FUSION AND ALIGNMENT: THE CASE OF MALAY[1]

Ann Delilkan
New York University
<ad17@nyu.edu>

1 Issues

1.1 *The Original Problem*

The facts of prefixation in Malay, in the case of nasal-final prefixes (e.g 'meN-' (transitive)), follow in (1). (Roots are underlined.)

1.	a. məN + pukol	[məmukol]	'hit', v.t.
	b. məN + kupas	[məŋupɑs] ('mengupas')	'peel', v.t.
	c. məN + təntu + kan	[mənəntukɑn]	'determine', v.t.
	d. məN + buwat	[məmbuwat]	'make', v.t.
	e. məN + garam	[məŋgaram] ('menggaram')	'salt', v.t.
	f. məN + dəsu	[məndəsu]	'resound', v.i.

Rule-ordering analyses of this data would claim that that nasal substitution in Malay involves the prefix-final nasal (N) assimilating to the place of articulation of the first sound of the following root (Rule 1), followed by deletion of that onset obstruent in the case of voiceless-obstruent-initial roots (e.g. /p/, /t/, /k/) (Rule 2). Thus, for instance, [məmukol] ('mempukul', *hit, v.t)* is correct, but *[məmpukol] is ungrammatical.

1.2 *OT and the Problem*

Such an analysis of the facts raises the following concern. Assimilation and deletion are two separate and unrelated processes, and there is no lack of cross-linguistic data to demonstrate this fact. Stipulating that the Malay data represents the product of a two-step process is not explanatorily adequate because it fails to explain why the processes occur together.

Optimality Theory avoids such stipulation, since it does not support a processual analysis of the data. In Optimality Theory, possible pronunciations (output candidates) of a word (input) are evaluated simultaneously against a set of ranked constraints. The pronunciation that incurs the fewest violations of the constraints is selected as the 'optimal' one (See Appendix A for the fundamental tenets of Optimality Theory). A recent Optimality theoretic treatment (Pater 1996) involves claiming that 'fusion' is a more desirable account of the above prefixation facts. A fusion version of məN$_1$ + p$_2$ukol follows in (2), where a two-to-one mapping is seen to hold between the two input segments, N$_1$p$_{2,}$ and the single output segment, m$_{1,2}$.

[1] I am grateful to Keith Fernandes, Adamantios Gafos, Fran Gulinella, Harry van der Hulst, Young-Kook Kwon, and Nancy Ritter for their feedback and kind support, although I hasten to add that the material in this paper may not necessarily reflect their views.

Mark Alves, Paul Sidwell & David Gil, eds. *SEALSVIII: papers from the 8th meeting of the Southeast Asian Linguistics Society (1998)*. Canberra, Pacific Linguistics, 2007, pp.45-55.

2. input m əN_1 + $\underline{p_2\ u\ k\ o\ l}$
\ /
output m ə $m_{1,2}$ u k o l

Pater's one-step analysis is meant to eliminate the explanatory gap inherent in two-step analyses of the Malay facts. Pater sees fusion as one of many 'repair' strategies adopted by various languages to avoid the sequence Nasal + Voiceless Obstruent (henceforth, 'NT').[2] He cites the faithfulness constraint LINEARITY (McCarthy and Prince 1995), which stipulates that S1 reflect the precedence structure of S2, and vice versa, where S1 is the input and S2 the output. (See Appendix B for full statement of all constraints employed in this paper). Bearing this in mind, consider once again the facts of 'fusion', depicted in (3). (The segment resulting from fusion has two indices associated with it.)

3. məN_1 + $\underline{p_2\text{ileh}}$ (input)
\ /
mə $m_{1,2}$ ileh (output) (Pater 1995:6)

As explained by Pater, N_1 precedes p_2 in the input, but not so in the output. This means that the output violates LINEARITY. The ranking that Pater thus proposes is responsible for the output selected is stated in (4).

4. *NT > LIN (where LIN = LINEARITY)

The tableau in (5) derives from Pater 1996, and shows the result of this ranking. ('✓' denotes the optimal candidate.) The candidate in (5b), in which no fusion has occurred, violates *NT and is eliminated. Fusion, in (5a), violates LINEARITY but obeys the more high-ranked *NT, and is therefore the preferred option.

5. **Fusion: *NT >> LIN (Pater 1996:9, #7)**[3]

Input: mə N_1+$\underline{p_2\text{ileh}}$	*NT	LIN
a. mə$m_{1,2}$ ileh ✓		*
b. məm_1 p_2 ileh	* !	

Root-internal data in Malay does not, however, seem to involve fusion. (See Appendix C.c) Pater accounts for this fact by claiming that it is more undesirable to allow fusion root internally than it is to permit the undesirable sequence, but that the constraint against the sequence, *NT, is a more powerful constraint than a constraint (LINEARITY-ROOT) that militates against non-root-internal fusion. Thus the constraints responsible for

[2] Pater finds that other attested strategies are nasalization of the obstruent, denasalization of the nasal, and deletion of the nasal.

[3] For typological simplicity, I have chosen to use neither Pater's 1996 *N constraint nor his 1995 version, *NC̥ but *NT. The check denotes the optimal candidate

selection of the optimal candidate in this analysis include both a root-specific and non-root-specific version of a constraint against fusion. The root-specific constraint outranks *NT, which in turn outranks the general LINEARITY constraint. The relevant ranking is expressed in (6). The tableau in (7) illustrates the selection made. Candidate (7a) fails to get selected because the high-ranked root-specific LINEARITY is violated since fusion has occurred within a root. The non-fusion candidate, (7b), which violates the lower-ranked *NT, is selected instead.

6. LINEARITY-ROOT >> *NT >> LINEARITY

7. **Root-internal NT tolerance: LINROOT >> *NT>> LIN**

Input: <u>əm$_1$p$_2$at</u>	LINROOT	*NT	LIN
a. əm$_{1,2}$ at	* !		*
b. ə m$_1$p$_2$ at✓		*	

(Pater 1996:10, #8 and Pater 1995:8)

1.3 *Further Complications*

Pater writes, 'Clearly a lot of work needs to be done to determine the empirical coverage of root-specific LINEARITY constraints' (Pater 1995) — and it does. There is further data in Malay that fails to be accounted for by this analysis. Specifically, the boundaries a) between prefixes, b) between suffixes, c) between reduplicants, and d) between roots and suffixes do not pattern along the dichotomy Pater suggests. (See Appendix C for data in question.) Thus, for instance, fusion does not occur between prefixes, but there is nasal place assimilation, as shown in (8). (Again, italics denote the result of fusion.)

8. məN + pər + <u>səmbah</u> (+kan)
məmpərsəmbahkan, (*mə*m*ərsəmbahkan)
\ |
lab 'present (verb)'+ performative

These facts mirror the root-internal facts of the language. In the environments (b-d) (above), neither fusion nor place assimilation is in evidence. Taken together with the prefix-root boundary, which is the only spot where fusion does seem to occur, these facts indicate the need for some new explanation.

2 Analysis

2.1 *A New Dichotomy*

I maintain that parameterizing a fusion constraint according to environment would fail to capture the generalization underlying all the environments, and would be theoretically uneconomical. I suggest instead that there is a dichotomy between sequences of segments that are homogenous with respect to morphological category-type (e.g. root-internal, prefix-prefix, suffix-suffix, reduplicant-reduplicant) and those that are heterogenous with respect to morphological category type (e.g. prefix-root, root-suffix). Fusion is blocked in the former type of sequences, but may be permitted in the latter. I take my cue from

McCarthy and Prince (1995), extending their UNIFORMITY constraint against coalescence (fusion) by specifying the domain of it application as in (9b).

9. a. UNIFORMITY (M&P 19995): S2 may not have multiple correspondents in S1.
   b. UNIF-HMG: S2 may not have multiple correspondents in S1, where S1 an environment that is homogenous with respect to morphological category-type

The ranking of constraints that I propose follows in (10).

10. UNIF-HMG >> *NT

I omit irrelevant candidates (e.g. post-nasal voicing, epenthesis, etc.) in (11), which illustrates the result of the ranking in (10).

11. **Prefix-prefix juncture: No fusion**

Input: $m\text{ə}N_1$ + $p_2\text{ə}r$ + (<u>buwat</u>)	UNIF-HMG	*NT
a. $m\text{ə}m_1p_2\text{ə}r$.. ✓		*
b. $m\text{ə}\mathit{m}_{1,2}\text{ə}r$…	*!	

The fusion candidate, in (11b), violates UNIF-HMG and is eliminated. The selected non-fusion candidate, in (11a), honors this constraint although it violates the low-ranked *NT. All other environments are tested in (12), using the same constraint ranking as in (11). The ranking. selectes grammatical outputs in all three cases. Root-internally (in (12a)) and between suffixes (in (12c)), fusion is blocked because it would involves violating the high-ranked UNIF-HMG. The opposite holds in the heterogenous environment that obtains at the prefix-root juncture, fusion selected over an *NT violation.

12. a. **Root-internally: No fusion**

Input: <u>$t\text{ə}m\ {}_1p\ {}_2at$</u>	UNIF-HMG	*NT
i. $t\text{ə}m_1p_2at$ ✓		*
ii. $t\text{ə}\mathit{m}_{1,2}at$	*!	

b. **Prefix-root juncture: Fusion**

Input: $m\text{ə}N_1$+ p_2ileh	UNIF-HMG	*NT
i. $m\text{ə}\mathit{m}_{1,2}ileh$✓		
ii. $m\text{ə}m_1p_2ileh$		*!

c. **Suffix-suffix juncture: No fusion**

Input: ..+kan_1+ k_2ah	UNIF-HMG	*NT
kan_1k_2ah✓		*
$ka\mathit{ŋ}_{1,2}ah$	*!	

2.2 *Fusion And Alignment*

One section of the data still needs to be accounted for, as shown in (13).

13. ***EXCEPTION: Root-Suffix**

Input:<u>rakam</u>$_1$+k$_2$an	UNIF-HMG	*NT
a. rakam$_1$k$_2$an		*
b. *raka$\eta_{1,2}$an✓		

The posited ranking selects the ungrammatical candidate in (13). The fact that fusion fails to occur at the root-suffix boundary is, I suggest, the result of an interaction of the constraint against fusion with an alignment constraint that protects the integrity of left edges of affixes more than it does their right edges. The alignment constraint in question follows in (14).

14. ALIGN Affix L, PrWd R: The left edge of every affix must be aligned with the right edge of some prosodic word.

Fusion at the boundary referred to in (14) would constitute a violation of this constraint, as it would 'damage' the edge in question. For fusion to be prohibited at this juncture, *NT must be ranked lower than the alignment constraint, as shown in (15). (The comma between UNIFHMG and ALIGN AffixL, PrWdR indicates that these two constraints are not ranked with respect to each other.) The tableau in (16) shows the results of this ranking.

15. UNIF-HMG, ALIGN Affix L, PrWd R >> *NT

16. **Root-suffix juncture: No fusion**

input: <u>rakam</u>$_1$+ k$_2$an	UNIF-HMG	ALIGN Affix L,PrWd R	*NT
a. [[rakam$_1$] k$_2$an][4] ✓			*
b. [[raka]$\eta_{1,2}$ an]		*!	

The fusion candidate, in (16b), violates the alignment constraint and is dispreferred even though it avoids violation of UNIF-HMG and *NT. The preference is for candidate (16a), in which the root right edge does not interact with the suffix left edge, despite its violation of the low-ranked *NT.

The tableau in (17) now illustrates all environments tested against this ranking. (Roots are underlined.) Within roots (cf. (17a)), between prefixes (cf. (17c)), between

---

[4] I assume exhaustive prosodization (McCarthy and Prince 1993) of all morphemes into recursive prosodic words.

suffixes (cf. (17e)), and between roots (cf. (17f)) fusion is blocked by the high-ranked UNIF-Hmg. In each case, a low-ranked *NT violation is preferable. Between a root and a suffix (cf. (17d)), fusion is blocked to avoid a costly violation of the alignment constraint. Fusion is selected in (17b) because it avoids violaton of all three constraints, in preference of a *NT violation.

17. **All other environments**

	UNIFHMG,	ALIGN (Aff L, PrWd R)	*NT
a. Input: təm_1p_2 at			
i. tə$m_1p_2$at √			*
ii. tə$m_{1,2}$at	*!		
b. Input: məN$_1$ + p_2 ileh			
i. məm_1p_2 ileh			*
ii. mə$m_{1,2}$ ileh √			
	UNIF-HMG	ALIGN(Aff L,PrWd R)	*NT
c. Input: məN$_1$+ p_2ər (+ bu..			
i. məm_1p_2ər .. √			*
ii. mə$m_{1,2}$ər ..	*!		
d. Input: rakam_1 + $k_2$an			
i. rakam_1]k_2 an] √			*
ii. raka] ŋ$_{12}$an][5]		*	
e. Input: … + kan_1 + $k_2$ah			
i. …] kan_1]k_2 ah] √			*
ii. …]ka]ŋ$_{12}$ah]	*!	*	
f. Input: kawan_1 + $k_2$awan[6]			
i.[[kawan_1] $k_2$awan]√			*
ii. [[kawa] ŋ$_{1,2}$awan]	*!	*	

2.3 *Prosodic Implications*

If the left edge of an affix must be aligned perfectly alongside the right edge of a prosodic word, this suggests that the way in which prosodic words are built in Malay is [[prefix+root]+suffix], i.e. with the prefix and root forming a complex. This possibilty runs

5 I assume the insertion of prosodic word boundaries occurs to show the grouping I propose preliminarily in 2.3, namely that suffixes fall outside the prosodic word occupied by the root and prefix(es) I assume also that each suffix projects its own prosodic word.

6 I assume reduplication is suffixal.

counter to the widely-held view that prosodic words for such a language are built with the root and suffix forming a complex, as shown in (18a). The new prosodic structure I propose is repeated in (18b).

18. a. Traditional analyses: [prefix + [root+ suffix]]
    b. Proposed prosodic structure: [[prefix + root] + suffix]

Nasal assimilation facts in Malay seem to pattern along this tendency in the language to 'protect' the left edges of its affixes, lending credence to the implied asymmetry between left and right edges of affixes. The alignment constraints used thus far interact with not only *NT but also a constraint that penalizes codas that are not place-linked to their adjacent onsets (cf. CODA CONDITION (Ito and Mester 1994)). As stated in (19), the alignment constraint outranks both the Coda Condition and *NT constraints. The results of this ranking. appear in (20)

19. ALIGN AFFIX L, Pr Wd R >> CODACOND >> *NT

20. **Alignment outranks CodaCond**

Input: <u>rakam</u> + kan | | lab vel	ALIGN AFF L,PrWd R	CODA COND	*NT
a. [[rakam] kan] √ | | lab vel		*	*
b. [[raka ŋ] kan] \ | vel	*!		

Place assimilation, in candidate (20b), violates the high-ranked alignment constraint. The candidate in which no interaction occurs between the right edge of the lower prosodic word and the left edge of the suffix is preferable, therefore, despite its violation of CODACOND and *NT.

In (21), all other environments are tested against this ranking. ((In (21b), I omit candidates that involve variations in the juncture between the root and the first suffix.)

21. **Other environments: Alignment outranks CODACOND**

a. Input: məN + pər + <u>buku</u> + kan | lab	ALIGN AffL,PrWdR	CODA-COND	*NT
i. [[məm.pər.<u>bu.ku</u>] kan] √ \ | lab			*
ii. [[məŋ.pər. <u>bu.ku</u>] kan] | | vel lab		*!	
Input: <u>tuhan</u> + kan + kah | | cor vel			
i. [[[tuhan] kan]kah] √ | | cor vel		*	*
ii. [[[tuhan]kaŋ]kah] \ | vel	*!		

c. Input: <u>əmpat</u> | | lab lab	ALIGN Aff L,PrWdR	CODA- COND	*NT
i. əm.pat] √ | | lab lab			*
ii. [əŋ.pat] | | vel lab		*!	*

In (21a), nasal assimilation between the prefixes avoids violation of CODACOND and cannot be blocked by the alignment constraint as the relevant boundary for its application is not involved. The candidate in which, for instance, the prefix final nasal assumes a velar place specification, falls afoul of CODACOND and is therefore dispreferred. Candidate selection in (21c) parallels that in (21a). In (21b), the alignment constraint blocks selection of a place assimilated candidate (cf. (21bii), the candidate in which no interaction occurs between the suffixes selected instead, despite its CODACOND violation.

3 Conclusion, implications and direction for further research

The claims of this paper are two-fold. First, it shows that the data under investigation are less irregular than they seem. There is a dichotomy amongst the data that is captured by referring to a constraint like UNIF-HMG, which governs behavior of segments according to their morphological affiliation. This account of the data is achieved without parameterization of the constraint according to every environment involved.

Second, I have shown that apparent exceptions to this are the result of asymmetrical strength of edges. I maintain that the correct edges to consider require looking from the perspective of affixes, not roots.

Significantly, this last claim involves a commitment to the way in which prosodic words are built in Malay such that the prefix and root form a complex, with suffixes falling outside in separate prosodic words.. Nasal assimilation facts support the posited asymmetry between prefixes and suffixes, but an investigation of stress assignment would be vital to determine the strength of the current claim about prosodic word formation, and to motivate the boundaries involved more thoroughly. It would also be crucial to determine whether any other segmental processes in the language interact with the proposed boundaries. Certainly, too, much work still needs to be done to determine, for instance, whether a re-ranking of the constraints under review will account for dialect differences in Malay, before it can be ascertained how far an OT treatment of such data improves upon previous analyses.

Appendix A
Fundamentals of OT (From McCarthy and Prince 1993a)

a) Universality. Constraints are universal; and universally present in all grammars.
b) Violability Constraints are violable, but violation is minimal.
c) Ranking Constraints are ranked on language-particular basis: The notion of minimal violation is defined in terms of this ranking. A grammar is a ranking of a universal constraint set.
d) Inclusiveness The constraint hierarchy evaluates a set of possible candidates that are admitted by very general considerations of structural well-formedness.
e) Parallelism Best-satisfaction of the constraint hierarchy is computed over the whole hierarchy and the whole candidate set. There is no serial derivation.

An Optimality-Based Grammar, schematically:

Gen(in) = (cand 1, cand 2…)
Eval ({cand 1, cand 2..})→ cand n (the output)

GEN determines the set of candidate analyses consistent with a given input.(GEN may freely delete, insert or link segments as well as assign prosodic structure). GEN is conceived as a function that, for each particular input form, generates the range of all possible candidate linguistic analyses. These candidates are evaluated simultaneously by a function EVAL. EVAL rates the members of the candidate set in terms of their relative harmony, or degree of success with regard to the language's ranking of the constraints, imposing an ordering on them. A maximally harmonic member of the candidate set is optimal.

Appendix B
Constraint list

*NC̥ (Pater 1995): No nasal-voiceless obstruent sequences (Replaced in Pater 1996 by *N, and referred to in this paper as *NT)

LINEARITY (McCarthy and Prince 1995): S_1 reflects the precedence structure of S_2, and vice versa.

MAX (McCarthy and Prince 1995): Every element of S1 has a correspondent in S2. "MAX allows an interpretation of fusion as a two-to-one mapping from Input to Output: Two Input segments stand in correspondence with a single Output segment" (Pater 96: 7)

UNIFORMITY (McCarthy and Prince 1995): No element of S2 has multiple correspondents in S1, S1 the input, S2 the output.

UNIF-HMG: No element of S2 has multiple correspondents in S1, where elements of S1 are homogenous with respect to morphological category-type, S1 the input, S2 the output.

ALIGN Affix L, PrWd R: The left edge of every affix must be aligned with the right edge of some prosodic word.

ALIGN Affix R, PrWd L: The right edge of every affix must be aligned with the left edge of some prosodic word.

CODA COND (Ito and Mester 1994): Codas are disallowed unless linked to a following onset.

Appendix C
DATA (a-b, Pater 1995/6)

a. Nasal substitution (NT) b. Failure of substitution (ND)

məN+ pileh	məmileh	'choose', v.t.	məN + bəli	məmbəli	'buy', v.t
məN + tules	mənules	'write', v.t.	məN + dapat	məndapat	'get', v.t
məN+ kesah	məŋesah	'relate', v.t.	məN + ganti	məŋganti	'(ex)change', v.t

c. Failure of substitution root-internally (N + any T)

təmpat 'place', n.	tambah 'add on', v.t.
hantar 'send', v.t.	təndaŋ 'kick', v.t
muŋken 'poss./maybe', adj.	tiŋgal 'live/remain/die/leave', v.i.

d. Failure of substitution elsewhere (new DATA)

i) məN + pər + hamba (+kan)	məmpərhambakan, *mə*mə*rhamba.. 'enslave', v.t.
ii) məN + tər + balek + kan	məntərbalekkan, *mə*nə*rbalekkan 'overturn', v.t.
iii) rakam +kan	rakamkan, *raka*ŋ*an 'record', v.t.
iv) di + tuhan + kan + kah	dituhankankah, *dituha*ŋ*a*ŋ*ah 'deified', pass/attrib. + interrogative

References

Cohn, Abigail & John McCarthy. 1994. *Alignment and parallelism in Indonesian*. MS, Cornell University and University of Massachussetts, Amherst.

Ito, Junko and Armin Mester. 1994. Reflections on Coda Conditon and Alignment. In Jason Merchant, Jaye Padgett and Rachel Walker (eds.) *Phonology at Santa Cruz (PASC)* III:27-46.

Ito, Junko and Armin Mester, and Jaye Padgett. 1995. Licensing and Underspecification in Optimality Theory. *Linguistic Inquiry* 26:571-614.

McCarthy, John and Alan Prince. 1995. Faithfulness and Reduplicative Identity. In Jill Beckman, Laura Walsh Dickey and Susan Urbanczyk (eds.) *University of Massachusetts Occasional Papers: Papers in Optimality Theory* 18:249-384.

-----. 1993a. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ, MS. [Forthcoming MIT Press].

-----. 1993b. Generalized Alignment in Geert Booij -Jaap van Marle (eds.), *Yearbook of Morphology 1993*. Dordrecht, Kluwer: 79-153

Onn, Farid M. 1980. *Aspects of Malay Phonology and Morphology*. Bangi: Universiti Kebangsaan Malaysia. [Published version of 1976 University of Illinois dissertation].

Pater, Joe. 1996. Austronesian Nasal Substitution and other N effects. To appear in Kager, René, Harry van der Hulst, and Wim Zonneveld (eds.), *The Prosody Morphology Interface*. Cambridge University Press

Pater, Joe. 1995. *Austronesian Nasal Substitution and other NC effects.* Draft of paper presented at proceedings of the Utrecht Prosodic Morphology Workshop.

Prince, Alan S. and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. MS, Rutgers University and University of Colorado at Boulder. [Forthcoming MIT Press].