Is Malay Grammar Uniform? A Constraint-Based Analysis

Sharifah Raihan Syed Jaafar

National University of Malaysia
s_raihan@ukm.my

Abstract. This paper presents phonological systems of Malay grammar. Focusing on prefixation, which includes single and multiple prefixation in Malay, this study claims that the grammar of Malay is not completely uniform. The occurrence of nasal and voiceless obstruent clusters is not always resolved by nasal substitution, as claimed by previous Malay scholars regarding the clusters. Based on evidence from one million words obtained from the DBP-UKM corpus database, I further claim that Malay has co-existent grammars, one of which allows nasal and voiceless obstruent clusters, while another does not. This paper proposes that the co-existent grammars in Malay can satisfactorily be explained by adopting a constraint-based analysis named Optimality theory (Prince and Smolensky, 1993).

Keywords: Malay, prefixation, nasal substitution, Optimality theory.

1 Introduction

It has long been observed that the phonological patterns of a language are not completely uniform (e.g. Inkelas and Zoll 2007). This means that the grammar of a language can possibly have more than one phonological pattern. As stated in Inkelas and Zoll (2007), a language can vary systematically like in social register, lexical stratum (native vs. non-native), and morphological category (e.g. stem vs. affix, reduplicant vs. base).

According to what Inkelas and Zoll have claimed above, this paper intends to discuss the phonological system of Malay. By focusing on the issue of nasal and voiceless obstruent clusters in Malay prefixation, I am in agreement with the statement made by Inkelas and Zoll. Transformation and innovation happen continuously in languages. 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. 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 language therefore has undergone much transformation and innovation

G. Morrill and M.-J. Nederhof (Eds.): FG 2012/2013, LNCS 8036, pp. 109–125, 2013.

which have affected the language systems. One of the language systems that was affected is the phonological system of the language. In this paper, I am going to focus on nasal and voiceless obstruent clusters in Malay prefixation.

2 Previous Studies on Malay

It has been widely claimed by previous Malay scholars (e.g. Hassan, 1974; Omar, 1986; Koh, 1981; Othman, 1983; Ahmad, 1993; Karim et al., 1989, 1994; Karim, 1995; and many others) that a nasal segment is always homorganic to the following consonant. In the case of prefixation, when nasal final prefixes are attached to roots, the nasal segments in the prefixes must be homorganic to the initial consonants of the bases. Besides the homorganic nasal, previous Malay scholars have also claimed that bases beginning with voiceless obstruents following nasal segments undergo deletion. Voiced obstruents after nasal segments however are retained – no deletion occurs. This is because the language disallows that a sequence of nasal and voiceless obstruents would emerge in the surface representation.

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-/. At the same time, 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).

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/	[məŋinar]
ACT.PRF-ray ‘to ray’	

In rule-based analyses, two rules have been postulated to ensure the phonological restriction mentioned above is obeyed. The two rules are: (1) Nasal Assimilation, and (2) Voiceless Obstruent Deletion. These two rules have to be applied in order. I show below how these two rules apply:

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 (The Institute of Language and Literature, National University of Malaysia) corpus, there are counter-examples where the clusters emerge in the surface representations, as shown below. This 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.

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

The voiceless obstruents [t] and [p] in the above examples remain undeleted after the assimilated nasal. The rules: 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.

Besides the aforementioned examples, there is another case where we can find the occurrence of nasal and voiceless obstruent clusters in the language. It occurs in multiple prefixation in Malay, i.e. when two prefixes are attached to a root. 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). I shall now demonstrate how the analyses postulated by these scholars pose a problem when accounting for nasal and voiceless obstruent clusters in multiple prefixation.

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, the outputs are:

(i)	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]
(ii)	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 (i). These rules, however, fail to account for the data in (ii), as *[mə.mər.lu.as.kan] is not the right output, although 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 prefixation.

Although some of the examples given fulfil the descriptive rules, 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.

I shall discuss how the rule-based analysis poses a problem when accounting for Malay prefixation, particularly nasal final prefixes. We will then see that the problem can be accounted for by constraint-based analysis i.e. Optimality theory.

3 Data and Methodology

In order to investigate the actual process of prefixation in Malay, corpus data from the DBP-UKM corpus database were collected. As many as one million words were collected for this study. Corpus data were chosen to prove the existence of the peculiar phonological behaviour of nasal and voiceless obstruent clusters in Malay grammar, i.e. in its process of prefixation. The data are essentially needed to verify what previous studies have claimed regarding the clusters. Furthermore, corpus data were chosen because the data comprise examples of real usage of the language. As was claimed by previous Malay scholars, nasal substitution is the regular phonological process applied to break up nasal and voiceless obstruent clusters in Malay prefixation. They further claim that the occurrence of the clusters in some Malay prefixed words as listed above are exceptional cases in the language. This paper argues against this claim. By adopting a constraint-based analysis, i.e. OT, this paper claims that the cases are due to different lexical strata, i.e. native vs. non-native. The claim then reveals that Malay has co-existent grammars.

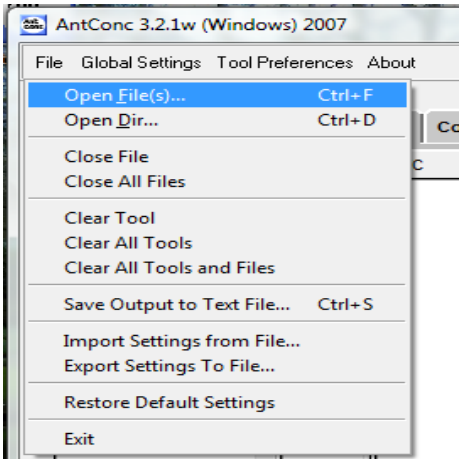
Since the data accessed from the DBP-UKM corpus are raw data, they need to be categorised according to the type of prefixes they belong to. It would be difficult to categorise one million data manually. I have therefore used corpus software named ‘AntConc’ to do the categorisation.

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. There 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 *NC$_o$, the markedness constraint. I now explain how those groups, i.e. voiced

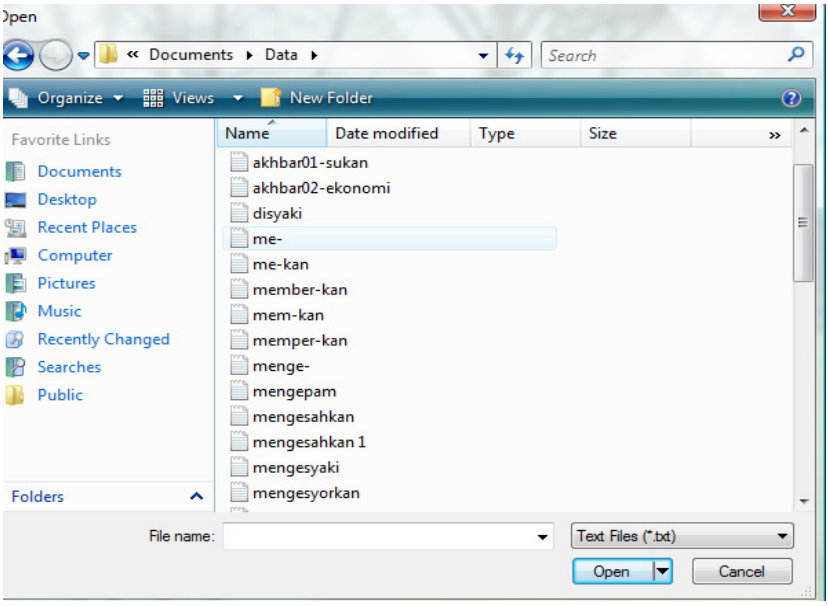
and voiceless obstruents (with or without nasal substitution), are categorised using AntConc software.

(1) Voiced Obstruents

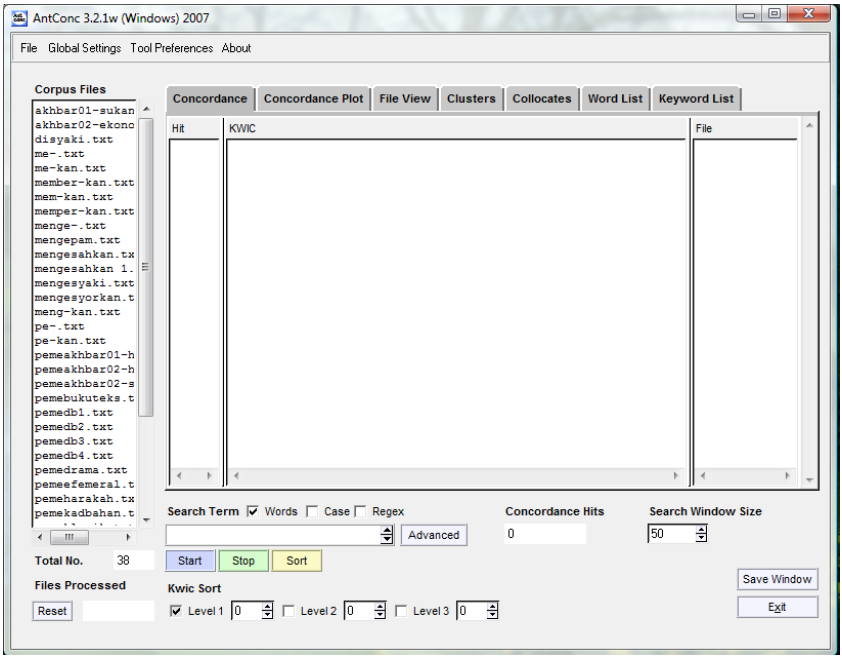
Before the relevant data for this group can be generated, we must choose a text file(s) where the data are stored by clicking on the **file** which is located at the top left of the software page and then select **open file(s)**. It looks like this:



A standard file-open will then appear. We can 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:

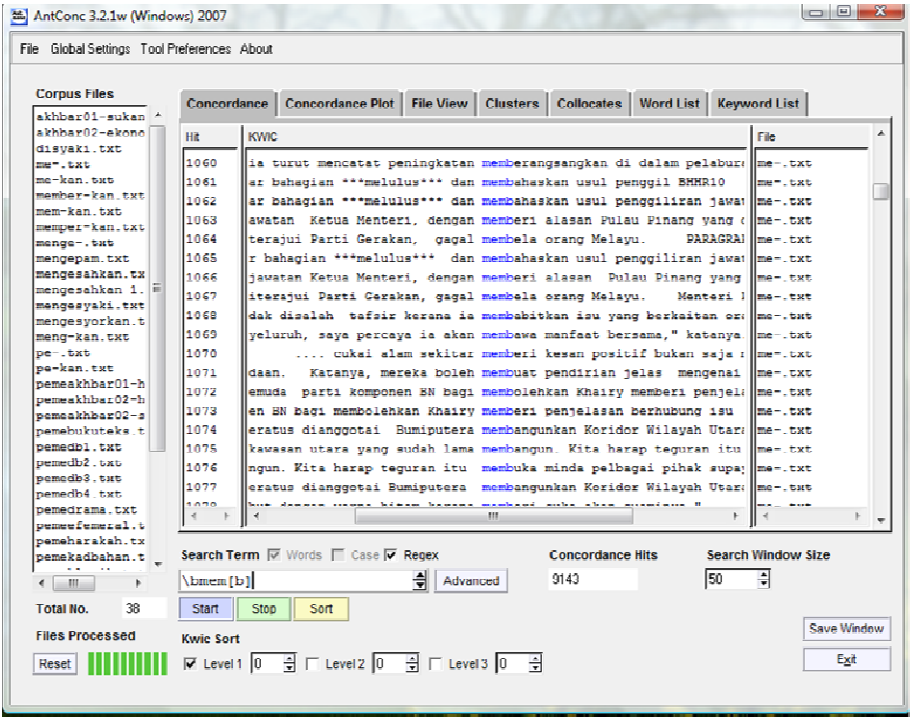


After all the text files have been selected, the data are now listed on the main page of AntConc, as shown below:



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 the same 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 undergo 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 words where the root does not actually begin with /p/ combined with a single nasal final prefix but is a sonorant consonant instead, such as [pəminat] where the underlying form is /pəŋ+minat/; (3) nominal multiple prefixes /pəŋ+pər/ →

[pəmər], as in /pəŋ+pər+badan+an] → [pə.mər.ba.da.nan]. 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/ → [pəməgaŋ]. 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 but 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 that of 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 has to be done manually whereby all the examples are categorised according to their phonological character. This means that we have five types of data, two of which are only useful for our analysis, i.e. (1) and (3). 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) 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

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/ → [məmpər], 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) 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 Malay Co-existent Grammars: Constraint-Based Analysis

Observations from the DBP-UKM corpus show that the claim regarding nasal substitution postulated by previous Malay scholars on prefixation does not hold for the whole dataset. The generalization postulated by previous studies can only explain some of the output derived from the process of prefixation. This shows that the proposed rule-based analysis does not adequately explain the real process of prefixation in Malay. I am going to discuss the two patterns that occur in Malay i.e. (1) outputs with nasal and voiceless obstruent clusters, and (2) outputs without nasal and voiceless obstruent clusters. Why do these two patterns occur in the language? Supposedly, outputs with nasal and voiceless obstruent clusters should not emerge in the surface representation as the language precludes such clusters.

The occurrence of the two patterns in single prefixation is analysed in terms of different strata of Malay words, according to their etymology: native or non-native (Itô and Mester 1999). Based on the corpus data, I thus postulate the following lexical strata for Malay:

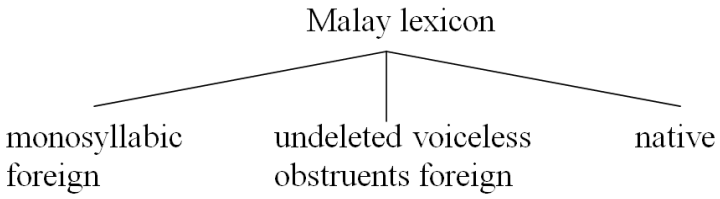


Fig. 1. The three strata of Malay lexicon (Syed Jaafar 2010)

Constraints are ranked differently in each lexical stratum according to the role played by the crucial constraints in the ranking, i.e. the markedness constraint *NC$_o$,

which bans a sequence of nasal and voiceless obstruents in the surface and the faithfulness constraint, which requires the output to be as faithful as possible to the input, i.e. UNIFORMITY.

In what follows, I am going to present how the three lexical strata proposed for the Malay lexicon are analysed. As we see below, the phenomenon of inconsistency of the occurrence of nasal and voiceless obstruent clusters is analysed by the same set of constraints but they are ranked differently. Each lexical stratum has its own constraint ranking. As was mentioned, 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 the following tableau analyses, nasal substitution as the regular strategy to eliminate the clusters, as claimed by previous Malay scholars, only applies when the roots are Malay native words. This means that *NC$_{o}$, 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.

> ***NC$_{o}$**
> No nasal/ voiceless obstruent sequences.

As we will see, *NC$_{o}$ is violated by foreign words as nasal substitution is not the way to resolve the clusters. Nasal and voiceless obstruent clusters in monosyllabic foreign and undeleted voiceless obstruent foreign are resolved by vowel epenthesis and nasal assimilation, respectively. The relevant constraints that play an important role to explain these phonological processes are DEP-IO and NASAL ASSIMILATION. When the clusters undergo nasal assimilation, the two segments, i.e. the nasal segment and the initial voiceless obstruent, are preserved. The preservation segments can be explained by a faithfulness constraint named UNIFORMITY. All the three constraints are defined below:

> **DEP-IO**
> Every segment in the input must have a correspondent in the output.
>
> **NAS ASS** (cf.: Jun, 1995; Padgett, 1995; Boersma, 1998; Pater, 2001)
> A nasal must share place features with a following consonant.
>
> **UNIFORMITY** ('No Coalescence') (McCarthy and Prince, 1999: 296)
> No element of S_2 has multiple correspondents in S_1.

For the monosyllabic foreign lexicon, a nasal and voiceless obstruent occurring in the input representation is resolved by vowel epenthesis. Nasal substitution, which is claimed to be the regular strategy to eliminate the clusters in the language, is not applied however. I briefly exemplify some of the relevant data from the corpus:

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'

Bringing together all the constraints introduced thus far, I establish the following tableau to account for the monosyllabic foreign words. The relevant constraint ranking is: NASAL ASSIMILATION >> *NC̥ >> UNIFORMITY >> DEP-IO.

/məŋ$_1$+p$_2$am/	NAS ASS	*NC̥	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 now see how words in the group of undeleted voiceless obstruent foreign words are analysed. Before I establish a tableau analysis for this group, let us observe first some of the relevant data below:

Nasal final prefixes (from the DBP-UKM corpus)

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

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

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

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

With the same set of constraints in the monosyllabic foreign lexical strata, I establish the following tableau for undeleted voiceless obstruent foreign words. Observe that the constraints are ranked differently from monosyllabic foreign words. The *NC̥ constraint which bans the clusters to emerge in the surface is ranked lower as this group allows nasal and voiceless obstruent clusters.

/məŋ$_1$+p$_2$roses/	NAS ASS	DEP-IO	UNIFORMITY	*NC̥
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		*!		

The third group is that of native words. Before I start the analysis, let us first consider some relevant examples of this group:

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

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

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

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

The tableau analysis for this group is shown below:

/məŋ$_1$+p$_2$otoŋ/	NAS ASS	*NC$_o$	DEP-IO	UNIFORMITY
a. ☞ məm$_{12}$otoŋ				*
b. məm$_1$p$_2$otoŋ		*!		
c. məŋ$_1$p$_2$otoŋ	*!			
d. məŋ$_1$əp$_2$otoŋ			*!	

In multiple prefixation, the co-existent grammars occur at prefix-prefix boundaries when two prefixes end with nasal segments attached to voiceless obstruent initial roots. At this morphological boundary, the clusters emerge in the surface representation in verbal multiple prefixes /məŋ+pər/. The clusters however undergo nasal substitution in nominal multiple prefixes, /pəŋ+pər/. I exemplify some of the data taken from the corpus:

a) 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**-taʧam-kan
VERBL.PRF-VERBL.PRF-sharp- CAUS.SUF
'to cause to sharpen for'

b) 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'

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 (a). On the other hand, when the multiple prefixes form a verbal word, as in (b), nasal substitution is blocked. In this analysis, 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.

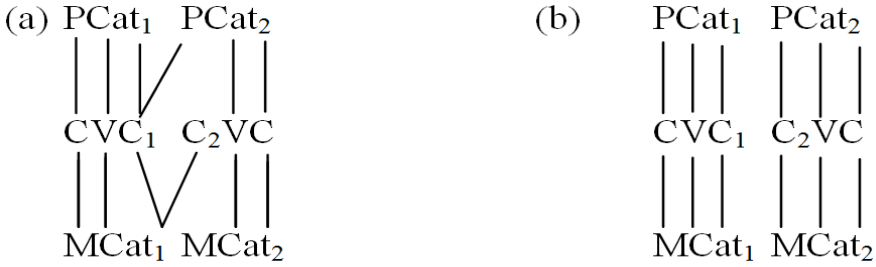
As already noted, the language does not allow nasal and voiceless obstruent clusters in the surface representation. Therefore, voiceless obstruents following nasals regularly undergo nasal substitution, as claimed by previous scholars. One question that can be asked here is: Is it obligatory for a sequence of nasal and voiceless obstruents 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. Considering both factors, I claim that the process of multiple prefixation should differentiate between the verbal and the nominal prefixes. As nasal substitution is blocked in the verbal prefixes, the EDGE INTEGRITY constraint thus plays a crucial role to account for the blocking process of nasal substitution.

> **EDGE INTEGRITY** (McCarthy and Prince, 1995)
> Edge segments in the input preserve their segments at the edge of the corresponding prosodic structure.

As defined, EDGE INTEGRITY requires that the morphological unit preserves 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, 2002).

In the following diagram, we see the structure in (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 (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.

/CVC + CVC/ ('+' stands for a morphological boundary) (from Kang, 2002).



I now establish the following constraint ranking for verbal multiple prefixes: EDGE-INTEGRITY >> NASAL ASSIMILATION >> *NC$_o$ >> DEP-IO >> UNIFORMITY.

a) Verbal prefixes

/məŋ$_1$+p$_2$ər+bəsar/	EDGE INTEG	NAS ASS	*NC$_o$	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 *NC$_o$. 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 *NC$_o$, as the candidate does not undergo nasal substitution. Since the markedness constraint *NC$_o$ 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 >> *NC$_o$, can thus account straightforwardly for why nasal substitution does not occur in the environment of the prefix-prefix juncture.

b) Nominal prefixes

/pəŋ$_1$+p$_2$ər+kasa/	NAS ASS	*NC$_o$	EDGE INTEG	DEP-IO	UNI
a. ☞pəm$_{12}$ərkasa			*		*
b. pəm$_1$p$_2$ərkasa		*!			
c. pəŋ$_1$p$_2$ərkasa	*!				

With a rule-based analysis, two rules, (1) nasal assimilation and (2) voiceless obstruent deletion, would be 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:

	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]

The above derivation 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:

	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/. From the above derivation we derive an output with nasal substitution. This is incorrect since /məŋ+pər/ does not undergo nasal substitution.

5 Conclusions

The above discussion has presented some important points about the grammar of Malay. As we saw, nasal and voiceless obstruent clusters are not entirely prohibited in 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 do not.

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. However, as we saw, nasal substitution applies to nominal 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, as opposed to any other model.

Those phonological processes occurring in Malay prefixation clearly show that Malay has co-existent grammars. As a result of transformation and innovation, Malay allows nasal and voiceless obstruent clusters in the language, as occurring in foreign words, next to an absence of nasal and voiceless obstruent clusters in native words.

References

1. Ahmad, Z.: Fonologi generatif: teori dan penerapan. Institute of Language and Literature, Kuala Lumpur (1993)
2. Boersma, P.: Typology and acquisition in functional and arbitrary phonology. Ms. University of Amsterdam (1998)
3. Hassan, A.: The morphology of Malay. Institute of Language and Literature, Kuala Lumpur (1974)
4. Inkelas, S., Zoll, C.: Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. Linguistics 45, 133–171 (2007)
5. Ito, J., Mester, A.: The phonological lexicon. In: Tsujimura, N. (ed.) The Handbook of Japanese Linguistics. Oxford Blackwell (1999)
6. Jun, J.: Perceptual and articulator factors in place assimilation: an Optimality theoretic approach. Doctoral dissertation. UCLA, Los Angelas (1995)
7. Kang, E.: Edge integrity and the syllable structure in Korea. In: The Proceedings of the 16th Pacific-Asia of Language. Informations and Computation. The Society of Korean Information Society, Seoul (2002)
8. Karim, N.S., Onn, F., Musa, H., Mahmood, A.H.: Tatabahasa dewan, 2nd edn. perkataan. Institute of Language and Literature, Kuala Lumpur (1989)
9. Karim, N.S., Onn, F., Musa, H., Mahmood, A.H.: Tatabahasa dewan. Institute of Language and Literature, Kuala Lumpur (1994)
10. Karim, N.S.: Malay grammar for academics and professionals. Institute of Language and Litrature, Kuala Lumpur (1995)
11. Koh, B.B.: Pengajaran bahasa Malaysia. Utusan Publications & Distributors Sdn. Bhd, Kuala Lumpur (1981)
12. McCarthy, J.J., Prince, A.S.: Faithfulness and Identity in Prosodic Morphology. In: Kager, R., van der Hulst, H., Zonneveld, W. (eds.), pp. 218–384 (1999)
13. Omar, A.: Nahu Melayu mutakhir, 2nd edn. Institute of Language and Literature, Kuala Lumpur (1986)
14. Othman, A.: Imbuhan me- dan pentingnya dalam pengajaran tatabahasa bahasa Melayu. Institute of Language and Literature (1983)
15. Padgett, J.: Partial class behaviour and nasal place assimilation. In: Proceedings of the Arizona Phonology Conference: Workshop on Features in Optimality Theory. Coyote Working Papers. University of Arizona Department of Linguistics, Tuscon (1995)
16. Pater, J.: Austronesian nasal substitution revisited. In: Lombardi, L. (ed.) Segmental Phonology in Optimality Theory: Constraints and Representations, pp. 159–182. Cambridge University Press (2001)
17. Prince, A., Smolensky, P.: Optimality theory: constraint interaction in generative grammar. Available on Rutgers Optimality Archive, ROA-537 (1993)
18. Jaafar, S., Raihan, S.: Newcastle Working Papers in Linguistics, Newcastle University, vol. 16 (2010)