless obstruent clusters at prefix-root junctures in single prefixation. It also discusses the occurrence of clusters at prefix-prefix junctures in multiple prefixation. As we saw, nasal and voiceless obstruent clusters at prefix-prefix junctures somehow undergo nasal substitution, as in the nominal prefixes /pəŋ+pər/. However, in verbal prefixes /məŋ+pər/, nasal substitution fails to occur. As discussed, the non-application of nasal substitution in the verbal prefixes /məŋ+pər/ was analysed by previous scholars (e.g.: Asmah, 1986; Karim et al., 1989) as an exception because the regular process of nasal substitution, which is applied to avoid nasal and voiceless obstruent clusters, fails to occur.

In contrast to those previous studies, in my analysis the inconsistency of nasal substitution in multiple prefixes is analysed as the application of nasal substitution at prefix-prefix junctures is mainly determined by a morphological factor, i.e. the domain, rather than a phonetic factor. Thus, I claim that nasal substitution is blocked in the morphological domain of prefix-prefix junctures in Malay. To account for this, a morphology-phonology interface constraint called EDGE INTEGRITY, which requires edge segments in the input to preserve their segments at the edge of a corresponding prosodic structure, is essential to explaining why clusters occur in this morphological domain. This means nasal and voiceless obstruent clusters are not always resolved by nasal substitution, even if the phonetic factor is satisfied. The phonological process of nasal substitution is driven by a morphological factor and can satisfactorily be accounted for in constraint-based theory, OT. But why does nasal substitution apply in the nominal prefixes /pəŋ+pər/? This is analysed in this study as a case of analogy, whereby the application of nasal substitution in the nominal prefixes /pəŋ+pər/ has been analysed by previous studies in the same way as in single prefixation.

In the reduplication chapter, there are analyses of three types of reduplication, i.e. total, partial and affixal. By referring to five million words of corpus data for SM, this chapter criticizes the notion of reduplication being best treated as self-compounding rather than affixation. Based on the idea of MBT, that the typical size of prosodic morphemes should fall out from their morphological categories, I claimed that total and affixal reduplication should be considered as compounding. Since the size of the reduplicative morphemes in total and affixal reduplication is larger than monosyllabic, the typical size of affixes, they cannot be said to constitute affixation.

In this chapter, I also analysed why nasal segments in the prefixes are copied into the reduplicative morphemes, as in: (i) /məŋ-ə-lap/ → [məŋ-ə-lap-ŋə-lap] and (ii) /məŋ-təkan-i/ → [mə-nəkan-nəkan-i], but not in (iii) /məm-bandiŋ-bandiŋ/ → [məm-bandiŋ-bandiŋ]. In this thesis, I claim that the nasal segements in the prefixes in (i) and (ii) are copied into the reduplicative morphemes for two reasons. The nasal segment is copied: (1) to satisfy the disyllabicity minimality condition of the language, as in (i), and (2) to provide an onset to the initial syllable of the reduplicative morpheme due to nasal substitution which is applied to break up a sequence of nasal and voiceless obstruents, as in (ii). In contrast to (iii), a nasal segment in the prefix is not copied into the reduplicative morpheme because the word minimality of the reduplicated word has already been satisfied. Moreover, the initial syllable of the reduplicative morpheme contains an onset, therefore copying the nasal segment is not significant. Such a problem was never discussed in previous studies, however it has been given attention in this thesis since it arises in the data. From the ideas of the canonical shape of an affix and the word minimality requirements proposed in MBT, those problems can be explained straightforwardly.

Having referred the previous study of Ahmad (1991) on partial reduplication in PD, this study discovers that not only does Malay have one pattern of reduplicative morphemes, i.e. light, in its process of reduplication, but also heavy reduplicative morphemes which are found in PD as one of the patterns. Since Malay has two patterns of reduplicative morphemes, light and heavy, I have claimed in this thesis that the CV analysis of Ahmad (2000) in autosegmental analysis cannot be used to account for both patterns. The CV analysis is only applicable to explain light reduplicative morphemes. As I have demonstrated in (27), the CV analysis fails to account for CVC reduplicative morphemes as the result produces an incorrect output.

In this analysis, the two patterns, light and heavy, are accounted for in co-phonology analysis. The heavy reduplicative morpheme is explained by the tendency for prosodic constituents to be of maximal size, as stated in the MAXIMALITY CONDITION (see 164). The light reduplicative morpheme on the other hand is explained by the opposing tendency where some prosodic morphemes have unmarked structure. In a co-phonology analysis, both patterns are associated with a distinct constraint ranking. As shown, a co-phonology analysis can account for light and heavy reduplicative morphemes in Malay by reversing the ranking of the relevant markedness constraint with the faithfulness constraint in the hierarchy. This was done by ranking NOCODA above the faithfulness constraint in heavy reduplication, while ranking this constraint lower in light reduplication.

Chapter 6 offered an OT account of the study of variation. Variation occurring in three Malay dialects, Perak, Kelantan and NS, was analysed as each dialect has a particular constraint ranking. The analysis reveals some interesting points that have been missed in previous studies. As we saw in the discussion, although nasal and voiceless obstruent clusters are claimed by previous studies (e.g.: Hassan, 1974; Omar, 1986; Hassan, 1987; Karim et al., 1989, 1994; Karim, 1995; and many others) to be disfavoured in the language, clusters do however occur in one of the non-standard dialects of Malay discussed, i.e. Perak. The clusters occur in root-internal position of the dialect. The occurrence of clusters within the roots has been analysed by using the root-faithfulness constraint called UNIFORMITY-ROOT. This faithfulness constraint ensures that nasal substition is only blocked within roots. Therefore, this root-faithfulness constraint is ranked high in the constraint ranking of Perak. In contrast, nasal and voiceless root-internal obstruent clusters in Kelantan and NS dialects are strictly prohibited when compared to Perak as the clusters are resolved by nasal deletion in both dialects.

Apart from that, the analysis of dialectal variation also reveals another important point regarding nasal substitution in the language. Voiced obstruents also undergo nasal substitution, as do voiceless obstruents. Since both voiceless/ voiced obstruents undergo nasal substitution, the *NC̥ constraint cannot be used to explain both situations because this constraint cannot account for voiced obstruent nasal substitution. Therefore, I proposed that CRISP-EDGE[σ] is a better constraint than the *NC̥ to account for both voiced and voiceless obstruent nasal substitution.

7.2 Contribution of the research

This thesis has offered a new analysis in its investigation of prefixation and reduplication in Malay which have not been satisfactorily accounted for in previous studies. Reanalysing those topics with reference to a current theory in Prosodic Morphology, MBT, has allowed the thesis to explain all the problems arising in both topics. It is apparent from the reanalysis performed that MBT developed within OT can satisfactory account for prefixation and reduplication, in contrast to other earlier approaches applied by scholars of Malay. In what follows, I am going to list the contributions that the thesis has made:

The thesis has made an important contribution to understanding a number of theoretical aspects of Malay linguistics. It reveals some important facts about the grammar of Malay. First, as we saw, nasal and voiceless obstruent clusters are not entirely prohibited in the language. There are three instances where nasal and voiceless obstruent clusters emerge in the surface representation of the language, namely, in root-internal position, prefix-root and prefix-prefix junctures. In agreement with Mascaró (1976) and Kiparsky (1982, 1993a) cited in Kager (1999: 75), who claim that root internal segments tend to be more resistant to other phonological processes than segments in other word positions, nasal substitution fails to apply inside a morpheme in Malay. A root-specific faithfulness constraint, UNIFORMITY-ROOT, blocks this phonological process from occurring in this environment, therefore it has been added into the constraint ranking of Malay. Faithfulness constraints that require inputs to be as faithful as possible to outputs are only applicable in OT and cannot be found in any other alternative model.

The occurrence of clusters at prefix-root junctures is claimed as the roots are non-native. Thus, the regular phonological process to avoid clusters can be applied, as non-native words cannot be subject to the same phonological requirements as those imposed on native words. By adapting Itô and Mester's (1999b) co-phonology analysis, where the markedness constraint, *N$\c{C}$ can be ranked reversely in the same language, non-native words with no application of nasal substitution have their own constraint ranking. In the constraint rankings of non-native words, *N$\c{C}$ is ranked lower than in the ranking of native words. As demonstrated in Chapter 4 subsection 4.2.1.2, the blocking of nasal substitution could not be resolved by the rule ordering proposed in rule-based analyses. As was mentioned, the occurrence of nasal and voiceless obstruent clusters at prefix-root junctures has never previously been discussed by scholars of Malay. This study has thus come out with new data showing that there is another environment where nasal and voiceless obstruent clusters can be found. The heavy usage of corpus data for the analysis validates our findings regarding the occurrence of clusters at prefix-root junctures.

The non-application of nasal substitution at prefix-prefix junctures on the other hand is due to a morphological factor, i.e. domain. Domain is the main factor which determines when nasal substitution should occur while the phonetic factor comes afterwards. By this, nasal substitution is actually blocked at prefix-prefix junctures since the multiple prefixes are outside the domain where nasal substitution can occur, as in the multiple verbal prefixes /məŋ+pər/. This morphology-conditioning phonology that occurs in the language can be accounted for by a morphology-phonology interface constraint named EDGE INTEGRITY (McCarthy and Prince, 1995). This constraint requires edge segments in the input to preserve their segments at the edge of corresponding prosodic structures. In other words, EDGE INTEGRITY requires the integrity of the morphological constituent.

As was pointed out, this morphology-phonology interface constraint is absent in previous approaches, such as those in rule-based and non-linear autosegmental analyses. Thus, it is impossible that this case can be captured in earlier approaches. This study thus supplies a more plausible theoretical motivation to account for the blocking of nasal substitution in multiple prefixes than those previously offered by scholars in rule-based analyses, such as Omar (1986) and Karim et al. (1989), who simply claimed that the non-application of nasal substitution in /məŋ+pər/ was an exception.

The three environments where nasal and voiceless obstruent clusters occur in the language prove the fact that the phonological system of the language is not completely uniform. The language thus has co-existing distinct phonological systems, or so-called co-phonology. The diversity of how nasal and voiceless obstruent clusters emerge in the language has been captured by the morphological constructions with distinct phonological grammars, i.e. different constraint rankings. In the case of Malay, the situation where one constraint ranking allows nasal and voiceless obstruent clusters in the surface representation while the other one does not can be captured with two different co-phonologies. As mentioned, 'Markedness Reversals', as allowed in co-phonology, enable us to rank the markedness constraint *NÇ in one ranking in some morphological construction(s), but the opposite ranking in others.

Second, the idea of word minimality requirements proposed in MBT, which has been put forward for prefixation and reduplication, can explain a number of problems in the grammar of the language. As was discussed in Chapter 4, the idea can account for monosyllabic roots in the language. Malay, as claimed by previous scholars, is a disyllabic language, however the requirement cannot be met in all words. As explained in the analysis, monosyllabic roots like [pam], [bom] and [lap] surface as they are. No additional segment can be added to them to satisfy disyllabicity minimality requirements. I have argued that the disyllabic template for monosyllabic roots in Malay, i.e. V.CVC and CV.CVC, as postulated by Teoh (1994) and Ahmad (2000b) respectively, cannot be accepted as schwa is only added when prefixes that end with a nasal segment are attached to monosyllabic roots. Otherwise, monosyllabic roots remain monosyllabic. MBT, with its idea of word minimality requirements, i.e. derived words are distinct from underived words to satisfying the word minimality requirement, can satisfactorily account for this matter. Since monosyllabic roots remain monosyllabic on the surface, they therefore have to satisfy the minimum size of roots, i.e. HEAD BRANCH as proposed by MBT. Such a clear explanantion provided by MBT, regarding the satisfaction of word minimality in monosyllabic roots, works well for Malay. This sort of explanation can however never be found in any earlier approaches. Therefore, no previous Malay scholar has given an accurate explanation of the disyllabicity minimality requirements of the language.

Besides that, the idea of word minimality can also account for why a nasal segment is copied into the reduplicative morpheme when monosyllabic bases are reduplicated. Based on that idea, the copying of a nasal segment into the reduplicative morpheme is to satisfy disyllabic minimality requirements. This theoretical justification provided by MBT works well for Malay since a nasal segment is not copied when disyllabic bases (except for voiceless obstruent initial bases) are reduplicated. Observations from the huge corpus dataset confirm the case. Therefore, I claim that a nasal segment in the prefix is only copied by reason of satisfying the disyllabic minimality requirements of the language. Such a matter of copying nasal segments into reduplicative morphemes when monosyllabic bases are reduplicated has never been addressed in any previous studies. The claim I make is thus the first explanation for the grammar of Malay concerning reduplication based on an appropriate theory, i.e. MBT.

Third, the investigation of dialectal variation in the three selected non-standard dialects provides another new fact about Malay that has also never been covered in previous studies. Nasal and voiced obstruent clusters in two of the Malay dialects examined in this study, i.e. Perak and NS, can also undergo nasal substitution, as nasal and voiceless obstruent clusters do. As demonstrated, voiced obstruent nasal substitution can be explained well within the theoretical framework of OT by implying a CRISP-EDGE[σ] constraint which is able to rule out an element in a syllable which has multiple linking to an adjacent syllable. A sequence of nasal and voiced obstruents can be ruled out by this constraint, thus allowing a voiced obstruent to undergo nasal substitution. Therefore, CRISP-EDGE[σ] is ranked high for those two dialects of Malay, while it is not in Kelantan as the dialect disallows voiced obstruents from undergoing nasal substitution. It should be noted here, when discussing voiced obstruent nasal substitution, that this study finds that the use of the *NC̥ constraint is limited to voiceless obstruents only. This constraint cannot explain voiced obstruent nasal substitution. In the analysis, I presented the limited role of *NC̥ when accounting for both voiceless and voiced obstruents in the three selected Malay dialects. The study of Malay dialects discussed in this thesis can thus be another example of how voiceless and voiced obstruents in the Perak and NS dialects are out of the reach of *NC̥, as well as Pater's (2001) analysis of Muna.

REFERENCES

Abas, Lutfi. 1975. *Tatabahasa perbandingan: bahasa-bahasa Malaysia, Cina, Tamil dan Inggeris*. Institute Language and Literature: Kuala Lumpur.

Abd Ghani, Rusli, Mohamed Husin, Norhafizah, and Lee Yim, Chin. Available at: http://dbp.gov.my/korpus/korpus_DBP.pdf.

Abdul Malik, Sharifulnizam. 2004. *Fonologi vokal rangkap dan kepelbagaian dialek Melayu: analisis teori Optimality*. Ph.D Thesis. National University of Malaysia.

Abdul Rahman, Zaiton. 1987. *Kertas rancangan projek analisis teks secara komputer. cawangan penyelidikan*. Institute of Language and Literature: Kuala Lumpur..

Ahmad, Zaharani. 1989. *Syarat koda suku kata dalam bahasa Melayu*. Jurnal Dewan Bahasa. 33(11).

Ahmad, Zaharani. 1991. *The phonology and morphology of the Perak dialect*. Kuala Lumpur: Institute of Language and Literature.

Ahmad, Zaharani. 1993. *Fonologi generatif: teori dan penerapan*. Kuala Lumpur: Institute of Language and Literature.

Ahmad, Zaharani. 1999. Menangani kelegapan asimilasi nasal bahasa Melayu: suatu analisis teori fonologi leksikal. *Jurnal Dewan Bahasa* 43(9).813-24.

Ahmad, Zaharani. 2000a. Penggandaan separa bahasa Melayu: suatu analisis aotusegmental. *Jurnal Dewan Bahasa* 44(7).722-36

Ahmad, Zaharani. 2000b. Pemanjangan dan penyisipan vokal bahasa Melayu. *Jurnal Dewan Bahasa* 44(12).1202-21.

Ahmad, Zaharani. 2005. *The phonology-morphology interface in Malay: an Optimality theoretic account*. Pacific Linguistics: The Australian National University.

Ahmad, Zaharani. 2007. *Aspek nahu praktis bahasa Melayu*. The Publisher of National University of Malaysia, Bangi.

Ahmad, Zainal Abidin. 1964. *Pelita bahasa*. Institute of Language and Literature: Kuala Lumpur.

Akinlabi, Akinbiyi. 1984. *Tonal underspecification and Yoruba tone*. Doctoral dissertation. University of Ibadan.

Anttila, Arto. 1997. *Deriving variation from grammar*. In Frans Hinskens, Roeland van Hout and Leo Wetzels (eds.), Variation, change and phonological theory. Amsterdam, John Benjamins. 35-68.

________. 2002. *Variation and phonological theory*. In Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), Handbook of language variation and change. Blackwell, Oxford, U.K., and Malden, Massachusetts. 206-43.

_________. 2007. *Variation and optionality*. In Paul de Lacy (ed.), The Cambridge Handbook of Phonology. Cambridge: Cambridge University Press, Cambridge. 519-36.

Ao, Benjamin. 1991. Kikongo nasal harmony and context-sensitive underspecification. *Linguistic Inquiry* 22.193-96.

Archangeli, Diana. 1983. The root CV template as a property of the affix: evidence from Yawelmani. *NLLT1*. 348-84.

Archangeli, Diana, and Langendeon, Terence. 1997. Optimality theory: an overview. University of Arizona. Blackwell Publishing.

Aripin, Tajul. 2005. Penggandaan separa dalam dialek Perak: suatu analisis autosegmental. *Jurnal Bahasa* 5 (3).162-88.

Baeur, Laurie. 1988. *Introducing Linguistics morphology*. Edinburgh: Edinburgh University Press.

Bar-el, Leora. 2000. Reduplicants are roots in Skwxwú7mesh (Squamish Salish). *Proceedings of WECOL* 12, 81-99.

Boersma, Paul. 1998. *Typology and acquisition in functional and arbitrary phonology*. Ms. University of Amsterdam.

Che Kob, Ajid. 1981. Reduplikasi dalam bahasa Malaysia: tinjauan tentang proses beberapa jenis daripadanya. *Jurnal Dewan Bahasa* 25(3).15-22.

Che Kob, Ajid. 1985. *Dialek geografi Pasir Mas*. The Publisher of National University of Malaysia.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass: MIT Press.

Chomsky, Noam and Morris, Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.

Clements, G. N, and Keyser, S. J. 1983. *From CV phonology: a generative theory of the syllable*, ed. by Goldsmith, John Anton, 185-200. Blackwell Publishers Ltd.

Coetzee, Andries, and Pater, Joe. 2008. *The Place of Variation in phonological theory*. University of Michigan and University of Massachusetts. Available in ROA 946. http://roa.rutgers.edu/files/946-0108/946-COETZEE-0-0.PDF.

Collins, James T. 1983. *Dialek Pahang: rangka pengenalan (bahagian pertama)*. Kuala Lumpur: Institute of Language and Literature.

________. 1986. Dialek Melayu Selangor: tinjauan di Jugra. *Jurnal Persatuan Linguistik* 2.74-94.

Dell, Gary S. 1984. Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory and Cognition* 10. 222-33.

———. 1986. A spreading activation theory of retrieval in sentence production. *Psychological Review* 93. 283-321.

Davenport, Mike, and Hannahs, S. J. 1998. *Introducing phonetics and phonology*. Oxford University Press Inc.

Dereau, Léon. 1955. *Cours de Kikongo*. Namur: A. Wesmael-Charlier.

Downing, Laura J. 2006. *Canonical forms in prosodic morphology*. Oxford University Press.

Downing, Laura J. 2008. *Co-phonologies and morphological exponence in OT*. 3rd exponcence network meeting and workshop on theoretical morphology 4, University of Leipzig. 20-21 June 2008.

Dresher, Elan, and Hulst, H.G. van der. 1998. Head-dependent asymmetries in prosodic phonology: Visibility and complexity. *Phonology* 15/3. 317-52

Durand, Jacques, and Katamba, Francis. 1995. *Frontiers of phonology*. Longman Linguistics Library.

Dziwirek, Katarzyna. 1989. *Malagasy phonology and morphology*. Linguistics Notes from La Jolla 15. 1-30.

Edwards, Jane, and Shriberg, Elizabeth. 1983. *Phonology*. College-Hill Press, Inc. San Diego, California.

Eulenberg, John B. 1971. *Conjunction reduction and reduplication in African languages*. In Chin-Wu Kim and Herbert Stahlke (eds.), Papers in African Linguistics. Edmonton: Linguistic Research, Inc., 71-80.

Féry, Caroline. 1991. *German schwa in prosodic morphology*. Zetschrift für Sprachwissenschaft 10. 65-85.

Fromkin, Victoria A. 1971. The non-anomalous nature of anomalous utterances. *Language* 41. 27-52.

Golston, Chris. 1991. Minimal word, minimal affix. InnT. Sherer (ed.), *proceedings of NELS* 21.95-110. Amherst, MA: GLSA.

Goswami, Upendra. 1955. Onomatopoetic and echo-words in Kamrupi. Indian *Linguistics* 17.161–4.

Guy, Gregory R. 1997. *Competence, performance and the generative grammar of variation*. In Frans Hinskens, Roeland van Hout and Leo Wetzels, eds.

Variation, Change and Phonological Theory. Amsterdam: John Benjamins. 125-43.

Hassan, Abdullah. 1974. *The morphology of Malay*. Kuala Lumpur: Institute of Language and Literature.

Hassan, Abdullah. 1987. *Penerbitan kata dalam bahasa Malaysia*. Petaling Jaya: Fajar Bakti.

Hayes, Bruce, and Abad, May. 1989. Reduplication and syllabification in Ilokano. *Lingua* 77.331-74.

Hayes, Bruce. 1986. Assimilation as spreading in Toba Batak. *Linguistic Inquiry* 17. 467-99.

Hendricks, Sean Quillan. 1999. *Reduplication without template constraints: a study in bare consonant reduplication*. Available at: http://roa.rutgers.edu/files/511-0402/511-0402-HENDRICKS-0-0.PDF.

Henry, Jules. 1948. The Kaingang language. International Journal of American *Linguistics* 14. 194-204.

Huffman, Marie K. 1993. *Phonetic patterns of nasalisation and implication for feature specification*. In Huffman and Krakow. 303-27.

_______ and Rena A. Krakow (eds.). 1993. Nasals, Nasalization, and the Velum. *Phonetics and Phonology* 5: San Diego: Academic Press.

Hussein, Ismail. 1973. Malay dialects in the Malay peninsular. *Nusantara* 3. 69-79.

Inkelas, Sharon. 1998. *The theoretical status of morphologically conditioned phonology: a case study of dominance effects*. In Geert Booij and Jaap van Marle, eds.Yearbook of Morphology 1997. Dordrecht: Kluwer. 121-55.

Inkelas, Sharon. 2005. *Reduplication doubling in morphology*. Cambridge: Cambridge University Press.

________. 2008. *The morphology-phonology connection*. Talk given at Pomona College, 1 April 2008.

Inkelas, Sharon, and Orgun, Orhan. 1998. *Level (non)ordering in recursive morphology: evidence from Turkish.* In Steven G. Lapointe, Diane K. Brentari and Patrick M. Farrel (eds.), Morphology and its relation to phonology and syntax. Stanford, Calif.: CSLI: 360-92.

Inkelas, Sharon, and Zoll, Cheryl. 2000. *Reduplication as morphological doubling.* Available at ROA 412-0800.

________. 2005. *Reduplication: doubling in morphology.* Cambridge: Cambridge University Press.

________. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45. 133-71.

Ito, Junko, and Mester, Armin. 1990. *Prosodic minimality in Japanese.* CLS 26-II: papers from the parasession on the syllable in phonetics and phonology. 213-39. http://people.ucsc.edu/~ito/papers/ito_1990_prosodic_minimality.pdf.

________. 1999a. *Realignment*, in R. Kager, H. v.d. Hulst and W. Zonneveld (eds.), 188-217.

________. 1999b. *The Phonological lexicon.* In McCarthy, John. 2004. Optimality theory in phonology. Blackwell Publishing. 552-68.

Iverson, Gregory K., and Shinsook Lee. 1994. *Variation as optimality in Korean cluster reduction.* Paper presented at ESCOL-94 (Eastern States Conference on Linguistics), University of South Carolina.

Jun, Jongho. 1995. *Perceptual and articulator factors in place assimilation: an Optimality theoretic approach.* Doctoral dissertation. UCLA. Los Angelas.

Kager, René. 1999. *Optimality theory.* Cambridge: Cambridge University Press.

_______, van der Hulst, Harry and Zonneveld, Wim (eds.). 1999. *The prosody-morphology interface.* Cambridge: Cambridge University

Kang, Eungyeong. to appear. *Edge integrity and the syllable structure in Korea.* Cornell University. Available at: http://www.aclweb.org/anthology/Y/Y02/Y02-1013.pdf.

Karim, Nik Safiah, Onn, Farid, Musa, Hashim, and Mahmood, Abdul Hamid. 1989. *Tatabahasa dewan 2nd edition: perkataan.* Kuala Lumpur: Institute of Language and Literature.

__________. 1994. *Tatabahasa dewan.* Kuala Lumpur: Institute of Language and Literature.

Karim, Nik Safiah. 1995. *Malay grammar for academics and professionals.* Kuala Lumpur: Institute of Language and Litrature.

Kaye, Jonathan. 1990. Coda licencing. *Phonology* 7:301-30.

Keenan and Polinsky. 1998. *Malagasy (Austronesian).* In Spencer and Zwicky.

Kim, Soohee. 1996. *Korean partial reduplication.* CLS 32.179-89.

Kiparsky, Paul. 1979. *Panini as a Variationist.* MIT Press and Poona University Press.

________. 1982. *From cyclic phonology to lexical phonology.* In H. van der Hulst and N. Smith (eds.), The structure of phonological representatitons, part 2. Dordrecht: Foris 131-76.

________. 1993a. *Variable rules.* Handout distributed at the Rutgers Optimality Workshop (ROW1).

________. 1993b. *Blocking in non-derived environments.* In S. Hargus and E. Kaisse (eds.), Phonetics and phonology IV: studies in lexical phonology. San Diego: Academic Press. 277-313.

Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1, 291-306.

Koh, Boh Boon. 1981. *Pengajaran bahasa Malaysia.* Kuala Lumpur: Utusan Publications & Distributors Sdn. Bhd.

Lamontagne, Greg, and Rice, Keren. 1995. A correspondence account of coalescence. In Jill N. Beckman, Laura Walsh Dickey and Suzanne Urbanczyk, eds. *Univeristy of Massachussets occasional papers in Linguistics* 18: Papers in Optimality Theory. Amherst, MA: Graduate Linguistic Student Association.

Lapoliwa, Hans. 1981. *A generative approach to the phonology of bahasa Indonesia.* Canberra: Pacific Linguistics D 34.

Levin, Juliette. 1983. *Reduplication and prosodic structure.* MIT Manuscript.

_______. 1985. *Reduplication in Umpila.* MIT Working papers in Linguistics 6.133-59.

MacKay, Donald G. 1970. Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia* 8. 323-50.

———. 1987. *The Organization of perception and action: a theory for language and other cognitive skills.* New York: Springer.

Marantz, Alec. 1980. *Reduplication and CV skeleta.* In R.Lieber. 1981. On the organization of the lexicon. Indiana Uni.

________. 1982. Re Reduplication. *Linguistic Inquiry* 13.483–545.

Maris, Yunus. 1980. *The Malay sound system.* Kuala Lumpur: Fajar Bakti.

Mascaro, Juan. 1976. *Catalan phonology and the phonological cycle.* PhD thesis. MIT Distributed by Indiana Linguistics Club.

Matthews, P.H. 1991. *Morphology (2nd ed.).* Cambridge: Cambridge University Press.

McCarthy, John J. 1995. *Extensions of faithfulness.* Rotuman revisited, Ms., University of Massachusetts, Amherst. Available at ROA 110.

________. 2002. *A thematic guide to Optimality theory.* Cambridge: Cambridge University Press.

__________. 2004. *Optimality theory in phonology: a reader.* Malden, Mass. Oxford: Blackwell Publishing.

__________. 2008. *Doing Optimality theory: applying theory to data.* Blackwell Publishing.

McCarthy, John J. and Alan S. Prince. 1986. *Prosodic Morphology.* MS, University of Massachusetss, Amherst and Branddeis University, Waltham, Mass.

________. 1990. Foot and Word in Prosodic Morphology: The Arabic Broken Plural. *Natural Language and Linguistic Theory* 8. Pg. 209-284.

____________. 1993, *Prosodic morphology: constraint interaction and satisfaction.* University of Massachusetts, Amherst.

__________. 1994a. The emergence of the unmarked: Optimality in prosodic morphology. *Proceedings of the North East Linguistic Society* 24.333–79.

__________. 1994b. *Prosodic Morphology: an overview.* Talks presented at the OTS/HIL Workshop on Prosodic Morphology, University of Utrecht.

__________. 1995. *Faithfulness and reduplicative identity.* Available in ROA-60. http://ruccs.rutgers.edu/roa.html

__________. 1999. *Faithfulness and Identity in Prosodic Morphology.* In Kager, René, Harry van der Hulst and Wim Zonneveld (eds.). pg: 218-384.

___________. 1995. *Prosodic morphology.* In John A. Goldsmith (ed.), The handbook of phonological theory, 318-66. Cambridge, MA: Blackwell.

McEnery, Tony, and Wilson, Andrew. 2001. *Corpus linguistics: an introduction.* Second edition. Edinburgh Textbook in Empirical Linguistics.

Meinhof, Carl. 1932. *Introduction to the phonology of the Bantu languages* (trans. by N. J. van Warmelo). Berlin: Dietrich Reimer/Ernst Vohsen.

Mills, Roger F. 1975. *Proto South Sulawesi and Proto Austronesian phonology.* Ph.D dissertation, University of Michigan.

Moravcsik, Edith. *Reduplicative constructions.* In J. H. Greenberg (ed.). Universals of human language: Word structure 3.297–334. Stanford, CA: Stanford University Press

Musa, Hashim. 1974. *Morfemik dialek Melayu Kelantan.* M.A Dissertation, Malaya University of Malaysia.

Mutaka, Ngessimo, and Hyman, Larry. 1990. Syllables and morpheme integrity in Kinande reduplication. *Phonology* 7. 73-119.

Myers, Scott, and Troi, Carleton. 1996. Tonal transfer in Chichewa. *Phonology* 13.39-72.

Niepokuj, Mary K. 1991. *The historical development of reduplication with special reference to Indo-European.* Ph.D dissertation, University of California-Berkeley.

Nooteboom, S. G. 1967. *Some regularities in phonemic speech errors.* IPO Annual Progress Report II. 65-70.

Odden, David. 1994. Adjacency parameters in phonology. *Language* 70. 289-330.

Omar, Asmah. 1975. *Essays on Malaysian linguistics.* Kuala Lumpur: Institute of Language and Literature.

________. 1982. *Nahu Melayu mutakhir.* Kuala Lumpur: Institute of Language and Literature.

_________. 1985. *Susur galur bahasa Melayu.* Kuala Lumpur: Institute of Language and Literature.

__________. 1986. *Nahu Melayu mutakhir(2nd edition).* Kuala Lumpur: Institute of Language and Literature.

__________. 1991. *Kepelbagaian fonologi dialek-dialek Melayu.* Kuala Lumpur: Institute of Language and Literature.

___________. 1993. *Nahu Melayu mutakhir.* Kuala Lumpur: Institute of Language and Literature.

Onn, Farid. 1980. *Aspect of Malay phonology and morphology: a generative approach.* Bangi: National University of Malaysia.

Orgun, Orhan. 1996. *Sign-based morphology and phonology: with special attention to Optimality theory*. PhD dissertation. UC-Berkeley.

_______. 1998. *Cyclic and noncyclic phonological effects in a declarative grammar*. Yearbook of Morphology 1997. 179-218.

Othman, Arbak. 1976. *Penggandaan dalam morphology bahasa Melayu*. Institute of Language and Literature: Kuala Lumpur.

_________. 1981. *Tatabahasa Malaysia*. Kuala Lumpur: Sarjana Enterprise.

_________. 1983. *Imbuhan me- dan pentingnya dalam pengajaran tatabahasa bahasa Melayu*. Institute of Language and Literature.

_________. 1985. *Mengajar tatabahasa*. Institute of Language and Literature: Kuala Lumpur.

Padgett, Jaye. 1995. *Partial class behaviour and nasal place assimilation*. Proceedings of the Arizona Phonology Conference: workshop on features in Optimality Theory. Coyote Working Papers. Tuscon: University of Arizona Department of Linguistics.

Pater, Joe. 1996. **NC*. Proceedings of the North East Linguistics Society 26. 227-39.

________. 1999. *Austronesian nasal substitution and other NC effects*. In René Kager, Harry Van Der Hulst and Wim Zonneveld (eds.) 1999. The prosody-morphology interface. 310-43.

_________. 2001. *Austronesian nasal substitution revisited*. In Linda Lombardi. (ed.). 2001. Segmental phonology in Optimality theory: constraints and representations. Cambridge University Press. 159-82.

Piggot, Glyne L. 1995. *Feature dependency in Optimality Theory: optimizing the Phonology of sonorants*. Ms., McGill University.

________. 1996. Implications of consonant nasalization for a theory of harmony. *CJL* 41. 141-74.

Poedjosoedarmo, Soepomo. 1982. *Javanese influence on Indonesian.* Canberra: Pacific Linguistics D 38.

Prince, Alan. 1985. *Improving tree theory.* Proceedings of BLS 11.471-90.

Prince, Alan, and Smolensky, Paul. 1993. *Optimality theory: constraint interaction in generative grammar.* Available on Rutgers Optimality Archive, ROA-537.

________. 2004. *Optimality theory: constraint interaction in generative grammar.* Malden, Mass: Blackwell.

Pulleyblank, Douglas. 1988. *Vocalic underspecification in Yoruba.* Linguistic Inquiry 19.233-70.

Roca, Iggy. 1994. *Generative phonology (Linguistic theory guides).* London and New York: Routledge.

_________(ed.). 1997. *Derivations and constraints in phonology.* Clarendon Press, Oxford.

Rose, Sharon. 1995. *Dialect variation and transparency: the role of constraints.* Paper presented at the annual meeting of the Canadian Linguistics Association, Université du Québec á Montréal.

Rose, Sharon and Walker, Rachel. 2004. *A typology of consonant agreement as correspondence.* http://muse.jhu.edu/journals/language/v080/80.3rose.pdf.

Rosenthall, Samuel. 1989. *The phonology of nasal obstruent sequences.* McGill University.

Rufus, Hendon. 1966. The Phonology and Morphology of Ulu Muar Malay. *Yale Publication in Anthropology* 70. Yale University.

Russell, Kevin. 1997. *Optimality theory and morphology*. In Diana Archangeli and D. Terence Langendoen (eds), Optimality theory: an overview. Oxford: Blackwell, 102-33.

Shattuck-Hufnagel, Stefanie, and Dennis Klatt. 1979. The limited use of distinctive features and markedness in speech production. *Journal of Verbal Learning and Verbal Behavior* 18. 41-55.

Spencer, A. 1991. *Morphological Theory*. Basil Blackwell Ltd.

Stemberger, Joseph P. 1985. An interactive activation model of language production. In Andrew W. Ellis, ed. *Progress in the Psychology of Language*, Vol. 1. London: Lawrence Erlbaum Associates

Steriade, Donca. 1982. *Greek Prosodies and the Nature of Syllabification*. PhD dissertation, MIT.

________. 1988. Reduplication and Syllable Transfer in Sanskrit and Elsewhere. *Phonology* 5.

Stonham, John T. 2004. *Linguistics theory and complex words. Nuuchahnulth word formation*. New York: Palgrave Macmillan.

Syed Ahmad, Sharifah Saharbi. 1998. *Reduplication in bahasa Melayu: the National language of Malaysia*. M.A Dissertation. Lancaster University.

Syed Jaafar, Sharifah Raihan and Ahmad, Zaharani. 2009. *Penggandaan separa bahasa Melayu: analisis teori Optimaliti*. In Kepelbagaian dimensi Linguistik. Siri Rampai Bahasa dan Sastera Melayu. 66-83. The Publisher of National University of Malaysia, Bangi.

Teoh, Boon Seong. 1994. *The sound system of Malay revisited*. Kuala Lumpur: Institute of Language and Literature.

Teoh, Boon S, and Ahmad, Zaharani. 2006. *Fonologi autosegmental penerapannya pada bahasa Melayu*. Kuala Lumpur: Institute of Language and Literature.

Topping, Donald M. 1973. *Chamorro reference grammar*. Honolulu: University of Hawaii Press.

Uhrbach, Amy. 1987. *A formal analysis of reduplication and its interaction with phonological and morphological processes*. Ph.D dissertation. University of Texas.

Urbanczyk, S. 1996. *Pattern of reduplication in Lushootsheed*. Ph.D dissertation, University of Massachusetts – Amherst.

Ussishkin, Adam P. 1999. The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence. *Phonology* 16. 401-42.

Ussishkin, Adam P. 2000. *The emergence of fixed prosody*. Ph.D dissertation, University of California-Santa Cruz.

Venneman, Theo. 1988. *Preference laws for syllable structure*. Berlin, NY, Amsterdam: Mouton de Gruyter.

Webb, Nancy. 1965. *Phonology and noun morphology of the Kindibu dialect of Kikongo*. MA thesis, UCLA.

Wilbur, Ronnie. B. 1973. *The phonology of reduplication*. Indiana University. Linguistics Blub.

Yeoh, C. K. 1988. *Kajian semula tentang morfem dasar penanda aktif dalam bahasa Melayu*. In Farid M. Onn (eds.) 1988b. Bunga Rampai Fonologi Bahasa Melayu. Petaling Jaya: Fajar Bakti.

APPENDICES

Appendix A

Voiceless obstruents without nasal substitution (non-deletion obstruents)

1. yang ja yang akan *mentadbirkan* pemerentahan kita dengan
2. demokrasi dan memp zi in yang *mentadbir* wilayah pemerintahan sendiri
3. di PTO yang *mentadbir* wilayah pemerintahan sendiri di
4. menguji kesetiaa tu mendengar dan *mentaati* (pemerintah)
5. Seterusnya, taat kepada pemerintah, iaitu *mentaati* semua unda
6. ditanya mengapa kewajipan rakyat *mentaati* pemerintah.
7. iban membina sistem pemerintahan yang mampu *mentadbir* dan
8. British dan India pun datang ka- Malaya dan *mentadbirkan* negeri ini.
9. chekap yang di-amalkan oleh pegawai2 yang *mentadbir* Kerajaan waktu
10. iap hari sudah dapat menyaksikan berbagai tempat *mentadbirkan*
11. merentah dan Ahli2 Majlis Permesuaratan, bagi *mentadbir* muslihat
12. bererti bahasa yang di- gunakan untok *mentadbirkan* negara dalam
13. jawab di bidang pemerintahan, kawalan dan *mentadbir*
14. 4782 bukan hanya *mentadbir* soal pemeriksaan, pemantauan dan
15. orporat terhadap nilai pemegang saham telah *mentafsirkan*
16. il bahagian dalam proses demokrasi, yaitu proses *mentadbirkan*
17. ereka juga mendakwa agensi yang *mentadbirkan*
18. yang setia kepada pemerintah Inggeris dan *mentafsirkan*
19. yang sama kedudukannya dan *menterjemah* kelemahan
20. bahawa pemerintah diamanahkan untuk *mentadbir* angka masa tiga tahun,
21. pembesar yang menyokong pemerintahan dan *mentadbirkan* kerajaa B00404
22. nubuhan Badan $$158 Pemerintah Negara bagi *mentadbir* nega
23. Pengetahuan bertujuan *mentadbir* dan mengurus
24. pula bagai tempat *mentadbirkan* pemerentahan; orang2 yang
25. dan sebal aitu proses *mentadbirkan* pemerintahan mereka sendiri
26. Tentera Inggeris Britis *mentadbirkan* negeri ini. Pemerentahan
27. Dunia Arab *mentafsirkan* pemencilan
28. Yang di-Pertuan atur dan *mentadbir* urusan pemerintahan negara."
29. Harta sepencarian Seksyen 2 EUKIS *mentafsirkan* harta sepencarian
30. hingga kini. Makna `mat' pula bagi *menterjemahkan* maksud kepada
31. strategi terbaik ke arah *mentransformasikan* universiti ini menjadi
32. Mereka sengaja *mentafsirkan* galakan pengajaran bahasa
33. wajib dilaksanakan oleh Muslim selepas *mentauhidkan* Allah dan berbuat
34. dilaksanakan oleh Muslim selepas *mentauhidkan* Allah dan berbuat baik
35. Saya sebagai masyarakat biasa *menterjemahkan* hasrat Perdana Menteri
36. sudahkah kita bersungguh-sungguh *mentaati* Allah? sendiri yang
37. sudahkah kita bersungguh-sungguh *mentaati* Allah? Kalau belum, jangan
38. Graduan harus *menterjemahkan* idea dan cita-cita murni melalui pende
39. memberi kebebasan kepada pembaca untuk *mentafsir* kepentingan di

40. Terdahulu kelmarin, Blatter *mengkritik* beberapa pengadilan dalam
41. uga mengetahui mengenai kewujudan songket Melayu, *mengkomersialkan*
42. 84 ajar matlamat kerajaan untuk memoden dan *mengkomersilkan* sektor i
43. matlamat kerajaan untuk memoden dan *mengkomersilkan* sektor itu. ..
44. kerajaan untuk memoden dan *mengkomersilkan* sektor itu.
45. peruntukan RM3.6 bilion bagi memoden dan mengkomersilkan sektor i
46. peruntukan RM3.6 bilion bagi memoden dan *mengkomersilkan* sektor itu.
47. peruntukan RM3.6 bilion bagi memoden dan mengkomersilkan sektor itu.
48. ganjarannya sama seperti *mengkhatam* al-Quran pada bulan lain.
49. milik seorang senator dipercayai *mengklon* AP kereta sehingga mengaut
50. Kejayaan dalam *mengkomersialkan* biodiesel sebagai bahan api komersial
51. Pengarang *mengkategorikan* hampir semua karya Pramoedya Ananta Toe
52. Pengarang *mengkategorikan* hampir semua karya Pramoedya Ananta To
53. beberapa orang jurutera lain yang *mengkhusus* dari segi peralatan, M&E
54. beliau selanjutnya *mengklasifikasikan* persaudaraan kepada empat
55. mereka untuk mengupas dan *mengkritik* sesuatu agama tidak kira Islam,
56. apabila diberi amanah suka *mengkhianati* amanah.
57. dan seterusnya *mengkomersialkan* produk dihasilkan," katanya.
58. Perdana Menteri yang *mengkritik* dengan begitu keras dan lantang,
59. Kejayaan dalam *mengkomersialkan* biodiesel sebagai bahan api komersial
60. Dr Mahathir *mengkritik* Abdullah sehari selepas perjumpaan empat ma
61. kepemimpinan Melayu, *mengkritik* parti Melayu dan kerajaan yang
62. kepemimpinan Melayu, *mengkritik* parti Melayu dan kerajaan yang
63. "mengapa Dr Mahathir *mengkritik* kerajaan yang pernah diterajuinya
64. beliau selanjutnya *mengklasifikasikan* persaudaraan kepada empat
65. beragama lebih suka *mengkritik* seni tetapi tidak mahu mengambil
66. beragama lebih suka *mengkritik* seni tetapi tidak mahu mengambil
67. akhbar juga *mengkritik* tindakan kita setiap hari (orang melaporkan
68. *mengklasifikasikan* kes itu sebagai kematian mengejut, k
69. mulanya *mengklasifikasikan* kes itu kematian mengejut,
70. berhasrat *mengkontrakkan* keluar mana-mana jaminan BHBR46 S
71. pembuat yang berhasrat *mengkontrakkan* keluar mana-mana jaminan
72. yang berhasrat *mengkontrakkan* keluar mana-mana jaminan tersedia in
73. "Kita perlu *mengkelaskan* pusat hiburan kerana sese BHCR80 SE
74. "Kita perlu *mengkelaskan* pusat hiburan kerana sesetengah pusat h
75. "Kita perlu *mengkelaskan* pusat hiburan kerana sesetengah pusat
76. lebih meluas ke arah *mengkomersialkan* sektor berkenaan serta BHCR12
77. lebih meluas ke arah *mengkomersialkan* sektor berkenaan serta
78. meluas ke arah *mengkomersialkan* sektor berkenaan serta penggunaan t
79. kemasukan kereta import *mengklonkan* AP untuk mendapatkan keunt
80. angan KDRM kerana tindakan *mengklonkan* AP dikatakan berlaku tanpa
81. perinci berhubung tindakan *mengklonkan* AP berkenaan
82. kereta import *mengklonkan* AP untuk mendapatkan keuntungan segera.
83. KDRM kerana tindakan *mengklonkan* AP dikatakan berlaku tanpa disedari
84. siasatan terperinci berhubung tindakan *mengklonkan* AP berkenaan.
85. kereta import *mengklonkan* AP untuk mendapatkan keuntungan segera.

86. kerana tindakan *mengklonkan* AP dikatakan berlaku tanpa disedari pem
87. berhubung tindakan *mengklonkan* AP berkenaan. "Saya sudah
88. berjaya membangun dan *mengkomersialkan* tiga teknologi sehingga BHDR34
89. berjaya membangun dan mengkomersialkan tiga teknologi sehingga
90. membangun dan *mengkomersialkan* tiga teknologi sehingga Februari la
91. hawanan dan tabung bagi *mengkomersialkan* hasil R&D. Oleh keran
92. keusahawanan dan tabung bagi *mengkomersialkan* hasil R&D
93. pembangunan keusahawanan dan tabung bagi *mengkomersialkan* hasil R&D.
94. contoh, *mengkelaskan* kehidupan seks mereka tin
95. sebagai contoh, *mengkelaskan* kehidupan seks mereka tinggi manakala t
96. sebagai contoh, *mengkelaskan* kehidupan seks mereka tinggi manakala
97. berfungsi mengurus dan *mengkomersialkan* produk penyelidikan uni
98. Mustapha, UPM bersedia *mengkomersialkan* lebih 40 produk penyeli
99. mengurus dan *mengkomersialkan* produk penyelidikan
100. UPM bersedia *mengkomersialkan* lebih 40 produk penyelidikan
101. berfungsi mengurus dan *mengkomersialkan* produk penyelidikan
102. Menurut Nik Mustapha, UPM bersedia *mengkomersialkan* lebih 40
103. usaha *mempopular* dan *mengkomersialkan* kain itu bertujuan seba
104. usaha *mempopular* dan *mengkomersialkan* kain itu bertujuan
105. usaha *mempopular* dan *mengkomersialkan* kain itu bertujuan
106. Bagaimana Shanon *mengkelaskan* genre album itu? Shanon
107. tetapi saya *mengkelaskan* album itu sebagai pop kre
108. Bagaimana Shanon *mengkelaskan* genre album itu? Shanon:
109. Terasnya tetap pop, tetapi saya *mengkelaskan* album itu sebagai
110. Bagaimana Shanon *mengkelaskan* genre album itu? Shanon:
111. tetapi saya *mengkelaskan* album itu sebagai pop kreatif. Mungkin
112. 43 bidang itu dan membantu mengkomersialkan produk mereka.
113. sebagai platform bagi *mengkomersialkan* produk yang dihasilkan
114. itu dan membantu *mengkomersialkan* produk mereka. Beliau juga
115. sebagai platform bagi *mengkomersialkan* produk yang dihasilkan
116. dan membantu *mengkomersialkan* produk mereka. "Antara bidang
117. bagi *mengkomersialkan* produk yang dihasilkan daripada keg
118. 46 rhatian penting sebelum *mengkomersialkan* produk baharu ini.
119. tiga perhatian penting sebelum *mengkomersialkan* produk baharu
120. tiga perhatian penting sebelum *mengkomersialkan* produk baharu
121. elidikan dan seterusnya *mengkomersialkan* produk dihasilkan,"
122. dan seterusnya *mengkomersialkan* produk dihasilkan,"
123. dan seterusnya *mengkomersialkan* produk dihasilkan," katanya.
124. Amerika Syarikat yang *mengkelaskan* Hamas sebagai pertubuhan
125. Amerika Syarikat yang *mengkelaskan* Hamas sebagai pertubuhan
126. Amerika Syarikat yang *mengkelaskan* Hamas sebagai pertubuhan
127. 51 idikan, pembangunan dan *mengkomersialkan* sumber tenaga boleh
128. dan *mengkomersialkan* sumber tenaga boleh diperbaharui,"
129. dan *mengkomersialkan* sumber tenaga boleh diperbaharui,"
130. ada pihak tertentu *mengklonkan* kad mereka. Adalah lebih
131. kerana kemungkinan ada pihak tertentu *mengklonkan* kad mereka.

132. ada pihak tertentu *mengklonkan* kad mereka. Adalah lebih baik
133. lah SWT turun ke bumi dan *mengkabulkan* setiap doa yang
134. Allah SWT turun ke bumi dan *mengkabulkan* setiap doa yang
135. Allah SWT turun ke bumi dan *mengkabulkan* setiap doa yang
136. "Usahawan wanita mampu *mengkomersialkan* hasil penyelidikan dan
137. "Usahawan wanita mampu *mengkomersialkan* hasil penyelidikan dan
138. "Usahawan wanita mampu *mengkomersialkan* hasil penyelidikan dan
139. 57 erharap pihak media tidak *mengkelaskan* artis tanah air yang
140. "Saya berharap pihak media tidak *mengkelaskan* artis tanah air
141. "Saya berharap pihak media tidak *mengkelaskan* artis tanah air
142. 61 nan serta memajukan dan *mengkomersialkan* paten berasaskan
143. serta memajukan dan *mengkomersialkan* paten berasaskan
144. serta memajukan dan *mengkomersialkan* paten berasaskan teknologi
145. 65 wujudan songket Melayu, *mengkomersialkan* tekstil itu dan
146. songket Melayu, *mengkomersialkan* tekstil itu dan dikenali pada
147. kewujudan songket Melayu, *mengkomersialkan* tekstil itu dan
148. 66 rajaan untuk memoden dan *mengkomersilkan* sektor itu.
149. matlamat kerajaan untuk memoden dan *mengkomersilkan* sektor
150. matlamat kerajaan untuk memoden dan *mengkomersilkan* sektor
151. hubung pendiriannya dalam *mengkelaskan* definisi filem nasional.
152. pendiriannya dalam *mengkelaskan* definisi filem nasional.
153. berhubung pendiriannya dalam *mengkelaskan* definisi filem
154. pada aktiviti ini ialah *mengkomersialkan* projek R&D. Bagi
155. perlu dilaksanakan bagi *mengkomersialkan* 10 peratus daripada
156. Sektor swasta pula mahu *mengkomersialkan* produk yang berpotensi
157. berfungsi mengurus dan *mengkomersialkan* produk penyelidikan. La
158. UPM diberitakan sedia *mengkomersialkan* 40 produk penyelidikan
159. berpotensi untuk *mengkomersialkan* projek terhasil. Kalau
160. daripada aktiviti ini ialah *mengkomersialkan* projek R&D. Bagi
161. perlu dilaksanakan bagi *mengkomersialkan* 10 peratus daripada
162. Sektor swasta pula mahu *mengkomersialkan* produk yang berpotensi
163. mengurus dan *mengkomersialkan* produk penyelidikan.
164. UPM diberitakan sedia *mengkomersialkan* 40 produk penyelidikan
165. syarikat berpotensi untuk *mengkomersialkan* projek terhasil.
166. penting daripada aktiviti ini ialah *mengkomersialkan* projek
167. pelbagai usaha perlu dilaksanakan bagi *mengkomersialkan* 10
168. Sektor swasta pula mahu *mengkomersialkan* produk yang berpotensi
169. berfungsi mengurus dan *mengkomersialkan* produk penyelidikan.
170. UPM diberitakan sedia *mengkomersialkan* 40 produk penyelidikan
171. berpotensi untuk *mengkomersialkan* projek terhasil. Kalau tidak
172. Kejayaan dalam *mengkomersialkan* biodiesel sebagai
173. Kejayaan dalam *mengkomersialkan* biodiesel sebagai bahan BHHR86
174. Kejayaan dalam *mengkomersialkan* biodiesel sebagai bahan api
175. Kejayaan dalam *mengkomersialkan* biodiesel sebagai bahan api
176. Kejayaan dalam *mengkomersialkan* biodiesel sebagai bahan api
177. Kejayaan dalam *mengkomersialkan* biodiesel sebagai bahan api

178. juga akan membantu *mengkomersialkan* hasil landskap mereka u
179. "Kita juga akan membantu *mengkomersialkan* hasil landskap mereka
180. "Kita juga akan membantu *mengkomersialkan* hasil landskap
181. saing; membangun serta *mengkomersialkan* bidang bioteknologi.
182. membangun serta *mengkomersialkan* bidang bioteknologi.
183. membangun serta *mengkomersialkan* bidang bioteknologi.
184. bilion bagi memoden dan *mengkomersilkan* sektor itu. Ini
185. mkan peruntukan RM3.6 bilion bagi memoden dan *mengkomersilkan*
186. mkan peruntukan RM3.6 bilion bagi memoden dan *mengkomersilkan*
187. alam bioteknologi serta *mengkomersialkan* harta intelek dalam
188. dalam bioteknologi serta *mengkomersialkan* harta intelek dalam
189. dalam bioteknologi serta *mengkomersialkan* harta intelek dalam
190. Dalam *mengkuduskan* ilmu seorang budiman meng BHIR69
191. Dalam *mengkuduskan* ilmu seorang budiman menganggap: "Pendi
192. Dalam *mengkuduskan* ilmu seorang budiman menganggap: "Pend
193. bagi membangun dan *mengkomersilkan* sistem itu yang dikataka
194. membangun dan *mengkomersilkan* sistem itu yang dikatakan bakal
195. dengan P21 bagi membangun dan *mengkomersilkan* sistem itu yang
196. 96 yang mana mereka sudah *mengkomersialkan* produk dan
197. yang mana mereka sudah *mengkomersialkan* produk dan perkhidmatan
198. yang mana mereka sudah *mengkomersialkan* produk dan
199. peribadi mereka yang cuba *mengkelaskan* bahan bacaan kepada dua
200. cuba *mengkelaskan* bahan bacaan kepada dua kelompok
201. peribadi mereka yang cuba *mengkelaskan* bahan bacaan kepada dua
202. Kemas yang pakar dalam *mengkomersialkan* produk makanan kampung,
203. Kemas yang pakar dalam *mengkomersialkan* produk makanan
204. Kemas yang pakar dalam *mengkomersialkan* produk makanan
205. 102 ia menjanjikan pasti akan *mengkabulkan* permohonanmu itu.
206. Dia menjanjikan pasti akan *mengkabulkan* permohonanmu itu.
207. serta Dia menjanjikan pasti akan *mengkabulkan* permohonanmu
208. gambil kesempatan untuk *mengkomersialkan* Ramadan dan Aidilfitri
209. mengambil kesempatan untuk *mengkomersialkan* Ramadan dan
210. mengambil kesempatan untuk *mengkomersialkan* Ramadan dan
211. 105 bersedia membeli serta *mengkomersialkan* pengeluarannya
212. membeli serta *mengkomersialkan* pengeluarannya sebanyak 100,000
213. bersedia membeli serta *mengkomersialkan* pengeluarannya sebanyak
214. Melaka untuk *mengkomersialkan* asam pedas Melaka, seka BHKR43
215. Duyung, Melaka untuk *mengkomersialkan* asam pedas Melaka, sekali
216. Melaka untuk *mengkomersialkan* asam pedas Melaka, sekali gus
217. bersedia membeli serta *mengkomersialkan* pengeluarannya sebanyak
218. serta *mengkomersialkan* pengeluarannya sebanyak 100,000 uni
219. China bersedia membeli serta *mengkomersialkan* pengeluarannya
220. "Tetapi kami mahu *mengkomersialkan* lagi dewan ini dengan

Appendix B

Multiple prefixes

(a) Verbal prefixes

1. gi *memperolehi* kuasa selaku pemegang hak komersil.
2. barulah McRae *memperlihatkan* ana dia hanya tunduk kepada
3. paksa *mempercepatkan* tarikh pemergian mereka untuk membe
4. *mempercayai* pemimpin dan pemerintah tentera sendir
5. selain *memperincikan* rancangan pengeboman itu.
6. dan *memperkukuhkan* Program Pemasaran Satu
7. Peringkat Daerah (MPUD) dan *memperkukuhkan* Program
8. kita akan *memperoleh* laporan yang boleh diterima
9. setelah lebih 30 tahun *memperjuangkan* ialah Setiausaha
10. pencinta anggerik untuk *memperluaskan* pengetahuan dan
11. menjalankan perniagaan, *memperkenalkan* produk baru, serta
12. menjalankan perniagaan, *memperkenalkan* produk baru, serta berupaya
13. dan masih mempertahankan media.
14. tidak boleh dijadikan asas untuk *mempertikaikan* kesahihan Artikel 121
15. Artikel 3 yang *memperuntukkan* Islam sebagai agama rasmi
16. Maka, *mempertikaikan* artikel ini mungkin boleh disamakan
17. boleh disamakan seperti *mempertikaikan* Artikel 3 berkenaan.
18. pekebun sayur, kini mereka mampu *memperoleh* pendapatan sehingga
19. mereka mula *mempersoalkan* apakah yang difahami
20. ini dan *mempersiapkan* pasukan wanita dengan baik."
21. mengeluh dan *memperbesarkan* isu remeh temeh tanpa
22. kepada pelajar yang mahu *memperoleh* kemudahan
23. Tinggi ini dapat *mempertingkatkan* lagi kualiti pendidikan
24. Perdana Menteri yang *mempertahankan* Perkara 121(1A)
25. ketegasan Perdana Menteri yang *mempertahankan*
26. Iran mempunyai hak untuk *mempertahankan* dirinya.
27. dan berharap untuk terus *mempertahankan* jawatan Ahli Parlimen
28. ke arah memaju dan *memperkasakan* sektor per
29. langkah ke arah memaju dan *memperkasakan* sektor pertanian
30. ke arah memaju dan *memperkasakan* sektor pertanian tidak perlu
31. sedangkan kerajaan *memperuntukkan* beratus juta setahun
32. terpaksa menanggung kos tinggi *memperbaiki*nya.
33. pasukan lelaki akan dapat *mempertahankan* emas perseorangan
34. terbaik mengenai program *memperkemas* laluan domestik
35. Pada bahagian ini juga Ginsberg *memperkatakan* mengenai teknik
36. strategi untuk memudah dan *mempercepatkan*
37. yang sentiasa *memperbaharui* serta berterusan *memperkenalkan* standard
38. percutian termasuk semua', *memperkenalkan* Mini Trident

39. Loloq kekal digunakan apabila dia *memperoleh* taraf pemastautin tetap
40. cerpen ini *memperlihatkan* Putu juga seolah-olah mahu menggambarkan
41. "Kami juga berjanji akan *memperketat* dan *memperkemas* proses kawalan
42. berjanji akan *memperketat* dan *memperkemas* untuk
43. menjaga, memartabat dan *memperkasa* institusi k
44. menjaga, memartabat dan *memperkasa* institusi keluarga.
45. tanggungjawab menjaga, memartabat dan *memperkasa* institusi keluarga.
46. kaedah ini membantu *mempercepatkan* pembayaran.
47. Malaysia bagaimanapun sudah *memperjelaskan* perkara itu kepada
48. sentiasa melalui budaya *mempertingkatkan* kualiti
49. "Satu lagi, bank akan *memperkenalkan* pembiayaan baru
50. dan Sarah Parker mula *mempersoalkan* hubungan perasaan mereka.
51. sungguh bernilai untuk *memperjuangkan* agama dengan cara menuntut
52. terus dilaksana termasuk *memperketatkan* pembinaan sistem kumbahan,
53. di pesisiran pantai serta *memperketat* peraturan pembuangan sampah.
54. titik awal iaitu dengan *memperkemaskan* mekanisme perundangan yang
55. sebelumnya pernah lantang *mempertikaikan* mengapa `orang tempatan'
56. Kementerian Pelajaran bersetuju *memperkenalkan* sistem ini
57. kaedah yang akan *memperbaiki*
58. kaedah yang akan *memperbaiki* kedudukan Tenaga Nasional Bhd (TNB).
59. tetapi untuk *memperbaiki* perjanjian sedia ada
60. Airocom turut *memperkenalkan* produk terbarunya yang dikenali sebagai
61. Selain itu, syarikat itu juga *memperkenalkan* LinguaMobile,
62. budayawan Indonesia itu *memperkatakan* Sastera dan Politik
63. irama dangdut untuk *memperkenalkannya* di persada seni.
64. kita mula *mempersoalkan* apakah pengarang boleh bertindak
65. Pada majlis *memperkenalkan* kenderaan dan anggota kedua-dua kumpulan
66. Mungkin ada golongan *mempersoalkan* sama ada masih
67. tanah rakyat dan *mempersoalkan* pindaan yang
68. tempat ini untuk *memperingati* Umno, parti yang kita
69. tempat ini untuk *memperingati* Umno, parti yang kita sayangi.
70. rakan peguamnya tidak *mempertikaikan* sama ada Rozana
71. memudah dan *mempercepatkan* proses mendapatkan bantuan daripada Tabu
72. dua hala kerana saling tidak *mempercayai* sejak hubungan diplomatik
73. melihat sesama sendiri sebagai pesaing, *memperkuatkan* angkatan
74. "Tetapi kami kembali dan *memperbaiki* kesilapan kami.
75. Dan Kalinichenko kemudian *memperoleh* golnya, menjaringkan hasil
76. Simao meluru di sebelah kiri, *memperdayakan* pertahanan di penjuru
77. karya yang secara halus *mempersenda* hukum Islam dengan
78. halus *mempersenda* hukum Islam dengan menyeleweng
79. menyumbang kepada usaha *mempertingkatkan* tahap kesihatan masyarakat di
80. pelbagai khasiat kini *memperkenalkan* tiga pilihan roti dalam
81. kajian serta tidak jemu *memperbaiki* diri.
82. Justeru, dia cuba *mempertingkatkan* penyampaian selagi terdaya.
83. tawar kering atau salai di samping *memperbanyakkan* promosi masakan
84. TLDM yang banyak berjasa dalam *mempertahankan* kedaulatan negara.

85. bukan saja bagi *memperkukuhkan* projek itu, masyarakat
86. bukan saja bagi *memperkukuhkan* identiti masyarakat itu
87. tetapi juga *memperkaya* khazanah budaya masyarakat Malaysia.
88. tujuan ritual seperti *mempertahankan* nyawa mereka di laut.
89. bukan saja bagi *memperkukuhkan* identiti masyarakat itu
90. tetapi juga *memperkaya* khazanah budaya masyarakat Malaysia.
91. perubahan dengan *memperkenalkan* makanan segera yang baik untuk
92. matlamat kami yang ingin *memperkenalkan* makanan yang sihat untuk
93. Siti tidak mahu *mempermainkan* media dan semua pihak
94. kerana dia tidak *memperoleh* undi tertinggi,
95. gerak kerja *memperkasa* ekonomi Melayu.
96. Malaysia-India dapat *memperkukuhkan* jalinan kerjasama strategik
97. rumput dipadang *memperoleh* cahaya matahari
98. Bank Rakyat kini mengorak langkah bagi *memperluaskan* perkhidmatan
99. pelanggan akan *memperoleh* prakelulusan ini
100. yang membabitkan usaha *memperkasa* modal insan.
101. kenyataan Dr Mahathir yang *mempersoalkan* mengenai Islam
102. tujuan utama ialah membantu memaju dan *memperkasakan* ekonomi ne
103. pelbagai program pembangunan bagi *memperbaiki* taraf hidup
104. perubahan besar yang saya lakukan untuk *memperbaiki*nya.
105. Rozi turut *mempersoalkan* dakwaan itu kerana menurutnya sehingga
106. Wajahnya *mempersona* Tiada lagi tandingan Bagai dewi menjelm
107. Nadzri pula berkata, BKN Selangor *memperuntukkan* RM200,000
108. saya yakin ramai yang akan *mempercayainya*, "
109. Keghairahan stesen radio ini *memperkenalkan* rancangan baru dan
110. pasaran lebih meluas dan *memperkembangkan* pelanggan MSC
111. Perisian baru itu dapat *memperluaskan* keupayaan simulasi
112. jumlah pemilik lori yang *memperbaharui* cukai jalan iaitu
113. generasi sekarang dalam usaha *memperbanyakkan* lagi ahli
114. kebaikan kerana terpaksa tidak *memperoleh* balasan kebaikan
115. anaknya selamat tetapi sukar *mempercayai* cara anaknya paling
116. Asean sedang *memperhalusi* draf Piagam Asean untuk diputuskan
117. Hari Pengguna juga bertujuan *memperkasakan* tahap kepenggunaan.
118. ini diharap akan *memperbetulkan* lagi struktur pengurusan dan
119. selain *memperhalusi* isu kebajikan pemain dan mendefinisikan se
120. sebelum kemudiannya *memperbetulkan* kenyataan betapa ia akan
121. Messi menyumbat golnya apabila *memperdaya* terpaan Conway
122. "Satu lagi langkah ialah *memperkenal* serta menggunakan kaedah
123. Bagaimanapun, yang pasti apabila *memperkatakan* hak dan
124. Jabatan Penerangan yang *mempersembahkan* lagu dan tarian
125. Kita perlu *memperkukuhkan* sistem pertahanan negara kita agar
126. Kita perlu *memperkukuhkan* sistem pertahanan kita untuk masa
127. *memperjuangkan* nasib anak bangsa tetapi mencari keselesa
128. pusat bandar Kuala Lumpur dan *memperluaskan* laluan Putra-LRT
129. aduan itu turut *mempersoal* tindakan tidak adil media kerana
130. tindakan tidak adil media kerana *mempersoalkan* petugas itu.

131. katanya sambil *mempertahankan* anak muda itu pasti akan menjadi
132. Hotel Legend *memperkenalkan* hidangan untuk Businessmen
133. Bagi musim kedua, penerbit *memperkenalkan* elemen baru,
134. ketika majlis *memperkenalkan* himpunan baru berkenaan kepada
135. Pada sesi *memperkenalkan* peserta lelaki yang berlangsung di
136. saya akan *memperkenalkan* teknik Magic Perms yang dijangka
137. Diva Popular akan *memperkenalkan* nama saya dalam industri
138. peringkat global serta *mempergiatkan* pelaburan syarikat
139. semua yang dirancang, kerajaan *memperuntukkan* RM159.4 bilion
140. proses *memperoleh* beberapa hartanah utama di beberapa lokasi
141. Kerajaan prihatin terhadap perlunya *mempermudahkan* lagi urusan
142. Bagi mempercepat dan *mempermudahkan* lagi proses permohonan
143. semoga usaha *mempertingkatkan* mutu menjadi satu budaya kerja.
144. Rata-rata pertubuhan dapat *mempertingkat* mutu kerja jika ia
145. Apa yang penting, aspek saling *mempercayai* antara kedua-dua
146. Sesiapa *memperbanyakkan* selawat kepadaku (Nabi SAW) pada bulan
147. Usaha *memperkukuhkan* pembangunan modal insan dalam industri pe
148. mengenal pasti, *memperkuat* serta membimbing potensi
149. latihan seni *mempertahankan* diri dan bersenam di gimnasium
150. walaupun ramai orang *memperkatakannya*. Nama sebenar saya dieja
151. "Saya terpaksa *memperlahankan* larian ketika turun kerana laluan
152. Kita juga akan *memperkenalkan* program francais sistem pajak
153. Kita juga akan *memperkenalkan* program francais sistem pajak
154. Justeru, katanya, syarikat *memperuntukkan* antara RM45 juta dan
155. kempen *mempertahankan* kejuaraan dunia di Belfast, Ireland Utar
156. toksin dan *memperbaiki* pered BHJR19 763 phrasia.
157. hotel *memperkenalkan* promosi set Banquet Ramadan dengan ma
158. A Window On Asia pula akan *mempersembahkan* 39 filem oleh
159. Jadi, saya berasa perlu *memperbetulkan* keadaan dengan mata
160. DIOR *memperkenalkan* produk anti penuaan bagi mendapatkan waj
161. Dari Siplis Ke Pararuhi *memperakukan* proses kreativiti
162. masalah kebakaran adalah *memperketatkan* hukuman, menjalankan
163. manusia rugi kerana tidak *memperoleh* sebarang manfaat atau
164. pendokumentasian dan *memperkenalkan* pengkaryaan ulama dan
165. peningkatan terbaru bagi *memperkasakan* pekerja - aset paling
166. Saya *memperjuangkan* nasib artis supaya mendapat pembelaan
167. tisu terhadap insulin serta *memperbaiki* metabolisme
168. Watsons dalam penjagaan wajah dengan *memperkenalkan*, Watsons
169. calon yang seolah-olah *memperlekehkan* usaha Berita Harian.
170. medan untuk *memperkenalkan* Anugerah Spa dan Kesejahteraan
171. baru, sekali gus *memperkasakan* bahasa Melayu sebagai lambang
172. Tidak perlu lagi *mempertajamkan* kebolehan dan keupayaan
173. RMK-9 iaitu *memperkasakan* sekolah kebangsaan supaya menjadi
174. kerajaan dijangka *memperuntukkan* kira-kira RM40 bilion bagi
175. ir selaras usaha kerajaan *memperkasakan* sekolah sebagai pusat
176. selaras usaha kerajaan *memperkasakan* sekolah sebagai pusat

(b) Nominal prefixes

1. *pemerhatian* itu dibuat, *pemerhatian* itu masih sah.
2. tung kepada pembacaan dan *pemerhatian* itu saja."Betul itu,
3. Ahmad, 1991). Sebetulnya pemerhatian itu sekadar menunjukkan
4. 1160 an lisan berkaitan dengan *pemerhatian* itu, yang seterusnya
5. (Dalam *pemerihalan* di atas beberapa perkara
6. Berdasarkan *pemerihalan* di atas, dilihat dari tuj
7. akan juga membolehkan *pemerihalan* dibuat mengenai agensi-age
8. jika dibandingkan *pemerihalan* morfologi dongengan darip
9. Daripada *pemerihalan* perkembangan sejarah Remb
10. linguistik sebagai proses *pemerolehan* bahasa tanpa konteks
11. biasanya berakhir dengan *pemerolehan* dan ada kalanya dengan
12. bahasa sebagai alat asas *pemerolehan* dan penyebaran ilmu.
13. (kesemuanya berasal dari *pemerolehan* faedah budaya) berserta
14. di samping sebagai ideologi *pemersatu* rakyat dalam perjuangan
15. bahasa Melayu menjadi unsur *pemersatu* apabila diwujudkan sebuah
16. Kalau fungsi *pemersatu* bahasa Melayu berlaku di
17. tidak dapat saya elakkan *pemertalian* Sumpah Pemuda dengan baha
18. berrpunca khususnya daripada *pemeruapan* hasil sampingan dan
19. Contohnya - *pemeruapan* komponen hidroperoksida
20. Saya menoleh keruang *pemerhati*. Ruang ini kelihatan pada
21. sebagai teknik penilaian *pemerhati*.
22. ke mesyuarat itu sebagai *pemerhati*. Johan memperlihatkan min
23. Bagi *pemerhati-pemerhati* politik, inilah
24. utama, di samping alat *pemersatu* dan pemadu bangsa yang
25. unifying force atau tenaga *pemersatu* hingga sekarang belum juga
26. Fungsi *pemersatu* itu amat jelas juga dengan
27. berperanan sebagai tenaga *pemersatu* kepada umat kepulauan Melayu
28. gerakan sosial, maka *pemerihalan* mengenainya berkeadaan
29. pemasaran menitikberatkan *pemerihalan* ciri-ciri pasaran.
30. bidang kajian merangkumi *pemerihalan* dan perbincangan mengenai
31. mempunyai BN sebagai bahasa *pemersatu* dan bahasa komunikasi
32. itu dikenal sebagai fungsi *pemersatu* dan fungsi pembeda.
33. bahasa sebagai alat asas *pemerolehan* dan penyebaran ilmu.
34. (kesemuanya berasal dari *pemerolehan* faedah budaya) berserta
35. bahasa lain itu memerlukan *pemerolehan* identiti lain selain dari
36. Malah, dia menganggap *pemerolehan* ketrampilan pengurusan
37. Hanya aku yang menjadi *pemerhati*. Aku tidak tahan bermain
38. ibunya bertindak sebagai *pemerhati*. Dia tidak berani mencelah
39. itu juga dihantar sebagai *pemerhati*. Katanya dia orang Amerika
40. percubaan memaksakan *pemersatuan* akan menimbulkan keretaka etaka
41. ukan pribumi sebagai alat *pemersatuan* dan penyatuan bangsa. $$13

Appendix C

Monosyllabic prefixed words in SM

1. berasa aman dan tenang untuk *mengecapi* percutian dengan lebih
2. rakyat terlepas *mengecap* nikmat pembangunan mengikut jadual.
3. satu ketika pernah *mengecam* cara pengendalian pasukan
4. menuju ke arah mencapai kejayaan dan *mengecapi* kejayaan sebenar.
5. pemain yang pernah *mengecapi* saat manis mendakap piala
6. sekiranya terus-menerus *mengecam* Pak Lah yang mempunyai personaliti
7. ia adalah lebih elok berbanding *mengecamnya.*
8. mereka yang dipalsukan dijual di pasaran dan cara *mengecam* barangan
9. bertindak *mengecas* atau mengionkan bahan api secara magnetik
10. bertindak *mengecas* atau mengionkan bahan api secara magnetik
11. pada hos minyak, bertindak *mengecas* atau mengionkan bahan api secara
12. mereka bertindak *mengecat*
13. mereka yang rosak, boleh *mengepam* lebih banyak darah berbanding
14. budak kecil yang tidak tahu *mengepam* tandas selepas menggunakannya
15. ri untuk keluar selaku pemenang, sekali gus *mengesahkan* tem
16. cuba *mengebom* sebuah pusat pemeriksaan tetapi alat jan
17. Datuk Seri Musa Hassan *mengesahkan* polis masih belum dapat
18. balan Ketua Polis Negara, Datuk Seri Musa Hassan *mengesahkan* polis
19. Minta doktor anda *mengesyorkan* pakar untuk keadaan itu,
20. "Cabaran yang dihadapi itu memaksa saya *mengehadkan* jumlah tidur dan
21. biasanya pihaknya akan *mengesyorkan* pelanggan lelaki melakukan proses
22. sunat sebagai garisan yang *mengehad* atau memisahkan mereka dari
23. garisan yang *mengehad* atau memisahkan mereka daripada berusaha u
24. amalan wajib dan sunat sebagai garisan yang *mengehad* atau memisahkan
25. urusan aset secara digital, *mengekod* video dan penerbitan
26. pengurusan aset secara digital, *mengekod* video dan penerbitan
27. Apabila terbaca berita itu, kami tidak *mengesyaki*
28. eminggu setiap bulan menyebabkan doktor berkenaan *mengesyaki* Shafique
29. setiap broker dikehendaki *mengesyorkan* hanya saham yang sesuai
30. rkemahiran tinggi," katanya. k, bertindak *mengecas* atau mengionkan
31. dipasang pada hos minyak, bertindak *mengecas* atau mengionkan bahan api
32. *pengewapan* dan pemejalwapan merupakan haba yang
33. termasuk mesin basuh dan alat *pengecas* telefon bimbit solar.
34. basuh dan alat *pengecas* telefon bimbit solar.
35. kaca mata O ROKR; *pengecaj* mudah alih P790; dan set
36. kaca mata O ROKR; *pengecaj* mudah alih P790; dan set kepala
37. berminat dan tanggungjawab *pengesahan* dan merekod resit
38. Untuk *pengesahan,* pihak lain terutama organisasi bebas yang
39. masalah kerosakan jentera *pengepam* air sejak lebih sembilan bulan
40. termasuk mesin basuh dan alat *pengecas* telefon bimbit solar.

Appendix D

Native words prefixation

1. Amrozi yang *mengenal* pasti mereka yang disyaki selain
2. siasatan sedang dijalankan untuk *mengesan* orang yang disyaki
3. nukilan Kamarul AR, yang *mengetengahkan* musyawarah sembilan daerah
4. 13 Mei, *mengenangkan* kepayahan para petani di bendang
5. harapan tinggi tampil *mengecewakan* apabila tumpas
6. Jika pelajar kita hanya *mengongkong* skop pergaulan sudah tentu ini
7. dan membenarkan *pemeriksaan* dijalankan pihak berkuasa BHBC12 478
8. perlu menjalani *pemeriksaan* ISN (Institut Sukan Negara
9. ukar untuk *menentukan* siapa pemenang di peringkat pertama i
10. *menewaskan* Pakistan 5-2. Pemenang pingat perak
11. kafe siber yang disyaki *menawarkan* permainan video termasuk berbentuk
12. kes Blackman selepas *menahannya* atas tuduhan disyaki menderanya
13. kami bersusah payah *menangkap* orang yang disyaki mereka dengan senang
14. Paling *menarik* di Adelaide ialah pasarnya, Adelaide Central
15. Selain *menembual* penjual kami bersarapan pagi di situ.
16. ke Kuala Lumpur untuk *menemui* keluarganya.
17. mereka yang *menaja* lumba haram boleh dikenakan tindakan
18. biro itu untuk *menangani* gejala lumba haram di kalangan anak muda di
19. hubungan dengan orang lain, *menunaikan* tanggungjawab serta berlaku
20. Islam *menekankan* bahawa usaha pemulihan pesalah mestilah di
21. Selain sukan *memanah* sukan lain yang boleh dijalankan di pusat ini
22. jika tidak lagi *memegang* jawatan dalam sukan memanah.
23. pemaju sudah mula *memasang* zink sebagai penghadang bagi melindungi
24. tindakan segera mungkin dengan *memindahkan* kami dalam waktu terdekat
25. pemaju sudah mula *memasang* zink sebagai penghadang
26. *memuji* lakonan Eja sebagai Mya Zara kerana dia seorang
27. seseorang untuk *memilih* pakaian, asalkan tidak melanggar kesopanan, m
28. beliau tidak mungkin dapat *memenuhi* harapan kita.
29. langkah berdaya maju untuk *memulihkan* keadaan ekonomi negara dan
30. Pemimpin berkesan mampu *menggalakkan* kakitangannya menghasilkan kerja
31. Indra Putra Yadi Aprin, *menggambarkan* kekuatan generasi muda Selangor
32. kakaknya yang *menggunakan* kerusi roda menuju tempat penginapan, iait
33. majikan di Malaysia *menggunakan* alasan ini untuk mengurangkan kos
34. alasan ini untuk *mengurangkan* kos dengan menggantikan pekerja
35. alasan untuk mereka dapat *menggantikan* tenaga kerja tempatan dengan
36. tidak mahu campur tangan dengan me*nggariskan* dasar memajukan industri
37. Dengan itu, ia *menggalakkan* pembuangan bahan kumuh.
38. Ketika remaja seusianya yang lain *menggemari* lagu berentak moden,
39. Advocate, *menggesa* superhero biru merah berkenaan untuk
40. persekitaran perniagaan kondusif dalam *menggalakka*n pertumbuhan

41. kelmarin, sekali gus *menggagalkan* cita-cita untuk mencapai perdagangan
42. dunia musnah kelmarin, sekali gus *menggagalkan* cita-cita untuk
43. lesen teknikal *membenarkan pemegang* francais, Euro Truck & Bus
44. Kad ini *membolehkan pemegang* polisi dimasukkan ke mana-m
45. *pembelian* seperti seorang *pemegang* polisi membeli polisi
46. sekali gus *membolehkan pemegang* polisi *menggunakan* kemudaha
47. Bagi *pemegang* polisi yang tidak *membuat* a
48. untuk *menggilap* potensi 'wanita perkasa' dalam *membuat* keputusan
49. Bukankah ini masa terbaik untuk puan *membalas* budi baik suami dan
50. puan *membalas* budi baik suami dan *membuktikan* kesetiaan kepada suami?
51. kapal terbang *membawa* ubat-ubatan *mendarat* kerana disyaki *membawa*
52. tuduhan disyaki *menderanya* sehingga mati dan
53. terkena jangkitan SARS daripada *menderma* darah meskipun tidak
54. penjenayah disyaki juga berjaya *mendapatkannya*, kata Gonggrijp.
55. menghirup udara segar di pantai serta *mendengar* deruan ombak Laut
56. lambaian pohon nyiur memang begitu *mendamaikan* fikiran.
57. memastikan setiap ahli keluarga *mendapat* kepuasan, gembira dan
58. di pantai serta *mendengar* deruan ombak Laut China Selatan
59. terutama kebersihan, gerai haram dan *pembangunan* cerun tinggi.
60. *penduduk* yang menjadi *pembayar* cukai di Sepang.
61. premis dan cawangan bank; *pemberian* cenderamata *menggunakan* produk
62. bermutu dan menjadi agen kutipan *pembiayaan* yang disalurkan Jabatan
63. berasaskan epal untuk sajian *pembuka* selera, hidangan utama hinggalah
64. ada kelahiran *memerlukan pembedahan* Caesarean atau sebaliknya, serta
65. sementelah *pembabitan* mereka dalam industri seni suara masih
66. untuk *mengurangkan* gejala berkaitan *pembesaran* kelenjar prostat. Saw
67. Penggunaan *pembayang* mata juga tidak seharusnya dikecualikan dal
68. sesiapa saja *menubuhkan* institusi *pendidikan* di sini dan kita harus
69. Oleh ZAM KEBANJIRAN *pendatang* asing ke Malaysia sememangnya dialu-
70. Tetapi, adakah kebanjiran pendatang asing *memberi* manfaat kepada
71. Semua ini *memberi* faedah ekonomi kepada *penduduk* setempat yang
72. Portugal *menggandakan pendahuluan* mereka selepas tindakan kurang
73. kebakaran juga boleh dikurang dengan *pendekatan* sesuai.
74. yang boleh *membawa* bencana dan *penderitaan*
75. kemungkinan *penggabungan* Bank Muamalat dengan bank lain,
76. beliau yang *menuntut* kebebasan *penggunaan* dadah yang kononnya
77. "Memilihnya (Abdullah) sebagai *pengganti*,
78. *mengelak* daripada dikaitkan dengan rangkaian '*pengganas*.
79. dan *membahaskan* usul *penggiliran* jawatan Ketua Menteri,
80. *penggiat* seni tanah air, kami tidak dapat *memenuhi*
81. beliau sebagai *pemeriksa* dan penggubal kertas
82. selepas dikesan mengalami *pendarahan* dalam air kencing akibat
83. Seperti diakui Korrie, betapa *pendekatannya* dalam kumpulan ini adalah
84. bagi *mendekatkan* stsen radio kami dengan *pendengar*," katanya.
85. perkhidmatan, pertanian, *pendidikan*, belia, kebudayaan, sains dan

Appendix E

Nasal and sonorant clusters in SM

1. ela minoriti *merujuk* kepada pemegang saham minoriti Gadek Cap
2. kajian beban tugas guru oleh NUTP yang *merumuskan* guru menanggung
3. di kawasan itu ketika kerja *meroboh* dijalankan.
4. kawasan itu ketika kerja *meroboh* dijalankan.
5. abitkan pengolahan secara *meromantiskan* kisah berkenaan.
6. membabitkan pengolahan secara *meromantiskan* kisah berkenaan.
7. pertumbuhan tinggi *merosot* sembilan peratus.
8. Lee Hsien Loong *merujuk* kepada kebuntuan di Taiwan
9. mengaut keuntungan dengan *merobek* pasaran negara
10. ahli masyarakat untuk *merubah* minda dan tindakan
11. garam dan gula), berhenti *merokok* dan kurangkan
12. *merupakan* tangan- tangan pemerintah untuk mengu
13. Jakarta *merundingkan* pemerintahan autonomi bagi Timor T
14. Chen yang *merujuk* kepada pemerintahan Beijing yang jauh d
15. yang *merupakan* penyokong pemerintahan beraja.
16. larisme telah *meruntuhkan* pemerintahan kita. Jika Islam dapat
17. pemerintahan lemah, ekonomi *merosot* da
18. ke Madinah, pemerintahan *merupakan* perkara utama y
19. Sebelum ini, ini *merupakan* kesalahan pemerintahan Order Baru
20. kepentingan industri dan *mewujudkan* kepelbagaian jenis perniagaan,"
21. pasti perniagaan dan perusahaan, *mewujudkan* kaedah baru menjalankan
22. Kelana *mewakili* Petaling, Permatang (Kuala Selangor),
23. turut menyampaikan ucapan *mewakili* Pengerusi Jawatankuasa Tetap
24. untuk mereka yang ingin *mewarna* rambut di rumah menerusi tiga
25. boleh memulakan proses *mewarna* rambut menggunakan krim pewarna
26. azhab Shafie sekiranya puan *mewasiat* melebihi kadar itu hukumnya
27. mengikut mazhab Shafie sekiranya puan *mewasiat* melebihi kadar itu
28. mazhab Shafie sekiranya puan *mewasiat melebihi* kadar itu
29. Datuk Dr Cyrus Das yang *mewakili* Fawziah Holdings ketika ditemui
30. Situasi ini akan *mewujudkan* suasana tidak seimbang
31. Oleh yang demikian, kerajaan perlu *mewujudkan* satu kerangka baru
32. Pramoedya Ananta Toer *mewacanakan* revolusi kemerdekaan,
33. undang-undang *mewajibkan*
34. semakin goyah semata-mata mahu *mewajarkan* peningkatan
35. Selain itu, ada juga sebab yang *meyakinkan* bagi syarikat
36. menangani isu itu bagi *meyakinkan* rakyat pada peringkat akar
37. Antaranya tugasnya ialah *memandi*, menjaga dan memantau
38. Arsenal yang dulunya bermegah *memiliki* penyerang tulen,
39. bagi melaksanakan tugasnya *memaju* dan menjayakan misi pertubuh
40. wujud seperti menaiki bot laju, *memudik* sungai, menyaksikan

41. Melayu memahami dan masih menganggap memiliki akhlak mulia akan
42. Seterusnya, *memaju* bangsa sehingga mampu berperanan besar kepad
43. ke arah perubahan positif demi *memajukan* ummah.
44. mengajak bagaimana *memuhasabah* dan menilai diri,
45. kita bersama-sama membantu *memeriah* dan menjayakan penganjuran
46. aman webnya bahawa dia mahu *memanfaat* segala kesempatan yang
47. ada segolongan pihak yang *memanipulasi* idea diilhamkan Friedan i
48. makanan atau minuman yang *memabuk* dan memudaratkan diri seperti
49. makanan atau minuman yang *memabuk* dan memudaratkan diri seperti
50. Melihat dan *memandikan* jenazah.
51. di sini antara lain *memohon* memadam semua rujukan berhubung
52. matlamat kerajaan *memaju* dan memodenkan sektor pertanian berhasil
53. selepas seminit, anda sudah boleh *memulakan* proses *mewarna* rambut
54. individu, tidak boleh berjudi, *meminum* minuman keras atau arak dan
55. kehambaan yang tinggi bakal *memerdekakan* negara kita daripada
56. lih berganti *menaiki* pentas pemenang.
57. pegawai penguat kuasa akan *menasihati* orang yang disyaki
58. Begitu juga malam esoknya kami *menikmati* hidangan malam
59. Selepas sekian lama *menanti* kita, akhirnya kita *memiliki* beberapa
60. Beliau tidak *menafikan* soal trend atau aliran dalam muzik tidak da
61. Untuk *menampakkan* kuku sihat dan bersih gosok permukaan kuk
62. ukar *melakonkan* watak Dhani apabila saya disuruh *menangis*.
63. kerana saya mahu penonton juga *menangis* bersama saya apabila mereka
64. *melindungi* haknya sebagai pemegang saham setelah
65. saham syarikat *meluluskan*
66. jauh *melebihi* sasaran *melihat* kemajuan tekn
67. *melihat* kemajuan teknologi pemesinan yang terbar
68. *melepasi* paras 802 mata.
69. mendakwa dirinya boleh *melakukan* pembedahan,
70. menempatkan, *meletakkan* atau *melupus* bahan buangan berjadual
71. kesalahan menempatkan, *meletakkan* atau melupus bahan buangan
72. Berita Minggu semalam *melaporkan* Siti Nazirah meninggal dunia pada 4
73. jangan rambang mata *melihat* pelbagai lauk tersedia.
74. sebuah lagi jambatan bagi *melengkapkan* sistem keluar masuk empat
75. hanya membenarkan mereka keluar *melontar* pada masa yang ditetapkan.
76. Kokanun sudah *meluluskan* pinjaman berjumlah hampir RM1 juta
77. sebelum dipasarkan secara *meluas* di seluruh negeri ini.
78. penggunaan NGV secara *meluas* wajar dilaksanakan
79. jadi popular dan ia akan *melicinkan* putaran roda ekonomi kita.
80. kedua-dua pihak *melahirkan* keyakinan terhadap rundingan damai
81. menghentikan kesuraman yang *melanda* kanak-kanak," katanya
82. setelah *melakonkan* watak itu, Rusdi menegaskan, dia tidak per
83. berkesinambungan berikutan kurang *pewaris* tani
84. kandungan sesuatu bahan pengawet, *pewarna* dan kandungan gula
85. erlakulah *pewarisan* dalam pemerintahan dan pemalsuan had
86. Para *pewakilan* dan pemerhati muktamar Pemuda PAS

Appendix F

Total reduplication in SM (disyllabic roots)

1. "Sejak *akhir-akhir* ini, permintaan terhadap perkhidmatan bengkel
2. Pejabat Daerah dan Tanah LMS, *baru-baru* ini.
3. itu sejak beberapa bulan lalu. *Baru-baru* ini, majlis
4. di kawasan rumah *masing-masing*," katanya. Selain itu, Mustaffa
5. tanah terbiar di kawasan *masing-masing* untuk meningkatkan pengeluaran
6. menjaga keselamatan kilang, *baru-baru* ini. Ketua Polis Daerah Kerian,
7. harus menyerahkan *bulat-bulat* kepada polis untuk membuat
8. serta memastikan jaga *benar-benar* bertugas.
9. di bekas tanah pertanian di *tengah-tengah* bandar raya Ipoh yang cukup
10. berharga RM320,000 dengan keluasan *kira-kira* 6 meter x 2
11. sementara saham *masing-masing* pada harga urus niaga
12. Tenaga dan Maybank *masing-masing* jatuh 10 sen
13. memperkenalkan akta dan undang-undang baru yang menjurus
14. dengan Tuhan, alam raya dan *makhluk-makhluk* sesamanya..."
15. Ajaran Nabi Ibrahim atau \`penemuan' beliau *benar-benar* lembaran baru
16. kaitan ibadat haji dengan *nilai-nilai* kemanusiaan
17. Ia mencakupi *nilai-nilai* luhur yang seharusnya menghiasi jiwa
18. Kemanusiaan yang menyebabkan *anak-anak* Adam menyedari arah yang dituju
19. *Makna-makna* itu diamalkan dalam pelaksanaan ibadah haji,
20. ia tetap *sama-sama* di hadapan Tuhan.
21. perbezaan antara mereka dalam *hal-hal* lainnya.
22. mewujudkan situasi *menang-menang* iaitu mampu dibayar pelanggan dan
23. Amerika dan Pesuruhjaya Tinggi *Bangsa-bangsa* Bersatu bagi Pelarian
24. Namun, *sumber-sumber* memberitahu Berita Harian Pengerusi Eksekutif RHB
25. kegawatan ekonomi akan menjejaskan *cita-cita* kita untuk menjadi pusat
26. yang dibenarkan Syari'ah *saham-saham* BSKL. Di bawah kemudahan
27. anggota yang terperangkap di *tengah-tengah* kegawatan mata wang dan
28. *Lebih-lebih* lagi, pelaburan IDB di Malaysia adalah bukti kepe
29. mendapatkan penjelasan *butir-butir* bekalan air yang akan dibekalkan
30. selain berbincang mengenai *butir-butir* yang dijanjikan, Hashim turut
31. James Wolfensohn (tengah) mendengar *kata-kata* Pengarah Urusan Tabung
32. WASHINGTON, Jumaat - *Negara-negara* Asia Tenggara yang digoncang dengan
33. Ramli berkata beliau dan *rakan-rakan* yang bertanding telah mendapat
34. sokongan memberangsangkan daripada *ahli-ahli* dewan dan jika dipilih,
35. By Ainuddin Dahlan SEPERTI di *tahun-tahun* yang lalu, ribuan rakyat
36. saya yang teraniaya hanya kerana *gara-gara* cara pembekalan air yang
37. Faktor ini terus menekan *saham-saham* terutamanya saham berwajaran
38. ata berlatar belakangkan keadaan yang lembap itu, *dana-dana* dikatakan
39. Ada stesen TV dari *awal-awal* lagi sudah senaraikan 10 syarikat pilihan
40. ceritanya entah *apa-apa*, skrip longgar, lakonan statik, pengarah tidak

41. "Dalam filem itu, Rahimah dan Rahman jadi *adik-adik* saya yang sangat
42. patan sektor korporat Malaysia tetapi belian pada *saat-saat* terakhir
43. Saya akan memandu *gila-gila* sikit. Begitulah kuat pengaruhnya.
44. hari ini. Esok, lusa mereka mungkin terlantar *tiba-tiba*," katanya. o
45. hidup tanpa bekalan ubat Pertubuhan *Bangsa-Bangsa* Bersatu (PBB).
46. menanti belas ikhsan dunia, si ibu dan *anak-anak* ini sudah putus asa.
47. Mereka tidak meminta yang *bukan-bukan* kecuali ubat. Ketua pediatrik
48. di hospital itu untuk membantu *kanak-kanak* daripada mati tanpa
49. membawa ubat ini untuk *kanak-kanak* yang lemah
50. khidmat *bank-bank* tempatan. Di bawah pengawasannya, *bank-bank*
51. ditawarkan oleh bank mengikut *syarat- syarat* yang berpatutan.
52. Obuchi memberitahu *Menteri-menteri* Luar Asean dan rakan sejawatan
53. Abdullah berkata, beliau nampak *rakan-rakan* dialog amat bersedia untuk
54. langkah yang diambil oleh *rakan-rakan* dialog terutama Amerika Syarikat
55. makanan ringan seperti *gula-gula* yang dapat dilihat di pasar raya
56. Kabinet sebagai Menteri *Tugas-tugas* Khas
57. hujah peguam bahawa bahawa *saksi-saksi* tidak akan diganggu apabila
58. "*Saksi-saksi* pendakwa kemungkinan boleh diganggu walaupun pada
59. Kes ini seharusnya dianggap sama dengan *kes-kes* yang lain dan tidak
60. aimana kamu boleh berkata begitu, terdapat banyak *kes-kes* yang lain
61. Juga adakah *kesan-kesan* seminal (air mani) daripada
62. habis dikotori dengan *tangan- tangan* kita dan adakah
63. Sampai bila, sampai *tasik- tasik* itu kering kontang
64. mberita di Mesyuarat Asia-Eropah (ASEM) di London *baru- baru* ini,
65. Namun apapun masalah dan penafian *kata-kata*, masyarakat haruslah peka
66. tetapi biarlah *rebung-rebung* yang ada dilentur dulu,
67. penjenayah ditangkap *hidup-hidup*, 6,205 daripadanya warga negara
68. Yang lebih menyedihkan lagi ialah *soal-soal* kematian. Bukan itu saja,
69. Dengan itu `ugutan' itu tidak timbul kecuali ada *pihak-pihak* tertentu
70. Dokumen Hakmilik Keluaran *tanah-tanah* seharusnya dikembalikan kepada
71. rombongan demi rombongan *wakil-wakil* pembeli mendatangi mereka?
72. Oleh itu, saya berharap *unsur-unsur* `mengugut' `menekan', `memberi
73. omi Melaka, Sulaiman Hermein, bahawa taklimat itu *sia- sia*.
74. yang dihadiri Mohd Zin *sia-sia* kerana Setiausaha Kerajaan
75. pelabur yang masih *ragu-ragu* akan keadaan pasaran saham negara tetapi
76. Menurutnya, Gates sangat seronok dengan *idea-idea* ini, malah
77. Ia mendapati bahawa *sekolah-sekolah* boleh menarik remaja daripada
78. Notebook untuk membuat *kerja-kerja* sekolah mereka.
79. digunakan dengan menentu dan *bukti-bukti* corroborative jarang diterima
80. Dr Maznah berkata wanita yang *rata-rata* menggunakan mahkamah
81. penonton kerana lebih kelihatan *suka-suka* saja.
82. Mereka lebih banyak berceloteh yang *bukan-bukan*.
83. Malah, Skrin juga kenalah perkenal *filem-filem* agung seperti Pather
84. "Fikirlah *baik-baik*, *dua-dua* berlawan .
85. Tidak usah bimbanglah, saya tidak akan *marah-marah*.
86. terlalu ramai yang datang tu, *faham-faham*, *macam-macam* benda yang

87. beliau (Abu Bakar) menang," kata *sumber-sumber* itu yang mentafsirkan
88. *Calon-calon* yang disebut-sebut sebagai berkemungkinan akan
89. Tentu saja tidak, malah *badan-badan* dunia yang mengelarkan diri mereka
90. penjaga keamanan seperti *masa-masa* yang lalu tentu akan mengadakan
91. sekolah serta menyumbangkan *idea-idea* yang berguna untuk dimanfaatkan
92. Walaupun *kadang-kadang* terdapat harimau ganas, ia mesti bersebab
93. aga, pelabur juga mengambil sikap berhati-hati di *tengah- tengah*
94. sebut oleh senator tadi, mungkin sebilangan kecil *bapa-bapa* yang kaki
95. saya bukan mahu menyebelahi *mana-mana* pihak tetapi apa yang pernah
96. bangsa sendiri hingga kita lupa di *celah-celah* kehodohan yang
97. selain memaparkan *cerita-cerita* pahit yang benar.
98. San Francisco dihadiahkan pula *ole-ole* kecil cenderahati daripada
99. Barangkali *ole-ole* cenderahati begini belum lagi popular di Malaysia
100. dalam proses penyediaan dan kelulusan terutama *agensi-agensi*
101. "Kita masih meneliti *cabang-cabang* peruntukan yang harus
102. malah ia tidak ubah seperti *budak-budak* yang bergaduh tanpa
103. *Ura-ura* untuk kembali berlatih bersama juga mendapat samb
104. tentang penyebutan *morfem- morfem* atau *kata-kata* dalam ujaran,
105. penyebutan sebenar *morfem- morfem* atau *kata-kata* dalam ujaran,
106. tumpuan ialah mengenal pasti *rumus-rumus* fonologi yang
107. di peringkat output, kecuali terdapat *bukti-bukti* yang boleh
108. perlu lagi kita berdolak-dalih dengan *dalil-dalil* yang tidak
109. dikomunikasikan itu harus mempunyai *sifat-sifat* yang jelas dan
110. kognitif bagi individu adalah *set-set* fakta yang dapat
111. membingkiskan *set-set* andaian yang mudah diproses oleh
112. bukan sahaja terhad kepada pendidikan peringkat *kanak-kanak*
113. atiri bayi akan terkena penyakit sawan. *Kanak-kanak* (a) jangan
114. *Kanak-kanak* (a) jangan main *sembunyi-sembunyi* pada waktu senja
115. teori relevans ialah memberi ruang kepada *bahasa- bahasa*
116. pola umum penyebaran *vokal- vokal* ini dapat dikaitkan dengan
117. menilai sama ada proses harmoni vokal itu *benar-benar* wujud
118. mengenal pasti konsep dan *ciri-ciri* perilaku harmoni vokal dari
119. gejala fonologi ini, yang berikut disenaraikan *contoh-contoh* ka
120. Kedua, *vokal-vokal* yang berharmoni itu harus mempunyai ciri
121. vokal tengah [e, o], seperti yang terlihat pada *contoh- contoh*
122. Dalam analisis fonemik, *bunyi-bunyi* yang didapati mempunyai
123. terbaik harus berlandaskan *asas-asas* teori fonologi yang kukuh.
124. dikemukakan oleh beliau ialah berdasarkan *ciri-ciri* sejagat
125. *Rangkap-rangkap* vokal ini terdiri daripada gabungan voka
126. *Hitam-hitam* tahi minyak dimakan juga, *putih- putih* hempas ke
127. *Buruk-buruk* kayu gaharu, dibakar berbau juga.
128. jaran biasa, *kanak-kanak* ini juga diajar dengan *kerja-kerja*
129. Malam itu juga aku mengumpul segala *surat-surat* Amar kepadaku
130. termasuklah hadiah dan *gambar-gambar* pemberiannya.
131. kembali *baris-baris* ayat yang penuh dengan *janji-janji* manisnya
132. atap kembali baris-baris ayat yang penuh dengan janji-janji

133. Apatah lagi dengan menggunakan sentuhan *warna-warna* alam yang
134. Apatah lagi *unsur-unsur* tradisional turut diserapkan menerusi
135. Ruang makan utama disediakan di *tengah-tengah* ruang yang
136. "Anis, tak sanggup dengar *gosip-gosip*, ia amat menakutkan
137. Malaysia dan menghasilkan *produk-produk* paling bermutu untuk
138. Maklumlah untuk menghiburkan *orang-orang* kita yang sedang
139. Keasidan ini melindungi kulit daripada *kuman-kuman* merbahaya
140. pH meningkat, kulit mudah dijangkiti oleh *kuman-kuman*
141. sepatutnya menanggalkan kotoran, *sel-sel* mati, *kuman-kuman*,
142. wangi sikit, saya nak jumpa orang, mestilah *elok-elok*
143. "Ha, ni balik *tengah-tengah* malam ni, mesti jumpa
144. balik *jauh-jauh* malam ,tak hubungi rumah beritahu *awal-awal*,
145. Tak payah *keluar-keluar* rumah,"ujar Masri bagaikan merajuk.
146. sakit jantung, asma kronik, kesan sampingan *ubat-ubat* lain
147. Bagaimanapun kesan *ubat-ubat* ini agak kurang berbanding dengan
148. senaman pasif bagi meningkatkan keupayaan *otot-otot*.
149. perlu menukar *komponen-komponen* tertentu untuk meningkatkan
150. Mutu *buku-buku* impot ini memuaskan.
151. Ia meliputi *aspek-aspek* perkakasan, perisian dan penyelesaian
152. Bagaimana mengakses *laman-laman* web yang menarik di Malaysia?
153. Satu lagi bahagian menyenaraikan *laman-laman* web di Malaysia.
154. Tahukah anda yang *akhbar-akhbar* harian terkemuka di Malaysia
155. The Malaysia Internet Book juga memaparkan *topik-topik* yang
156. Ia membariskan pencalonan *nama-nama* besar dan *filem- filem*
157. Warga kota yang dikelilingi *klinik-klinik* atau hospital pakar
158. *benar-benar* akan tunjukkan kepada mereka *jalan-jalan* Kami...
159. Dari kenikmatan itu Allah menurunkan dua jenis *manfaat-manfaat*
160. Allah menjauhkan $$57 *mafsadah-mafsadah* dan berbagai mudharat.
161. kita juga sering melihat dan mendengar *ceramah-ceramah* agama,
162. sekadar memenuhi masjid dan *majlis-majlis* agama.
163. Betapa indah pula *dinding-dinding* rumah dihias dengan ayat-
164. Kadangkala mangsa terdiri daripada *gadis-gadis* bawah umur
165. kain putih dan memanggil *nama-nama* makhluk halus yang dapat
166. ketenangan dan kesunyian Empangan Batu dengan *mitos-mitos*nya
167. surat pekeliling menetapkan *syarat-syarat* tertentu.
168. makhluk pilihan antara *makhluk-makhluk* Allah.
169. ilmu fardu kifayah pula merangkumi *ilmu-ilmu* kemahiran hidup,
170. mesti bermula dengan menghayati *prinsip-prinsip* ilmu nakli yang
171. Sehubungan itu, dapat dibuat kesimpulan bahawa *ekonomi-ekonomi*
172. Sebaliknya Vietnam dan *askar-askar* berpakaian pijama hitam yang
173. perniagaan yang dijaja di *tepi-tepi* jalan, di bawah pokok, di
174. di bawah pokok, di *lorong-lorong* belakang sudah berakhir.
175. Sebagai pewaris risalah kepada perjuangan *nabi-nabi* ulama harus
176. kita sudah jauh ketinggalan dalam *bidang-bidang* fardu kifayah
177. melanggar *lunas-lunas* yang ditetapkan oleh syarak.

Appendix G

Total reduplication in SM (monosyllabic roots)

1. puas dan sempurna - tanpa *was-was*.
2. Pak Malau *was-was* sama ada HVD sudah bersedia dari segi pasukan ten
3. Saham kewangan di BSKL pula *was-was* selepas ada ura-ura menyatakan
4. By Zainuri Misfar Pix:- Adnan *was-was*.
5. tapi saya *was-was* adakah Gafim pertubuhan
6. tapi saya *was-was* kerana fungsi dan pemimpinnya hampir sama. Suara
7. a mereka sentiasa memilih Harakah tanpa ragu dan *was-was*.
8. ternyata lebih berdisiplin dengan berkurangnya *kes-kes* melanggar
9. dianggap sama dengan *kes-kes* yang lain dan tidak dianggap satu kes
10. berkata begitu, terdapat banyak *kes-kes* yang lain yang jaminannya
11. dengan baik kerana pihaknya tidak menerima aduan *kes-kes* yang serius
12. Dalam *kes-kes* yang melibatkan pinjaman pula, kelulusan muktamad
13. "*Kes-kes* seperti ini juga dikesan di kebanyakan kawasan
14. Sebaliknya mereka tumpukan kepada *kes-kes* tertentu yang ternampak
15. kegiatan jenayah semakin meningkat dan *kes-kes* yang dilaporkan dalam
16. statusnya lebih tinggi, malah dalam *kes-kes* tertentu peranannya
17. sepanjang masa dan tanpa *pam-pam* tersebut ia kan menyebabkan banjir
18. pernah berjanji untuk membahagikan *lot-lot* tanah kepada penduduk sejak
19. kkannya, ditambah 's' sahaja _ sepatutnya menjadi *lot-lot*.
20. ntuk melancarkan 1,000 unit hartanah, terutamanya *lot-lot* banglo dan
21. *Lot-lot* banglo serta kedai pejabat tiga dan empat tingkat
22. Dunia Ke-2 pedagang-pedagang Cina telah memiliki *lot-lot* tanah yang
23. an petani Melayu. Pada keseluruhannya saiz purata *lot-lot* tanah
24. Mereka memiliki 300 *lot-lot* yang
25. i jumlah keluasan 12354 ekar berbanding dengan 43 *lot-lot* kepunyaan
26. 5 buah rumah kedai dan 20 *lot-lot* industri kepunyaan Syarikat dengan

Appendix H

Partial reduplication in SM

1. bukan main lagi *bebudak* sekarang ni.
2. bukannya nak marah sangat dengan *bebudak* sekolah yang suka hisap rokok
3. bagi merujuk kepada *lelaki* atau wanita.
4. terjadi kerana dominasi *lelaki* dalam masyarakat kebapaan (pairarchal)
5. Adik-beradiknya *lelaki* belaka. Ada seorang lagi perempuan masih
6. bagi merujuk kepada *lelaki* atau wanita.
7. dikatakan terjadi kerana dominasi *lelaki* dalam masyarakat kebapaan (
8. merujuk kepada jantina *lelaki* atau perempuan.
9. merujuk kepada *lelaki* yang gagah dan berani,
10. contoh seperti *lelaki*, *kekuda*, dan *kekura*. Penggandaan jenis ini
11. dililitkan pada buaian,pada kayu *kekuda* buaian dan juga pada bahagian-
12. kategori lelaki menerusi acara *kekuda* pelana mata
13. ataupun dalam keadaan memasang *kekuda*. Untuk mengelakkan diri
14. jejarak, *kekisi*, cahaya *gegeli*, *gegebar*, dan *pepaku*
15. memakai *kekisi* pada sebuah rumah berusia tua di Kg. Seri Lemak,
16. di bawah: *rerambut*, *gegelang*, *cecair*, *pepejal* dan *kekaca* (bidang Kim
17. itu terdapat *rerambut* yang berfungsi seperti sesungut pada serangga
18. merah di kaki dengan *rerambut* darah yang memancar daripadanya.
19. berbentuk corong dan ada *rerambut*. Dalam badan periuk kera ini
20. diperbuat daripada *rerambut* yang terdapat pada batang pokok ijuk.
21. alui kerjasama pendidikan antara kedua-dua negara *tetangga*.
22. antara dua buah negara *tetangga* dalam usaha menyeragamkan istilah
23. enerimaan orang-orang kampung yang menjadi jiran *tetangga* Milah
24. kita, sibuk dengan perbualan kita tentang jiran *tetangga*.
25. rakan, jiran *tetangga*, saudara mara, dan *tetamu* istimewa.
26. campur tangan dan *jejari* mereka kerana wujud perimbangan kekuatan
27. Di rantau ini, tangan-tangan dan *jejari* AS bertambah jelas dan
28. *pepohon*. 2. berserat-serat. 3. *jejari*. 4. terapung-apung. 5. beramai-ramai. 6. *dedaun*.
29. Apabila *jejari* setiap ruang dipanjangkan seluruh cakerawala di
30. disebut *pepenjuru* (diagonals). Kesemuanya terdapat 26 *pepenjuru* di
31. Petak-petak pada *pepenjuru*- agi berwarna
32. berusaha dalam tempoh *sesuku* abad akan datang bagi mencapai empat
33. berusaha dalam tempoh *sesuku* abad akan datang bagi mencapai empat
34. Syed Bakar yang masih ada *sesuku* lagi isinya dan diserahkannya
35. terbakar, orang *sesuku* tidak mahu berkongsi dengan dia, gadis-gadis
36. Kejadian berlaku di sebuah *jejantas* di Mina, dekat sini, di mana
37. dari *jejantas* itu. India berkata, sekurang-kurangnya
38. Arab Saudi melebarkan *jejantas* sepanjang satu

39. dari sebuah *jejantas* ketika jemaah melontar jamrah pada hari
40. sekarang dalam contoh seperti *lelaki, kekuda*, dan *kekura*.
41. dengan bahagian *kekura* kaki. - Tendang di tengah-tengah
42. dengan bahagian *kekura* kaki. - Tendang di bahagian bawah
43. dalam kaki, luar kaki, dan *kekura* kaki untuk mengawal
44. Misalnya, papan *kekunci*, pemacu cakera, pangkalan data, dan
45. monitor, dan papan *kekunci* yang tersendiri dan biasanya digunakan
46. dibawah papan *kekunci*, dan menggunakan televisyen sebagai
47. *bebola* ibu jari) yang dibina di dalam papan *kekunci*.
48. racun untuk menghalang pembiakan *jejentik* dalam larutan baja.
49. rumah pula dibubuh ubat *jejentik*. Menurutnya, kerajaan berpendapat
50. tempat serta *tetamu* yang sedang ditemu bual.
51. Mohamed Rahmat adalah antara *tetamu* kehormat yang dijangka
52. Tun Omar Yoke Lin Ong sebagai *tetamu* khas untuk menghadiri acara
53. terik tidak menghalang *tetamu* dan peminat Ziana Zain, 30,
54. bermula jam 12 tengah hari, ramai *tetamu* datang lebih
55. hal yang sukar dilakukan dengan imej batu atau *pepohon*.
56. telah kusimpulkan *pepohon* setia di pangkal hatiku
57. dipenuhi dengan *pepohon* dan rimbunan hijau menyejukkan mata.
58. "Di bawah lindungan *pepohon* ru, dekah ketawa kerabat berjalan
59. Malah, sesetengah beranggapan *pepijat* (bug) itu yang berpeluang cerah
60. Ia adalah masalah `*pepijat* alaf' (Y2K) yang dijangka menyukarkan
61. yang dibincangkan, dan inilah yang dinamakan *pepijat* alaf atau
62. lnilah *pepijat* alaf. Pepijat ini akan menjadi bom yang tidak diu
63. Sebagai percubaan, kayu *kekabu* yang ringan digunakan.
64. kanak-kanak yang diperbuat daripada kayu *kekabu*, pulau, jelutong,
65. kaki palsu yang diperbuat daripada kayu *kekabu* adalah yang paling
66. satu lagi dibuat daripada kayu *kekabu* untuk berjalan-jalan.
67. komputer, papan *kekunci* dan bahagian rangka komputer.
68. hanya duduk menekan papan *kekunci* komputer.
69. pengeluaran injap pintu *kekunci* paip air, menetapkan sasaran untuk
70. hanya perlu menekan *kekunci* komputer untuk membuka fail-fail yang
71. Kesan *Lelabah*. Tidak ramai yang tahu bahawa vena *lelabah* tida
72. Tidak ramai yang tahu bahawa vena *lelabah* tidak sama dengan vena
73. Istilah perubatannya ialah naevus *lelabah*, dan ia menyerang kira-kira
74. Namun bagi kehamilan seterusnya, vena *lelabah* ini `
75. lubang hidung dan terus ke *peparu*.
76. dan kambing juga bernafas melalui *peparu*. Belalang pula bernafas
77. A. Jantung. B. Hati. C. *Peparu*. D. Ginjal.
78. awah menunjukkan bahagian A. Liang pernafasan. B. *Peparu*. C. Insang.
79. bergerak menjauhi fokus melalui media *pepejal* dan separa *cecair* di
80. sistem pengurusan sisa *pepejal*.
81. dikultur dalam medium *pepejal* agar-agar.
82. sehinggalah ibu memberikannya makan *pepejal*.
83. iaitu serbuk dan *cecair*, yang mana kedua-duanya saling
84. baja organik bio-aktif bentuk *cecair*, yang dicampur

85. yang mengangkut Gas Asli *Cecair* (LNG) serta terbabit

86. Pukal *cecair* mengekalkan kedudukan sebagai komposisi kargo

87. tercekik - Anggur. - *Kekacang*. - Hotdog. - Air batu. 7

88. engaja oleh tupai yang menyembunyikan biji benih *kekacang* dan

89. terjadi kerana rata-rata penduduknya menjadikan *kekacang* dan sayuran

90. bahan lain terutama *kekacang* menggantikan makanan berkolesterol tinggi

91. terdapat sepasang *sesungut* dan sepasang mata.

92. sepasang *sesungut* dan abdomen. B. kepala, sepasang *sesungut*

93. epasang *sesungut* dan abdomen. B. kepala, sepasang *sesungut* dan

94. - Tisu. - *Bebola* kapas. - Air suam, losyen bayi atau tuala basah

95. minyak mandian, minyak urut, losyen, loofah atau *bebola* kayu,

96. burger, sosej, friter, *bebola* ayam, belogna, koktel

97. yang berbentuk *bebola* (kebanyakannya) selain peluru

98. Tiba-tiba terterjah *gegendang* telinganya oleh satu

99. akrab di *gegendang* telinga Tok Anjang.

100. kejam yang tertingkah-tingkah memecah *gegendang* telingaku pada

101. pembinaan bangunan Unsur-unsur: gelombang, *gegendang*, getaran

102. kesihatan manusia, misalnya *gegendang* telinga akan terganggu

Appendix I

Chiming reduplication in SM

1. Apot memerhatikan pengganas *mundar-mandir*.
2. Ketika sedang *mundar-mandir* dengan pakaian tidurnya, Arbuckle
3. Selepas puas *mundar-mandir*, saya singgah di pawagam Rex dekat
4. Sungai Paxaco yang *bengkang-bengkok* di pinggir tanah kabilah ini.
5. ganjil, dan *bengkang-bengkok*. Pada rumput yang berdebu berhampiran
6. menggunakan jalan lama yang *bengkang-bengkok* lagi merbahaya.
7. kita akan mendapati keadaan *huru-hara* cuaca melanda muka bumi dengan
8. keluarga yang mereka bina nanti *huru-hara* jadinya.
9. keputusan tersebut, hidupnya *huru-hara*. Harta dan tanah banyak yang
10. Sebenarnya Malaysia berada dalam keadaan *serba-salah* kerana dalam
11. keadaan dan menimbulkan rasa *serba-salah* atau implikasi yang tidak
12. Dengan itu menjadikan kita *serba-salah* kerana mungkin e-mel yang
13. dipersendakan melalui sifat, *tindak-tanduk* dan pemikiran/pendapatnya.
14. yang diberi sifat serta *tindak-tanduk* tertentu untuk diherotkan,
15. membawa penonton menyaksi *tindak-tanduk* lima watak utamanya mengej
16. apabila *cucu-cicit* kita membaca naskhah hari
17. Kuala Lumpur dan *cucu-cicit* setahun sekali kembali menjengah, saya
18. raja Suran dan *cucu-cicit* raja Iskandar Dzulkarnain.
19. memiliki anak yang cemerlang dalam *serba-serbi* lapangan.
20. membasuh pakaian, memasak dan *serba-serbi* tanpa pembantu rumah.
21. *Serba-serbi* tidak selesa. Duduk tidak kena, berdiri lagi par
22. *TEKA-TEKI* tayangan umum dua filem arahan U-Wei Saari, Jogho
23. 31 Disember 1998, pasti menjawab *teka-teki* mengenai prestasi ASN untuk
24. *Teka-teki* melengahkan perlantikan Timbalan.
25. gajak b. *gerak-geri* Takrif: *gerak-geri*, tingkah laku, sikap dan
26. 3. Ketahui *gerak-geri* mereka. Apabila kamu sudah lama bercampur gaul
27. Menurutnya, memandangkan *gerak-geri* pekerjanya dicurigai penduduk
28. tropika yang menghijau dan *bukit-bukau* terdapat kira- kira 20,000
29. lurah pergunungan serta *bukit-bukau* menjadi tumpuan laluan masyarakat
30. tiba-tiba muncul dari kawasan *bukit-bukau* Afghanistan, sebenarnya
31. batu pada *bukit-bukau* dan *gunung-ganang*, iaitu dikatakan bahawa dunia
32. Fairuz melintasi lautan, *gunung-ganang*, beberapa gurun akhirnya
33. Sebaliknya, *gunung-ganang*, tempat kediaman Tuhan mereka.
34. pasal apa? Hendak bersatu macam mana kalau hidup *nafsu-nafsi*,
35. sendiri-sendiri, penuh dengan *nafsu-nafsi*.
36. Alam semesta *gelap-gelita*, tiada manusia global yang dapat
37. Restoran jadi gelap. *Gelap-gelita*.
38. keadaannya berada dalam *gelap-gelita* yang sekali-kali tidak dapat
39. Bagi yang boleh berfikir apakah tindakan *gopoh-gapah* dia ini, serangan
40. Selalu katakan kepadanya bahawa *gopoh-gapah* ini datangnya daripada

41. Kedengaran *riuh-rendah* di halaman itu.
42. kawasan ini menjadi pesta ria yang *riuh-rendah* dengan pemandangan
43. Semua teman *riuh-rendah* dengan penuh rasa kesal memikirkan saya
44. Di sini juga tidak ada sungai untuk *mandi-manda*.
45. hutan serta berkelah dan *mandi-manda* di Lata Berkoh dan Lubok Simpon,
46. Bagi pengunjung yang ingin beriadah sambil *mandi-manda* boleh
47. kehidupan keluarga kita dan *suku-sakat* bangsa kita.
48. kepentingan kekeluargaan dan *suku-sakat* kalangan terpilih itu.
49. tapah cukup berbisa dengan *suku-sakat* keturunannya.
50. amat sibuk dengan kenderaan *lalu-lalang*, sepanjang hari ini nampak
51. Malah ada kalanya kambing *lalu-lalang* di dalam kelas tanpa perlu
52. Contoh: *lalu-lalang*. gunung, gelap, riuh, sayur, kuih, rumput, sim
53. keluarga yang *porak-peranda* dan persekitaran yang buruk, maka ia
54. Namun realitinya rumah tangga mereka *porak-peranda*.
55. Dan kini, berhadapan dengan *gegak-gempita* gejala sosial,
56. Pada ketika itu seluruh negara sedang *gegak-gempita* menanyakan
57. Sebahagian besar rekod dijual kepada *muda-mudi*.
58. sehinggakan menjadi siulan *muda-mudi* di sana.
59. permainan yang digemari oleh *muda-mudi* khususnya.
60. Perkahwinan di kalangan *muda-mudi* Orang Asli berlaku di kalangan suku
61. seruan terhadap nama-nama '*dewa-dewi* Hindu'.
62. Selanjutnya, penglibatan unsur '*dewa-dewi* Hindu' adalah untuk menakut
63. Oleh itu pemujaan *dewa-dewi* minor seperti Dewa Muniandi, Dewa
64. beliau bahawa keadaan *kucar-kacir* masyarakat Melayu meruntuhkan
65. yang tidak terancang serta *kucar-kacir*, sama ada ala diktator atau
66. Dalam suasana yang penuh *kucar-kacir*, tubuh Hussain pula dipanah.
67. Seni mempertahankan diri *turun-temurun* yang diwarisi pahlawan Melayu
68. mempunyai wang baik berupa warisan *turun-temurun* mahupun kekayaan
69. lupa budaya dan adat yang kita warisi sejak *turun-temurun*.
70. Pusat *Beli-Belah* Tangs, Edmonds bertugas sebagai Pengurus Besar
71. Black, dilancar di pusat *beli-belah* Tangs Ogos ini, sesuai lelaki dan
72. Dijual di semua pusat *beli-belah* sejak Jun lalu. - NZ.
73. Dijual di semua pusat *beli-belah*, UVLMC 50 g berharga

Appendix J

Rhyming reduplication in SM

1. perunding yang memahami *selok-belok* perniagaan di
2. Bukankah ketua kita tu mahir dengan *selok-belok* di bandar itu?
3. Jakun cukup masak dengan *selok-belok* sosial masyarakat
4. boleh membunuh *anak-pinak* HIV. Ini satu ancaman besar.
5. mendapati masih ada *saki-baki* keguguran dalam rahim menerusi
6. Setelah pengesahan dibuat (*saki-baki* keguguran dalam rahim)
7. Saya percaya masih ada *saki-baki* yang 'teringin'
8. keluarga, kenduri, *gotong-royong* dan lain-lain lagi tempat atau
9. kita melihat tradisi *gotong-royong* umpamanya hidup dengan subur.
10. Demang Lebar Daun. *Anak-cucu* Demang Lebar Daun sedia
11. hamba dan patuh kepada *anak-cucu* Sang Sapurba
12. Tanya-menanyakan *asal-usul* sehingga dapat diketahui pertemuan darah,
13. wang dan kekayaan, manusia lupa *asal-usul* dan saudara.
14. malah mempelajari *adat-resam* antara satu sama lain. Di
15. mencetuskan situasi dalam *adat-resam* Melayu iaitu
16. Elak tetamu *cirit-birit*. DAGING beku atau daging import biasanya
17. ia boleh menyebab tetamu *cirit-birit*. Untuk mengatasi masalah ini,
18. *Sebab-musibat* seluruhnya berada dalam kepala
19. jadikan masalah *remeh-temeh* sebagai kengkangan guru untuk
20. terhad kepada persoalan yang *remeh-temeh* semata-mata.
21. kerjanya pula tidak pula *remeh-temeh* dan banyak.
22. Setelah berbincang dengan *saudara-mara* dan keluarga mereka
23. Salman al-Farisi, Buraidah dan *saudara-mara* dari kalangan
24. berbaik-baik dengan *saudara-mara* kedua-dua belah pihak
25. Fikirannya kembali *orak-arik*.
26. mendapatkan maklumat secara *serta-merta*.
27. Kini kami diarah keluar dengan *serta-merta*.
28. dan melaksanakan tindakan *serta-merta* oleh semua agensi, besar dan
29. pelanggan terhadap perbuatan dan *tingkah-laku* syarikat
30. terhadap perbuatan dan *tingkah-laku* kami sebagai sebuah perniagaan.
31. Ada yang guna bahasa *tunggang-langgang*, syok sendiri, kurang adil
32. Bukan saja bahasa *tunggang-langgang* ketika berceloteh, malah sewaktu
33. kerana hanya apabila "*akar-umbi*" asas pemikiran itu
34. miskin dan negara *kacau-bilau*.
35. ''Hari ini kewangan negara-negara ini *kacau-bilau* dan mereka rugi
36. jika negara kita menjadi *kacau-bilau* suatu hari nanti.

Appendix K

Multiple prefixes reduplication

1. termasuk media luar negara, *memperbesar-besarkan* isu ini."
2. berkenaan dengan *memperbesar-besarkan* isu penyebaran bahasa Melayu
3. jika tidak mereka akan *mempermain-mainkan* ekonomi sesebuah negara itu.
4. kita sendiri yang *mempermain-mainkan* aqidah. Kementerian dan
5. dengan maksud *mempermain-mainkan* penonton (Ku Seman Ku Hussain, 1996:
6. memperalatkan agama, *mempermain-mainkan* ayat-ayat Allah, kaki judi,
7. menggunakan kekuatan kekayaan untuk *memperkecil-kecilkan* ekonomi
8. mengkritik, menegur, *memperkecil-kecilkan*, menghina, mengancam,
9. sombong dan suka *merendah-rendahkan* atau *memperkecil-kecilkan* orang
10. "Bagaimanakah *mempersia-siakan* amanah itu, wahai Rasulullah?"
11. pelajaran dan tidak *mempersia-siakan* masa terluang. Bagi Ahmad
12. diingini di samping tidak *mempersia-siakan* pelaburan besar tersebut.
13. Insya-Allah, Allah tidak *mempersia-siakan* orang yang benar-benar
14. dalam Islam ditinggalkan, *dipermain-mainkan* dan diperlekehkan.
15. umatnya hari ini boleh *dipermain-mainkan* dan diperlakukan oleh
16. pelanggan kami tidak lagi *dipermain-mainkan* oleh mereka yang menju
17. daripada diserang dan *dipermain-mainkan* oleh kekuasaan individu kaya
18. dan pendapat tanpa *dipermain-mainkan* oleh orang lain.
19. dia senang diherdik dan *dipermain-mainkan*. Oleh kerana itu dia sering
20. adalah kebenaran yang *dipermain-mainkan* oleh orang yang menyalinnya.
21. Sutomo berasa semakin *dipermain-mainkan*. "Tidaklah."
22. negara kini sudah berani *mempersenda-senda* (istihza') hukum-hukum

Appendix L

Single prefix reduplicated words

1. Kamal Shukri Abdullah Sani, turut *mengalu-alukan* arahan Menteri
2. berkata pihaknya *mengalu-alukan* langkah memperluaskan penggunaan
3. Dalam mengelamun *mengingat-ingat* sejarah Babylon, tidak sedar sudah
4. bijak berbicara dan *mengungkit-ungkit* zaman Rasulullah tanpa
5. kerajaan mesti dibuat secara *berhati-hati* dan berpilih supaya hanya
6. masyarakat *berlumba-lumba* untuk menghentam
7. Kementerian Pendidikan melahirkan *beribu-ribu* guru muzik dan
8. instrumen lain seperti waran, *bersama-sama* bon yang dilanggani.
9. satu serta boleh ditembak *berturut-turut* dengan pantas.
10. kepada tanah air tercinta *berkobar-kobar* menyebabkan saya melepaskan
11. Ini adalah kali kedua *berturut-turut* Dr Mahathir menghadiri dialog itu
12. Pengunjung ekspo memang *berebut-rebut* membeli perisian cetak rompak
13. teguh berikutan sokongan yang *berpilih-pilih* oleh pelabur institusi
14. kepelbagaian definisi yang *berbeza-beza*.
15. bangsa yang bertamadun dan bermaruah sejak *berkurun-kurun*
16. belajar secara sendirian ataupun *bersama-sama*.
17. lelangit lembut dan keras secara *berturut-turut*.
18. Saya tidak pernah *bermewah-mewah* dengan wang.
19. mengajar mereka untuk *bermewah-mewah* dengan wang.
20. pernah diabaikan selama *bertahun-tahun* sebelum ini.
21. kelingking ini begitu popular dan *berebut-rebut* dibeli,
22. rambut lebih bermaya dan *bersinar-sinar*.
23. Datuk Punch Gunalan yang sudah pun *berjinak-jinak* dengan dunia
24. negara begitu bersemangat tinggi dan *berkobar-kobar* untuk membawa
25. menunjukkan sikap *berbalah-balah* Amerika Syarikat
26. negara ini dilempar dengan *berbagai-bagai* krisis kepimpinan yang
27. menganggarkan *beribu-ribu* lagi mungkin telah mengadakan hubungan seks
28. Di pasaran terdapat *berpuluh-puluh* buku dari beratus hingga beribu
29. kami melakukannya *berkali-kali* dari malam hingga siang.
30. Carlo terpaksa menguruskan *beratus-ratus* orang pekerja,
31. Ada juga yang terpaksa mengambil masa *bertahun-tahun* untuk berlakon
32. Mariana secara *beransur-ansur* dengan kebaikan.
33. Adiknya, Mariana suka *berpoya-poya* dan tidak mempunyai pegangan hidup
34. Liga Perdana bagi musim kedua *berturut-turut* yang kian menepati
35. Secara *berperingkat-peringkat*, kaum India di negara
36. pasukan negara kita masih *terkial-kial* mencari tempat untuk melayakkan
37. FAM tidak perlu *terburu-buru* membubarkan pasukan kebangsaan kerana
38. perlu dibahaskan secara *tergesa-gesa*.
39. tahap serius dan kita *tertanya-tanya*, mengapa ia berleluasa dan sejauh
40. supaya tidak *terburu-buru* meningkatkan pengeluaran masing-masing.

41. "Kita tidak perlu *terikut-ik*ut untuk menghasilkan video
42. Saya masih *tertanya-tanya*, kenapa keadaan ini harus terjadi di negara
43. lidah *terjelir-jelir*, mata terpusing- pusing.
44. Jangan pula kau *terbawa-bawa* sehingga ke alam rumahtangga.
45. daripada membiarkan kau *tertanya-tanya*.
46. Lin *tersengih-sengih* mohon simpati.
47. As menekup mukanya lalu menangis *tersedu-sedu*.
48. Kenapa tidak dulu?" suaraku *tertahan-tahan*.
49. Dengan *terpinga-pinga*, Boboi menjerit,"
50. In *tercari-cari* di mana suaminya pergi.
51. peminat bagai *tercari-cari* Malek.
52. Ada yang menganggap Malek bersara secara *tergesa-gesa*.
53. Norjinah terus masuk tanpa *teragak-agak*.
54. masih *terdengar-dengar* tengkingan Puan Jah dua tahun lalu
55. Jika fikiran anda *terlayang-layang*, cuba kembalikan tumpuan anda
56. saya tidak *teragak-agak* untuk menambah peruntukan,'' katanya.
57. jadual penerbangan kerap *tertangguh-tangguh* tanpa kepastian.
58. di negara ini supaya jangan *teragak-agak* untuk menutup lapangan
59. pesisir pantai sambil *terbuai-buai* dilambung ombak,
60. beliau terpaksa berjalan *terhincut-hincut* dalam padang yang
61. ini media asing cuba *membesar-besarkan* peristiwa ini dengan
62. peristiwa ini dengan *memburuk-burukkan* Kerajaan Malaysia.
63. media Barat yang *memburuk-burukkan* negara ini dilakukan golongan
64. Tuan Khatijah *mengeleng-gelengkan* kepalanya tanda tidak setuju.
65. mereka mahu *menggembar-gemburkan* penganjuran majlis tersebut
66. masalah 'double standard' dan *membanding-bandingkan* karya antara
67. cara ia *membahagi-bahagikan* bantuan kecemasan makanan.

Appendix M

Voiceless obstruent initial roots

1. keluar tingkap pecah, *menimbang-nimbang* apa lagi yang boleh dijualnya
2. para sahabat nabi. Cuba juga *mengorek-ngorek* ingatan
3. Baginda memaksudkan dari tangan itu ialah *memukul-mukul* dada dan pipi,
4. ombak kuat yang sedang *memukul-mukul* pantai.
5. Sambil bermain dam, dia *memukul-mukul* meja dan dari situ juga dia
6. mendidik anak-anak dengan cara *menakut-nakutkan* adalah tidak
7. adalah dengan *menakut-nakutkan*. Waktu senja adalah waktu yang paling
8. Mesej untuk *menakut-nakutkan* anak-anak akan menjadi
9. masalah ini adalah dengan *menakut-nakutkan* mereka kerana dengan cara
10. Tegahan yang dilakukan akan dapat *menakut-nakutkan* pendengar dan
11. seksyen tersebut tidak akan *menakut-nakutkan* orang ramai daripada
12. meluka hati mereka, *menakut-nakutkan* dan menghina mereka.
13. bukti, informatif, tidak *menambah-nambah* maklumat, pastikan
14. Dia *memicit-micit* kepalanya. Hairan, bagaimana kaftan yang dijemur
15. Ariw mula *memicit-micit* kepala. Mouna mendewasakannya setelah
16. Si penulis mengangguk sambil *memicit-micit* dahi.
17. Tangannya masih *memicit-micit* dahi.
18. penulis masih duduk bersandar sambil tangannya *memicit-micit* dahi.
19. lalu *menepuk-nepuk* paha anaknya dengan tangan kirinya.
20. Musa *menepuk-nepuk* belakang Ramli ketika dia sampai ke atas dan ber
21. Esther cuba menenangkan Ailina sambil tangannya *menepuk-nepuk* lembut
22. Yang kecil melonjak-lonjak, yang besar *menepuk-nepuk*.
23. Akhirnya dia *menepuk-nepuk* bahu Balkis yang hampir terlena itu.
24. air besar atau kecil dia akan *menepuk-nepuk* punggungnya. Apabila haus
25. Akhirnya dia *menepuk-nepuk* bahu Balkis yang hampir terlena itu.
26. Malam terakhir itu dia mimpikan isterinya *memanggil-manggil* namanya
27. Suara itulah yang sering *memanggil-manggil* Jeman pulang setiap kali
28. dia terus *memanggil-manggil* nama Ramli beberapa kali.
29. mereka anak beranak menangis sambil *memanggil-manggil* nama pakcik
30. pintu rumahnya dan menjerit sekuat hati *memanggil-manggil,* penghuninya
31. anak beranak menangis sambil *memanggil-manggil* nama pakcik Amat.
32. Asma *memanggil-manggil* Fatimah, namun tidak disahut. Hancur luluh
33. Kanak-kanak kecil itu hanya mampu berteriak *memanggil-manggil*: "Mak
34. I sambil *mengetuk-ngetuk* tongkat panglimanya ke tapak tangannya.
35. Dia *mengetuk-ngetuk* tubuhku seolah-olah hendak menguji kekuatanku. S
36. Dia *mengetuk-ngetuk* tubuhku seolah-olah hendak menguji kekuatanku. S
37. ditujunya lantas menggoncang dan *mengetuk-ngetuk* objek.
38. Pantang melihat Hitam, anjing itu akan *mengejar-ngejarnya*.
39. tidak prihatin dan ingin *menunjuk-nunjuk* pada ketika banyak sektor
40. dianggap tidak prihatin dan ingin *menunjuk-nunjuk* pada ketika banyak

41. kepimpinan negara ingin *menunjuk-nunjuk* kepada negara luar menjadikan
42. ''Kami bukannya mahu *menunjuk-nunjuk* tetapi ingin menjadikan sebagai
43. berhutang dengan niat ingin *menunjuk-nunjuk* atau berlawan dengan
44. sikap kasihan # bukan sengaja diada-adakan # *menunjuk-nunjuk* di
45. limi orang lain. Ilmu yang dipelajari bukan untuk *menunjuk-nunjuk*,
46. selalunya suka hendak *menunjuk-nunjuk*, suka nak berlawan-lawan, suka
47. beberapa orang senior dan *menunjuk-nunjuk* ke arahnya.
48. Seperti mahu *menunjuk-nunjuk*. Demi Allah, tidak sebesar kuku pun ada
49. patut bersyukur kerana dengan *menunjuk-nunjuk*, syaitan berpeluang
50. cuba mengatasi masalah ini dengan *menunjuk-nunjuk* diri, mengada-adakan
51. Kanak-kanak sering memperlihatkan kelakuan *menunjuk-nunjuk* ini atau
52. kadar perhiasan apa lagi berniat untuk *menunjuk-nunjuk* maka mereka
53. kamu bersikap *menunjuk-nunjuk* dan dan serba tahu.
54. Sikap masyarakat yang mementingkan "budaya *menunjuk-nunjuk*" dengan
55. burung yang suka *menunjuk-nunjuk*? Walau apa-apa pun, merak pongsu
56. Seolah ingin *menunjuk-nunjuk*. Itukan sikap remaja sekarang?
57. kadar perhiasan apa lagi berniat untuk *menunjuk-nunjuk* maka mereka ini
58. bukan tujuan untuk *menunjuk-nunjuk* atau bersifat riak.
59. beribadah siang dan malam untuk *menunjuk-nunjuk* kononnya dialah orang
60. bunuh sebab biadap mulutnya *mengata-ngata* negeri Melaka di bawah
61. Mengumpat Mengumpat ialah amalan "*mengata-ngata*" rang.
62. Amalan *mengata-ngata* akan meletihkan orang yang mendengarnya.
63. Buku sama seperti mereka pergi *meninjau-ninjau* barangan di pasaraya
64. dari arah jalan dan *meninjau-ninjau* keadaan di ruang rumah itu.
65. Dalam pada itu kita *meninjau-ninjau* sejauh mana karya sastera negara
66. Lelaki yang kehilangan kaftan itu bangun, *meninjau-ninjau* sekitar
67. Rumah besar Tuk Cu sunyi sepi. Juah *meninjau-ninjau* ke arah pelantar
68. Sambil merehatkan badan, Ismail *meninjau-ninjau* kawasan hutan di
69. Puas lelaki itu *meninjau-ninjau* di luar jendela namun tiada siapa
70. Sambil langkah diperlahankan, Sabariah *meninjau-ninjau* di
71. Dalam esak tangisnya itulah Ahamd *meninjau-ninjau*, mencari-cari
72. *Meninjau-ninjau* beberapa ketika, imam nampak si suami sedang
73. Bukit Safar ke Bukit Marwah sambil *meninjau-ninjau* punca air.

Appendix N

Word list of nasal and voiceless obstruent clusters in three of Malay dialects

a) Nasal and voiceless obstruent root-internally

SM	Perak	Kelantan	Negeri Sembilan
kampoŋ	kampoŋ	kapoŋ	kapoŋ
sampan	sampan	sapɛ̃	sapan
sampai	sampai	sapɛ	sapa
kuntum	kuntum	kutum	kutom
kelantan	kelantan	kelatɛ̃	kelatan
teŋkat	teŋkat	tekaʔ	tekaʔ
təmpat	təmpat	təpaʔ	təpa
tampar	tampar	tapa:	tapo
bantal	bantal	bata:	bata
sampul	sampul	sapo	sapo
ʤəmpot	ʤəmpot	ʤəpoʔ	ʤaput
gantoŋ	gantoŋ	gatoŋ	gatoŋ
pintu	pintu	pitu	pitu
simpan	simpan	sipɛ̃	sipan
toŋkat	toŋkat	tokaʔ	tokaʔ
səmpit	səmpit	səpeʔ	sopeʔ
taŋkap	taŋkap	takaʔ	takaʔ
guntiŋ	guntiŋ	gutiŋ	guteŋ
sampah	sampah	sapoh	sapah
hanciŋ	hanciŋ	haciŋ	aceŋ
kəlonsoŋ	kəlonsoŋ	kəlosoŋ	kəlosoŋ
pantaŋ	pantaŋ	pataŋ	pataŋ
təntu	təntu	tətu	tətu
pantai	pantai	patɛ	patay
ʔuntuŋ	ʔuntuŋ	ʔutoŋ	ʔutɔŋ
taŋkat	taŋkat	takaʔ	takaʔ
caŋkat	caŋkat	cakaʔ	cakaʔ

b) Nasal and voiced obstruent sequences at prefix-root juncture

SM	Perak	Kelantan	Negeri Sembilan
məmbagi	magi	mmbagi	mmagi
məmbasoh	masoh	mmbasoh	mmasoh
məmbəli	məli	mmbəli	mməli
məmbaca	maca	mmbaca	mmaca
məmbunoh	munoh	mmbunoh	mmunoh
məndapat	napat	nndapaʔ	mnapeʔ
məndukoŋ	nukoŋ	nndukoŋ	mnukoŋ
məndaki	naki	nndaki	mnaki
məndiri	niɣi	nndiɣi	mniɣi
məŋgosoʔ	ŋosoʔ	ŋŋgosoʔ	mŋosoʔ
məŋgunteŋ	ŋunteŋ	ŋŋgunteŋ	mŋunteŋ
məŋgali	ŋali	ŋŋgali	mŋali
məŋgores	ŋores	ŋŋgores	mŋores
məɲdʒawab	ɲawab	ɲɲdʒawaʔ	mɲawab
məɲdʒəmput	ɲəmput	ɲɲdʒəmpoʔ	mɲaput
pəmbaca	pəmaca	mbaca	pəmacɔ:
pəmbasoh	pəmasoh	mmbaso	pəmasoh
pəmbawa	pəmawa	mmbawa:	pəmawa
pəndapat	pənapat	nndapaʔ	pənapat
pəŋgali	pəŋali	ŋŋgali	pəŋali
məmbədil	məmədil	mbəde:	məmədil
məmbawa	mawa	mbata:	mmawɔ:
məŋgatal	ŋatal	ŋgata:	mŋgatɛ:
məmbunuh	munoh	mmbuno	mmunoh
pəmbuka	pəmukaʔ	mbuka	pəmuka
məndidik	mənidiʔ	nndideʔ	mnideʔ
məndukuŋ	nukoŋ	nndukoŋ	mndukoŋ
məŋgoreŋ	ŋoreŋ	ŋŋoreŋ	mŋgoreŋ
məmbariŋ	mareŋ	mbareŋ	mmareŋ

c) Nasal and voiceless obstruent sequences at prefix-root juncture

SM	Perak	Kelantan	Negeri Sembilan
məmilih	milih	mmile	mmilih
məmukul	mukol	mmũkɔ	mmuko
mənarik	nareʔ	nnareʔ	mnareʔ
məŋətuk	ŋətoʔ	ŋŋətoʔ	mŋotoʔ
mənoloŋ	noloŋ	nnuloŋ	mnuloŋ
məmikir	mike	mmike	mmikɛ
mənukar	nuka	nnũka	mnuka
məŋipas	ŋipas	ŋŋipah	ŋipah
mənumbuk	numboʔ	nnũmboʔ	mnumboʔ
məŋacau	ŋacau	ŋŋacɔ	mŋacau
mənutup	nutop	nnũtop	mnutop
mənapis	napis	nnãpis	mnapeh
mənikam	nikam	nnikam	mnikam
məŋuroŋ	ŋuɣoŋ	ŋŋuɣoŋ	mŋuɣoŋ
məŋaet	ŋaet	ŋŋãet	mŋaet
məŋupas	ŋupɛh	ŋŋũpɛh	mŋupɛh
məɲcuci	ɲuʧi	ɲɲũʧi	mɲuʧi
məɲcabot	ɲabot	ɲɲãboʔ	mɲabuit
məɲʧukor	ɲukɔ	ɲɲũkɔ	mɲukɔ
pənakot	nakot	nnãkoq	mnakoq
pənari	nari	nnãri	mnari
məŋanduŋ	ŋandoŋ	ŋŋãnduŋ	mŋandɔŋ
pəŋasih	ŋasɛh	ŋasɛh	mŋasɛh
mənari	nari	nnari	mnari
məmotoŋ	motoŋ	mmotoŋ	mmotoŋ
məŋojak	ŋojaʔ	ŋŋojaʔ	mŋojaʔ
mənanam	nanam	nnãnɛ̃	mnanam