*Ingo Plag*

## Morphological haplology in a constraint-based morpho-phonology

1. Introduction[*]

Morphological haplology is generally viewed as a dissimilatory process that interacts in some way with morphological structure. In spite of a long research tradition on this pervasive phenomenon, the exact nature of morphological haplology has remained obscure. The current state of affairs can still be characterized in the words of Maurice Grammont, published more than a century ago:

> The topic is not new. Everyone has talked about dissimilation, everyone has seen examples of it and has cited cases, but no one has ever established what it really is, dissimilation, under which conditions it occurs, and which rules govern it. (Grammont 1895: 9, my translation[1]).

More recent approaches to haplology have stressed that it is best described as the avoidance of identical phonetic or phonological material in morphologically complex words. Basically, two versions of this idea have been developed, the so-called Repeated Morph Constraint and so-called Stem-End Haplology. Consider the definitions in (1) and (2):

(1)     Repeated Morph Constraint (e.g. Menn/McWhinney 1984)
        the avoidance of adjacent identical morphs

(2)     Stem-End Haplology (e.g. Stemberger 1981: 791)
        the absence of "an affix or clitic […] when the adjacent part of the stem is homophonous to it"

Several problems are still unresolved. The first problem concerns the universality of morphological haplology. Haplology, in one form or another, seems to occur in almost any language with enough morphology to create phonetically identical sequences. For example, in an impressive survey of cross-linguistic data, Dressler (1976) posits the strong universal constraint given in (3a), which is illustrated with two examples from German in (3b):

(3)     a. *...$V_i C_j$] $V_i C_j$...]
        b. *honig-ig*, *Zauber-er-in*

Although apparently pertinent in a large number of different languages, (3a) appears to be too strong, since counterexamples can easily be found. In German adjectival inflection for instance, such sequences are readily admitted (cf. *eigen-en* 'own, pl.'). Thus morphological haplology seems to be universal in some sense but language-specific and even affix-specific in another sense. The latter has been argued for by Booij (1983: 257), who advocates the existence of affix-specific haplology rules also on the grounds that sometimes identity is avoided by the deletion of material (as in G. *Zauberin, *Zauber-er-in* 'scorcer-ess') and sometimes by the ungrammaticality of both the full complex form and the deleted form (cf. G. **honig-ig*, **hon-ig*).

    Another problem concerns the interaction of haplology with other phonological properties, e.g. prosodic structure. In particular, I will show below that haplology sometimes depends on the suprasegmental properties of the complex word in question, such as syllabic structure and stress.

    A third problem is which of the definitions in (1) and (2) is the more adequate one, i.e. more useful for the description and explanation of the data. The Repeated Morph Constraint restricts haplology to strictly morphological units whereas Stem-End haplology also includes cases where phonological identity of stem and an adjacent morpheme is sufficient to trigger haplology, irrespective of the morphological status of the material of the stem. This relates to the fourth problem, namely whether morphological haplology is essentially a morphological or a phonological phenomenon. Finally, it is unclear whether identity avoidance in complex words is best described in a rule-based or in a constraint-based model.

    In the following I will try to find possible solutions to these problems by reexamining a number of pertinent phenomena from English, German and Dutch. The morphological categories involved are English *-ize* derivatives, German and English feminine person nouns, and the Dutch suffix *-er* in person nouns and comparatives. The analysis is couched in the framework of Optimality Theory (e.g. Prince and Smolensky 1993).

## 2. Proposal

I propose that morphological haplology results from a family of universal constraints on the repetition of identical phonological elements, OCP ('Obligatory Contour Principle', Leben 1973), which interact with other, phonological and morphological, constraints. By introducing universal violable constraints that are ranked in a language-specific and affix-specific way, the apparent paradox can be resolved that haplology seems to be universal and language- and affix-specific at the same time. The variability of haplology both across and within languages follows naturally from the interaction and the individual rankings of the constraints.

    OCP constraints have been proposed before by Yip (to appear) in order to account for the Repeated Morph Constraint. She puts forward the following constraints:

(4)    OCP (feature), OCP (segment), OCP (affix), OCP (stem)

I propose the following constraints instead:

(5)    OCP (feature), OCP (segment), OCP (onset), OCP (nucleus), OCP (onset, coda)

My proposal differs from Yip's in important respects. I eliminate the two constraints making reference to morphological instead of phonological structure, i.e. Yip's OCP (affix) and OCP (stem), and elaborate the kinds of phonological structure OCP constraints can operate on.[2] This has two main consequences. First, in my approach, morphological haplology is exclusively triggered by phonological constraints, which, as we will see, is both empirically and theoretically preferable. The increase in the number of different OCP constraints is independently needed, as we will see below. Second, by eliminating OCP (affix) and OCP (stem) as constraints, I argue against the Repeated Morph Constraint. It will be shown that both the Repeated Morph Constraint and Stem-End-Haplology are the consequence of the same mechanisms. My claims are summarized in (6) and (7):

(6)    Haplology results from the operation of the OCP constraints in (5), in interaction with other prosodic and morphological constraints relevant for the morphological category in question.

(7)    The ranking of the pertinent constraints is language-specific.

In the following section I will present evidence for these claims on the basis of data from English, German and Dutch.


## 3. Illustration and evidence


### 3.1. English *-ize* derivatives

English verbs derived by the suffixation of *-ize* feature some peculiar and apparently variable phonological properties. Thus base-final segments are often deleted (as in *emphasis - emphasize*), base-final consonants may change (as in *Celtic - Celticize*), vowels may alternate (as in *gentile - gentilize*), consonants appear to be inserted between the base and the

---

2    Although the account of haplology proposed below crucially rests on the notion of adjacency, nothing will be said about the representation of superficially non-adjacent subsyllabic constituents within one syllable (cf. OCP (onset, coda)) or across two adjacent syllables (cf. OCP (onset) and OCP (nucleus)). My model of haplology and the idea of OCP as such suggest that on some level or tier of the phonological representation the elements in question must be adjacent. It seems that current models of representation need to be revised to be able to account for the phenomena discussed in this paper. How exactly this may be done remains a matter of future research.

suffix (as in *stigma - stigmatize*) and even stress may be shifted (as in *cátholic - cathólicìze[3]*) or reduced (as in *géntile - géntilìze*).[4] The phenomenon with which I will deal here is the truncation of base-final segments if the base word ends in a syllable in which onset and rhyme are identical. The data in (8) illustrate this regular phenomenon.

(8)      *femininize                    feminize
          *minimumize                   minimize
          *metathesisize                metathesize

In her account of these facts, Raffelsiefen (1996) proposes an output-oriented constraint which prohibits identical onsets in adjacent syllables in the derived word: $*O_iRO_i$. My term 'OCP (onset)' is only a different name for this constraint, which reflects that it is part of a larger family of related constraints. Raffelsiefen's account works for most of the data she presents, but the picture seems to be more complicated. Thus, in my sample of 284 20th century neologisms I have found 6 words that do not feature haplology, contra to Raffelsiefen's predictions. Consider the forms in (9):

(9)      strýchninìze                  *strychnize
          clássicìze                    *classize
          dilletántìze                  *dilletize
          mírrorìze                     *mirrize
          pótentìze                     *potize
          térrorìze[5]                  *terrize

On closer inspection, the contradictory data reveal a striking regularity, namely that all base words of the forms cited by Raffelsiefen in favour of the constraint are polysyllabic (*emphasize - emphasis, metathesize - metathesis, feminize - feminine, maximize - maximum*, etc.), whereas all the counterexamples (except one, *dilletante*, to be discussed shortly) are disyllabic. At first sight the constraint appears to be sensitive to the number of syllables. However, the violators and conformers also differ in the stress pattern. Thus, the base words of the haplological forms all have antepenultimate stress with two unstressed syllables following, i.e. they are dactyls, while all violators do not exhibit a stress lapse. That this stress-based description is superior to the simple counting of syllables is corroborated by the behavior of *dilletántìze*, whose base word has main stress on the final syllable (which all the other polysyllabic forms cited lack). In sum, the operation of OCP (onset) needs to be restricted to those cases in which a base with two unstressed syllables precedes *–ize*. By way of illustration, a form like *\*fémininìze* exhibits a stress lapse, which would only be tolerable if the last two onsets were not identical. Since they are identical,

---

[3]   Following established conventions, I use acute accent to indicate primary stress, grave accent to indicate secondary stress.

[4]   See Plag (1997: chapter 7.2) for a comprehensive account of the stem allomorphy of *-ize* derivatives.

[5]   Raffelsiefen considers *terrorize* a French borrowing. According to the *OED* this is incorrect. Semantically, this word is also completely regular (see Plag 1997, in press for a detailed account of the semantics of *-ize* derivatives).

*\*fémininìze* is an illicit formation. Derivatives like *térrorìze* do not exhibit a stress lapse, hence they behave according to the generalization.

In Plag (1997: 155-191) I have proposed a number of phonological constraints that govern the segmental and prosodic make-up of *-ize* derivatives. Of these constraints, the following are relevant for haplology:

(10)   a. interacting constraints (see Plag 1997 for details):
             OCP (onset): no identical onsets in adjacent syllables
             IDENT HEAD: The prosodic head of the base is the prosodic head of the derivative (= no stress shift).
             CLASH HEAD: Bases must not have a final prosodic head (= *-ize* may not attach to iambic bases).
       b. constraint ranking:
             IDENT HEAD >> CLASH HEAD >> OCP (onset)

IDENT HEAD is one of the constraints that ensures the recoverability of the base in a derived word by not allowing a stress shift. CLASH HEAD expresses the well-known fact that languages tend to avoid sequences of adjacent stressed syllables. Assuming that *-ize* carries secondary stress, iambic bases would lead to a clash of the two stresses. Such clashes are only tolerated with *-ize* derivatives in order to satisfy a higher ranked constraint. Consider the interaction of these constraints in the following tableau:

(11)   OCP (onset) and *-ize* derivatives

| Candidates | IDENT HEAD | CLASH HEAD | OCP (onset) |
|---|---|---|---|
| ☞ féminìze | | | |
| fémininìze | | | *! |
| ☞ metáthesìze | | | |
| metathésisìze | *! | | * |
| metáthesisìze | | | *! |
| ☞ str̃chn[ɪ]nìze | | | * |
| str̃chn̄ze | | *! | |
| ☞ clássicìze | | | * |
| clássìze | | *! | |
| ☞ dilletántìze | | * | * |
| dílletìze | *! | | |
| dillétantìze | *! | | * |

On the one hand, derivatives on the basis of dactylic base words (*feminize, metathesize*) exhibit haplology because they would otherwise violate OCP (onset).[6] As shown in detail in Plag (1997), constraints against base truncation are ranked lower than the constraints in (10). On the other hand, derivatives on the basis of trochees are never truncated because this would lead to a violation of the higher-ranked CLASH HEAD. The candidate with an iambic base *dilletántìze* is optimal because the two lower-ranked constraints are violated in order to satisfy the highest constraint, IDENT HEAD. The haplological candidate *\*dílletize* is ruled out because of the stress shift.

The consistency of the patterns emerging from the operation of OCP (onset) is strong evidence not only for the reality of this constraint, but also for its interaction with other prosodic constraints. It is this interaction that triggers the apparent variability of haplology in *-ize* formations.

The proposed analysis has strong theoretical implications because it seriously challenges two widely-shared views on morpho-phonological processes. First, a number of people (e.g. Aronoff 1976, Booij 1977) have claimed that the truncation of stems always involves morphological constituents. Similarly, the Repeated Morph Constraint predicts that haplology can only occur if the adjacent elements are morphs. In my model, haplology effects are explained without making reference to the adjacency of identical morphs, i.e. the morphological status of the truncated strings does not play a role. In the following I will argue that restricting truncation to morphological entities is ill-advised.

Morphological truncation rules have been proposed, for example, in Aronoff (1976) or Booij (1977). The classic example of such a rule is the truncation of the verbal ending *-ate* when *-able* is attached (e.g. *demonstrate - demonstrable*). According to Aronoff (1976), this rule must make reference to the morphemic status of *-ate* since non-morphological *-ate* (as in *debate*) does not truncate (*\*debable*). As pointed out by Anderson (1992: 280), these facts can equally well be captured by a purely phonological generalization: only secondarily stressed *-ate* can be deleted, irrespective of the morphological status of the string [eɪt].

In general, it can be observed that truncation does often not involve morphological constituents. Corbin (1987: 345), for example, provides a whole range of French data which demonstrate that the deleted sequence at the end of the base cannot be a morpheme. The same is true for *-ize* derivatives. For example, the deletion of base-final [ɪ] in many of the neologisms under discussion can hardly be analyzed as the truncation of a putative morpheme *-y*, since [ɪ] does not seem to have morphemic status in words like *anthropólogy, (Madame) Bóvary, fántasy, mediócrity, Nórmandy* or *vaséctomy*, all of which form *-ize* derivatives under loss of *-y*. In particular, I am unaware of any claims that *-logy* or *-ity* are bimorphemic suffixes, consisting of *-log-* and *-y*, or *-it-* and *-y*, respectively. What the six words really have in common is not the morphology of putative *-y*, but their prosody: they end in or are dactyls that end in a vowel. As demonstrated in Plag (1997), such a phonological structure leads to the truncation of the final vowel across the board, ignoring morphological structure: *anthropologize, bovarize, fantasize, mediocritize, Normandize, vasectomize*. This point is corroborated by the form *patinize*, where the final schwa of the base word *patina* does not represent a suffix either, but is nevertheles truncated because the

---

6   Consonant-final dactylic bases words that do not lead to an OCP violation are never truncated, cf.
    *federal - federalize*. See Plag (1997) for detailed discussion.

base conforms to the kind of prosodic structure that necessitates the deletion of the final vowel when *-ize* is attached. Similar arguments hold for the haplology cases, where *-ite* in *appetite*, or *-is* in *metathesis* can hardly be regarded as morphemes (if morphemes are units of sound and meaning).

Paul Kiparsky (personal communication, August 1997) pointed out to me that a form like *parallelize* is counter-evidence to my claims, because my model incorrectly predicts that the truncated form **parallize* be optimal. If, alternatively, only morphemic [əl] could be truncated, the existence of *parallelize* would be naturally accounted for.

This argumentation is not entirely convincing, however. First, *parallelize* is a very old form, first attested 1610, which means that it perhaps does not reflect the prosodic constraints of present-day productive morphology. This is corroborated by the fact that today's native speakers generally prefer the synonymous converted verb *parallel* to *parallelize* (see also the respective entries in the *OED*). Second, it seems that suffixal *-al* is equally almost never truncated (cf. *federalize* vs. **federize*), so that the argument in favor of morphological truncation collapses. Third, even if we allow *parallelize* to be a *possible* derivative (and not only an *actual* one), it is still conceivable that OCP (onset) needs to be further specified for phonological features. For example, [l] might be allowed to appear in identical onsets whereas other consonants might not. Unfortunately, there are no data that could show this, because all stems which can take adjectival *-al* and which end in [l] take *-ar* instead of *-al* as an adjectival suffix (cf. **polal* vs. *polar*).[7] In principle, the featural specification of OCP constraints is not unusual[8] and could also be used to solve the problem at hand. In view of the arguments just presented it is however preferable to assign idiosyncratic status to the word *parallelize*.

In summary, the idea of strictly morphological truncation has two main flaws. It necessitates the postulation of otherwise unmotivated morphological structure, and it cannot explain the robust phonological generalizations that hold across morphologically diverse derivatives.

We may now turn to other kinds of derivational processes which feature haplology.

## 3.2. German (and English) feminine person nouns

In German, derivatives ending in the person noun forming suffix *-er* may combine with the feminine suffix *-in* to denote female persons. The semantics of *-er* derivatives is rather flexible, they can for example denote agents or inhabitants (*Läufer* 'runner', *Marburger* 'person from Marburg'), hence the label 'person noun' used in this paper.[9] In the following we will consider some well-known phonological peculiarities of some of these forms when combined with the feminine suffix *-in*.

---

[7]  This is an interesting fact by itself, because it shows the operation of another OCP constraint (to be discussed below) which does not allow identical onset and coda in a single syllable.

[8]  For example, the occurrence of epenthetic schwa in English verbs and nouns involving the inflectional suffix *-s*, has been explained by a constraint which prohibits adjacent sibilants (see, for example, Yip (to appear), Russell 1997).

[9]  Note that *-er* derivatives may also denote instruments, as in *Öffner* 'opener'. See Booij (1986) for a semantic account of similar forms in Dutch.

Consider the data in (12):

(12)   a.      Zauber-er              Bewunder-er            Haiger-er
       b.      *Zauber-er-in          *Bewunder-er-in        *Haiger-er-in
       c.      %Zaubrerin             %Bewundrerin           ?Haigrerin
       d.      Zauberin               Bewunderin             Haigerin
       'magician (m./f.)'            'admirer (m./f.)'       'male/female from Haiger'
       e.      Lehr-er-in, Fahr-er-in, Kassier-er-in
       'teacher', 'driver', 'cashier' (all f.)

While it is clear that the forms in (12b) are ungrammatical, the status of the derivatives in (12c) and (12d) is somewhat controversial. Some sources only allow (12d), others maintain that both patterns can be found. No matter how one wants to account for this variability, it seems that the data in (12c) and (12d) crucially have in common that they avoid the phonetic sequence *[ɐʁɐʁɪn] at the end of the derivatives. In essence, the sequence ✓[ɐʁɐ] is allowed (see 12a), the sequence ✓[ʁɐʁɪn] is allowed (see 12e), but not a combination of the two *[ɐʁɐʁɪn] These facts are usually analyzed as a haplology effect. A similar phenomenon can be observed in English with the corresponding suffixes -*er* and -*ess*:

(13)   sorceress/*sorcereress, murderess/*murdereress

We will focus in our analysis on the fully productive pattern in German, since -*ess* is no longer productive in English. In principle, a parallel account could be devised for the English data, but an explanation in terms of lexicalized patterns seems preferable.

   To return to the German data, one of the problems traditional rule-based accounts have to face is the nature of the deleted elements. Plank (1981) and Fleischer/Barz (1995) assume that the agentive morpheme -*er* is deleted, as in *Zauber#[er]#in*. However, a truncation of syllables works equally well: or *Zau.be.[re].rin*. I am not aware of any empirical arguments that could be adduced in favor of either solution. As we will shortly see, in an output-oriented model this problem disappears. I propose the following constraints on German feminine person nouns:

(14)   Constraints on German feminine person nouns
       OCP (nucleus): no identical nuclei in adjacent syllables
       MAX: every segment of the input has an identical correspondent in the output (=no deletion)
       DEP: every segment of the output has an identical correspondent in the input. (=no epenthesis)
       COMPL ONS: no complex onsets in schwa syllables

The constraints in (14) should be uncontroversial, except perhaps COMPL ONS. In German schwa syllables, complex onsets are usually prohibited. The reason for this restriction is unclear, but it can be suspected that it is the combined effect of a number of general constraints on German syllable structure. Be that as it may, COMPL ONS can be seen either as an abbreviation of these constraints or as a constraint in its own right. For the purposes of this paper we can assume the latter, because the details of the prohibition of complex onsets

in schwa syllables need not concern us here. The ranking of the constraints varies according to the permissibility of forms such as those in (12c) above. Let us first consider the grammar that rules out derivatives like *Zaubrerin*. I propose the following constraint ranking for this 'grammar 1':

(15)    OCP (nucleus) >> COMPL ONS >> MAX (grammar 1)

(16)    Haplology and German feminine person nouns, grammar 1

| Candidates /tsaubr/-/r/-/in/ | OCP (nucleus) | COMPL ONS | MAX | DEP |
|---|---|---|---|---|
| ☞ Zau.be.rin | | | * | * |
| Zau.be.re.rin | *! | | | ** |
| Zau.bre.rin | | * | | * |

In accordance with most analyses in the literature (e.g Wiese 1996) I assume that schwa syllables do not contain a vowel underlyingly. Under this assumption, the derivative *Zauberin* involves a DEP violation (schwa has no correspondent in the input) and a MAX violation (one /r/ does not have a correspondent in the output). *Zauberin* is more optimal than *Zaubrerin* because the latter derivative violates the higher-ranked constraint prohibiting complex onsets in schwa syllables. In other words, in this grammar the loss of input /r/ is more easily tolerated than a complex onset in a schwa syllable.

Speakers who prefer *Zaubrerin* over *Zauberin* have a grammar ('grammar 2') where the ranking of COMPL ONS and MAX is reversed.[10] Consider (17):

(17)    Haplology and German feminine person nouns, grammar 2

| Candidates /tsaubr/-/r/-/in/ | OCP (nucleus) | MAX | COMPL ONS | DEP |
|---|---|---|---|---|
| Zau.be.rin | | *! | | * |
| Zau.be.re.rin | *! | | | ** |
| ☞ Zau.bre.rin | | | * | * |

The haplology data so far do not provide evidence for the exact ranking of DEP in either grammar, since all pertinent candidates violate this constraint, and the non-haplological candidate *\*Zaubererin* is ruled out anyway by OCP (nucleus).

---

[10] There are also speakers like one of the editors of this volume, Wolfgang Kehrein, who equally accept *Zaubrerin* and *Zauberin* (*\*Zaubererin* is impossible even with these speakers). In this kind of grammar ('grammar 3'), MAX and COMPL ONS are crucially non-ranked, yielding two optimal candidates.

It might be argued that the rejection of *Zaubererin* is not due to a violation of OCP (nucleus) but due to a violation of OCP (onset). This solution would work for the candidates in (16) and (17), but not for the forms in (12e), such as *Fahrerin* or *Lehrerin.* Thus, any account of haplology must be able to predict that in both grammars, *Fah.re.rin* is more optimal than *Fah.rin.* Consider the tableau in (18), which also features OCP (onset).

(18)    Haplology and German feminine person nouns

| Candidates /far/-/r/-/in/ | OCP (nucleus) | MAX | OCP (onset) | DEP |
|---|---|---|---|---|
| ☞   Fah.re.rin |  |  | * | * |
| Fah.rin |  | * |  |  |

The ungrammaticality of *Fah.rin* indicates that MAX must be ranked higher than OCP (onset), and that the decisive constraint in (16) and (17) must indeed be the one prohibiting identical nuclei.[11] For obvious reasons, COMPL ONS does not play a role in the evaluation of the two candidates *Fahrerin* and *Fahrin.*

Note that the truncation behavior of feminine person nouns cannot be accounted for by making reference to stress-related constraints instead of OCP. At first sight it looks as if a form like \*Zaubererin could be ruled out because of an undesired stress lapse, i.e. a violation of a constraint prohibiting adjacent stressless syllables (\*LAPSE, see, for example, Plag 1997). Under this assumption *Fahrerin* would be optimal because there is no stress lapse, and truncation would result in an undesired stress clash. The problem is that with agent and instrument nouns the analysis is equivocal because two adjacent schwa syllables exhibit both identical nuclei and a stress lapse. There is, however, a set of data which show that a stress-related explanation of haplology effects is ill-advised, namely nouns denoting inhabitants. The haplological behavior of inhabitant nouns parallels the one of agent nouns (see, for example, the data in (12) above). Female inhabitant nouns whose base ends in two unstressed syllbales now illustrate that only identical nuclei trigger truncation, whereas non-identical nuclei do not. In other words, stress lapses are tolerated if the nuclei of the pertinent syllables are different. Thus, a female inhabitant of the town *Haiger* in Hessia is called *Haigerin,* not \*Haigererin, whereas a female from *Trebur* (stress only on the first syllable) is called *Treburerin,* not \*Treburin. This pattern is unexpected if \*LAPSE (or CLASH) played a significant role. Under the OCP account developed above, however, the pattern is predicted.

To summarize, we can say that identity avoidance in some German feminine person nouns (and the lack thereof in others) can be accounted for by the same type of constraint that are at work with *-ize* derivatives, OCP. What makes the difference in behavior between the two morphological categories in the two languages is the overall ranking of the con-

---

[11]  Of course, \*Zaubererin also violates OCP (onset), but this violation is not decisive, and was therefore not indicated in the tableaux (16) and (17).

straints involved. Furthermore, slight differences in the rankings can straightforwardly account for the variability found among German speakers.

We may now turn to our last case study, the Dutch suffix *-er*.


### 3.3. Dutch *-er*

In Dutch the suffix *-er* is used to derive denominal and deverbal agent and inhabitant nouns, similar to German and English. Unlike in English and German, however, the Dutch suffix exhibits a peculiar kind of phonologically conditioned allomorphy. There are three allomorphs whose distribution and orthographic and phonetic representation is given in (18):

(19)   a.  *-der*   [dər]        after stressed syllable ending in /r/
       b.  *-er*    [ər]         after stressed syllables not ending in /r/
       c.  *-aar*   [aar]        after unstressed syllable (i.e. elsewhere)

These generalizations are illustrated in (19):

(20)   a.  Bijlmerméer      Bijlmerméerder/*-aar/*-er       'person from B.'
           zéur-            zéurder/*-aar/*-er              'nagger'

       b.  Hoogevéen        Hoogevéener/*-aar/*der          'person from H.'
           Amsterdám        Amsterdámmer/*-aar/*der         'person from A.'

       c.  Díemen           Díemenaar/*-er/*-der            'person from D.'
           lúister-         lúisteraar/*-er/*-der           'listener'

How can this distribution be explained? It seems that any account of the allomorphy must make reference to the prosodic structure of the derivatives and to the fact that in general the sequence [rər] is avoided. The latter generalization can be interpreted as a haplology effect, since onset and coda of the last syllable may not be identical (if the nucleus is a schwa).

Booij (1998) assumes that there are three different suffixes available, of which the correct one is selected on the basis of five constraints: OCP-Place, Superheavy, FootMax, ParseSyll, FootMin, ranked in this order. OCP-Place prohibits an OCP violation on the place tier. Since schwa has no place features, sequences involving a schwa between two identical consonants are ruled out. The high ranking of this constraint ensures that /r/-final bases do not combine with /ər/, but only select one of the other allomorphs. In those cases, the lower-ranked constraints regulate the distribution of *-er* and *-aar* according to the prosodic structure of the derivative.

Booij's account of the haplology facts faces three problems. The first is that in Dutch, schwa is sometimes allowed between identical consonants. Thus, sequences involving /n/ as the consonant are readily allowed. As pointed out by Booij, underspecification of /n/ does not solve the problem, because /n/ must be specified at least for the feature [coronal], even if it is the default consonant in Dutch. He therefore proposes that OCP-place "will not

take the place features of the default consonant of a language into account" (1998: 157). I fail to see any independent reason why that should be the case.

The second problem is that the avoidance of /rər/ is part of a larger generalization which includes schwa and all short vowels. For example, the sequences */rɔr/, */rɛr/, */rɑr/, */rur/ are impossible in Dutch (Booij 1995: 43). One can thus formulate a general constraint prohibiting the sequence */rVr/. Obviously, this haplological constraint cannot be explained by OCP-Place, because short vowels do have place features. Notice also that the patterning of schwa together with short vowels is unexpected, because in most, if not all, other environments, schwa behaves like a long vowel (Booij 1995: 19f).

The third problem concerns the nature of the allomorphy involved. Contrary to his earlier analysis (Booij 1995), Booij (1998) assumes that the three suffixes [dər], [ər], and [a:r] are phonologically unrelated and that the allomorphy is best analyzed in terms of allomorph selection. Under this approach, the close phonological similarity between the suffixes remains coincidental. One of Booij's arguments for the assumption of three different underlying representations is the claimed autonomy of the individual suffixes. Thus, there are a number of derivatives where the base does not end in /r/ and still /dər/ is chosen (e.g. *diender* 'policeman'), or where inspite of the stress clash, /-ar/ is chosen (e.g. *leeraar* 'teacher'). In my view, the existence of such lexicalized exceptions is not a compelling argument against a single underlying representation. For example, the fact that some English verbs choose the wrong one of the three phonologically conditioned allomorphs [d], [t], [əd] (e.g. *learn*[t] vs. *learn*[d], *spel*[t] vs. *spel*[d]) does not refute the existence of a phonologically conditioned allomorphy in the default case.

In view of these problems, I suggest an alternative solution, which makes use of an OCP constraint that prohibits syllables in which onset and coda are identical: OCP (onset, coda).

I assume that there is only one underlying representation of the suffix deriving inhabitant and agent nouns, not three. I take it that this underlying representation is /ar/. This assumption departs from earlier approaches which standardly assumed /ər/ as the underlying representation. The main reason for choosing /ar/ is that [ər] can be seen as the destressed realization of /ar/. In the form *Amsterdammer*, for example, the place features of the suffix vowel are lost in order to satisfy a constraint that prohibits adjacent stressed syllables, CLASH. This constraint is well-known in the OT literature and is functionally equivalent to Booij's Superheavy constraint, which prohibits superheavy syllbales in weak positions. (21) illustrates the computation of the optimal candidate:

(21)

| Candidates | CLASH | MAX (feature) |
|---|---|---|
| ☞   Amsterdámmer | | * |
| Amsterdámàar | *! | |

With trochaic base words the derivative in [-aar] is optimal because there is no stress clash. But let us turn to haplology. I suggest that the occurrence of [d] should be analyzed as a case of epenthesis, i.e. as a violation of DEP (see (14) above). DEP is violated in order to satisfy OCP (onset, coda). The latter constraint can be provisionally defined as in (22):

(22)    OCP (onset, coda): Onset and coda of a single syllable may not be identical. (provisional)

Consider the form *zeurder* 'nagger' which is more optimal than *\*zeurer*, because lower-ranked DEP is violated in order to satisfy higher-ranked OCP (onset, coda).

(23)

| Candidates | CLASH | OCP (onset, coda) | MAX (feature) | DEP |
|---|---|---|---|---|
| ☞  zéurder | | | * | * |
| zéurer | | *! | * | |
| zéuràar | *! | * | | |
| zéurdàar | *! | | | * |

Possible candidates in [-aar] are ruled out by CLASH. A potential problem for this kind of analysis arises with trochaic bases, however, which, contrary to the prediction, should end in [-daar] in order to avoid an OCP violation. Consider (24):

(24)

| Candidates | OCP (onset, coda) | MAX (feature) | DEP |
|---|---|---|---|
| ☜ [12]lúisteràar | * | | |
| ☞  lúisterdàar | | | * |
| lúisterer | *! | * | |
| lúisterder | | *! | * |

We are thus faced with contradictory constraint rankings. To make *luisteraar* the optimal candidate, DEP would have to be ranked higher than OCP (onset, coda). On the other hand, stress-final and monosyllabic bases require the opposite ranking (see (23) above). To solve this dilemma we must take into consideration the fact that in Dutch, long vowels are in general allowed following and preceding /r/, but short vowels are not (✓/rVVr/ vs. \*/rVr/, see e.g. Booij 1995: 43). This indicates that OCP (onset, coda) must be sensitive to the structure of the nucleus, which leads me to a reformulation of this constraint in terms of moraicity. Assuming that long vowels are bimoraic and that short vowels, schwa and /r/ (in the rhyme) are monomoraic, I propose the following constraint:[13]

---

[12] Hands pointing to the left are used to indicate attested outputs that run counter to the predictions of the model.

[13] Following Hayes (1989), I assume that trimoraic syllables are possible in some languages. For Dutch, others (e.g. van der Hulst 1984, Kager 1989) have already argued that there are superheavy

(25)    OCP(onset, coda)/$\mu\mu$: Onset and coda of a bimoraic syllable may not be identical.

This constraint is ranked above MAX (feature) and DEP, which is illustrated in the following tableau:

(26)

| Candidates | OCP (onset, coda) / $\mu\mu$ | MAX (feature) | DEP |
|---|---|---|---|
| ☞  lúisteràar | | | |
| lúisterdàar | | | *! |
| lúisterer | *! | * | |
| lúisterder | | *! | * |

The candidate *luisteraar* violates none of the constraints in (26), hence nothing can be gained by epenthesis or destressing.

CLASH is ranked higher than MAX. For stress-final or mono-syllabic bases ending in /-r/, the tableau looks as follows:

(27)

| Candidates | CLASH | OCP (onset, coda) / $\mu\mu$ | MAX (feature) | DEP |
|---|---|---|---|---|
| ☞  zéurder | | | * | * |
| zéurer | | *! | * | |
| zéuràar | *! | | | |
| zéurdàar | *! | | | * |

In general, epenthesis is only allowed with /r/-final bases, because with non-/r/-final bases OCP does not play a role. Instead, epenthesis leads to an additional DEP violation, making each non-epenthesized candidate more optimal than its epenthesized counterpart. Thus the potential candidate *\*Amsterdamder* is ruled out because it violates DEP, which the optimal candidate *Amsterdammer* does not.

To summarize, the account presented above makes correct predictions concerning the allomorphy of the suffix and has the additional advantage of being able to explain the phonological similarity of the suffix variants. Notice that the proposed constraints yield correct results both under the allomorph selection assumption (à la Booij 1998) and under the phonologically triggered allomorphy approach advocated here. Thus, my model does not only rule out competing attested allomorphs, but also conceivable others (as in *\*roerdaar, \*luisterdaar*, or *\*Amsterdamder*).

---

syllables, in other words, that the sequence /VVr/ is trimoraic in Dutch. From this assumption it follows that OCP(onset, coda)/$\mu\mu$ rules out bimoraic /rVr/ but allows trimoraic /rVVr/.

There are, however, three problems involved. First, as mentioned above, not all consonants trigger the same haplological effects as does /r/. In terms of the model developed in this paper, this fact calls for a further parametrization of OCP (onset, coda) with respect to the kinds of consonants that fill onset and coda. Such parametrization is not at all unusual. In many languages constraints on possible structures of syllabic constituents must refer to sonority hierarchies or even to individual segments (e.g. Hammond 1997 on English syllables). Thus we can imagine that additional constraints such as OCP (onset, coda)/r and OCP (onset, coda)/n, in the appropriate ranking, regulate the differences between haplology effects when these consonants are involved.[14]

The second problem lies in the fact that *-aar* does not appear after non-coronal sonorants (cf. *bezem-er* vs. *\*bezem-aar* 'sweeper', Booij 1998). Neither in Booij's account nor in mine can this be derived in a straightforward manner from prosodic constraints, but must be stated as an idiosyncratic property.

Finally, the nature of the epenthetic consonant in Dutch is controversial. It is usually assumed that [n] is the default consonant in Dutch. The postulation of an epenthetic [d] runs counter to this assumption. However, putative epenthetic [d] does not only occur with the person noun forming suffix just discussed but also with the comparative suffix /ər/, the adjective-forming suffixes *-erig* (*zeur-d-erig* 'nagging') and the nominal suffix *-erij* (*boer-d-erij*). This can be interpreted in such a way that [d] has some kind of default status at least in certain morphological categories, if not in Dutch morpho-phonology in general. Under an allomorph selection approach, the generality of [d] is unaccounted for. Further research is necessary to substantiate arguments on the status of [d] in these derivatives.

Turning briefly to haplology in comparatives, an analysis similar to that proposed for person nouns is possible. The major difference between the person noun suffix and the comparative lies in the absence of [-aar] as a variant of the comparative suffix. I propose that this is the direct consequence of a different underlying representation of the comparative, namely /ər/ (as against /ar/ for person nouns). The two attested allomorphs have the distribution given in (`28a), examples are given in (28b):

(28)   a.   suffix allomorphs of comparatives
           *-der* after stem-final /r/, *-er* elsewhere

       b.   duister 'dark'              duisterder
            groen 'green'              groener

The pertinent constraints are the same as above, with the interesting difference that LAPSE does not play a role, i.e. it is ranked below the constraints under discussion:

---

[14] A similar solution suggests itself for another peculiarity of Dutch. According to Booij (1995), Dutch does not allow sequences of /l/, long vowel and /l/ (i.e. *\*/lVVl/), but readily admits /lVl/, i.e. a short vowel between two /l/s. This is exactly the reverse of what we observed with /r/. This fact could be accounted for by assuming that there is also an OCP (onset, coda) constraint that is parametrized for bimoraic syllables, OCP (onset, coda)/$\mu\mu$. Given the correct ranking of the two mora-sensitive OCP constraints and the consonant-sensitive OCP constraints, it should be possible to predict the correct optimal forms. The details of such an account remain to be worked out.

(29)    Haplology and Dutch comparatives: underlying representation /ər/

| Candidates | OCP (onset, coda) / $\mu\mu$ | DEP |
|---|---|---|
| ☞  gróener | | |
| gróender | | *! |
| ☞  dúisterder | | * |
| dúisterer | *! | |

Therefore, the difference in allomorphy between person nouns on the one hand and the comparative on the other does not lie in the inexplicable absence of a third underlying allomorph, but in a different underlying representation.

To conclude the discussion of haplology with Dutch person nouns and comparatives, we have seen again that a small number of phonological constraints is responsible for haplology effects as well as for other kinds of allomorphy.


4. Conclusion

In this paper I have made a case for an output-oriented approach to morphological haplology. The proposed constrained-based model of haplology was illustrated by the (re-)analysis of a small number of well-known and less well-known phenomena in three Germanic languages. I have argued that, in general, the purely phonological constraints of the OCP family are responsible for morphological haplology effects. These constraints, which rule out adjacent identical phonological structures, are extended in the present model to syllabic contituents such as onset and coda, but not to morphological constituents such as affix or stem. The main advantage of the model lies in the motivation of seemingly contradictory properties of haplology, namely its universality and its variability.

It was shown that the similarities of haplology phenomena across and within languages can be explained by a set of universal phonological constraints against adjacent identical elements. At the same time, the observable variability of haplology among different morphological categories in one language and across languages turns out to be the consequence of affix-specific and language-specific rankings of the OCP constraints. The interaction of OCP constraints with other prosodic and morphological constraints on the other can also explain the variability of haplology among derivatives of the same morphological category (cf. *strychninize* vs. *feminize*).

According to the arguments presented above, the Repeated Morph Constraint and Stem-End Haplology both result from essentially the same mechanisms. Sometimes the phonologically identical elements happen to be exponents of morphemes, sometimes they do not. The former would traditionally be subsumed under the Repeated Morph Constraint, the latter under Stem-End Haplology. In my model this distinction becomes epiphenomenal.

If it is true that morphological haplology is triggered exclusively by OCP constraints, morphological haplology must be seen as a purely phonological phenomenon. It is only the affix-specific ranking of OCP constraints and other prosodic and morphological constraints that makes haplology appear to be a morphological or morpho-phonological phenomenon.

## References

Anderson, Stephen R. (1992): *A-morphous Morphology*. Cambridge: CUP.

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

Booij, Geert E. (1977): *Dutch Morphology. A Study of Word Formation in Generative Grammar*. Lisse: de Ridder.

— (1983): Rezension von: Plank, Frans *Morphologische (Ir-)Regularitäten. Aspekte der Wortstrukturtheorie*, Tübingen: Narr, 1981. *Zeitschrift für Sprachwissenschaft* 2, 251-264.

— (1986): Form and Meaning in Morphology: The Case of Dutch 'Agent Nouns'. *Linguistics* 24, 503-517.

— (1995): *Dutch Phonology*. Oxford: Clarendon.

— (1998): Phonological Output Constraints in Morphology. In: Wolfgang Kehrein and Richard Wiese (eds.): *Phonology and Morphology of the Germanic Languages.* Tübingen: Niemeyer, 143-163.

Corbin, Danielle (1987): *Morphologie dérivationelle et structuration du lexique*. 2 Vols. Tübingen: Niemeyer.

Dressler, Wolfgang (1976): Phono-morphological Dissimilation. W. U. Dressler, O. E. Pfeiffer (eds.): *Phonologica 1976*, 41-48. Innsbruck: Institut für Sprachwissenschaft.

Fleischer, Wolfgang, and Irmhild Barz (1995): *Wortbildung in der deutschen Gegenwartssprache*. 2. Aufl. Tübingen: Niemeyer.

Grammont, Maurice (1895): *La dissimilation consonantique dans les langues indo-européennes et dans les langues romanes*. Dijon: Darrantière.

Hammond, Michael (1997): Vowel Quantity and Syllabification in English. *Language* 73, 1-17.

Hayes, Bruce (1989): Compensatory Lengthening and Moraic Phonology. *Linguistic Inquiry* 20, 253-306.

Hulst, Harry van der (1984): *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.

Kager, René (1989): *A Metrical Theory of Stress and Destressing in English and Dutch.* Dordrecht: Foris.

Leben, William (1973): *Suprasegmental Phonology*. PhD thesis, MIT, Cambridge.

Menn, Lise, and Brian McWhinney (1984): The Repeated Morph Constraint: toward an Explanation. *Language* 60, 519-541.

Plag, Ingo (1999) *Morphological Productivity. Structural Constraints in English Derivation.* Berlin, New York: Mouton de Gruyter.

— (in press): The Polysemy of *-ize* Derivatives. The Role of Semantics in Word-formation. Geert Booij, Jaap van Marle (eds.): *Yearbook of Morphology 1997*. Dordrecht et al.: Kluwer.

Plank, Frans (1981): *Morphologische (Ir-)Regularitäten*. Tübingen: Narr.

Prince, Alan, and Paul Smolensky (1993): *Optimality Theory*. Ms., Rutgers University.

Raffelsiefen, Renate (1996): Gaps in Word-formation. In: Ursula Kleinhenz (ed.): *Interfaces in Phonology*. Berlin: Akademie-Verlag, 193-208.

Russell, Kevin (1997): Optimality Theory and Morphology. In: Diana Archangeli, Terence Langendoen (eds.): *Optimality Theory: An Overview*, 102-133. Oxford: Blackwell.

Stemberger, Joseph P. (1981): Morphological Haplology. *Language* 57, 791-817.

Wiese, Richard (1996): *The Phonology of German.* Oxford: Clarendon Press.

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