To appear in:
de Lacy, Paul & Anita Nowak (eds.) (1999) *University of Massachusetts Occasional Papers: Papers from the 25th Reunion*. Amherst, MA: GLSA.

# Morphological Haplology and Correspondence*

Paul de Lacy

University of Massachusetts, Amherst

## 1 Introduction

When the Arabic feminine singular morpheme *ta* attaches to the verbal prefix *ta*, the resulting form is not the expected **tata*, but instead *ta*:

(1) ta + ta + kassaru → takassaru *it (fem.sg.) breaks*, *tatakassaru (Wright 1971: 65)

This is a typical example of morphological haplology: while there are two phonologically identical morphemes underlyingly, only one phonological string appears in the surface form (Stemberger 1981, Menn & MacWhinney 1984).[1] There are two fundamental questions about morphological haplology (hereafter 'haplology') that any analysis of the phenomenon must address:

- What is the nature of haplology?
- What triggers haplology?

* I am particularly grateful to John McCarthy for his many comments, and to Ingo Plag, Moira Yip, and Brett Hyde for their input. I hasten to add that they are not to be blamed for the proposals in this paper as I frequently failed to take their advice. Also thanks to Emmon Bach, Patrik Bye, Ed Keer, Wayne Lawrence, Steve Parker, Lisa Selkirk, Jennifer Smith, and the audience at the UMass 25th Reunion poster session. This work was supported by the National Science Foundation under grant SBR-9420424.

[1] Phonological haplology is basically a diachronic process (Hock 1991:109, Grammont 1933), and is the idiosyncratic deletion of similar sequences of morpheme-internal material.

In response to the first question, it is proposed that haplology is coalescence. In other words, underlying phonological material of different morphemes merges in the output – there is no deletion.

The most popular answer to the second question is that there is a OCP-like constraint that bans adjacent identical strings (e.g. Menn & MacWhinney 1984, Golston 1995, Plag 1998, Yip 1998, forthcoming). This view is rejected in this paper. Instead, it is argued that haplology can triggered by any relevant markedness constraint.[2] Furthermore, it is argued in §4 that there is no OCP-like constraint that directly bans adjacent identical strings.

These two ideas – that haplology is coalescence and it is triggered by any relevant markedness constraint ℂ – are given formal expression in terms of the Optimality Theoretic constraint ranking ‖MAX » ℂ » UNIFORMITY‖. This ranking requires underlying segments to be preserved (obeying MAX), while allowing them to coalesce (violating UNIFORMITY) in order to produce an output with less structure (so satisfying ℂ). A variety of constraints interact with this ranking to produce a typology of haplological processes.

This paper is organised as follows: In §2, a proposal for dealing with haplology in Optimality Theory is presented. §3 provides empirical evidence that haplology is coalescence. In §4, the constraint ℂ that triggers haplology is discussed, and alternatives are examined. This is followed by a typological investigation of haplological processes in §5. It is shown that there are many different types of haplology, including partial-identity haplology (§5.1), reduplicative haplology (§5.2), stem-edge and non-local haplology (§5.4.1,5.4.3), and coextensive haplology (§5.4.2). Alternative theories are examined in §6, and conclusions are presented in §7.

## 2 Haplology and Correspondence

One of the principal proposals of this paper is that haplology is coalescence. In other words, when two input strings such as Arabic /ta/ {fem.sg.} and /ta/ {verbal prefix} haplologize to form [ta], neither actually deletes; instead, the output [ta] is the simultaneous realisation of both input strings. To formally express this coalescence, McCarthy & Prince's (1995, 1997) Correspondence Theory will be employed.

In Correspondence Theory, a relation called 'correspondence' holds between input and output segments. A number of constraints regulate correspondence relations. For coalescence, the following are crucial:

(2) MAX "Every segment in the input corresponds to a segment in the output."
UNIFORMITY "No segment in the output corresponds to more than one segment in the input."

[2] 'Relevance' is defined in §4.3. In short, a constraint is relevant if it is violated by the haplologizing morpheme. For example, *coronal is relevant for *ta*, and so can trigger *ta* haplology, but *labial is not.

As established by McCarthy (1995), McCarthy & Prince (1995), and Lamontagne & Rice (1995), if MAX and a coalescence-triggering constraint ℂ are ranked above UNIFORMITY, coalescence occurs:

(3) **The Basic Ranking**

| | $/t_1a_2/ + /t_3a_4/$ | MAX | ℂ | UNIFORMITY |
|---|---|---|---|---|
| | (a) $t_1a_2t_3a_4$ | | x! | |
| | (b) $t_1a_2$ | x x! | | |
| ☞ | (c) $t_{1,3}a_{2,4}$ | | | x x |

The subscript numbers indicate correspondence relations. For example, the *t* of the leftmost input string corresponds to the first segment in candidate (a), as indicated by the subscript '1's. Similarly, both input /t/'s correspond to the [t] in candidate (c), as indicated by the subscript 1 and 3.

Candidate (a) fails because it has not undergone coalescence, fatally violating the coalescence-triggering constraint ℂ (the identity of which will be discussed below). In candidate (b), one of the input strings has no correspondent (i.e. has been deleted). Because of this, not every input segment has an output correspondent, so MAX is fatally violated. Candidate (c)'s only failing is that an output segment corresponds to more than one input segment, violating UNIFORMITY. However, since UNIFORMITY is ranked below the other constraints, its violations are inconsequential. This means that (c) is the most harmonic form.[3]

From this tableau, it is evident why MAX and ℂ must dominate UNIFORMITY: if the rankings were reversed candidate (c)'s UNIFORMITY violations would be fatal, and either deletion – (b) – or full-realisation – (a) – would result, depending on the ranking of MAX and ℂ. So, the ranking of MAX and ℂ above UNIFORMITY is crucial for coalescence, and therefore for haplology.

This leaves the identity of the constraint that triggers haplology – ℂ. The position advocated here is that ℂ is any markedness constraint, including constraints such as *F (F is a feature), ALIGN (McCarthy & Prince 1993a), and others. This is discussed in detail in §4.3. For the moment, to give some substance to ℂ, the constraint *STRUC will be used (Prince & Smolensky 1993: fn.13, Zoll 1993, 1996). *STRUC militates against structure, incurring a violation for every node in the output form. It is used here because it is the markedness constraint *par excellence* – in banning structure it directly expresses a property that is shared by almost all markedness constraints. The difference between it and other markedness constraints is the extent of discrimination: while *STRUC bans all structure, other markedness constraints are more selective (e.g. *labial only bans instances of the feature [labial]). While *STRUC will be used in the following discussion, it must be remembered that *any* markedness constraint can be a potential trigger for haplology, as

[3] Keer (1999) proposes that UNIFORMITY does not exist. If so, there would have to be a constraint on morphological affiliation to prevent haplology in every situation: a segment cannot belong to more than one morpheme in the output – i.e. MORPHDIS (McCarthy & Prince 1995:§3.8). With this constraint, the present approach translates straightforwardly into Keer's terms.

long as it is relevant (see §4.3). This proposal competes with one in which ℂ *directly* induces haplology: such a constraint would prohibit adjacent identical phonological strings (i.e. an OCP-like constraint). It is argued in §4 that no such constraint exists.

In the tableaux in following sections, only violations of *STRUC that candidates do not have in common are shown in order to conserve space.

The haplology-triggering constraint has an effect on the basic ranking. MAX must outrank *STRUC, otherwise deletion would occur:

(4)

| /kæt/ | MAX | *STRUC |
|---|---|---|
| (a) ∅ | x x x! | |
| ☞ (b) kæt | | x |

Candidate (a) fails because it has no phonological content (symbolised by ∅); since no input segment has an output correspondent, it fatally violates MAX. Candidate (b) is the most harmonic form since it satisfies MAX, despite violating *STRUC. It is obvious that reversing the rankings would have a very harmful effect: if *STRUC outranked MAX, no form would ever be able to surface with phonological content. Constraints other than *STRUC have similar effects: ranking the trigger constraint above MAX causes deletion.

In summary, the Basic Ranking for haplology is as follows:

(5)

Basic Ranking for Morphological Haplology:
MAX » ℂ » UNIFORMITY
where ℂ is a relevant markedness constraint.

This account is not quite complete, though. As it stands, the Basic Ranking allows haplology between any two affixes, whether they are identical or not. This is because MAX simply requires an input segment to have an output correspondent – it does not require the correspondents to be identical. Identity is enforced by another constraint – IDENT-F:

(6) IDENT-F If an input segment is αF then its output correspondent is αF.
(i) F is a feature
(ii) α is a featural specification (+ or -)

With IDENT-F ranked above ℂ, only identical strings can coalesce. This is illustrated in the tableau below:

(7) **The Basic Ranking**

| ❶ | $/t_1a_2/ + /t_3a_4/$ | MAX | IDENT-F | ℂ | UNIFORMITY |
|---|---|---|---|---|---|
| | (a) $t_1a_2t_3a_4$ | | | x! | |
| | (b) $t_1a_2$ | x x! | | | |
| ☞ | (c) $t_{1,3}a_{2,4}$ | | | | x x |
| ❷ | $/t_1a_2/ + /d_3i_4/$ | | | | |
| ☞ | (a) $t_1a_2d_3i_4$ | | | x! | |
| | (b) $t_1a_2$ | x x! | | | |
| | (c) $t_{1,3}a_{2,4}$ | | x x! | | x x |

Haplology between /ta/ and a featurally identical string is shown in ❶. As shown in tableau (3), the coalesced candidate (c) fares better than any other here. The most important point is that it does not violate IDENT-F; even though $[t_{1,3}]$ in candidate (c) corresponds to both $/t_1/$ and $/t_3/$ in the input, IDENT-F is satisfied since the output segment has the same featural specifications as its input correspondents. The same is not true of the coalesced candidate in ❷.

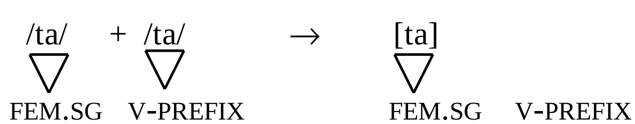
A case of non-identical strings is presented in ❷. In contrast to ❶, the coalesced candidate (c) in ❷ violates IDENT-F. This is due to the fact that the candidate does not preserve all the features of both its input correspondents. The crucial feature here is [voice]: while output $[t_{1,3}]$ preserves the [-voice] feature of its input correspondent $/t_1/$, it disagrees in voicing with its [+voice] input correspondent $/d_3/$. This demonstrates why coalescence of non-identical strings does not occur: doing so necessarily entails violation of IDENT-F. The workings of IDENT-F are discussed in detail in §5.1.

This section has presented the outline of a Correspondence-based theory of haplology. Before considering its typological implications (§5), the two premises on which it is based will be examined: the following section will provide evidence that haplology is coalescence, and §4 will advocate the idea that haplology can be triggered by any relevant markedness constraint, and not one that directly bans adjacent identical strings.

## 3 Haplology is Coalescence

The deletion and coalescence approaches to haplology make different predictions about the morphological affiliation of output segments. In the deletion approach, phonological material of one of the morphemes is not present in the output while the other's is preserved. Hence, there is no phonological realisation of one of the underlying morphemes in the surface form. This is illustrated in the diagram below; this shows the output of the concatenation of the Arabic morphemes *ta* {fem.sg} and *ta* {verbal prefix}:

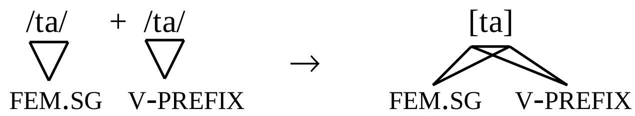
(8)



Here, the phonological content of the verbal prefix is not present in the output (the *fem.sg.*'s phonological material could equally have been deleted – the choice is arbitrary).

In comparison, with coalescence the phonological contents of both morphemes are present in the output. However, since they have merged, the resulting phonological string is affiliated to both morphemes:

(9)




The different output structures produced by deletion and coalescence provide a way to determine which strategy is being used. Specifically, constraints that make reference to a segment's morphological affiliation are predicted to apply differently in the two cases. This is shown to be a testable prediction in the following sections, with evidence adduced from Japanese (§3.1) and French (§3.2).

### *3.1 Japanese*

Lawrence (1997) argues that haplology involving the Japanese Classical Predicative (CP) suffix *-si* [ʃi] is coalescence, and not deletion.[4,5] The following reproduces the essentials of his argument, with slight modifications.[6] The reader is referred to Lawrence (1997) for further details.

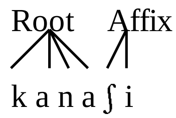
As established in the preceding section, the deletion and coalescence approaches produce different output structures. These two proposals result in three possible outcomes for haplology: (1) deletion of root material, (2) deletion of affixal material, and (3) coalescence. These are graphically illustrated below by the haplologized form [kanaʃi], from /kanasi/ 'sad' + /si/:

[4] Thanks to Wayne Lawrence for bringing this to my attention, and for subsequent discussion. Thanks to Jennifer Smith for elucidating various aspects of Japanese accentuation. For further discussion of Lawrence's approach to haplology, which is somewhat different from the present proposal, see §6.

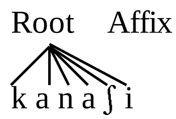
[5] In Japanese, coronal sibilants palatalise before /i/, hence /si/ → [ʃi], and /zi/ → [ʒi].

[6] The modification is in regard to the conception of the accented *vs* non-accented distinction in Japanese roots. Lawrence divides roots into those that permit accentuation *vs* those that prohibit it. The view taken here is that roots that permit accentuation have an underlying accent, while those that prohibit accentuation do not.

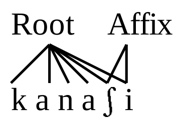
(10) Root-Deletion:




Affix-Deletion:




Coalescence:




Accent placement allows these three alternatives to be distinguished. In Japanese, words may surface as either accented or unaccented. Surface accents have two sources: either they exist underlyingly, or they are supplied by an affix. The CP suffix is such an affix: when suffixed to an underlyingly unaccented root, -*si* induces an accent to appear on the syllable that immediately precedes it (i.e. preaccentuation): e.g. /aka/ 'red' + /si/ *CP* → [akǎʃi].

The Root-Deletion, Affix-Deletion, and Coalescence approaches make different predictions about where the CP's accent will fall when it haplologizes. In the case of the underlyingly unaccented /kanasi/, for example, the accent is predicted to fall on the word-penultimate syllable in the Root-Deletion and Coalescence cases. This is evident from the representations above since the word-penultimate syllable immediately precedes the affixal material. This is the correct prediction, producing the attested form [kanǎʃi].

In comparison, the Affix-Deletion approach does not predict this outcome. If the affix's phonological material is deleted no accent should appear at all, producing *[kanaʃi]. The alternative is to suppose that this is a case of opacity: preaccentuation takes place before the affix is deleted. However, this approach makes the wrong predictions:

| | | | |
|---|---|---|---|
| (11) | /kanasi/ + /si/ | → (1) concatenation: | [kanaʃiʃi] |
| | | → (2) preaccentuation: | [kanaʃìʃi] |
| | | → (3) haplology: | *[kanaʃì] |

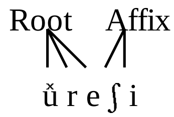
As shown, this predicts that the accent should be on the final syllable, not on the penult. So, the Root-Deletion and Coalescence approaches both predict the correct outcome while the Affix-Deletion proposal does not.

Another accentual process provides evidence that the Root-Deletion approach is incorrect. When -*si* is suffixed to an underlying accented root, the underlying accent appears on the penultimate vowel of the root: e.g. /siro/ 'white' + /si/ → [ʃìroʃi], where *siro* is underlyingly accented.[7] The Root-Deletion and Coalescence approaches make different predictions about where the accent should fall when haplology has taken place.

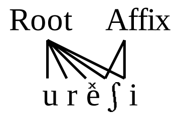
[7] Note that with underlying accented roots, -*si* does not cause preaccentuation.

The following structures show the output of the concatenation of underlyingly accented /uresi/ 'happy' with the CP suffix /si/ as predicted by the Root-Deletion and Coalescence proposals:

(12) Root-Deletion:




Coalescence:




Because accent falls on the penultimate vowel of the root, it is predicted to fall on [u] in the Root-Deletion approach, and on [e] in the Coalescence approach. The actual output is [urěʃi], so confirming the Coalescence prediction.

In summary, both Root- and Affix-deletion make false predictions about accent placement. In order for accent placement to be explained, the [ʃi] that results from -*si* haplology must belong to both the root and the affix simultaneously, so supporting the Coalescence approach.

### *3.2 French*

Another case of haplology that supports the coalescence view is found in French with the suffix -*iste* /ist/. The following data is taken from Corbin & Plénat (1992):

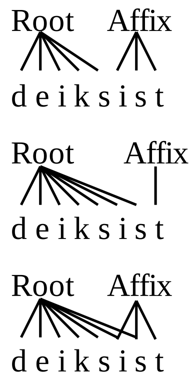
(13) /bodis/ *Baudis*{name} + /ist/ → [bodist], *[bodisist]
/deiksis/ *deixis* + /ist/ → [deiksist], *[deiksisist]
/analiz/ *analyse* + /ist/ → [analist], *[analizist][8]
/elefantiasis/ *elephantiasis* + /ist/ → [elefãtiasist], *[elefãtiasisist]
/maʁini/ {*Name*} + /ist/ → [maʁinist], *[maʁiniist]

This case differs from the ones discussed above in that only part of the affix haplologizes. Similar cases will be discussed in §5. For the moment, it is enough to recognise that –*iste* haplologizes with stem-end material.

Again, there are three possibilities: (1) deletion of root material, (2) deletion of affix material, or (3) coalescence. Representations of these three proposals are given below, showing the output [deiksist] from the input /deiksis/ + /ist/:

[8] This example shows that -*iste* haplologizes with forms ending in /iz/, not just /is/. Cases where the two haplologizing strings are not quite identical are common, and are discussed in §5.1.

(14) Root-Deletion:
Affix-Deletion:
Coalescence:



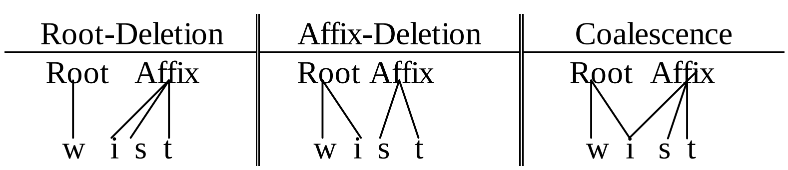

There are two restrictions related to this haplology that allow these alternatives to be distinguished. The first is a limit on the size of the target of *-iste* haplology:[9]

(15) (a) /kʁiz/ *crise* + /ist/ → [kʁizist], *[kʁist]
(b) /gi/ *Guy* {name} + /ist/ → [gi(j)ist], *[gist]
(c) /wi/ *oui* + /ist/ → [wi(j)ist], *[wist]

The data shows that *-iste* cannot haplologize with a root that is 'too small'. More precisely, the root must consist of more than a single mora (i.e. a vowel) if haplology is to apply to it.

The restriction can only be formulated naturally under two of the approaches to haplology. This is shown by considering the ill-formed haplologized output of /wi/ 'yes' + /ist/ → *[wist] as it would appear under the three different possibilities:

(16)




Why are these ill-formed? With Root-Deletion there is a readily identifiable reason: there is not enough material left in the root. Here, French requires a lexical morpheme (i.e. a root) to contain at least one moraic element (i.e. a vowel), a condition that Root-Deletion violates. This is a plausible restriction since many languages have minimal size requirements on lexical morphemes (McCarthy & Prince 1986 and many others).[10]

A similar restriction can be stated for coalescence. This time, though, the size restriction must be more narrowly stated: a lexical morpheme must contain at least one moraic element that is *unique* to it (i.e. not an exponent of another morpheme). The plausibility of this restriction derives from related constraints such as MORPHREAL, which require every morpheme to have some phonological realisation (Samek-Lodovici 1993 and many others).

[9] The term 'target' will be used to refer to the string with which the haplologizing morpheme coalesces.
[10] See Plénat (1992) for some discussion of minimal word effects in French.

A restriction is not so easily formulated with Affix-Deletion. Unlike the other two approaches, the root *does* contain a mora that is unique to it in the ill-formed output (as shown above). This raises a conundrum: what is the essential quality that differentiates the ill-formed **wist* from the well-formed *wiist* in the Affix-Deletion approach? The well- and ill-formed structures are shown below for comparison:

(17)

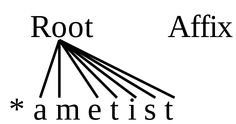


Assuming that an output-oriented size requirement on the root restricts the application of this haplology, the forms above present a puzzle since the root has the same content in both the ill- and well-formed outputs. This poses a significant problem in an output-oriented theory of grammar such as Optimality Theory. From a theory-internal point of view, the inability to identify an adequate output constraint implies that the premise is incorrect – in other words, -*iste* haplology does not involve deletion of affixal material. In short, only the Root-Deletion or Coalescence approaches present plausible explanations for the minimal size requirement.

The second restriction decides between the root-deletion and the coalescence approach. As shown in the following data, -*iste* cannot entirely haplologize:

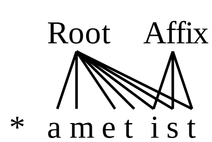
(18) (a) /ametist/ *amethyst* + /ist/ → [ametistist], *[ametist]
(b) /evaʁist/ *Évarist* {*Name*} + /ist/ → [evaʁistist], *[evaʁist]

Again, a natural output-oriented restriction can only be formulated in two of the three approaches. One of these is the Affix-Deletion approach. This restriction is evident from the output of the failed *[ametist] with affix-deletion:

(19)



In this form, there is no output segment affiliated to the affix. This restriction can be formulated as a requirement that an input affix must have at least one phonological exponent in the output. Similar minimum size requirements on affixes are found in a variety of languages (Greek, English – Golston 1991, Italian prefixes – Peperkamp 1999:225).

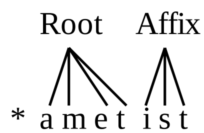
A similar restriction can be stated for the coalesced form, as shown by the ill-formed output:

(20)


The restriction can be stated as "An affix must have an output exponent that is unique to it". The plausibility of these restrictions is supported by the fact that roots have a very similar requirement, as demonstrated above.

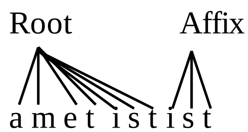
When it comes to Root-deletion, though, it is difficult to formulate an adequate restriction:

(21)




The form of the affix in this ill-formed output is identical to that in a well-formed one:

(22)




The same argument as used in the minimum-size requirement can be employed here: if the restriction on this type of haplology is an output-oriented size requirement on the affix, Root-Deletion presents a conundrum since the affix has the same form in both the ill- and well-formed outputs. As above, this eliminates the Root-deletion approach as a plausible solution to this problem.

So, the minimal-size restriction on roots eliminated the Affix-deletion approach as viable, while the size restriction on affixes gave evidence against Root-deletion. This leaves coalescence as offering the only plausible account of size requirements on French -*iste* haplology.

In conclusion, evidence from Japanese and French supports haplology as coalescence. This raises a significant issue from a theoretical point of view: if haplology is coalescence, then an adequate theory must not only permit haplology as coalescence but be unable to produce haplology as deletion. This issue is addressed in the next section.

## 4 What Triggers Haplology?

It was proposed in §2 that any markedness constraint is able to trigger haplology. This proposal is a significant departure from previous work, where it is assumed that there is some OCP-like constraint that specifically bans adjacent identical strings (Menn & MacWhinney 1984, Golston 1995, Plag 1998, Yip 1998, forthcoming). The aim of this section is to examine the virtues of the present proposals, and argue against the alternatives.

One virtue of the present approach is its theoretical economy: no new constraints are needed to deal account for haplology. The faithfulness constraints used – MAX, UNIFORMITY, IDENT – have been shown to have many applications elsewhere (McCarthy & Prince 1995, 1997 and references cited therein), and the trigger of haplology can be any markedness constraint – a point discussed in detail in §4.3.

A more empirically significant aspect of the present approach is that using a markedness constraint to trigger haplology instead of an OCP-like one cannot result in haplology as deletion. This is argued to be the case in the following section. In fact, the existence of an OCP-like constraint that bans adjacent identical strings is explicitly denied in §4.2. The nature of the triggering constraint is examined in §4.3.

## *4.1 Haplology is not Deletion*

If haplology is coalescence and not deletion, an adequate theory must be able to produce the former but not the latter. In other words, it is not enough to show that one's theory can result in haplology as coalescence – it must also be unable to produce haplology as deletion. The aim of this section is to show that a theory of haplology that employs a markedness constraint as the trigger of haplology can only result in coalescence. Again, *STRUC will be used to represent markedness constraints. This is only for the sake of concreteness: the results of this section are the same no matter what markedness constraint is used (see §4.3).

Half of the aim of this section has already been achieved: it was shown in §2 that *STRUC can be used to trigger haplology as coalescence. It remains to show that it cannot cause haplology as deletion. To do this, let us proceed *reductio ad absurdum* and assume that *STRUC *can* be used to motivate haplology as deletion.

Since haplology is a morpheme-specific process there must be a separate version of MAX for every morpheme; for a haplologizing morpheme M, then, there is a constraint MAX-M.[11]

In order for deletion to occur, the trigger of haplology (*STRUC in this case) must outrank MAX-M. If this were not so, preservation of underlying segments (due to MAX-M) would take precedence over minimisation of structure (due to *STRUC), and deletion would not occur. This was shown in §2 – if MAX outranks the triggering constraint, coalescence occurs.

However, if *STRUC outranked MAX-M, the triggering environment for haplology would be lost: M would delete in all contexts, not just next to an adjacent identical string. This is illustrated in the following tableau. Again, this presents the case of Arabic haplology with the leftmost morpheme /$t_1a_2$/ as the haplologizing morpheme and the constraint MAX-*ta* which refers to it. The rightmost morpheme /$t_3a_4$/ has a MAX-M constraint that is ranked above *STRUC (i.e. it is non-haplologizing), so it cannot be deleted (hence no candidates in which it is deleted are presented):

[11] The morpheme-specific nature of haplology has been discussed by Stemberger (1981). If there are no morpheme-specific MAX constraints but just one MAX for all morphemes, deletion would occur in every morpheme, not just the haplologizing one. This is illustrated in tableau (4) above.

(23)

| ❶ $/t_1a_2/ + /t_3a_4/$ | *STRUC | MAX-*ta* |
|---|---|---|
| $t_1a_2t_3a_4$ | x x! | |
| ☞ $t_3a_4$ | | x |
| ❷ $/t_1a_2/ + /d_3i_4/$ | | |
| $t_1a_2d_3i_4$ | x x! | |
| ☞ $d_3i_4$ | | x! |

In ❶, *ta* deletes when adjacent to an identical string. However, it also deletes when next to a non-identical string such as *di* in ❷. This situation cannot be remedied by permuting the ranking of other constraints. In particular, ranking IDENT-F above *STRUC has no effect since IDENT-F only requires *correspondents* to have identical features and there are no relevant correspondents in the output form: if an input *x* segment has no correspondent in the output, IDENT is not violated if *x*'s features do not appear (for related discussion of this point see §5.1).

In summary, assuming that *STRUC can motivate haplology as deletion leads to a contradiction: in order for deletion to occur *STRUC must outrank MAX, but in order for haplology to occur, MAX must outrank *STRUC. Of course, this result generalises to whatever constraint is used as the trigger (see §4.3).

This result is achieved because of the way identity between the haplologizing affix and its target is enforced: there is no direct comparison of identity between the haplologizing affix and its target by the triggering constraint. Instead, the comparison of featural identity of the affix and its target is undertaken by IDENT-F. However, IDENT-F only comes into play when there is coalescence; when there is deletion, IDENT-F is irrelevant, so identity between the haplologizing affix and its target are likewise irrelevant. This irrelevance of featural identity under deletion is what results in indiscriminate deletion in every environment. In short, the division of labour employed in the present approach – with the identity requirement of haplology the responsibility of IDENT-F and the triggering the responsibility of the markedness constraint – precludes haplology as deletion.

The argument presented in this section is only true if the triggering constraint does not directly compare the featural identity of the haplologizing affix and its target. Specifically, if haplology is triggered by a constraint that bans adjacent identical strings, haplology as deletion can result. This is demonstrated in the next section.

### *4.1.1 The OCP and Deletion*

The main alternative to the present proposal is to employ a constraint that directly bans adjacent identical strings. This approach has been implemented by a number of authors in various ways (Golston 1995, Russell 1997, Plag 1998, Yip 1998, forthcoming), and is seen as an extension of the OCP (Leben 1973, McCarthy 1986). This constraint will be termed the 'generalised OCP' here to differentiate it from the traditional OCP.

One of the problems with such a proposal is that the generalised OCP allows haplology as deletion.

This is illustrated in the following tableaux with the Arabic prefixes *ta* {fem.sg.} and *ta* {verbal prefix}. OCP(*ta*) bans adjacent sequences of /ta/, and MAX-*ta* requires the contents of the feminine singular morpheme *ta* – represented as /$t_1a_2$/ – to be realised. All other MAX-M constraints are high-ranked, so nothing but /$t_1a_2$/ can be deleted.

(24) OCP with deletion:

| | | OCP(*ta*) | UNIFORMITY | MAX-*ta* |
|---|---|---|---|---|
| ❶ | /$t_1a_2$/ + /$t_3a_4$/ | | | |
| | (a) $t_1a_2t_3a_4$ | x! | | |
| ☞ | (b) $t_3a_4$ | | | x |
| | (c) $t_{1,3}a_{2,4}$ | | x x! | |
| ❷ | /$t_1a_2$/ + /$d_3i_4$/ | | | |
| ☞ | (a) $t_1a_2d_3i_4$ | | | |
| | (b) $d_3i_4$ | | | x! |
| | (c) $t_{1,3}a_{2,4}$ | | x x! | |

This tableau shows that the generalised OCP can produce a deletion version of haplology. This is achieved by ranking the generalised OCP and UNIFORMITY above MAX. In ❶, the fully realised candidate (a) fatally violates the generalised OCP, leaving the coalesced form (c) and the deletion form (b). The deletion form is chosen since it does not violate UNIFORMITY.

In comparison, ❷ shows that deletion does not take place with non-identical strings. This is due to the fact that the fully realised candidate (a) does not violate the OCP. Since (a) does not violate any other constraint, it is the most optimal form.

The following tableau shows that the OCP can also be used to motivate haplology as coalescence:

(25) OCP with coalescence:

| | | OCP(*ta*) | MAX-*ta* | UNIFORMITY |
|---|---|---|---|---|
| ❶ | /$t_1a_2$/ + /$t_3a_4$/ | | | |
| | $t_1a_2t_3a_4$ | x! | | |
| | $t_3a_4$ | | x! | |
| ☞ | $t_{1,3}a_{2,4}$ | | | x x |
| ❷ | /$t_1a_2$/ + /$d_3i_4$/ | | | |
| ☞ | $t_1a_2d_3i_4$ | | | |
| | $d_3i_4$ | | x! | |
| | $t_{1,3}a_{2,4}$ | | | x x! |

The difference between this tableau and the preceding one is that MAX outranks UNIFORMITY. As shown, this produces haplology as coalescence.

So, an OCP-like constraint can produce both haplology as deletion and as coalescence. The reason for this has to do with identity. Haplology has two aspects: (1)

it is an attempt to minimise structure (or more generally to avoid constraint violation) and (2) it will only occur if all underlying featural specifications are retained in the output form (the identity requirement). In the approach advocated in this paper, the identity requirement is actuated by IDENT-F, while the structure-minimisation part is due to the triggering constraint. In comparison, the OCP approach amalgamates both of these aspects into one constraint: the generalised OCP requires minimisation of structure only if there is identity. It is the amalgamation of these two aspects into one constraint that allows haplology as deletion.

Because of this, the claim in this paper is that a constraint that bans adjacent identical strings is undesirable and unnecessary. There are further reasons for denying that such a constraint exists in Universal Grammar. These are discussed in the next section.

### *4.2 Against the OCP*

Using a generalised version of the OCP has been a popular way to explain haplology. One implementation of this idea is Plag (1998) and Yip's (1998, forthcoming) version of the OCP, which bans adjacent identical constituents. Yip argues that these constituents can be both morphological and phonological (e.g. affix)[12] and Plag that they can only be phonological (e.g. onset, nucleus).[13] For example, OCP(affix) bans adjacent affixes with identical phonological content.

One difficulty with both these approaches is the assumption that the haplologizing affix and its target are always identical in terms of some constituency. For example, OCP(affix) bans adjacent identical affixes with the same phonological content. This comes up against difficulties with cases where affixes haplologize with stem material (see §5.4.1). For example, the Japanese Classical Predicative *–si* haplologizes, so presumably there is a constraint of the OCP(affix) variety – OCP(affix-*si*), banning adjacent identical instances of an affix *–si*. However, this approach incorrectly predicts that there should be no haplology in the case of *kanasi* 'sad' + *si*. This is because the non-haplologized candidate [kanasisi] does not violate OCP(affix-*si*). In order to violate OCP(affix-*si*) there must be two adjacent identical *affixes* containing /si/; this is not found in [kanasisi] since [si] is part of a root, not an affix. This problem arises in all cases of affixal haplology with root material.

The same issue arises for phonological constituents. For example, Plag argues OCP(onset) bans identical onsets in adjacent syllables: i.e. *sisa. The other constraints Plag proposes are OCP(nucleus) which bans identical nuclei in adjacent syllables, and

---

[12] Yip (1998) and Brentari (1998) suggest that a more general constraint – *REPEAT – may subsume the OCP's role in this area, and perhaps generally. I believe that the same arguments against Yip's OCP presented here apply to *REPEAT. Apart from this, Yip's proposal shares some relation to the present one in that it employs a constraint very much like UNIFORMITY (MORPHDIS). The reader is referred to Yip's work for further discussion.

[13] Apart from using the OCP as the trigger of haplology, Yip's is similar to the present approach in several respects: Yip's approach uses simultaneous realisation which is somewhat akin to coalescence – see §6 for discussion on this point. The reader is referred to Yip's work for further details.

OCP(onset, coda) which bans an identical onset and coda in the same syllable (i.e. $*[C_iVC_i]_\sigma$).

The difficulty with this approach is that in many cases the target and haplologizing element do not form coherent phonological constituents. This is found in French, where -*iste* /ist/ haplologizes with roots ending in –*is* (§3.2). For example, *Baudis* + *iste* forms [bo.dist], not *[bo.di.sist]. But which OCP constraint does the non-haplologized *[bo.di.sist] violate? It does not violate OCP(onset) since no two onsets are alike, nor OCP(onset,coda), since the onset [s] of the final syllable and its coda [st] differ. This leaves OCP(nucleus) alone, which this form does violate since the final syllable's nucleus and the one preceding it are identical – both have [i]. However, if only OCP(nucleus) is relevant in this process, then the identity of the stem-final consonant should be irrelevant. This incorrectly predicts that *Lénine* /lenin/ + iste should form the haplologized [le.nist], not [le.ni.nist] since this latter form violates OCP(nucleus). The attested form, though, is [leninist] (Corbin & Plénat 1991:105).

One way to potentially avoid these difficulties is to deny that the OCP refers to constituents at all (contra Plag and Yip). Instead, it could simply refer to adjacent identical strings. This is initially appealing since a constraint that bans adjacent identical strings seems to be a more general version of the traditional OCP (Leben 1973, McCarthy 1986). However, while the original formulation of the OCP applies to adjacent identical *elements* on the same autosegmental tier; in haplology, it has to refer to adjacent identical *strings*. This latter sort of OCP will be termed the 'generalised OCP' to differentiate it from the traditional sort.

Non-total haplology also presents a problem for the generalised OCP. Non-total haplology occurs when only part of the affix coalesces. For example, French -*iste* haplologizes with stems ending in /i/ or /is/. Such a situation is impossible to describe with one OCP constraint. For example, if the constraint was OCP(*iste*) this should only ban sequences of [ist+ist]. To ban [i+ist] and [is+ist] another two OCP constraints are needed, each referring to *substrings* of the morpheme -*ist*: OCP([i]-*iste*), which bans adjacent [i]'s if one belongs to *iste*, and OCP([is]-*iste*), which bans adjacent [is] sequences as long as one belongs to *iste*. There are a number of similar examples:

- <u>Spanish Plural</u>: (i) paraguas + es → paragu<u>a</u>s *umbrellas*,
  (ii) casa + es → cas<u>a</u>s *houses*
  This would require OCP([e]-*es*) and OCP([es]-*es*).
- <u>Koine Greek *mega*</u>: (i) mega + garon → meg<u>a</u>ron,
  (ii) mega + eθos → meg<u>e</u>θos
  This would require OCP([ga]-*mega*) and OCP(V-*mega*)
- <u>French *ité*</u>: (i) /fɔʁtɥi/ *fortuit* + /ite/ *ité* → /fɔʁtɥ<u>i</u>te/
  (ii) /ilisit/ *illicite* + /ite/ *ité* → /illis<u>it</u>e/
  This would require OCP([it]-*ité*) and OCP([i]-*ité*).

At the very least, the generalised OCP must be made more complex to account for these cases.

In addition, haplology does not always apply to strictly local strings. In Swedish, the suffix /ar/ haplologizes with the suffix /are/ [arə]: e.g. *dom* ‘judge’ + *are* [arə] {agentive} + *ar* → *domare*, **domarear*. The formulation of the OCP would have to be made more complex to deal with such cases since the requirement of adjacency is central to it. An account of this pattern using the present approach is given in §5.4.3.

The observations made in preceding paragraphs suggest that the implementation of the OCP approach to account for haplology is far from straightforward. While it is true that haplology can involve adjacent identical strings, not all cases of haplology conform to this: haplology can occur between non-adjacent strings, and many cases of haplology involves only partially identical strings. In comparison, the present approach can account for both partial-identity haplology and non-adjacent haplology without introducing any new constraints; this is discussed in detail in §5.

Invoking the traditional OCP gives the generalised OCP a significant precedent. In fact, if the generalised OCP *is* the traditional OCP, the argument in this paper that the *STRUC approach is more theoretically economical loses its force. After all, if the generalised OCP is the traditional OCP, there is no need to add a new constraint to the grammar to trigger haplology (assuming that the traditional OCP is needed). This poses an even more significant problem from another point of view: if the generalised OCP is the traditional OCP and does exist in the grammar, then haplology as deletion is possible and there must be two available strategies for haplology – the OCP approach and the present approach. Of course, this all depends on the generalised OCP being the same as the traditional OCP.

It is not. The traditional OCP can be defined as requiring that a relation of non-identity hold between adjacent entities. The key formal difference between the traditional OCP and its generalised version centres on the notion ‘entity’. In the traditional OCP, the entities are individual elements, making the OCP a first-order relation. In the generalised OCP, though, those entities are *sets* of elements, making the OCP a second-order relation. In this sense, then, the traditional and generalised OCPs are formally distinct – they are not the same constraint, and the existence of one does not necessitate the existence of the other. So, the generalised OCP *is* an addition to the grammar – it is distinct from the traditional OCP. This conclusion means that the present proposal is more desirable from the point of view of Occam’s razor since nothing needs to be added to the grammar. This result also means that the generalised OCP can be eliminated as a viable constraint, so eliminating the possibility that haplology is deletion.

So, an OCP that refers to constituents is not adequate to trigger haplology. The alternative – i.e. allowing the OCP to range over adjacent strings without regard to constituency – is also undesirable since it constitutes an an addition to the grammar and cannot be derived from already constraints already known to exist. This leads to the conclusion that there is no such constraint. More generally, there is no constraint that directly compares the identity of two strings.

As a caveat, this claim may be a little too general. It not only bans OCP constraints on strings but also the traditional OCP, which bans adjacent identical elements. If the traditional OCP is retained, though, haplology as deletion can take place in a limited fashion: it can motivate haplology as deletion when the haplologizing morpheme is mono-segmental (e.g. Nisgha *t* – §5.4.2). This may suggest that the traditional OCP should also be dispensed with. Certainly, the validity of the OCP has recently been questioned (e.g. Ito & Mester 1996, 1998, see also Keer 1999). In any case, even if the traditional OCP is retained, haplology as deletion is only possible with mono-segmental affixes; in all other cases, it must be coalescence.

### *4.3 Identity of the Trigger*

One of the claims of this paper is that there is no constraint that is specifically designed to trigger haplology – in particular, there is no generalised OCP. Because there is no constraint that is specific to haplology, any constraint can in principle serve as a trigger. The aim of this section is to explore the consequences of this proposal.

While any constraint can in principle serve as the trigger of haplology, there is one (almost trivial) restriction: the constraint must be relevant. A constraint is relevant if it is violated more in a non-haplologized form than in its haplologized counterpart. For example, *STRUC is the ideal trigger; it will always be less violated in the haplologized form since that form always has less structure than its non-haplologized counterpart. Other constraints can also serve as the trigger of haplology, though. Brett Hyde (p.c.) has pointed out that most ALIGN constraints can also be used as a trigger. For example, ALIGN(σ, L, PrWd, L) requires the left edge of every syllable to be aligned with the left edge of a PrWd. This means that the more syllables a form has, the more it will violate this constraint. Since affixation often causes the addition of a syllable to the base form, a haplologized counterpart will violate ALIGN(σ, L, PrWd, L) less that the non-haplologized form. The only other caveat is that the triggering constraint cannot compare the identity of the haplologizing affix and its target. As shown in the preceding section, if a constraint does refer to identity there are undesirable effects. In other words, a generalised OCP cannot trigger haplology. No other markedness constraint (that I am aware of) has been proposed that refers to identity in this way, so all others are potential triggers of haplology.

One concern that this proposal raises is that of ‘delicacy’ (Cohn & McCarthy 1994): one would expect that being able to use any markedness constraint would predict a large variety of different types of haplology, each differing in some minute way. If the differences are too minute, this raises the danger of predicting unattested processes. In haplology, the danger has to do with the triggering environment. One would expect that the constraint *labial, for example, would trigger haplology in quite a different environment than ALIGN(σ, L, PrWd, L).

It is somewhat surprising, then, that the choice of triggering constraint has hardly any effect on where haplology applies. So far, *STRUC has been shown to produce haplology in the right environment: i.e. when the haplologizing affix is identical to an adjacent string. Most other markedness constraints have the same result. For example, an anti-feature constraint such as *labial causes haplology to apply only between identical strings. This is shown in the following tableau with *labial taking the place of *STRUC. The haplologizing affix in this case is /$b_3a_4$/:

(26)

| | | MAX | IDENT-F | *labial | UNIFORMITY |
|---|---|---|---|---|---|
| ❶ | /$b_1a_2$/ + /$b_3a_4$/ | | | | |
| | (a) $b_1a_2b_3a_4$ | | | x x! | |
| | (b) $b_1a_2$ | x x! | | x | |
| ☞ | (c) $b_{1,3}a_{2,4}$ | | | x | x x |
| ❷ | /$b_1i_2$/ + /$b_3a_4$/ | | | | |
| ☞ | (a) $b_1i_2b_3a_4$ | | | x x | |
| | (b) $b_1a_2$ | x x! | | x | |
| | (c) $b_{1,3}a_{2,4}$ | | x! | x | x x |

In ❶, *labial correctly triggers haplology with identical strings. In comparison, there is no coalescence with non-identical strings in ❷. The same results are obtained with all *F (F is a feature) constraints.[14]

This result is obtained because of the division of labour employed in the present approach: while markedness constraints provide the impetus for preferring the coalesced candidate over all others, the environment in which this can be actualised is determined by the faithfulness constraints – in particular, IDENT-F. IDENT-F requires featural identity, and so effectively prevents non-identical strings from coalescing (see §5.1 for further discussion of IDENT-F's effect).

Certainly, some markedness constraints do have an overt effect on their environment. This is true of some prosodic constraints. For example, a metrical constraint such as LAPSE will only motivate haplology when the non-coalesced candidate has adjacent unstressed syllables (hence, haplology only occurs when the two strings are identical and the alternative is to have adjacent unstressed syllables). The reason for this is that IDENT does not regulate metrical structure (since metrical structure is not present in the input), so cannot control this aspect of the environment. This is not an undesirable result, though: Plag (1998) discusses a number of cases of haplology which can be analysed as being triggered by constraints on metrical structure and syllabic form (esp. English -*ize* and German noun-forming -*er*).

In summary, the fact that the trigger of haplology can be any markedness constraint does not predict a vast number of minutely different types of haplology. This is due to the fact that the environment for triggering haplology is largely under the control of the faithfulness constraint IDENT-F. Since IDENT-F permits haplology only under identity and outranks the triggering constraint, there is little room for variation in the triggering environment for haplology.

[14] With *F constraints as triggers for haplology there is one proviso: the haplologizing affix must violate *F. For example, *coronal cannot be used to trigger haplology of the affix /ba/ since /ba/ does not contain a [coronal] feature. It is not enough for the *target alone* to violate the *F constraint: if the target violated *F but the haplologizing affix did not they cannot be featurally identical and so cannot coalesce (because of IDENT-F).

# 5 Typology

The purpose of this section is to show that a typology of haplological processes results by varying the ranking of other constraints with respect to the Basic Ranking ||MAX » ℂ » UNIFORMITY|| established in §2. These other constraints are of two main types: faithfulness and markedness. Faithfulness constraints regulate correspondence relations, while markedness constraints refer to the structure of the output form. The interaction of faithfulness constraints with the Basic Ranking is examined in §5.1-5.3. Of particular interest are the IDENT constraints, which require preservation of phonological features. Varying the ranking of these constraints determines how identical the affix must be to its target (§5.1). The effect of faithfulness constraints on dimensions other than the input-output one is also examined: it is demonstrated that haplology also applies in reduplication (§5.2). Other faithfulness constraints are also discussed, including those that preserve adjacency and precedence relations (§5.3).

The interaction of markedness constraints with the Basic Ranking is examined in §5.4. The markedness constraints that are of particular interest are those that regulate the position of morphemes. It is demonstrated that various rankings produce a range of results, accounting for the effect of morphological constituency on haplology.

Again, *STRUC will be used in this section as the triggering constraint. As discussed in the previous section, the choice of triggering constraint is largely irrelevant.

## *5.1 Identity*

In the prototypical case of haplology, the strings that coalesce must be featurally identical. This was explained in §2 as the effect of IDENT-F, which requires featural identity between an input segment and its output correspondents. The definition of IDENT-F is repeated here for convenience:

(27) IDENT-F If an input segment is αF then its output correspondent is αF.
(i) F is a feature
(ii) α is a featural specification (i.e. + or -)

IDENT-F is a actually family of constraints: there is an IDENT constraint for each feature (e.g. IDENT[coronal], IDENT[voice], etc.). For the sake of brevity, IDENT-F will be used below to refer to all IDENT constraints.

As shown in §2, to require complete featural identity between strings in haplology, IDENT-F must outrank the triggering constraint ℂ. This ranking ensures that the preservation of underlying feature specifications outweigh the need to avoid violating ℂ (i.e. to coalesce). This is illustrated in the tableau above with the Swedish plural suffix *-er*, which only haplologizes with completely identical strings. The first input form consists of the noun *kemiker* ‘chemist’ and the plural suffix *–er*. Due to haplology, these form *kemiker*, not **kemikerer* (Stemberger 1981:797). The second form, *rör* ‘choir’ + *er*, is given as a comparison; haplology does not take place here since *rör* does not end in an

-*er* sequence. Once again, *STRUC will be used as the triggering constraint, for the sake of concreteness.

(7) **Total Identity Haplology: IDENT » ℂ**

| ❶ /kemike$_1$r$_2$/+/e$_3$r$_4$/ | MAX | IDENT-F | *STRUC | UNIFORMITY |
|---|---|---|---|---|
| a. kemike$_1$r$_2$e$_3$r$_4$ | | | x x! | |
| b. kemike$_1$r$_2$ | x x ! | | | |
| c. kemike$_3$r$_4$ | x x ! | | | |
| ☞ d. kemike$_{1,3}$r$_{2,4}$ | | | | x x |
| ❷ /rö$_1$r$_2$/ + /e$_3$r$_4$/ | | | | |
| ☞ a. rö$_{1,3}$r$_{2,4}$ | | x! | | x x |
| b. rö$_1$r$_2$ | x x! | | | |
| c. re$_3$r$_4$ | x x! | | | |
| d. rö$_1$r$_2$e$_3$r$_4$ | | | x x | |

In ❶, the fully realised candidate (a) fails because of its excessive amount of structure. Candidates (b) and (c) are examples of deletion: in (b) the affix's phonological material has not been realised, while in (c) the root's has been deleted; these are both ruled out by the high-ranking MAX. This leaves only the coalesced candidate – (d).

In comparison, the coalesced candidate (a) in ❷ is not the most harmonic. This is due to the fact that both underlying /ö/ and /e/ have the same output correspondent, but the features of *both* /ö/ and /e/ are not retained. More specifically, the [-round] specification of the underlying /e/ does not appear in the output, fatally violating IDENT-F.[15] Again, the candidates with deletion – (b) and (c) – are both eliminated by MAX. This leaves the candidate with full realisation: (d). Although this candidate violates *STRUC more than any other form, this is irrelevant since all competing candidates have been ruled out by higher ranked constraints. Example ❷ illustrates the crucial ranking of IDENT-F over *STRUC. If the ranking was otherwise, the coalesced candidate (a) would be preferred over candidate (d). Hence, if a string only haplologizes with a featurally identical string, this is due to the ranking of IDENT-F over *STRUC.

In short, there is no need to trigger haplology by means of a constraint that directly bans identity between the haplologizing affix and its target (as does the generalised OCP). As noted in previous sections, the approach advocated here fragments the generalised OCP: while the generalised OCP both bans structure and requires identity, the present approach parcels this out to two different constraints – *STRUC bans structure and IDENT-F requires identity. This division of labour makes predictions about potential types of haplology. In particular, the status of IDENT-F as the identity-enforcing constraint allows the potential for variation in the exactness of identity required between the haplologizing affix and its target. This is the subject of the next section.

[15] The candidate [re$_{1,3}$r$_{2,4}$] would fare equally badly: with this the [+round] feature of underlying /ö/ is not present. A form with only a coalesced [r] – i.e. [rö$_1$r$_{2,4}$e$_3$] – fails for another reason: it does not preserve underlying precedence relations since r$_4$ follows e$_3$ in the input, but precedes it in the output. Preservation of precedence relations are discussed in §5.3.

(28)

| Identical Haplology: |
| --- |
| IDENT-F » ℂ |

### 5.1.1 Partial-Identity Haplology

Total-identity haplology results if every IDENT constraint outranks the triggering constraint. However, since constraints are freely rankable in Optimality Theory, there should be cases where the ranking is reversed with the triggering constraint outranking IDENT constraints. Such as situation is called 'partial-identity haplology', and is the subject of this section.[16]

In the prototypical case, the haplologizing affix and its target (the string it haplologizes with) are featurally identical. However, there are a number of cases where partial identity suffices. A few of these are listed below (others will be mentioned in later sections):

(29)

- Japanese Classical Predicative -*si* [ʃi]: haplologizes with both [ʃi] and [ʒi].
  e.g. /imizi/ *extreme* + /si/ → [imiʒi], *[imiʒiʃi] (Lawrence 1997)

- Nisgha 3rd person singular *t* [t]: with any voiceless coronal ([s], [t], or [ɬ]).
  e.g. /luutk'almiɬ/ *dies in the fire* + /t/ → [luutk'almiɬ], *[luutk'almiɬt]
  /naks/ *marry* + /t/ → [naks], *[nakst] (Russell 1995)[17]

- French Noun forming -*iste* /ist/: haplologizes with: /is/ and /iz/
  e.g. /analiz/ *analyse* + /ist/ → [analist], *[analizist]
  (Corbin & Plénat 1992)

Informally speaking, some features simply do not matter in partial-identity haplology. For Japanese and French, the feature [voice] is irrelevant in computing the identity of adjacent strings, while for Nisgha only [coronal] and [voice] (and perhaps [consonantal]) matter – all others are irrelevant.

To capture this is straightforward in the present proposal. To say that certain features 'do not matter' means that there is no need to preserve them. Since the preservation of featural identity is regulated by IDENT-F, this amounts to ranking IDENT-F low in the constraint hierarchy, where F is the relevant feature.

This is exemplified by Japanese -*si* [ʃi]. In this case, the voicing of the consonant is irrelevant in determining whether haplology takes place. This means that IDENT[voice] must be ranked low. In fact, the following tableau shows that it must be ranked below the triggering constraint – *STRUC in this case. In the following tableau IDENT-F' stands for the set of IDENT constraints that refer to every feature except [voice]:

[16] In the most extreme case, *STRUC would outrank all IDENT-F constraints. This means that an affix would coalesce in every environment, regardless of identity. This means that the affix would never have any independent realisation, effectively making it a zero morph.

[17] See Russell (1995) for arguments that this is not simply due to an OCP on adjacent coronal consonants.

(30) **Japanese *-si* Partial Haplology**

| | $/imiz_1i_2/$ + $/s_3i_4/$ | MAX | IDENT-F’ | *STRUC | IDENT[voice] | UNIFORMITY |
|---|---|---|---|---|---|---|
| | (a) $imiʒ_1i_2$ | x x! | | | | |
| | (b) $imiʒ_1i_2ʃ_3i_4$ | | | x x! | | |
| ☞ | (c) $imiʒ_{1,3}i_{2,4}$ | | | | x | x x |

The candidate with deletion – (a) – crucially violates MAX. This leaves candidates (b) and (c). Of the two, candidate (c) is the most harmonic since it has less structure. The fact that candidate (c) does not preserve /s/’s underlying [-voice] specification is irrelevant, due to the low ranking of IDENT[voice].

In short, the tableau shows that it is more harmonic to haplologize and be unfaithful to an underlying [voice] specification in Japanese than to violate the triggering constraint ℂ.

(31)

| |
|---|
| Partial-Identity Haplology:<br>If feature *f* is irrelevant, then ‖ℂ » IDENT-*f*‖ |

### 5.1.2 Root-Faithfulness

Cases of partial-identity haplology all have one thing in common: an underlying feature does not appear in the output. For example, in the Japanese case, underlying [-voice] does not appear while for French it is underlying [+voice]. This raises the question of prediction: which feature is preserved in partial haplology?

I propose that it is the root’s feature that is preserved.[18] This is due to root-specific faithfulness constraints of the type $\text{IDENT}_{\text{ROOT}}$F (McCarthy & Prince 1995, Urbanczyk 1996, Alderete 1997, Beckman 1998):[19]

(32) $\text{IDENT}_{\text{ROOT}}$F: If segment $x$ is αF and a member of a root then $x_C$ is αF.
(i) $x_C$ is the correspondent of $x$.[20]

---

[18] Stemberger (1981:792) makes a stronger claim: “it is always the affix, the outermost Z, that is absent.” Counter-evidence against this claim is found in Nisgha, where *t* always haplologizes with an adjacent morpheme in the same word whether it is preceded or followed by that morpheme (Russell 1995). Because of this, I take the weaker view that it is always the root’s features that are retained. This is weaker because this says nothing about which feature will be preserved when two affixes coalesce.

[19] Affix faithfulness constraints may also exist, in which case they must always be ranked below root faithfulness constraints. The alternative is to have is a root faithfulness constraint, a general faithfulness constraint, and no affix faithfulness constraints. It makes no difference how these are ranked - the same result is obtained.

[20] The proposals in this paper – that an output segment can be specified as both root and affix – allow another possible root-faithfulness constraint: $\text{IDENT}_{\text{OUTPUT-ROOT}}$[F]: “If $x$ is αF, then $x_C$ is αF iff $x_C$ is a member of a root.” This applies not only to underlying root elements, but to affix segments that have haplologized with root material. The effects of such a constraint will not be explored here.

For example, in the Japanese case of [imiʒi] + [ʃi] → [imiʒi], the root's [+voice] feature is preserved over the affix's [-voice] specification due to the constraint $\text{IDENT}_{\text{ROOT}}$[voice].

(33) **Japanese Root-Faithfulness**

| | /imiz$_1$i$_2$/ + /s$_3$i$_4$/ | IDENT-$_{\text{ROOT}}$[voice] | IDENT-[voice] |
|---|---|---|---|
| ☞ | imiʒ$_{1,3}$i$_{2,4}$ | | x |
| | imiʃ$_{1,3}$i$_{2,4}$ | x! | x |

(34) In Partial-Identity Haplology, the Root's feature will be preserved in deference to $\text{IDENT}_{\text{ROOT}}$F

However, this does not predict that the root's underlying feature will *always* be preserved. In some cases, a phonological constraint outranks $\text{IDENT}_{\text{ROOT}}$F, forcing unfaithfulness to the root. For example, *analyze* /analiz/ + /ist/ results in [analist], not *[analizt]: here, the [+voice] specification of the root does not appear in the output. However, this does not necessarily mean that the [-voice] specification of the affix has specifically been preserved. The fact that the [-voice] specification emerges receives a phonological explanation: [zt] codas are not permitted in French, due to voicing assimilation. This constraint, which I will term *zt]$_\sigma$ for convenience, outranks $\text{IDENT}_{\text{ROOT}}$[voice]:[21]

(35) **French Phono-Constraint » Root-Faithfulness**

| | /anali$_1$z$_2$/ + /i$_3$s$_4$t/ | *zt]$_\sigma$ | IDENT-$_{\text{ROOT}}$[voice] | IDENT-[voice] |
|---|---|---|---|---|
| | anali$_{1,3}$z$_{2,4}$t | x! | | x |
| ☞ | anali$_{1,3}$s$_{2,4}$t | | x | x |

(36) If a root feature *f* is not preserved in partial haplology, there is some phonological constraint Ᵽ that outranks $\text{IDENT}_{\text{ROOT}}$*f*.

Of course, this does not make any prediction about the preservation of features when two affixes haplologize. Which feature is preserved in such cases could be dependent on a variety of factors such as affix-specific faithfulness, other phonological processes, preservation of certain marked features or default to unmarked features. In other words, it is predicted that there will be no simple pattern in cases of affix-affix haplology.

## *5.2 Reduplicative Haplology*

Haplology seems to be the opposite of reduplication: reduplication creates two adjacent identical strings while haplology eliminates them (Wurzel 1976, Dressler 1977). It might be somewhat surprising, then, that reduplicants can haplologize. It is argued that

[21] As noted, *zt is properly expressed as voicing assimilation, and so should properly be replaced by some constraint to that effect. It is not the present concern of this paper to do so, however, so I leave the exact form of this constraint up to the reader's personal preferences and imagination.

a case of reduplicative haplology in Samoan can be explained by employing the same constraints as used for the input-output cases. The only difference is that the faithfulness constraints apply to reduplicants and their bases, not to the input-output dimension.

The Polynesian language Samoan has a foot-sized prefixal reduplicant. Some examples are given below, with the reduplicant underlined:

| | | | | | |
|---|---|---|---|---|---|
| (37) | fiti | *flick* | → | <u>fiti</u>fiti | *flick (pl.)* |
| | maʔai | *sharp* | → | <u>maʔa</u>maʔai | *sharp (pl.)* |
| | ŋaosa: | *untidy* | → | <u>ŋao</u>ŋaosa: | *in disorder* |

[Mosel & Hovdhaugen 1992: 224-5]

While the reduplicant is an exact copy of its base in most cases, in colloquial Samoan there is a class of exceptions:[22]

| | | | | |
|---|---|---|---|---|
| (38) | lele *fly* | → | <u>le:</u>lele, *<u>lele</u>lele | *fly* (frequentive) |
| | nini *apply, rub in* | → | <u>ni:</u>nini, *<u>nini</u>nini | *apply* (frequentive) |
| | pepe *butterfly* | → | <u>pe:</u>pepe, *<u>pepe</u>pepe | *flutter* |
| | tutu *light* | → | <u>tu:</u>tutu, *<u>tutu</u>tutu | *light* (frequentative) |

All the exceptions have bases of the form $C_iV_kC_iV_k$. In other words, outputs of the form $[\underline{C_iV_kC_iV_k}C_iV_kC_iV_k]$ are not permitted.

This can be explained as a type of haplology of the reduplicant. Instead of two identical *input* strings having a single *output* correspondent, two identical strings in the *base* have just one in the *reduplicant*. More concretely, *lele* ‘fly’ consists of two [le] strings, which correspond to a single [le] in the reduplicant.

To explain this, the same ranking that was used for other cases of haplology can be employed. The major difference is that the correspondence constraints refer to the Base-Reduplicant dimension:

(39) **Reduplicative Haplology in Samoan**

| RED + $/p_1a_2p_3a_4/$ | MAXBR, $\text{IDENT}_{BR}$F | *STRUC | $\text{UNIFORMITY}_{BR}$ |
|---|---|---|---|
| (a) $p_1a_2p_3a_4p_1a_2p_3a_4$ | | x x! | |
| ☞ (b) $p_{1,3}a_{2,4}$:$p_1a_2p_3a_4$ | | | x x |

MAXBR ensures that every element in the base has a correspondent in the reduplicant, while $\text{UNIFORMITY}_{BR}$ prohibits a reduplicant segment from having two base correspondents. As shown here, the fully reduplicated form (a) *<u>papa</u>papa* crucially violates the triggering constraint *STRUC, leaving the coalesced reduplicant (b) as the winner.

[22] The fully reduplicated forms are produced in formal registers of Samoan. The data presented here is from my own fieldwork, conducted in 1996 and 1997.

There is one additional ingredient to this ranking: LINEARITY-BR must be ranked below *STRUC. LINEARITY-BR preserves the precedence relations of the base in the reduplicant. If this was not violable, haplology could not take place since [$p_3$] follows [$a_2$] in the base, but the order is reversed in the reduplicant of candidate (b).[23] Cases where LINEARITY is violated on the IO dimension are discussed in the next section.

(40) **Reduplicative Haplology:**
MAX-BR, $\text{IDENT}_{BR}$-F » ℂ » $\text{UNIFORMITY}_{BR}$, $\text{LINEARITY}_{BR}$

While this ranking accounts for reduplicative haplology in outline, some other constraints are also needed to achieve the correct output in Samoan:

- That the reduplicant has a long vowel (i.e. [pa:], not *[pa]) is due to a constraint on the size of the reduplicant, which crucially outranks *STRUC (see McCarthy & Prince 1995:§4.3 for details).
- Coalescence of vowels without coalescence of consonants is prohibited:
  e.g. pata → $\underline{\text{pata}}$pata, *$\underline{\text{pa}_{1,2}}\text{tpa}_1\text{ta}_2$

This is due to a syllable wellformedness constraint prohibiting codas (NOCODA). This is ranked above *STRUC (and above every faithfulness constraint as well, since Samoan does not permit codas in any syllable).

- Coalescence of consonants without vowel coalescence is also prohibited:
  e.g. papi → $\underline{\text{papi}}$papi, *$\underline{\text{p}_{1,2}\text{ai}}\text{p}_1\text{ap}_2\text{i}$

There are few formal differences between the form *[$\underline{\text{p}_{1,3}\text{a}_2\text{i}_4}\text{p}_1\text{a}_2\text{p}_3\text{i}_4$] and [$\underline{\text{p}_{1,3}\text{a}_{2,4}\text{:}}\text{p}_1\text{a}_2\text{p}_3\text{a}_4$]. Both violate LINEARITY since the order of [$a_2$] and [$p_3$] in the base is reversed in the reduplicant and both admit UNIFORMITY violations. The main point in which they differ is in *adjacency* relations. In the reduplicant [$p_{1,3}a_{2,4}$] of [$p_1a_2p_3a_4$], all *adjacency* relations of the base are preserved: $p_1$ is adjacent to $a_2$, $a_2$ is adjacent to $p_3$, and $p_3$ is adjacent to $a_4$. This can be compared to the ill-formed reduplicant [$p_{1,3}a_2i_4$] of the base [$p_1a_2p_3i_4$]: in this base [$p_3$] and [$i_4$] are adjacent, whereas they are not in the reduplicant. So, a constraint on retaining adjacency relations must outrank *STRUC.

This results of this section are significant because they show that the Basic Ranking can apply on more than one faithfulness dimension. This generality lends support to the current approach.[24,25]

---

[23] In cases of haplology discussed in previous sections, LINEARITY was *not* violated since precedence relations do not hold between morphemes underlyingly.

[24] As a final note, Samoan is not an isolated case: in Nakanai (Johnston 1980) a foot-sized reduplicant also fails to fully copy bases of the form $C_iV_kC_iV_k$: e.g. lolo → $\underline{\text{lo}}$lolo *hearing*, *$\underline{\text{lolo}}$lolo, bebe → $\underline{\text{be}}$bebe *butterflies*, *$\underline{\text{bebe}}$bebe.

[25] Further support for reduplicative haplology and this analysis is presented in de Lacy (forthcoming), where certain circumscriptional effects are explained by reduplicants which haplologize with their bases.

### 5.3 Other Faithfulness Constraints

While limitations on space preclude a more in-depth examination of the effect of other faithfulness constraints on haplology, a couple more will be discussed briefly in this section.

Dressler (1977) observes that the German suffix *–en* haplologizes with roots ending in *-en*. However, it also haplologizes with roots ending in *–ern*: e.g. Eisern 'iron' + 'en' → eisern, *eisenern, *eisen 'made of iron'. This haplology is surprising since the two strings are not identical: one contains an *r*. In descriptive terms, the suffix *–en* has split in half in order to coalesce. This means that constraints that require retention of underlying adjacency (i.e. CONTIGUITY) must be ranked below *STRUC:

(41)

| | $eise_1r_2n_3 + e_4n_5$ | *STRUC | CONTIGUITY |
|---|---|---|---|
| ☞ | (a) $eise_1r_2n_3e_4n_5$ | x x! | |
| | (b) $eise_{1,4}r_2n_{3,5}$ | | x |

Candidate (a) has more structure than candidate (b), and so is not as harmonic. Candidate (b), on the other hand, violates CONTIGUITY since $e_4$ and $n_5$ are adjacent in the input, but this is not so in the output. This violation is not crucial though, since it is ranked below *STRUC. In standard cases of haplology, where such splitting does not take place, CONTIGUITY is ranked above *STRUC.

Another case involves the preservation of underlying precedence relations. Corbin & Plénat (1992:105) report the following patterns of haplology with the French suffixes -*iste* /ist/ and -*ique* /ik/:

(42) Boulganine + iste → boulganiste, *boulganiniste
oasis + ique → oasique, *oasisique
stalactite + ique → stalactique, *stalactitique
Thucidide + iste → thucidiste, *thucididiste.

The common characteristic of the roots is that they all end in a $C_1iC_1$ sequence. The explanation adopted here (akin to Corbin & Plénat's autosegmental approach) is that the underlyingly identical consonants coalesce. This allows the suffix to haplologize with the root-final /i/, thereby reducing *STRUC violations.

Crucial to this analysis is the fact that underlying precedence relations are not maintained: in the input, the root-final C follows /i/, but in the output it precedes it. This means that the faithfulness constraint that requires input precedence relations to be maintained (LINEARITY) must be violated – i.e. ranked below *STRUC:

(43)

| | $oas_1i_2s_3 + i_4que$ | *STRUC | LINEARITY |
|---|---|---|---|
| ☞ | (a) $oas_1i_2s_3i_4que$ | x x! | |
| | (b) $oas_{1,3}i_{2,4}que$ | | x |

The non-coalesced candidate (a) is eliminated due to its excessive structure. This leaves the coalesced form (b). Here, input $s_1$ and $s_3$ have coalesced, as have input $i_2$ and $i_4$. The coalescence of the *s*'s violates LINEARITY because $i_2$ precedes $s_3$ in the input, but not in the output. Other examples of haplology where LINEARITY is ranked low include the English suffix *–ize*, discussed in Plag (1998).

## *5.4 Markedness and Position*

The markedness constraints that have the most visible effect on haplology are those that regulate the position of affixes. This section will examine these constraints and their interaction with the Basic Ranking.

In almost all cases of haplology, an affix will only haplologize with material at a stem-edge. Specifically, a suffix will only haplologize with stem-end material, and a prefix will only haplologize with stem-initial material. The restriction to stem-edges will be shown to result from ALIGN constraints (McCarthy & Prince 1993a). Different rankings of these constraints will be shown to produce a variety of effects. Apart from accounting for standard stem-edge haplology (§5.2.1), non-stem-edge haplology (§5.2.3-4) and coextensiveness (§5.2.2) will be discussed.

### 5.4.1 Stem-Edge Haplology

In the theory of morphological concatenation proposed in McCarthy & Prince (1993a), the direction of attachment of a morpheme is regulated by an affix-specific ALIGN constraint. ALIGN constraints have the following form:

(44) ALIGN($\alpha$, $Edge_1$, $\beta$, $Edge_2$): For all $\alpha$ there is some $\beta$ s.t. $Edge_1$ of $\alpha$ and $Edge_2$ of $\beta$ coincide.
(i) Edges are 'Left' or 'Right'.

Constraints on the direction of attachment for affixes are of the sort ALIGN(*affix*, Edge, MCat, Edge) where MCat is a morphological category. As an example, for the Japanese suffix -*si* there is a constraint ALIGN(*si*, R, Stem, R) requiring the right edge of *si* to be aligned with the right edge of a stem. This is only satisfied if *si* appears after the root, as in [[aka]ʃi], where [ ] are stem boundaries. If *si* appeared before the root, its right edge would not be aligned with a right stem-edge: *[ʃi[aka]].

The requirement for *si* to appear at the right stem-edge can be maintained in stem-edge haplology. For example, in the haplologized form [[kanaʃi̲]], from *kanasi* + *si*, the affix's phonological material is still at the right-edge of a stem even though it has coalesced. This can be contrasted with a case of root-internal coalescence such as *[[ʃ̲iro]], from *siro* 'white' + *si*. In this form, the affix *si* has coalesced with root segments but it does not appear at the right edge of a stem, fatally violating ALIGN(*si*, R, Stem, R).[26] This requirement that the same-edge alignment constraint ALIGN(M, $\text{Edge}_i$, Stem, $\text{Edge}_i$) be unviolated effectively limits haplology to take place at the edge of stems.

(45)

| Stem-Edge Haplology (first formulation): ALIGN(same-edge) » ℂ |
|---|

In the summary above, ALIGN(same-edge) is an abbreviation for ALIGN(M, $\text{Edge}_i$, Stem, $\text{Edge}_i$). In following sections, different-edge alignment – i.e. ALIGN(M, $\text{Edge}_k$, Stem, $\text{Edge}_i$) where $i \neq k$ – will be abbreviated as ALIGN(different-edge).

### 5.4.2 Coextensive Haplology

In the cases discussed so far, the haplologizing affixes have coalesced with any (near-)identical string, whether that string constituted an entire morpheme or not. For example, French -*iste* haplologizes with stems ending in [is], and Japanese -*si* haplologizes with stems ending in [ʃi]. However, some morphemes will only haplologize if their target constitutes an entire morpheme. This is dubbed 'coextensive' haplology here. Several cases are listed below (see Menn & MacWhinney 1984: 522-523 for other examples):

- Arabic *ta* {verbal prefix}: only haplologizes before the feminine singular marker *ta* (Wright 1971: 65).
  e.g. ta + ta + kassaru → takassaru *it*(fem.sg.) *breaks* (instransitive), *tatakassaru
  *cf* ta + tabaʕa → tatabaʕa 'follow', *tabaʕa

- English possessive: haplologizes only with the plural (Stemberger 1981: 794-5, Russell 1997, Yip 1998).
  e.g. cat + *plural* + *possessive* → cats' [kæts], *[kætsəz]
  *cf* niece + *possessive* → niece's [nisəz], *[nis]

- Russian -*sk* 'inhabitant of': only haplologizes with the homophonous adjective-forming morpheme -*sk* (Aronoff 1976, Dressler 1977).
  e.g. Tom {Placename} + sk {adjective-forming} + <u>sk</u> + i → tomski, *tomskski
  *cf* Bask *Basque* + sk + i → baskski, *baski

- Classical Greek *e*- {past}: only haplologizes with the perfect *e*- (Stemberger 1981: 799).
  e.g. e + e + phthime:n → ephthime:n, *e:phthime:n
  *cf* e + ergon → e:rgon

---

[26] LINEARITY is not relevant here as I assume that morphemes are unordered in the input. Even if it was assumed that morphemes were so ordered, LINEARITY could not explain this case since *all* cases of haplology would have to violate it, not just non-local cases. Hence, stem-internal haplology would still be indistinguishable from the stem-edge variety.

The constraints used so far cannot distinguish between coextensive and non-coextensive haplology. For example, the ALIGN constraint discussed in the previous section is of no use here since it is satisfied equally well if the coalesced material does not consist of an entire morpheme. However, ALIGN constraints do offer a way to explain coextensive haplology. If the stem edges are indicated in a coextensive form, it is noticeable that both edges of the haplologizing affix coincide with a stem-edge:

(46) (a) Tom + sk + sk + i → [[[tom]sk]]i]
(b) Bask + sk + i → [[[bask]sk]i]], *[[[bask]]i]

In the haplologized form in (a), the left edge of the affix is adjacent to a right stem-edge. This can be compared with the ill-formed haplologized form in (b), *[[[bask]]i]; here, the left edge of the haplologized affix is not adjacent to any stem-edge.

This observation can be expressed as a constraint requiring the left edge of the suffix *sk* to align with the right edge of a stem. With this and the other ALIGN constraint outranking the triggering constraint ℂ (*STRUC in this instance), coextensiveness results (see Russell 1997 for a similar analysis of the English possessive and plural):

(47)

| ❶ Tom + sk + sk + i | ALIGN(*sk*, L, Stem, R) | ALIGN(*sk*, R, Stem, R) | *STRUC |
|---|---|---|---|
| [[[[Tom]sk]sk]i] | | | x x! |
| ☞ [[[[Tom]sk]]i] | | | |
| ❷ Bask + sk + i | | | |
| ☞ [[[bask]sk]i] | | | x x |
| [[[bask]]i] | x x! | | |

As shown in the most harmonic form in ❶, both edges of *sk* align with stem edges, so satisfying both ALIGN constraints. This is not the case in the failed candidate in ❷, though. Here, the left edge of *sk* is not aligned with the right edge of any stem.

(48) Coextensive Haplology:
ALIGN(different-edge), ALIGN(same-edge) » ℂ

Having introduced the different-edge alignment constraint ALIGN(M, $E_i$, Stem, $E_k$), the statement of stem-edge non-coextensive haplology as presented in the preceding section must be revised. Since the same-edge alignment constraint dominates *STRUC in both coextensive and non-coextensive haplology, the difference must lie in the ranking of the different-edge alignment constraint: in coextensive haplology, this constraint must outrank *STRUC, while in non-coextensive haplology the ranking is reversed. With *STRUC outranking the different-edge alignment constraint, there is no need for the internal edge of the haplologizing affix to align with the edge of a stem, so permitting haplology with root-internal material.

(49) Non-Coextensive Stem-Edge Haplology (final version):
ALIGN(same-edge) » ℂ » ALIGN(different edge)

The proposal in this section effectively means that every affix has two relevant ALIGN constraints – each referring to a different edge.[27] The introduction of the second ALIGN constraint – the one referring to opposite edges – has applications in other cases of haplology, as shown in the following section.

### 5.4.3 Stem-Internal Haplology

In the previous section it was shown that haplology is restricted to the ends of stems because of the ranking of the triggering constraint ℂ between two ALIGN constraints. Since ranking is freely variable, this makes a number of predictions. In fact, with the constraints ALIGN(same-edge), ALIGN(different-edge), and ℂ there are four rankings that are relevant to haplology. Two have already been investigated, and have been used to account for coextensive and non-coextensive stem-edge haplology. This section investigates one of the remaining rankings:

(50) ALIGN(different-edge) » ℂ » ALIGN(same-edge)

This ranking is the reverse of the one for non-coextensive stem-edge haplology. It is noteworthy that since ALIGN(same-edge) is ranked below ℂ, the haplologizing affix does not need to coalesce with stem-edge material. However, since ALIGN(different-edge) is ranked above ℂ, the internal edge of the affix must always appear at the right edge of a stem if it is a prefix, and at the left edge of a stem if it is a suffix.

An example of this ranking is found in Swedish with the non-neuter plural -*ar* (Stemberger 1981:797-8). This haplologizes after the agentive morpheme -are [arə]:[28]

(51) *dom* 'judge' + *are* {agentive} + *ar* → domare, *domarear, *domarar

The surprising fact here is that *ar* does not haplologize with a stem-final substring. Instead, it haplologizes with material *internal* to the stem. This fact means that ALIGN(*ar*, R, Stem, R) must be outranked by ℂ – *STRUC in this case. If this was not so, the form *domarear* should be most harmonic.

However, *ar* cannot haplologize with <u>any</u> *ar* sequence in a stem. Specifically, it can only haplologize with an affix beginning with *ar* (i.e. the agentive *are*). This fact can be explained when the stem location of stem edges is taken into account. In the haplologized [[[dom]<u>ar</u>e]], where <u>ar</u> is the haplologized string, the left edge of the non-

[27] This raises a potential situation where the relevant ALIGN constraints for an affix might be ALIGN(M, $E_i$, Stem, $E_k$) and ALIGN(M, $E_i$, Stem, $E_i$). As a more concrete example, let us suppose that *sk* had the constraints ALIGN(M, Right, Stem, Left) and ALIGN(M, Right, Stem, Right). This will not cause any difficulties in affix order, though, since the constraints are mutually conflicting. Since constraints are totally ranked, one of the two ALIGN constraints has to outrank the other. Which one prevails determines whether the morpheme is a prefix or a suffix.

[28] Haplology is optional with nouns ending in non-morphemic -*are* (e.g. hammare + ar → hammare *or* hamrar). However, it is questionable whether this is really haplology (see Stemberger 1981:798 for discussion).

neuter plural *ar* is at the right edge of the stem containing *dom*. This accords with the constraint ALIGN(*ar*, L, Stem, R), which essentially requires *ar* to be a suffix. With this constraint ranked *above* *STRUC the correct results are obtained:

(52)

| ❶ dom + are + ar | ALIGN(*ar*, L, Stem, R) | *STRUC | ALIGN(*ar*, R, Stem, R) |
|---|---|---|---|
| [[[dom]are]ar] | | x x! | |
| ☞ [[[dom]are]] | | | x |
| ❷ hamare + ar | | | |
| ☞ [[hamare]ar] | | x x! | |
| [[hamare]] | x! | | x |

In ❶, *domare* is the most harmonic candidate since the left edge of *ar* is aligned with the right edge of a stem and it has minimal structure. In ❷, however, the [are] sequence is part of a larger morpheme – *hamare*. Because of this, the haplologized *hamare* is not preferred since *ar* has haplologized with a substring internal to the stem, hence the left edge of *ar* does not appear adjacent to a right stem boundary. This crucially violates ALIGN(*ar*, L, Stem, R).[29]

(53)

Restricted Non-Edge Haplology:
ALIGN(different-edge) » ℂ » ALIGN(same-edge)

### 5.4.4 Unrestricted Haplology

The free ranking of ALIGN constraints with respect to *STRUC also predicts another type of morpheme: one which can haplologize with any string in the stem, regardless of its location. This occurs if all ALIGN constraints are ranked below *STRUC: the desire to minimise structure will outweigh any need to keep affixes in their proper position.

There are very few such examples, and none are entirely convincing.[30] It is possible that this is simply an artefact of the limited data available on morphological haplology and that such cases do exist. After all, cases where affixes do appear internal to a stem – i.e. infixes – do exist. Unrestricted haplology would be the haplologizing analogue of infixes. On the other hand, if such affixes do not exist, the current approach overpredicts.[31]

---

29 One potential objection to this proposal is that the final schwa in *are* [arə] is 'transparent' in some sense. However, there is no reason to think that it is epenthetic, and so does constitute part of the morpheme. It is difficult to see how the schwa can be ignored outright in this case.

30 A possible case of this is found in Old Provençal (Anglade 1921: 269-270, Platt 1981). The first person singular suffix -*i* is optionally omitted with verbs of the -*ar* conjugation: gir + i → giri *or* gir, tir + i → tiri *or* tir, vir + i → viri *or* vir. Anglade (1921:270) notes that this was attested in the Toulous dialect in the middle of the 14th century. This could be seen as non-local haplology, but equally (and probably more likely) it could be seen as apocope, or the OCP constraint operating on the vowel tier. Anglade does not provide any further information.

31 It is possible that there is an entirely different constraint that bans non-local coalescence in general. A theory of coalescence that is based on proposals that are very similar to the present ones can be found in Keer (1999). Like the present approach, Keer's system predicts non-local coalescence (when LINEARITY is

To conclude this section, the ranking of morpheme-specific ALIGN constraints provides an explanation for a variety of effects in haplology ranging from stem-edge haplology to coextensive haplology.

## 6 Alternative Theories

The thesis of this paper is that haplology is coalescence. However, a variety of other analyses have been presented: haplology has been argued to be (1) deletion (Jespersen 1954, Cardona 1968, and Harris 1980), (2) vacuous rule application (Stemberger 1981, Menn & MacWhinney 1984), and (3) simultaneous realisation (Stemberger 1981, Plénat 1991, Corbin & Plénat 1991, Lawrence 1997). In this section, each of these alternatives will be discussed in turn.

Of the variety of proposals regarding the nature of haplology, deletion can be rejected immediately; the cases discussed in §3 show that at least some instances of haplology are coalescence (see Stemberger 1981 for other arguments). This leaves vacuous rule application and deletion.

An alternative is vacuous rule application (Stemberger 1981, Menn & MacWhinney 1984). In vacuous rule application, affixation is seen as a 'checking device'. For example, the phonological entry of the English possessive suffix is simply stated as "Stem must end in [s]". When confronted with a stem that already ends in [s], such as a form with the plural, this requirement is already satisfied and the phonological structure of the stem does not need to be altered. Vacuous rule application has been implemented in Optimality Theory by the means of constraints (see esp. Russell 1995). For example, instead of the possessive suffix contributing phonological material to the input, it is seen as applying a constraint on the phonological form of the output. In the present paper, it has been assumed that affixes are not constraints – they are actual input elements. However, the conclusions made in this paper are equally consistent with an output-oriented model. In a sense, the output-oriented model also treats haplology as coalescence: in the constraint-based view of affixation, the [s] in the possessive plural of *cat* – i.e. [kæts] – is a realisation of *both* the possessive and the plural morphemes.

The theory of simultaneous realisation differs from the coalescence view. Stemberger (1981), Plénat (1991), Corbin & Plénat (1992), and Lawrence (1997) suggest that haplology is the association of segments of different morphemes to the same X-slots (or root nodes). This is possible since the featural material exists on different autosegmental planes. This is illustrated below with the Japanese morpheme *-si* and the root *kanasi* (Lawrence 1997):

ranked low) – a situation that does not seem to occur. I suspect that the resolution to Keer's theory will also solve the problems identified here.

(54)
```
k   a   n   a   ʃ   i
|   |   |   |   |   |
X   X   X   X   X   X
                |   |
                ʃ   i
```

As Lawrence points out, the fact that two X-slots have two sets of featural material each does not pose any interpretive difficulty since the material is identical. This explains why haplology occurs to identical strings – if two different sets of features were associated to the same timing slot this would create a fatal ambiguity in interpretation.[32]

This proposal is close to the coalescence view in that segments (i.e. root nodes/X slots) simultaneously belong to two different morphemes in the output. However, the point of divergence is that in the view of simultaneous realisation a segment has *two* sets of features; in the coalescence view, a segment only has *one*.

An objection to simultaneous realisation is that not all haplology involves identical strings (§5.1). However, Lawrence (1997) counters this argument by claiming that (at least for Japanese -*si*) underspecification plays a role. In Japanese, -*si* can haplologize with [ʒi], represented in the following structure (Lawrence 1997:383):

(55)
```
i   m   i   ʒ   i
|   |   |   |   |
X   X   X   X   X
            |   |
            ʃ   i
```

At this point, it seems that interpretive chaos would arise since one X slot has conflicting [voice] specifications – [ʃ] is [-voice] while [ʒ] is [+voice]. However, Lawrence argues that this can be explained if /s/ is underspecified for [voice] – i.e. /s/ has no [voice] specification at all. This means that the X slot is unambiguously marked as [+voice] (also see de Haas 1988).

Lawrence's proposal makes one interesting prediction: in cases of non-identical haplology only the specified feature will be preserved. This makes a different prediction from the proposal made herein – that the root's feature will be preserved over that of the affix's. The point of divergence between Lawrence's and the present proposal would arise when the affix has a non-underspecified segment and the root has an underspecified one: Lawrence's approach predicts that the affix's segment would surface while the root's would be obscured, whereas the present approach predicts the opposite.

Of all the cases of haplology that I have examined, none provide the necessary conditions to decide between the two proposals. This is unsurprising. Affixes usually contain unmarked phonological material (Beckman 1998). Hence, if an affix haplologized

[32] Lieber (1987) proposes a principle that explicitly prohibits a segment from bearing different featural specifications at the same time.

with a root, Lawrence's proposal would predict that the root's feature would be maintained because they are more marked (hence overtly specified) than the affix's. The present proposal would predict the same since faithfulness to Root features outweighs faithfulness to affix features. Nevertheless, in all likelihood there is a case that can decide between these two proposals. Unfortunately, this must be left for further research.

Lawrence's approach does raise one issue, though: it seems to predict that there should be no cases of total identity haplology. For example, Arabic *ta* could coalesce with a morpheme *da*. Since /t/ is (presumably) unspecified for [voice], its simultaneous realisation with /d/ should not cause any interpretive clash. The fact that there are cases of total identity haplology suggest that this approach is too powerful.[33]

In any case, there is no good evidence to suggest that the coalescence view is incorrect. In fact, it shares several features with the simultaneous realisation and vacuous rule application theories.

## 7 Conclusions

In this paper I have argued that morphological haplology is coalescence. Like other cases of coalescence, it can be explained by the following constraint ranking:

(56) MAX » ℂ » UNIFORMITY, where ℂ is any relevant markedness constraint.

The interaction of a variety of other constraints, especially IDENT-F and ALIGN, produces a typology of haplological processes. This accounts for a wide variety of attested processes, including partial-identity haplology, coextensive haplology, and reduplicative haplology.

One of the main points of this analysis is that there is no need to posit a generalised OCP – a constraint that bans adjacent identical strings – and that such a constraint has undesirable consequences. In fact, the proposals made herein have not introduced any new constraints into the grammar. This reduces the phenomenon of morphological haplology to a special case of correspondence.

[33] Of course, this argument is dependent on the theory of underspecification assumed (i.e. Radical Underspecification here – Archangeli 1984). Underspecification is generally agreed not to be an option in Optimality Theory (Prince & Smolensky 1993). A way out of this is to proposed that haplology and default-feature specification rules can have varying orders. If haplology occurred before default feature insertion, partial-identity haplology would result; if the order was reversed, this would be total identity haplology. It is difficult to square this approach with theories of underspecification, though, where it is standardly assumed that default features are filled in after the phonological component.

## Appendix: Constraints and Ranking Summaries

From McCarthy & Prince (1995, 1997):

MAX-αβ: Every segment *x* in α has a correspondent *x*' in β.

(i) α,β ∈ {Input, Output, Base, Reduplicant}

$\text{IDENT}_{\alpha\beta}$F: If segment *x* is γF in α then *x*'s correspondent *x* is γF in β.

(i) F is a feature. γ is a feature specification.

UNIFORMITY: A segment in the output may not have more than one correspondent in the input.

From McCarthy & Prince (1993a):

ALIGN(α, $\text{Edge}_1$, β, $\text{Edge}_2$): For all α there is some β s.t. $\text{Edge}_1$ of α and $\text{Edge}_2$ of β coincide.

(i) Edges are 'Right' or 'Left'.

**Rankings**:

Note: ℂ is any relevant markedness constraint (see §4.3)

Haplology: MAX » ℂ » UNIFORMITY

Total-Identity: IDENT-F » ℂ.

Partial-Identity: If feature *f* is irrelevant, then ℂ » IDENT-*f*

Root-Faithfulness: is due to $\text{IDENT}_{\text{ROOT}}f$

(i) If a root feature *f* is not preserved in partial haplology, there is some phonological constraint ℙ that outranks $\text{IDENT}_{\text{ROOT}}f$.

Stem-Edge (§5.4.1, 5.4.2): ALIGN(M, $E_i$, Stem, $E_i$) » ℂ » ALIGN(M, $E_k$, Stem, $E_i$)

Restricted Non-Edge (§5.4.3): ALIGN(M, $E_k$, Stem, $E_i$) » ℂ » ALIGN(M, $E_i$, Stem, $E_i$)

Coextensiveness (§5.4.2): ALIGN(M, $E_k$, Stem, $E_i$), ALIGN(M, $E_i$, Stem, $E_i$) » ℂ

Unrestricted Haplology (§5.4.4): ℂ » ALIGN

Reduplicative (§5.2): MAX-BR, $\text{IDENT}_{\text{BR}}$-F » ℂ » $\text{UNIFORMITY}_{\text{BR}}$, $\text{LINEARITY}_{\text{BR}}$

## References

Archangeli, D. 1984. *Underspecification in Yawelmani Phonology and Morphology*. PhD (MIT).

Alderete, John. 1997. "Root-Controlled Accent in Cupeño." ms. University of Massachusetts, Amherst.

Anglade, Joseph. 1921. *Grammaire de L'Ancien Provençal*. Paris: Librarie C.Klincksieck.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Beckman, Jill. 1998. *Positional Faithfulness*. PhD dissertation, University of Massachusetts, Amherst.

Brentari, Diane. 1998. "Comments on the paper by Yip." in Steve G.Lapointe, Diane K.Brentari, & Patrick M.Farrell (eds.) *Morphology and its Relation to Phonology and Syntax*. Stanford, CA: CSLI Publications, pp.247-258.

Cardona, George. 1968. *On Haplology in Indo-European*. Philadelphia: University of Philadelphia Press.

Cohn, Abigail and John McCarthy. 1994. “Alignment and Parallelism in Indonesian Phonology.” Rutgers Optimality Archive #25.

Corbin, Danielle & Marc Plénat. 1992. “Note Sur L’Haplologie des Mots Construits.” *Langue Française* 96: 101-112.

Dressler, Wolfgang. 1977. “Phono-morphological dissimilation.” in W.Dressler & O.Pfeiffer (eds) *Phonologica 1976*. Innsbruck. pp. 41-8.

Golston, Chris. 1991. "Minimal Word, Minimal Affix." NELS 21:95-109.

Golston, Chris. 1995. “Syntax outranks phonology: evidence from Ancient Greek.” *Phonology* 12.3: 343-368.

Grammont, Maurice. 1933. *Traité de Phonétique*. Paris: Librairie Delagrave.

de Haas, Wim. 1988. “A Formal Theory of Vowel Coalescence: A Case Study of Ancient Greek.” *Publications in the Linguistic Sciences* 30. Dordrecht: Foris.

de Lacy, Paul. forthcoming. “Parsing-Out Circumscription in Optimality Theory.” to be presented at *AFLA VI*, April 1999; to appear in Catherine Kitto & Carolyn Smallwood (eds) *TWPL (Toronto Working Papers in Linguistics): Proceedings of AFLA VI*.

Harris, J.W. 1980. “Nonconcatenative Morphology and Spanish Plurals.” *Journal of Linguistic Research* 1.1: 15-31.

Hock, H. H. 1991. *Principles of Historical Linguistics*. Berlin, New York: Mouton de Gruyter.

Ito, Junko & Armin Mester. 1996. “Rendaku 1: Constraint Conjunction and the OCP.” Rutgers Optimality Archive #144.

Ito, Junko & Armin Mester. 1998. “Markedness and Word Structure: OCP Effects in Japanese.” Rutgers Optimality Archive #255.

Jespersen, Otto. 1954. *A Modern English Grammar on Historical Principles, VI*. Copenhagen: Munksgaard.

Johnston, R.L. 1980. *Nakanai of New Britain: The Grammar of an Oceanic Language*. Pacific Linguistics Series B, No.70.

Keer, Edward. 1999. *Geminates, the OCP and Faithfulness*. PhD dissertation, Rutgers University.

Lamontagne, Greg & Keren Rice. 1995. “A Correspondence Account of Coalescence.” in Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk (eds.) *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. Amherst: GLSA.

Lawrence, Wayne P. 1997. “Haplology and Vowel Underspecification.” in *Report of the Special Research Project for the Typological Investigation of the Languages and Cultures of the East and West*., pp. 381-388.

Leben, W. 1973. Suprasegmental Phonology, MIT.

Lieber, Rochelle. 1987. *An Integrated Theory of Autosegmental Processes*. Albany, N.Y.: State University of New York Press.

McCarthy, John J. 1986. “OCP effects: Gemination and Anti-Gemination.” *Linguistic Inquiry* 17: 207-63.

McCarthy, John. 1995. “Extensions of Faithfulness: Rotuman Revisited.” Rutgers Optimality Archive #110.

McCarthy, John and Alan Prince. 1986. *Prosodic Morphology*. ms. University of Massachusetts

McCarthy, John and Alan Prince. 1993a. "Generalized Alignment." in G. Booij and J. van Marle (eds). *Yearbook of Morphology*. pp. 79-153.

McCarthy, John and Alan Prince. 1993b. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. University of Massachusetts, Amherst, and Rutgers University.

McCarthy, John and Alan Prince. 1995. "Faithfulness and Reduplicative Identity." in J.Beckman, L.W.Dickey, S.Urbanczyk (eds). *Papers in Optimality Theory*. UMOP 18: 249-384.

McCarthy, John & Alan Prince. 1997. "Faithfulness and Identity in Prosodic Morphology." Rutgers Optimality Archive #216.

Menn, Lise and Brian MacWhinney. 1984. "The Repeated Morph Constraint." *Language* 60: 519-541.

Mosel, Ulrike & Even Hovdhaugen (1992) *Samoan Reference Grammar*. Scandinavian Press.

Peperkamp, Sharon. 1999. "Prosodic Constraints in the Derivational Morphology of Italian." in G.Booij & J.van Marle, (eds). *Yearbook of Morphology* '99. Dordrecht: Kluwer Academic Publishers, pp. 207-244.

Plag Ingo. 1998. "Morphological haplology in a constraint-based morpho-phonology", in Wolfgang Kehrein and Richard Wiese (eds.) *Phonology and Morphology of the Germanic Languages*. Tübingen: Niemeyer. pp.199-215.

Platt, Diana. 1981. "Old Provençal: The Balance Between Regularity and Irregularity in Verbal Inflection." in Tracy Thomas-Flinders (ed.) *UCLA Occasional Papers in Linguistics 4: Working Papers in Morphology*, pp.41-71.

Plénat, Marc. 1991. "Le javanais: concurrence et haplologie" *Langages* 101: 95-117.

Plénat, Marc. 1992. "Observations sur le Mot Minimal Français: L'Oralisation des Sigles." ms. Université de Paris X-Nanterre & C.N.R.S.

Prince, A. and P. Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. New Brunswick, NJ, Rutgers University.

Russell, Kevin. 1995. "Morphemes and Candidates in Optimality Theory." Rutgers Optimality Archive #44.

Russell, Kevin. 1997. "Optimality Theory and Morphology." in Diana Archangeli and D. Terence Langendoen (eds.) *Optimality Theory: An Overview*. Malden, MA: Blackwell. pp.102-133.

Samek-Lodovici, V. 1993. "A unified analysis of crosslinguistic morphological gemination." *Proceedings of CONSOLE 1*.

Stemberger, Joseph. 1981. "Morphological Haplology." *Language* 57.4: 791-817.

Urbanczyk, Suzanne. 1996. *Patterns of Reduplication in Lushootseed*. PhD dissertation, University of Massachusetts, Amherst.

Wright, W. 1971. *A Grammar of the Arabic Language*. Cambridge: Cambridge University Press.

Wurzel, W.U. 1976. Zur Haplologie. *Linguistische Berichte* 41: 50-57.

Yip, Moira. 1998. "Identity Avoidance in Phonology and Morphology." in Steve G.Lapointe, Diane K.Brentari, Patrick M.Farrell (eds.) *Morphology and its Relation to Phonology and Syntax*. Stanford, CA: CSLI Publications, pp.216-246.

Yip, Moira. forthcoming. "Repetition and its avoidance: the case of Javanese." to appear in K.Suzuki and D.Elzinga (eds.) *Arizona Phonology Conference 4: Features in Optimality Theory*. Tucson: University of Arizona, Coyote Papers.

Zoll, Cheryl. 1993. "Directionless Syllabification in Yawelmani." in ROW1, Rutgers University.

Zoll, Cheryl. 1996. “A Unified Treatment of Segments and Features.” Rutgers Optimality Archive #137.

Paul de Lacy
University of Massachusetts, Amherst

<delacy@linguist.umass.edu>
http://www-unix.oit.umass.edu/~delacy
Department of Linguistics
South College
University of Massachusetts
Amherst, MA 01003
U.S.A.