Onset Satisfaction and Violation in Malay: An Optimality Account

Zaharani Ahmad
Universiti Kebangsaan Malaysia

1.0 Introduction

Malay generally requires that every surface syllable must have an onset. Underlying vowel sequences derived by morphemic concatenation cannot be syllabified heterosyllabically because it creates an onsetless syllable. The hiatus is then resolved by C-Epenthesis. Although this requirement is crucial, it can be violated in certain environments. There are two instances where the onset requirement is violated in this language, namely the word initial and root medial environments. This fact is observed and well attested in Yunus (1980)[1966], but it is not satisfactorily accounted for in the analyses of Farid (1980) and Teoh (1994).

In this paper, I attempt to demonstrate that the *satisfaction* and *violation* of *onset conditions* can be captured satisfactorily under the constraint-based approach of Optimality Theory (Prince and Smolensky 1993). The pivotal analytical proposal of Optimality Theory is that a grammar is a hierarchical ranking of well-formedness constraints. These constraints are specified in the Universal Grammar, and individual grammars are constructed by imposing a language-particular ranking of those universal well-formedness constraints. The distinguishing feature of OT with respect to other constraint-based approaches is that it allows violation of those universal constraints. Lower ranked constraints can be minimally violated in order to assure the satisfaction of higher ranked constraints.

2.0 Syllabification and Onset Satisfaction

It has long been observed that the basic structure of the Malay syllable is (C)V(C) (Abdullah 1974, Yunus 1980, *Farid 1980). Typologically, this language belongs to a class* of languages which Clements and Keyser (1983) refer to as a type IV group that has four basic syllable structures, namely V, VC, CV and CVC. This analysis, however, has been challenged by Teoh (1994) claiming that Malay is a type III language with a CV(C) syllable structure, suggesting that every syllable must have an onset.

In contrast to Teoh (1994), I am inclined towards the earlier proposal that the Malay syllable is (C)V(C). The language generally requires that every surface syllable must have an onset. Despite the fact that this requirement is crucial, it can be violated in certain environments, in particular word initially and root internally.

It is a well-established fact that the sonority hierarchy plays a major role in determining the nucleus and margins of the syllable. Vowels are more sonorous than consonants, and basically make more harmonic nuclei and less harmonic margins. In Malay only vowels are permitted in the syllable nucleus position, whereas consonants are invariably associated with the syllable margins, namely onsets and codas. Each syllabic constituent can only be occupied by a single segment, suggesting that the language disfavours segmental clusters.

As commonly accepted by most phonological theories, syllable structures are not present in the lexicon, and are derived in the course of phonological derivation. Within the OT framework, the process of syllabification is a matter of choosing the optimal output from among the possible analyses, rather than algorithmic structure building (Prince and Smolensky 1993:15). Syllable structure is generated in the same way as any other grammatical property by the function GEN, which produces a set of candidates with various possibilities of syllable parsing from each unsyllabified input. These possible candidates are then evaluated in parallel by the function EVAL based on a language particular constraint hierarchy. As expected, a candidate that minimally violates the constraints in the hierarchy is termed optimal and declared the true output.

In early OT (Prince and Smolensky 1993, McCarthy and Prince 1993a, *et seq.*), syllabification is construed as a process of incorporating segments into higher prosodic constituents. Phonological elements are said to be 'parsed' when they are associated and dominated by the appropriate node of the prosodic hierarchy (Selkirk 1980, McCarthy and Prince 1986, 1990ab), and this is controlled by a formal constraint called PARSE. As a family of constraints, PARSE provides a number of constraints that ensure parsing, such as PARSE-SEGMENT which requires that all segments must belong to moras and PARSE-μ which demands that all moras be parsed into syllables. The crucial idea about a constraint family is that a group of similar and related constraints are all built from a single broad concept (i.e. PARSE) but they are separately rankable in the hierarchy.

With the advent of Correspondence Theory (McCarthy and Prince 1995b), the earlier faithfulness constraint of the PARSE family has been subsumed under the MAX constraint family which requires that every segment of S_1 (input/base) has a correspondent in S_2 (output/reduplicant). PARSE-SEGMENT is now reformulated as MAX-IO, which demands that every segment of the input must have a correspondent in the output. A process of phonological deletion is reckoned as a *violation of MAX-IO*[1]*. Similarly, for PARSE-μ: it* can be reformulated as MAX-IO-μ.

The process of syllabification is primarily an interaction of the faithfulness constraint MAX-IO and the syllable structure constraints, such as ONSET, NO CODA and *COMPLEX, which are formally defined as follows:

1. Syllable structure constraints (Prince and Smolensky 1993)
   ONSET - Syllables must have onsets
   NO CODA - Syllables must not have a coda
   *COMPLEX - No more than one segment may associate to any one syllabic constituent (i.e onset, nucleus, coda)

Let us first consider the interaction between MAX-IO and the syllable structure constraint NO CODA. It is apparent that MAX-IO and NO CODA can be in a relation of conflict which means that there are pairs of competing candidates with conflicting constraints. Crucially, one of the candidates (the actual output form) must emerge as optimal.

As noted, MAX-IO demands that all the input segments must appear on the surface regardless of whether the form has an illicit syllable structure, for instance a syllable with a coda. This is to ensure that all underlying segments are parsed. On the other hand, NO CODA disfavours any coda element. Since Malay is a (C)V(C) language which optionally allows codas, the relevant ranking is: MAX-IO dominates NO CODA. This conclusion is illustrated in the following tableau (syllable boundaries are marked by a period '.').

2. MAX-IO >> NO CODA - i.e. /pasti/ 'sure, certain'

/pasti/	MAX-IO	NO CODA
a. pa.ti	*!	
b. ☞ pas.ti		

Tableau (2) shows that faithfulness to the underlying form by parsing all the input segments leads to a violation of a syllable structure constraint. Generally, such a violation can be avoided by epenthesis, which is one way of ensuring that all the input segments are parsed, and concurrently satisfy the syllable structure constraint.

In standard OT analysis, epenthesis is governed by another faithfulness constraint called FILL (Prince and Smolensky 1993, McCarthy and Prince 1994), which states that all nodes of syllable structure must be filled by underlying segments. In the Correspondence Theoretic approach, this constraint is subsumed under the DEP constraint family which demands that every segment of S_2 (output/reduplicant) has a correspondent in S_1 (input/base). FILL is now reformulated as DEP-IO, which requires that every segment of the output must have a correspondent in the input.

DEP-IO can also be in a conflict relation with NO CODA. The latter prefers a syllable without any coda, and this can be achieved by inserting an epenthetic schwa interconsonantally. The former, by contrast, favours a nonepenthetic form, even though it has an illicit syllable structure. In Malay, DEP-IO clearly outranks NO CODA. The interaction is shown in the tableau below.

3. DEP-IO >> NO CODA

/pasti/	DEP-IO	NO CODA
a. pa.sə.ti	*!	
b. ☞ pas.ti		

Another possible form that should be considered is [pa.sti]. In this candidate, the intervocalic consonant cluster /s/ and /t/ are both parsed to the second syllable, creating a complex structure in the onset node. Considering the available constraints developed in (2) and (3), this candidate obeys all their requirements, and thus it would be the most harmonic. However, this is not the correct surface form. It must then be the case that another constraint is crucially involved in ruling out this candidate, and this constraint must be more dominant. The relevant constraint that plays a crucial role here is *COMPLEX which bans the occurrence of clusters in any node of the syllable structure. This constraint is unviolated, therefore undominated in the hierarchy[2].

4. *COMPLEX >> NO CODA

/pasti/	*COMPLEX	NO CODA
a. pa.sti	*!	
b. ☞ pas.ti		

Malay loan phonology offers a good piece of evidence that *COMPLEX is highly respected in the language. Borrowed lexical items containing clusters are generally resolved by schwa epenthesis and C-deletion. For example, English words like *stamp, glass, class, club, post* are realised as [sətem], [gəlas], [kəlas], [kəlap] and [pos], respectively[3].

It must be mentioned that in principle, MAX-IO, DEP-IO and *COMPLEX are also in a conflict situation with respect to each other. For instance, *COMPLEX disfavours illicit syllable structure, and this can be resolved either by C-deletion (i.e [pa.ti]) or by V-epenthesis (i.e. [pa.sə.ti]). The satisfaction of *COMPLEX by the former compels a violation of MAX-IO, whereas the latter involves the DEP-IO violation. MAX-IO demands that all the input segments must appear on the surface regardless of whether the syllable contains an illicit consonant cluster (i.e [pa.sti]). Likewise, DEP-IO prefers a nonepenthetic form even though it has an illicit consonant cluster as well.

The potential conflict between the three constraints, however, is not significant because all their candidates are ill-formed. In this case the constraints at hand are not crucially ranked with respect to each other. Conventionally, this kind of interaction is indicated by a dotted line in the tableau.

5. *COMPLEX, MAX-IO, DEP-IO >> NO CODA

/pasti/	*COMPLEX	MAX -IO	DEP- IO	NO CODA
a. pa.sti	*!			
b. pa.ti		*!		
c. pa.sə.ti			*!	
d. ☞ pas.ti				

After evaluating NO CODA, let us assess the position of ONSET as it interacts with the faithfulness constraints MAX-IO and DEP-IO. Malay generally requires that every surface syllable must have an onset. Underlying vowel sequences derived by morphemic concatenation can never be faithfully syllabified. For instance, a morphological process of prefixation that brings together /V+V/ cannot be syllabified heterosyllabically as [V.V], since it produces an onsetless syllable, a clear violation of ONSET. The hiatus is then resolved by C-Epenthesis[4]. The examples in (6) show that the vowel clusters surface as a heterosyllabic sequence separated by a glottal stop.

6.	/di + ubah/	[diʔubah]	'to move (passive)'
	/di + ikat/	[diʔikat]	'to tie (passive)'
	/dʒuru + atʃara/	[dʒuruʔatʃarə]	'master of ceremony'
	/sə + indah/	[səʔindah]	'to be as beautiful as'
	/sə + elok/	[səʔeloʔ]	'to be as pretty as'
	/kə + ibu + an/	[kəʔibuwan]	'motherhood'
	/kə + əmas + an/	[kəʔəmāssan]	'golden'

As can be seen in (6), the occurrence of epenthetic glottal intervocalically is triggered by the ONSET requirement. Obedience to ONSET compels a violation of DEP-IO, as the output glottal stop has no correspondent in the input form. The two constraints conflict with each other, and evidently DEP-IO must be dominated by ONSET.

7. ONSET >> DEP-IO

/di + ubah/	ONSET	DEP-IO
a. di.u.bah	*!	
b. ☞ di.ʔu.bah		*

Potentially, there are two other possibilities for ensuring ONSET satisfaction. The first candidate is *[du.bah], where one of the vocalic segments in the input undergoes the process of deletion. This leads to a violation of the faithfulness constraint MAX-IO, which ensures that all the input materials must surface in the output. Deleting a vowel to ensure ONSET satisfaction is never permitted in this language, suggesting that MAX-IO is an unviolated constraint in the hierarchy.

The second candidate is *[dju.bah], where the high vowel is parsed in the onset. Although this candidate spares ONSET, MAX-IO and DEP-IO, it fatally violates the undominated syllable structure constraint *COMPLEX. The interaction between the four constraints is controlled by the following ranking: *COMPLEX, MAX-IO >> ONSET >> DEP-IO. The first two constraints do not conflict, and therefore they are not crucially ranked with respect to each other.

8. *COMPLEX, MAX-IO >> ONSET >> DEP-IO

/di + ubah/	*COMPLEX	MAX -IO	ONSET	DEP-IO
a. di.u.bah			*!	
b. du.bah		*!		
c. dju.bah	*!			
d.☞di .ʔu.bah				

In conclusion, the process of syllabification in OT involves choosing the most harmonic output between a set of candidates with various possibilities of syllable parsing based on the interaction of wellformedness constraints ranked on a language particular basis. For all the constraints we have discussed so far, we can establish the hierarchical ranking: *COMPLEX, MAX-IO >> ONSET >> DEP-IO >> NO CODA.

3.0 Onset Violation

Although Malay disfavours an onsetless surface syllable, ONSET is a dominated constraint in this language, and thus it is violable. There are two instances where *ONSET is violated in this language, namely word initially* and root medially.

Let us first examine word-initial ONSET violations. As noted in McCarthy and Prince (1993ab), it is quite common cross-linguistically for languages that otherwise demand strictly C-initial syllables to admit V-initial words. As observed in Farid (1980) and Yunus (1980), the initial syllable of Malay words can be onsetless. This seems to be the evidence that corroborates their claim that the basic syllable structure of Malay is (C)V(C). In (9), we lay out some examples which show that all the six underlying vowels can occur in this environment.

9.	/ubah/	[ubah]	'to change'
	/indah/	[indah]	'beautiful'
	/elok/	[eloʔ]	'pretty'
	/olah/	[olah]	'to beguile'
	/aŋkat/	[aŋkat]	'to lift'
	/əmak/	[əmāʔ]	'mother'

As was demonstrated in (6), when V-initial stems combine with the V-final prefixes, such as /sə-/, /kə-/ and /di-/, the underlying vowel sequences /V+V/ at the prefix juncture cannot be parsed heterosyllabically as [V.V], as it produces an onsetless syllable which consequently disobeys ONSET. This is then resolved by Glottal Epenthesis, and the price is a DEP-IO violation.

On the other hand, when those stems concatenate with consonant-final prefixes, such as /bər-/, /tər-/, /məŋ-/ and /pəŋ-/, the onsetless stems then get their onset from the preceding consonant in accordance with the Minimal Onset Satisfaction Principle (Roca 1994). In this case, ONSET can be fully satisfied without affecting the faithfulness constraint DEP-IO.

10. a. Vowel-final prefixes + Vowel initial stems

/di + ubah/	[diʔubah]	'to change (passive)'
/sə + indah/	[səʔindah]	'to be as beautiful as'
/kə + indah + an/	[kəʔindahhan]	'beauty'
/sə + elok/	[səʔeloʔ]	'to be as pretty as'
/di + olah/	[diʔolah]	'to beguile (passive)'
/di + aŋkat/	[diʔaŋkat]	'to lift (passive)'
/dʒuru + atʃara/	[dʒuruʔatʃarə]	'master of ceremony'

b. Consonant-final prefixes + Vowel initial stems

/məŋ + indah+kan/	[mə̃ŋĩndahkan]	'to beautify'
/pəŋ + aŋkat/	[pəŋãŋkat]	'lifter (instrument)'
/tər + elok/	[tərelok]	'most beautiful'
/məŋ + ubah/	[mə̃ŋũbah]	'to change (active)'
/bər +aŋkat/	[bəraŋkat]	'to depart'

As shown in (10), Malay freely tolerates onsetless syllables word initially. Although Glottal Epenthesis is potentially active as an alternative way to satisfy ONSET, this solution does not seem to be preferred in this particular environment. The violation of ONSET in V-initial stems is common in many languages, such as in Timugon Murut (Prentice 1971, McCarthy and Prince 1993ab), Tagalog and Axininca Campa (McCarthy and Prince 1993ab), and so it is not particularly remarkable to find it in Malay.

In McCarthy and Prince's (1993ab) analysis of Axininca Campa, the V-initial phenomenon arises from the interaction of ONSET and ALIGN-LEFT (11), an alignment constraint of the prosody-morphology interface which requires that the left edge of any stem must coincide with the left edge of a PrWd (Prosodic Word). ALIGN-LEFT is unviolated, and therefore it is undominated in the constraint hierarchy of Axininca Campa. ONSET is violated when it conflicts with ALIGN-LEFT, and the ranking is ALIGN-LEFT >> ONSET.

11. ALIGN-LEFT (McCarthy and Prince 1993b)
Align (Stem, L, PrWd, L)

It is apparent that the interaction ALIGN-LEFT >> ONSET can handle a similar phenomenon in Malay. However, in order to account for the Malay data adequately and satisfactorily, I will adopt a different definition of ALIGN-LEFT, as formalised in (14).

ALIGN-LEFT belongs to a family of well-formedness constraints, called GENERALISED

ALIGNMENT (henceforth GA), which is formalised in McCarthy and Prince (1993b) as in (12).

12. GENERALISED ALIGNMENT (McCarthy and Prince 1993b:80)

Align (Cat1, Edge1, Cat2, Edge2) $=_{def}$

$\forall$ Cat1 $\exists$ Cat2 such that edge1 of Cat1 and Edge2 of Cat2 coincide.

Where

Cat1, Cat2 $\in$ Pcat $\cup$ Gcat

Edge1, Edge2 $\in$ {Right, Left}

GA requires that a designated edge (i.e. left or right) of each prosodic or morphological constituent (i.e. Pcat and Gcat) of type Cat1 coincide with a designated edge (i.e left or right) of some other prosodic or morphological constituent (i.e. Pcat and Gcat) Cat2[5]. As demonstrated in McCarthy and Prince (1993b), GA is able to express a wide range of references to edges in the grammar of many languages via various types of alignment constraints. For instance, to account for stress patterns in Garawa, two alignment constraints are proposed, namely ALIGN-PRWD - Align (PrWd, L, Ft, L), and ALIGN-FT - Align (Ft, R, PrWd, R); Tagalog prefixation requires ALIGN-*um* - Align ($[\text{um}]_{Af}$, L, Stem, L); Ulwa suffixation needs ALIGN-TO-FOOT - Align ($[\text{Poss}]_{Af}$, L, Ft, R).

It is important to note that the term 'edge' in Alignment theory is interpreted as relational rather than categorical. According to McCarthy and Prince (1993b:89), the notion that we really need is relational, something like 'sharing an edge', rather than categorical, referring to edge *per se*. Two categories are aligned when they 'share an edge', and the Alignment constraint specifies the categories and which side of each is involved in 'sharing an edge'.

Now let us examine the V-initial phenomenon in Malay as shown in the data (10). Observe how, in the following examples, an onsetless syllable guarantees coincidence between the word stem and the edge of a syllable. While, Glottal Epenthesis locates the morphological word edge inside a syllable. The relevant word-edge is marked by '|' and the syllable boundary is shown by a period '.'.

13. Word-Syllable Alignment
Input: /ubah/ Output: a. [.|u.bah]
b.*[.ʔ|u.bah]

The distinction between matching and non-matching of word/syllable edges in (13) is regulated by a formal constraint called ALIGN-LEFT, which is formally defined in this study as in (14).

14. ALIGN-LEFT
Align (Word, Left, σ, Left)

Unlike (11), constraint (14) says that the left edge of any morphological word must coincide with the left edge of a syllable. In order for ALIGN-LEFT to be fully satisfied, the V-initial word must be parsed with an ONSET violation. If epenthesis were to apply, the presence of C-epenthetic segment which is not part of the morphological word will shift the syllable edge away from the word edge (13b). This causes a misalignment of the leading edges of the syllable and the word. Equivalently, deleting the initial vowel, a MAX-IO violation, as a way to avert an ONSET violation, can never bring a form into agreement with ALIGN-LEFT (cf. McCarthy and Prince 1993ab).

In short, obedience to ALIGN-LEFT can only be achieved, if the word-initial segments, vowels or consonants, occupy the word initial position. ALIGN-LEFT is unviolated, and therefore it is undominated in the constraint hierarchy. When ALIGN-LEFT conflicts with ONSET, inevitably the latter has to give way. This suggests that the ranking is ALIGN-LEFT, MAX-IO >> ONSET >> DEP-IO. The following tableau illustrates the arguments just made.

15. ALIGN-LEFT, MAX-IO >> ONSET >> DEP-IO

/ubah/	ALIGN-LEFT	MAX-IO	ONSET	DEP-IO
a. ☞.|u.bah				
b. |< >.bah	*!			
c. .ʔ |u.bah	*!			

As can be seen, in the losing candidates (15b) and (15c), the word edge and the syllable edge do not coincide due to the process of deletion (as shown by '< >') or the presence of epenthetic an glottal stop. In contrast, the optimal candidate (15a) is well-aligned, but minimally violates the syllable structural constraint ONSET.

Although ONSET can be violated in the bare forms, the situation is totally different in the prefixed forms, particularly in the case where V-final prefixes concatenate with V-initial stems (10a). The initial vowel of the stem presently appears in a word internal position, thus, ALIGN-LEFT is irrelevant and vacuously satisfied. Glottal Epenthesis then has to apply in compliance with the *ONSET requirement.*

16. ALIGN-LEFT, MAX-IO >> ONSET >> DEP-IO

/di+ubah/	ALIGN-LEFT	MAX-IO	ONSET	DEP-IO
a. .|di.u.bah			*!	
b. .|du.bah		*!		
c.☞.|di.ʔ u.bah				

All the candidates satisfy ALIGN-LEFT, since the edges of the word and the syllable coincide. The next constraints that should be consulted are MAX-IO and ONSET, which then rule out (16a) and (16b) respectively. Although the optimal candidate (16c) violates DEP-IO, this is irrelevant, since the victor has already been determined.

In the case where C-final prefixes concatenate with V-initial stems, obviously the rule of Glottal Epenthesis is not required. The final consonant of the prefix is readily available to fulfil the minimal ONSET requirement. Thus, the optimal candidate fully satisfies all the four given constraints. The following tableau demonstrates this fact.

17. ALIGN-LEFT, MAX-IO >> ONSET >> DEP-IO

/məŋ+ubah/	ALIGN-LEFT	MAX-IO	ONSET	DEP-IO
a. .|məŋ.u. bah			*!	
b. .|məŋ.ʔu.bah				*!
c.☞.| mə.ŋu.bah				

In conclusion, the emergence of a glottal stop preceding the V-initial stem in the prefixed forms is motivated by the wellformedness condition on syllable structure which requires that every syllable must have an onset. However, when the V-initial stems occur in isolation as independent words, Glottal Epenthesis can never apply due to the dominant ranking of ALIGN-LEFT. This readily explains the phonological alternation that has taken place.

In Farid's linear analysis (1980:48-50), the phenomenon of Glottal Epenthesis is captured by a rule called Glottal Insertion Rule, which is formalised as in (18)[6].

18. Ø → ʔ / V- ___ V

Condition: '-' designates a prefix boundary

In his non-linear analysis, Teoh (1994) suggests that the glottal stop at the prefix juncture is not rule-derived, but underlyingly present in the stems. Teoh (1994:89) writes,

> ... vowel-initial stems may be pronounced optionally either with or without the glottal stop when in isolation and that the same glottal stop resurfaces obligatorily when vowel-initial stems are prefixed to the vowel-final passive marker /di-/. In order to solve this particular problem we have again

assumed the [ʔ] as underlying in all so-called vowel-initial stems.

Postulating the glottal stop as an underlying segment word initially in the so-called vowel-initial stems (i.e. /ʔubahʔ, /ʔindah/) is consistent with his primary claim that no syllable in this language can begin with a vowel, as it is constrained by the syllable typology CV(C) of the language.

To support his analysis which has an underlying glottal stop word initially, Teoh (1994) offers a piece of language external evidence extracted from a language game inverting the syllables of each stem. For example, forms such as /batu/ 'stone' or /satu/ 'one' will be transformed into [tuba] and [tusa] respectively. However, in vowel-initial stems, such as /aku/ 'I' and /apa/ 'what', the words are inverted and rendered as [kuʔa] and [paʔa], and not as *[kuwa] and *[paa] which is what one would expect to be. The alternation in the language game can be accounted for more generally and simply if the glottal stop is postulated as part of the underlying representation (i.e. /ʔaku/ and /ʔapa/).

Teoh's analysis misses two important phonological generalisations. First, he fails to capture the regular process of cross-morphemic syllabification at the prefix-stem juncture, which is motivated by the principle of Minimal Onset Satisfaction (Roca 1994). This is represented in (10b) where the prefix-final consonant is syllabified in the following onset of the stems. In Teoh's analysis, a possible way of accounting for this fact is through a rule that first deletes the stem-initial glottal stop, followed by the resyllabification rule. Obviously, such a solution introduces complexity in the grammar, and therefore it should be discarded.

Second, the so-called underlying glottal stops only occur in this specific environment and never in any other positions in the word. When this distributional restriction is taken into account, the phonemic status of the glottal stop is suspect. It is worth noting that the occurrence of a glottal stop in other environments, such as in the stem syllabic coda and in the onset at the suffix boundary is not lexical, but rule-derived via Debuccalisation and Glottal Epenthesis (Zaharani 1997), respectively. Surely the language games data can also be understood in terms of epenthesis.

In contrast to Teoh, we assume that the so-called vowel initial stems lexically begin with vowel segments. This suggests that Malay basic syllable structure are onsetless. The occurrence of a glottal stop in intervocalic position at the juncture of a prefix and stem is interpreted as a result of C-epenthesis, which is phonologically motivated as a resolving mechanism for breaking up the hiatus (cf. Farid 1980, Durand 1987, McCarthy and Prince 1993b).

Another case that involves a violation of ONSET is within the root domain. The examples in (19) illustrate this situation[7].

19.	/kaen/	[kaen]	'cloth'
	/naek/	[nãẽʔ]	'to ascend'
	/maen/	[mãẽn]	'to play'
	/haos/	[haos]	'thirsty'
	/laot/	[laot]	'sea'

Notice that the underlying vowel sequences are parsed heterosyllabically preserving the hiatus in the surface output, a clear violation of ONSET. Apparently, ALIGN-LEFT is irrelevant in this context, since the position occupied by the onsetless syllable is not at the edge of the word. Given the schematic ranking ONSET >> DEP-IO established thus far, we would expect the rule of epenthesis will generally apply to resolve the conflict. Nevertheless, this is not the case here. This suggests that the preservation of hiatus root internally must be due to some other formal constraint. Before we identify that particular constraint, it is important to note that there is a disagreement among linguists with respect to both the input and the output representations of the data in (20).

20. a. Yunus (1980)

/kain/	[kaen]
/naik/	[naeʔ]
/laut/	[laot]

b. Farid (1980)

/kain/	[kae̯n]
/naik/	[nae̯ʔ]
/laut/	[lao̯t]

c. Durand (1987)

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

d. Teoh (1994)

/ kaʔin/	[kaʔen]
/naʔek/	[naʔeʔ]
/laʔut/	[laʔot]

In Yunus (1980), Farid (1980) and Teoh's (1994) analyses, the underlying high vowels /i, u/ in the closed final syllable are lowered to [e, o], respectively, by the so-called Vowel Lowering. However, Zaharani (1997) argues that this rule is not phonologically motivated, and therefore it is preferable to represent the underlying vowel as a mid-vowel.

Notice that in Farid's description the derived mid-vowels [e, o] are syllabified in the margin via Marginal Vowel Derived Rule, whereas in Yunus and Teoh's they are syllabified in the nucleus. As far as the syllable position of mid-vowels is concerned, I agree with Yunus and Teoh's description. The parsing of the mid-vowels in the margin creates complex codas, and this runs against the basic syllable structure (C)V(C) proposed in Farid (1980:24).

In Durand (1987), the high vowels in closed final syllable do not get lowered into [e, o]. They remain as high vowels, but tautosyllabically parsed in the rhyme, giving rise to complex codas as in Farid (1980). According to Durand (1987:98), the Malay syllable template allows for complex onsets and codas. This assumption contradicts the general view that the basic structure of the Malay syllable is simplex (cf. Yunus *1980*, Farid *1980*, Teoh *1994*).

Contrary to Farid (1980) and Yunus (1980), Teoh's (1994) surface forms contain an intervocalic glottal stop. *This interpretation is observationally inadequate.* Teoh (1994) misinterpreted the data arguably because he wants the forms to be in agreement with his primary claim that the Malay syllable structure is CV(C), which requires an obligatory onset.

Based on my observations, I agree with Yunus's (1980) analysis that the input vowel sequences in (20) are parsed heterosyllabically, preserving the hiatus in the surface output. This observation is supported by psychological evidence from the same language game discussed earlier which involves reversing a stem syllables. Thus, words like [nã.ẽʔ] and [la.ot] are reversed into [ẽʔ.nã] and [ot.la].

We have observed that C-epenthesis is a general mechanism that the language employs in order to break up an underlying hiatus. However, this is inapplicable for the case under discussion. In what follows, we attempt to determine the. relevant constraint that rules out this possibility.

We have seen that Glottal Epenthesis (6) is used to resolve underlying hiatus at the prefix-stem juncture. We then established the schematic ranking MAX-IO >> ONSET >> DEP-IO, as demonstrated in (8). However, this ranking fails to account for the phenomenon under discussion, since it yields an incorrect result, as the following tableau shows.

21. MAX-IO >> ONSET >> DEP-IO

/kaen/	MAX-IO	ONSET	DEP-IO
a. ☺ka.en		*!	
b.☞ *ka.ʔen			
c. kan	*!		

As can be seen, (21b) is chosen as the optimal candidate, as it minimally violates the lower-ranked constraint DEP-IO. Nevertheless, the correct surface form is (21a), the candidate is marked by '☺'. This means that there must be another constraint which is crucially involved in evaluating these candidates, and this constraint must be ranked higher than ONSET. The relevant constraint is CONTIGUITY, which demands that the input and the output strings must be contiguous.

It has been observed that, in most languages, there are many phonological processes that typically apply at the edge of a grammatical constituent rather than internally. For example, in Axininca Campa and Lardil, epenthetic augmentation is external to the root (McCarthy and Prince 1993a); in Chukchee, morpheme-edge epenthesis is favoured over morpheme-internal epenthesis (Kenstowicz 1994c, Spencer 1993); in Diyari, a prohibition on syllable codas causes all consonants to be deleted word finally, but not word medially. This situation is captured by a general constraint called CONTIGUITY, which is defined in McCarthy and Prince (1995b) as follows:

22. CONTIGUITY
I-CONTIG (No Skipping)
The portion of S_1 (input) standing in correspondence forms a contiguous string.
O-CONTIG (No Intrusion)
The portion of S_2 (output) standing in correspondence forms a contiguous string.

The constraints in (22) distinguish two types of contiguity. The constraint I-CONTIG rules out internal deletion in the input string. For instance, when a string /abc/ surfaces as [ac], this violates I-CONTIG because ac is not a contiguous string. This constraint, however, is not violated if the deletion rule applies at the edge, as in /abc/ →[ab], because ab is a contiguous string. Likewise, the violation of O-CONTIG is compelled if epenthesis were to apply internally to the input string, such as /ac/ → [abc]. By contrast, epenthesis at the edge, such as /ab/ → [abc] is not. For present purposes, we don't need to distinguish these two constraints. Both epenthesis and deletion are controlled by a single general constraint called CONTIGUITY.

The question is, does Glottal Epenthesis in Malay violate CONTIGUITY? The answer can either be yes or no, depending on the grammatical constituent which constitutes the domain of the application of the rule. CONTIGUITY is violated at the word level, since the rule applies internally to the word domain. However, at the root level, CONTIGUITY is fully satisfied, since the epenthesis rule only applies at the edge of the root domain.

It has been commonly observed that a large number of disparate phonological phenomena are subject to stricter faithfulness requirements within the root than elsewhere in the word, that is, from the relative markedness of roots (cf. McCarthy and Prince 1995b). The greater markedness of roots is undoubtedly driven by the demand to sustain more contrasts between roots than between affixes. McCarthy and Prince (1995b) formalise this difference in markedness by proposing a general ranking schema in which root-specific versions of faithfulness constraints are intrinsically ranked higher than the general, or affix-specific version of the same constraint.

For the case under discussion, we need a root-specific CONTIGUITY constraint called ROOTCONTIG which bans root-internal epenthesis and deletion. ROOTCONTIG is an unviolated constraint in Malay, therefore it cannot be dominated in the hierarchy. The relevant ranking to account for ONSET violation root internally is as follows: ROOTCONTIG, MAX-IO >> ONSET >> DEP-IO.

23. Onsetless syllable root internally

/kaen/	ROOT CONTIG	MAX-IO	ONSET	DEP-IO
a ☞ka.en			*	
b. ka.ʔen	*!			*
c. kan	*!	*		

It is evident now, despite the fact that ONSET is highly ranked, it is being disobeyed in this particular environment because other possible solutions, such as Glottal Epenthesis (23b) and Vowel Deletion (23c) fatally violate the undominated constraint ROOTCONTIG. Since there are no other possible competitors in the set, the candidate with ONSET violation (23a) emerges as the winner.

4.0 Conclusion

In conclusion, Malay generally requires that every surface syllable must have a single onset, implying that the syllable structure constraint ONSET is highly ranked in the language. Despite the fact that onsets are strongly preferred, there are two instances where a surface syllable can be onsetless, namely, in word-initial and root-internal positions. The regular rule of Glottal Epenthesis never applies in this environment. The violation of ONSET is admissible here to assure the satisfaction of two undominated constraints in the hierarchy, namely, ALIGN-LEFT and ROOTCONTIG, which militate against segmental epenthesis word-initially and root-internally.

ENDNOTES

[1]It must be noted that the crucial difference between violating PARSE-SEGMENT and MAX-IO is that in the former case the unrealised surface segment is not deleted, but remains unparsed (marked by an angle bracket < >). This is due to Containment which forbids any deletion of input materials. In the latter case, however, this is interpreted as phonological deletion.

[2]Candidate (4a) would also violate Sonority Sequencing Generalisation (or Sonority Sequencing Principle (Selkirk 1984).

[3]It must be noted that in literary Malay, particularly under the new spelling system 1975 (Pedoman Umum Bahasa Malaysia) borrowed words containing consonant clusters are lexically preserved. However, in the old spelling system (Ejaan Sekolah), such clusters are not permitted.

[4]Heterosyllabic parsing of vowel sequences within a morpheme is also disfavoured in the language. Underlying clusters with prevocalic, postvocalic and intervocalic high vowels cannot be syllabified heterosyllabically. For instance, underlying /HV, /VH/ or /VHV/ (i.e. V stands for vowel and H for high vowel) cannot be parsed as [H.V],

[V.H] or [V.H.V]. Unlike in the heteromorphemic case, the optimal way of resolving vowel sequences morpheme internally is not by Glottal Epenthesis, but by parsing the high vowels in the margin (see Zaharani and Roca 1999).

[5]GCat = Grammatical Category, among which are the morphological categories Mcat = Root, Stem, Morphological Word, Prefix, Suffix, etc. PCat = Prosodic Categories = μ, σ, Ft, PrWd, PhPhrase, etc. (McCarthy and Prince 1993a).

[6]According to Farid (1980:49), the rule of Glottal Epenthesis also applies between two identical vowels morpheme internally (i.e. /saat/ seconds and /peel/ behaviour become [saʔat] and [peʔel]). Teoh (1994:86) denies this, and takes the position that these words are borrowed from Arabic with the voiced pharyngeal fricative [ʕ] occurring in the medial position, and this consonant is replaced by [ʔ] in Malay. I am in agreement with Teoh in this respect. In another case, Teoh (1994) *postulates that Malay has an underlying glottal stop intervocalically* (i.e. /kaʔin/ cloth and /laʔut/ sea become [kaʔen] and [laʔot]). By contrast, *in Farid s (1980) work there is no intervocalic glottal stop* both in the underlying and surface forms (i.e. /kain/ and /laut/ become [kaen] and [laot]). I will discuss this in more detail in this paper as we proceed.

[7]Neither k ~ ʔ alternation nor Vowel Nasalisation are relevant here.

REFERENCES

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

Clements, George N. and S. Jay Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable.* Cambridge, MA: MIT Press.

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

Farid M. Onn. 1980. *Aspects of Malay Phonology and Morphology: A Generative Approach.* Bangi: Universiti Kebangsaan Malaysia.

Kenstowicz, Michael. 1994a. *Phonology in Generative Grammar.* Cambridge: Blackwell.

Kenstowicz, Michael. 1994b. Sonority-driven stress. Ms. MIT.

Kenstowicz, Michael. 1994c. Syllabification in Chukchee: A constraint-based analysis. In Alice Davison, Nicole Maier, Glaucia Silva and Wan Su Yan, eds., *Proceedings of the Formal Linguistics Society of Mid-*

America 4. Iowa City: Department of Linguistics, University of Iowa. Pp. 160-181.

McCarthy, John J. and Alan M. Prince. 1986. Prosodic morphology. Ms., University of Massachusetts, Amherst, and Brandeis University, Waltham, Mass.

McCarthy, John J. and Alan M. Prince. 1990a. Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8: 177-208.

McCarthy, John J. and Alan M. Prince. 1990b. Prosodic morphology and templatic morphology. In M. Eid and John J. McCarthy, eds., *Perspectives on Arabic Linguistics 2: Papers from the Second Annual Symposium on Arabic Linguistics*. Amsterdam: John Benjamins. Pp.1-54.

McCarthy, John J. and Alan M. Prince. 1993a. *Prosodic Morphology 1: Constraint Interaction and Satisfaction*. Ms., University of Massachusetts, Amherst, and Rutgers University, New Brunswick. [To appear Cambridge, MA: MIT Press.]

McCarthy, John J. and Alan M. Prince. 1993b. Generalized alignment. In Geert Booij and Jaap van Marle, eds., *Yearbook of Morphology 1993*. Dordrecht: Kluwer. Pp. 79-153.

McCarthy, John J. and Alan M. Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In Merc Gonz lez, ed., *Proceedings of the North East Linguistic Society 24*. Amherst, MA: Graduate Linguistic Student Association. Pp. 333-379.

McCarthy, John J. and Alan M. Prince. 1995a. Prosodic morphology. In John Goldsmith, ed., *The Handbook of Phonological Theory*. Oxford: Basil Blackwell.

McCarthy, John J. and Alan M. Prince. 1995b. Faithfulness and reduplicative identity. In Jill N. Beckman, Laura Walsh Dickey and Suzanne Urbanczyk, eds., *University of Massachusetts Occasional Papers in Linguistcs 18: Papers in Optimality Theory*. Amherst, MA: Graduate Linguistic Student Association.

Prentice, D. J. 1971. *The Murut Languages of Sabah.* Canberra: Australian National University. [Pacific *Linguistics, Series C, no. 18.]*

Prince, Alan M. and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar.* Ms., Rutgers University, New Brunswick, and University of Colorado, Boulder. [To appear, Cambridge, MA: MIT Press.]

Roca, Iggy. 1994. *Generative Phonology.* London: Routledge.

Roca, Iggy. 1997. There are no glides, at least in Spanish: An optimality account. *Probus* 9: 233-265.

Roca, Iggy. 1997. Derivations or constraints, or derivations and constraints? In Iggy Roca, ed., *Derivations and Constraints in Phonology.* Oxford: Clarendon Press.

Rosenthall, Samuel. 1994. *Vowel/Glide Alternation in a Theory of Constraint Interaction.* PhD dissertation, University of Massachusetts, Amherst.

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

Selkirk, Elisabeth. 1982. Syllables. In H. van der Hulst and N. Smith, eds., *The Structure of Phonological Representation, vol.2.* Dordrecht: Foris. Pp 337-383.

Selkirk, Elisabeth. 1984. On the major class features and syllable. In Mark Aronoff and Richard T. Oerhle, eds., *Language Sound Structure.* Cambridge, MA: MIT Press.

Spencer, A. 1993. The optimal way to syllabify Chukchee. Talk presented at Rutgers Optimality Workshop I, Rutgers University, New Brunswick, NJ.

Teoh Boon Seong. 1988. *Aspects of Malay Phonology Revisited: A Non-linear Approach.* Ph.D. dissertation, University of Illinois at Urbana-Champaign.

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

Yunus Maris. 1980 [1966]. *The Malay Sound System.* Kuala Lumpur: Fajar Bakti.

Zaharani Ahmad. 1997. *Phonology and Morphology Interface in Malay: An Optimality Account.* PhD dissertation, University of Essex.

Zaharani Ahmad and Iggy M. Roca. 1999. High vowel syllabification in Malay vowel sequences: an optimality-theoretic account. *Essex Research Reports in Linguistics* 24: 46-70.