ON THE PHONOLOGICAL BOUNDARY BETWEEN CONSTITUENTS OF MODERN ICELANDIC COMPOUND WORDS

Summary. If the word boundary is posited between constituents of Modern Icelandic compound words, a number of mutually unrelated phonological phenomena are accounted for without any extra machinery (say, in the structure of rules). However, I have not been able to prove that any Modern Icelandic phonological phenomena actually REQUIRE the word boundary between constituents of compound words. I could only demonstrate that certain phenomena require SOME boundary between constituents of compound words; if the boundary required in those cases is identified with the morpheme boundary, certain phonological rules of Modern Icelandic have to be complicated in ways which can be avoided when the said boundary is assumed to be a word boundary.

§ 1. The Modern Icelandic lexicon contains many compound words, such as bók-menntir 'literature', til-einkadur 'dedicated', gud-fræði 'theology'. From the phonological point of view these words behave differently from simplex words in some respects, notably as regards the quantity of their stressed vowels, the permitted consonant clusters, and their stress pattern. The differences could presumably be captured in several ways, none of which have so far been applied to the Icelandic situation, to the best of my knowledge. The method used in comparable situations in other languages is to posit, between constituents of compound words, a boundary of a kind that does not occur in simplex words. Applying this to Modern Icelandic, it should first be noticed that the examination of Icelandic phonology has so far recognized the need for not more than two phonological boundaries, the morpheme boundary (+) and the word boundary (# #). Until this limitation of the kinds

1 My thanks are due to Miss Margaret G. Davis, who has improved the style of the paper. All errors are my own. The theoretical framework and the terminology of this paper are those of generative phonology as expounded by Chomsky and Halle 1968. Non-phonetic representations are bounded by the obliques, //, except in phonological derivations, where the obliques are omitted. Very often the reader will find italicized, i.e. orthographic representations within obliques. They are used to represent phonological units in those cases where I am unwilling to commit myself on the exact nature of the segments involved.

51
of Icelandic phonological boundaries is disproved, the boundary which can occur in simplex words is by definition the morpheme boundary, and the one never to be found in simplex words is the word boundary. Compound words can contain morpheme boundaries, and presumably contain at least one word boundary. To facilitate the exposition, the word boundary occurring between constituents of compound words will be designated with the ad hoc symbol §. For example, the compound word böka-safn 'library', literally 'book collection', is phonologically (roughly) /bóka§safn/.

This kind of treatment, with which the initiated reader will be familiar from elsewhere (see, for instance, Chomsky and Halle 1968 for English) takes care of many mutually unrelated phonological phenomena which would otherwise have to be accounted for less generally. Here follows a sample of such phenomena.

(I) The phonological component of Modern Icelandic grammar contains a VOWEL SYNOCOPE RULE:

\[
\begin{align*}
V  & \\
\text{stressed} & \\
\text{»tense«} & \\
\text{'elidible'} & \\
\end{align*} \rightarrow \emptyset / \begin{align*}
C  & \\
\text{coronal} & \\
\text{tense} & \\
\end{align*} + V
\]

i. e. an unstressed non-»tense« 'elidible' vowel is deleted if followed by a coronal lax consonant, the morpheme boundary, and a vowel, in that order. The rule is discussed and formulated in Orešnik 1972. The feature »TENSE« (within quotation marks) is a common property of the segments which are in Icelandic phonetics usually referred to as breid sérhjöð (Einarsson 1949:11), i. e. the diphthongs and the monophthongs i, ũ, [y] (the last one as in hugi). The feature 'ELIDIBLE' is an ad hoc feature associated with all the vowels that actually undergo the vowel syncope rule (1). Thus the unstressed vowel of mikill 'great, large' is 'elidible', and that of heimill 'at free disposal' is 'inelidible' because of contrasts such as dat. pl. miklum vs. heimilum. The fact that the ad hoc feature 'elidible' is mentioned in (1) shows that the present formulation of the rule is provisional. The feature CORONAL refers to t r n d p t d s. TENSE, this time without quotation marks, is used in its accepted meaning.

Formulation (1) shows that the rule applies across a morpheme boundary (and in fact can apply only if a morpheme boundary immediately follows the vowel to be deleted). Example: jökull 'glacier' contains an 'elidible' ũ, cf. nom. pl. jöklar from /jökul + ar/. On the other hand, the rule never applies if the vowel to be deleted is immediately followed by a §-boundary and another vowel. Cf. jökul-alda 'moraine', from /jökul§alda/, not *jökł-alda, which would be the expected normal result if the phonological representation were /jökul + alda/.

(II) Benediktsson 1969:394 has formulated, for Old Icelandic, a VOWEL TRUNCATION RULE which deletes unstressed vowels immediately followed by the morpheme boundary and another vowel. This rule still operates in
Modern Icelandic, with the result that there are no segment clusters of unstressed lax vowel plus vowel in simplex words, on the phonetic level. The rule applies, e.g., in the 1p. pl. pres. köllum from /kalla + um/, and in the 2p. pl. pres. kallid from /kalla + id/, of kalla 'call'. (That the stem of kalla is phonologically bisyllabic, /kalla/, follows from the singular present indicative forms kalla(r), the imp. sg kalla, and from the dental stem kallad.) On the other hand, the rule does not apply in compound words like sögu-eyja 'saga island', from /sögu(+)eyja/. If the phonological representation were /sögu(+)(eyja/), we would expect *sögeyja on the phonetic level.

(III) A d is usually inserted between an n and the immediately preceding r. Einarsson 1949:21. The insertion is performed by the D-INSERTION RULE. The rule can apply even if the r and the n are separated by a morpheme boundary. Cf. dat. pl. förnum of farinn 'gone', from /far+in+um/, whereupon the umlaut rule and the d-insertion rule produce förnum, pronounced with d before n. On the other hand, the d-insertion rule does not apply across a §-boundary. Cf. stór-netla '(plant) urtica dioeca', without d before n, from /stör§netla/. If the phonological representation were /stór(+)netla/, the expected phonetic representation would contain d before n.2

(IV) The consonantal segments which are realized as non-preaspirated stops in intervocalic position, are preaspirated if they immediately precede n within the word. Einarsson 1949:23. The sandhi rule which affects this change can apply even across a morpheme boundary. Cf. dat. pl. sopnum, with preaspirated p, from /sop+in+um/, of sopinn 'drunk'; dat. pl. getnum, with preaspirated t, from /get+in+um/, of getinn 'begotten'; dat. pl. auknum, with preaspirated k, from /auk+in+um/, of aukinn 'increased'. (In these examples, the vowel syncope rule (1) first brings the lax stop and the n into contact, whereupon the stop is preaspirated across the intervening morpheme boundary). On the other hand, the rule which preaspirates the stops in the above examples cannot operate across the §-boundary. Cf. djúp-nökkvi '(seal) han-

---

2 In order to diminish the number of the unknowns, I have here and elsewhere in this paper avoided those phonological rules concerning which I cannot prove that the presence of the morpheme boundary does not block their operation. Thus there is a change of rl to rdl (with subsequent loss of r in a number of cases), very similar to the change of rn to rdn (and further to dn in a number of cases) described sub III. I know of no reliable examples proving that (r)dl ever comes from phonological /rl/ if /r/ and /l/ are separated by a morpheme boundary. Examples like the contracted cases ferl, pronounced with rdl, of ferill 'trace, path', show that the d-insertion rule must operate in the context r—l, but are not reliable instances of the substring /r+1/; the morpheme boundary after fer- is uncertain, for we lack (a) any phonological evidence for such a boundary in ferill and similar words. In fact, I do not know of any phonological phenomena of Modern Icelandic that would REQUIRE the positing of morpheme boundaries whose decisive motivation would come from facts of word derivation. As will be seen below, the situation is different with regard to word composition.

(b) a theory which would tell us whether jara 'go, travel' and ferill (historically the latter is a derivative of the former) are sufficiently similar to each other semantically and phonologically for the word-formational relation between them to exist synchronically, and thus help motivate a morpheme boundary after fer- in ferill.
leyja abyssorum', with non-preaspirated $p$, from /djúp$nökkvi/; mat$nìdingur 'one stingy with food', with non-preaspirated $t$, from /mat$nìdingur/; bak$naga 'slander', with non-preaspirated $k$, from /bak$naga/. If the three examples were phonologically /djúp(+)nökkvi/, /mat(+)nìdingur/, and /bak(+)naga/, respectively, their $p$, $t$, and $k$ would be preaspirated on the phonetic level. See also footnote 2.

(V) The phonological fricatives /$v$/ and /$q$/ (i.e. voiced velar fricative) become stops when immediately followed by $n$ and a vowel. Einarsson 1949:13—14. The process is not blocked if a morpheme boundary intervenes between the sounds involved in the change. Cf. nom. pl. masc. sofinir, with $b$ before $n$, from /sof+in+ir/, of the past part. sofinn 'slept'; nom. pl. fem. dregnar, with a velar stop before $n$, from /dreg+in+ar/, of the past part. dreginn 'drawn'. No such process applies across the $§$-boundary. Cf. of$næmi 'allergy', with [vn], from /of$næmi/; hag$nyta 'use', with [qn], from /hag$nyta/. If the phonological representation of the two compounds were /of(+)næmi/ and /hag(+)$nyta/, respectively, we would expect to find [bn] and [gn] in them on the phonetic level.3 See also footnote 2.

(VI) Whenever any number of segments from the set \{t, d, p, $l$, \} immediately follow each other in a simplex word on the phonological level, and such a consonantal group is not accompanied by additional consonants, a sandhi rule applies to coalesce such a consonantal cluster into a long preaspirated $t$ if at least one of the original segments is $t$, and into a long $d$ otherwise. The following examples show that the coalescing rule can apply even across a morpheme boundary (Einarsson 1949:54, 82—83):

(2) (a) Preterite and past participle stem of regular weak verbs:

root + dental suffix (± /t/)

- hréda 'frighten', pret. and past part. stem hrédd-
- mæta 'meet' mætt-
- brynda 'border, edge' brydd-
- hitta 'hit' hitt-

(b) Nom./acc. sg. ntr. of strong adjectives: root + $t$

- gladur 'glad', nom./acc. sg. glatt
- latur 'lazy' latt
- saddur 'satisfied' satt
- brattur 'steep' bratt

3 Phonetic [qn] exists in Modern Icelandic, cf. the contracted cases brugtn- [brYqn-] of the past part. brugtnn 'moved quickly'. Pending an examination of the Modern Icelandic internal sandhi rules I assume for the time being that the rule which changes /$q$/ into a stop before /$n$/ precedes the loss of /d/ in the context $q$—$n$. The following derivation results, for the strong nominative plural feminine:

- vowel syncope rule: brugd$+$in$+$ar
- $q \rightarrow g$ /$n$/ non-applicable
- $d \rightarrow \emptyset$ /$q$—$n$/ [brYqnar]

54
On the other hand, the coalescing rule does not apply across the §-boundary. Cf. *hvít-pinur* '(plant) abies alba', from */hvít$pinur/, not from */hvít(+)$pinur/, which would yield preaspirated *tt* instead of *tp*; *rit-deila* 'polemics (in the press, etc.)', from */rit§deila/, not from */rit(+)$deila/, which would yield preaspirated *tt* instead of *td*; *rótt-tækur* 'radical', from */rótt$tækur/, not from */rótt(+)$tækur/, which would yield preaspirated *tt* instead of *tđ*; *blóð-þyrstur* 'bloodthirsty', from */blóð$þyrstur/, not from */blóð(+)$þyrstur/, which would presumably yield *dd* instead of *tđ*; *støð-deyfing* 'local anaesthesia', from */støð§deyfing/, not from */støð(+)$deyfing/, which would yield *dd* instead of *tđ*. Etc.

(VII) The consonantal segments which are realized phonetically as the preaspirated labial stop (*pp*), as the non-preaspirated labial stop (*p*), as the preaspirated velar stop (*kk*), and as the non-preaspirated velar stop (*k*) become fricatives [*f*] and [*x*], respectively, in the context *V-*t in simplex words. Einarsson 1949:17, 20, 29. The process is not stopped by an intervening morpheme boundary, as the following examples show:

(3) imperative singular with postposed personal pronoun:

- *súptu* 'sip', with [*ft*], from */súp+tu/*
- *taktu* 'take', with [*xt*], from */tak+tu/
- *slepptu* 'let go', with [*ft*], from */slepp+tu/
- *slökkktu* 'quench', with [*xt*], from */slökkk+tu/

(VIII) In Orešnik 1971 b I mentioned the Modern Icelandic PALATALIZATION RULE, which accounts for the fact that velar consonants are almost invariably palatalized if immediately followed by *i*, *e*, or by diphthongs that begin with *i* or *e*. The palatalization rule is not blocked when it applies across a morpheme boundary, as shown by the weak nom. sg. masc. *hagi*, with palatalized /q/ and diphthongized root vowel, from */hag+i/, of *hagur* 'skilful; elaborate'. On the other hand, the palatalization rule cannot apply across the §-boundary. Cf. *hag-yrdingur* 'rimester', from */hag§yrdingur/, with non-fronted *g* before the §-boundary, and consequently without the diphthongization of the *a*.

(IX) Stressed vowels are phonetically long if immediately followed by just one lax consonant within the simplex word (or by certain consonant clusters, see Einarsson 1949:4; such clusters will be disregarded in this discussion). Otherwise stressed vowels are short. When the stressed vowel is separated by a morpheme boundary from the consonant which determines, or from (a part of) the consonants which determine, its quantity, this circumstance has

4 The non-phonetic representations of (3) within the obliques are the not necessarily phonological representations to which the sandhi rule discussed sub VII applies.
no bearing upon the quantity of that vowel. Cf. ást 'love', with phonetically short á and no morpheme boundary anywhere in the word; the middle infinitive ást of á 'rest and graze horses', likewise with a phonetically short á, and with a morpheme boundary before st; strong nom./acc. sg. ntr. fut'st. of júst 'willing', again with a phonetically short vowel, and with the morpheme boundary between s and t. On the other hand, if the stressed vowel and the immediately following st are separated by the §-boundary, the stressed vowel is phonetically long. Cf. á-stæda 'ground, reason', from /ªstæda/, with phonetically long á. If the phonological representation were /ª(+)-stæda/, the present rules, as formulated on the basis of the situation in simplex words, would produce, wrongly, phonetically short á.

§ 2. The discussion in I—IX has shown that the identification of the boundary between constituents of compound words with the word boundary takes care of many phonological phenomena — some of which have just been sketched — typical of compound words, without the need for any extra machinery. This of course speaks in favour of the said identification. However, it should be noted that the phenomena under discussion, in fact, as far as I know, Icelandic phonological phenomena in general, do not seem to REQUIRE this particular treatment. To see this, consider again the situations described in I—IX.

It suffices to emend the rules discussed sub I—VII with the stipulation that the first vowel after the segments to be affected by the rules in question be unstressed, and the operation of the rules is blocked in the relevant environments in compound words without further complications, as far as can be seen. The alteration of course presupposes that the stress is allocated to vowels before the operation of the rules discussed sub I—VII. This is the case anyway. One of the rules discussed, the vowel syncope rule, happens to be one of the earliest rules in the Modern Icelandic phonological component, and all the remaining rules discussed here follow it in the ordering.5

5 The relative ordering of the vowel syncope rule and of the vowel truncation rule has not yet been determined, but both are certainly «early» rules, and even considered subparts of the same rule by one scholar (Benediktsson 1969:394).

The sandhi rules discussed sub VI and VII follow the vowel syncope rule: the latter creates some of the segment clusters to which those sandhi rules apply. Cf. the derivations of the lp. sg. pret. subj. mætti of mæta 'meet', and sleppiti of sleppa 'let go':

<table>
<thead>
<tr>
<th>vowel syncope rule</th>
<th>sandhi rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>mæt+i+D+i</td>
<td>slepp+i+D+i</td>
</tr>
<tr>
<td>mæt+D+i</td>
<td>slepp+D+i</td>
</tr>
<tr>
<td>mæ[ht]+i</td>
<td>sle[ht]+i</td>
</tr>
<tr>
<td>[mæ[mh]t]</td>
<td>[sle[mh]t]</td>
</tr>
</tbody>
</table>

[D] is a cover symbol for a dental consonant from the set t, d, p, ð. For the justification of the representations [mæt+i+D+i] and [slepp+i+D+i] being tri-syllabic, see Orešnik 1971 a.

For the arguments concerning the ordering of the remaining rules mentioned in the main text, with respect to the vowel syncope rule, see Orešnik 1972.
It follows from formulation (1) that the stress must be determined — lexically and/or by a rule — by the time the vowel syncope rule applies, and consequently the distribution of the stresses can also be utilized when the remaining rules apply. The said alteration even abolishes the need for any kind of boundary between the constituents of compound words. On the other hand, the revision has two unpleasant consequences. First, the rules are now not only more complicated than they were before the reformulation, but also less general, because the adjustment, occasioned by the facts typical of compound words, does not have any positive effects outside compound words. Secondly, once the non-morpheme boundaries are barred from compound words, the stress can only be allotted lexically. To appreciate this point consider how the stress can be assumed to be placed if the constituents of compound words are allowed to be separated from each other by word boundaries. There must be an early rule (operating before the vowel syncope rule, as explained above), which assigns [+stress] to the first vowel after a word boundary. Given the phonological representation /jókulšalda/ of jókul-alda 'moraine', the rule would stress /ó/ and the constituent-initial /a/. Next there must be a rule which weakens certain stresses that are not word initial to different degrees, taking into consideration the syntactic structure of the compound in the process. For example, skóla-bóka-safn, literally ‘school book collection’, has one stress pattern when the word means ‘collection of school books’, and another when it means ‘school library’. Now, if the only boundaries permitted in compound words are morpheme boundaries, and the syntactic structure of the compound words is not indicated, this system of stress rules cannot produce the desired phonetic results. Consequently the stress must be allotted lexically in such a case.

However, while the modification of the rules sub I—VII can thus be seen to be somewhat disadvantageous, it is not unrealistic in the sense of not obeying the established constraints on the form of phonological components. It is in this sense that it can be asserted that the rules described sub I—VII do not require that there be non-morpheme boundaries in compound words. Moreover, those rules do not require any boundaries at all between the constituents of compound words.

A different situation obtains with the revision of the palatalization rule, discussed sub VIII. To prevent the fronting of g in hag-yrdingur 'rimester', it is not enough that the fronting vowel be stressed, for stressed vowels do front preceding velar consonants in simplex words. Cf. the singular present indicative forms geng-, skaf-, kem-, etc., with fronted velars before the stressed e, as against their respective present stems gang-, skaj-, kom-, of the strong verbs.

---

6 What follows builds on the analogy of the relevant aspects of the English stress rules as expounded by Chomsky and Halle 1968.

7 It is assumed here that strings such as ]+[, where the square brackets indicate the syntactic structure, are impossible on the phonological level. Since this assumption makes it more difficult, rather than easier, for me to prove the existence of the word boundary between constituents of compound words, I accept it for the sake of argument without further discussion.
ganga 'go', skafa 'scrape', and koma 'come'; cf. also the preterite subjunctive forms kæmi, skæfi, kefði, gysi, kynn, etc., with fronted velars before stressed vowels, as against the respective preterite indicative stems kom-/kóm-, skóf-, kajfð-, gus-, kunn-, of koma 'come', skafa 'scrape', kefþa 'suffocate', gjósa 'gush', kunna 'know how to'. It is therefore necessary to reformulate the palatalization rule so that a stressed vowel does not palatalize a preceding velar if the two are separated by any kind of boundary. Thus the palatalization rule offers precious positive evidence that there must be SOME boundary between the constituents of compound words in Icelandic.

The same important result is achieved if the vowel quantity rule, discussed sub IX, is altered to accommodate the hypothesis that compound words do not contain any non-morpheme boundaries. Suppose that the vowel quantity rule lengthens vowels before a single lax consonant (the proviso, mentioned sub IX, about certain consonant clusters as allowing lengthening, is still to be kept in mind). In this case the vowel quantity rule would fail to lengthen the word-initial vowel of the compounds í-stað 'stirrup' and íst-turn 'ice tower', for their phonological representation would be /i+stað/ and /is+turn/, respectively, the morpheme boundary would be disregarded (cf. ást, á+st, and fús+t, discussed sub IX above), and the rule blocked because of the cluster /st/ which follows the i to be lengthened. The desired phonetic output, with long i, would not be secured. And it does not help to stipulate that the vowel quantity rule is blocked if the next vowel in the compound word is stressed, for this stipulation would still fail to differentiate between, say, á-stæða 'reason, ground', with phonetically long á, and ást-úð 'lovable character, kindness', with phonetically short á. Obviously the rule must be formulated so that it counts the postvocalic consonants only as far as the first boundary, which is after á in á-stæða, and after t in ást-úð. This of course is tantamount to saying that the vowel quantity rule requires that there be a boundary between constituents of compound words, otherwise the correct phonetic results cannot be obtained without calling in hopelessly ad hoc machinery. — The same conclusion would have been arrived at if the vowel quantity rule were formulated, not as lengthening certain vowels (as has just been done above), but as shortening certain other vowels.8

Thus, while I still have not been able to present positive evidence for the word boundary between constituents of compound words, I hope to have succeeded in demonstrating that at least two Modern Icelandic phonological rules, viz. the palatalization rule and the vowel quantity rule, require that there be SOME boundary between the constituents of compound words.

If the boundary between the constituents of certain compound words which seems to be required by at least two phonological rules is to be more than an ad hoc device limited to just those compound words, some further motivation must be found for it. Such motivation could possibly be found in a theory 8The uninitiated reader should be warned that the facts presented here in the discussion of the vowel quantity rule by no means exhaust the problems connected with that rule. See, for instance, Bergsveinsson 1941:84—86, for the description of some additional relevant facts.
which would identify at least \((n-1)\) of the \(n\) constituents of any compound word with some word/stem to be found in the same lexicon. Thus, if the jökul- of jökul-alda 'moraine' is identified with the stem jökul- of the simplex word jökull 'glacier', the boundary between jökul- and -alda is motivated. Furthermore, the theory I have in mind would automatically extend this type of motivation beyond the examples which have originally stimulated the search for such motivation, to all compound words the \((n-1)\) of whose \(n\) constituents can be identified with other words/stems in the lexicon. It can be foreseen that semantics will play an important part in such a theory, seeing that the link between, say, jökull 'glacier' and jökul-alda 'moraine' is primarily semantic. Until that theory is constructed, we must try to establish the existence and the nature of the boundary between constituents of compound words on purely phonological grounds. While the present paper may have established the EXISTENCE of a boundary between the constituents of at least some Modern Icelandic compound words, the NATURE of that boundary remains a field for further research.

REFERENCES

Orešnik, Janez. »On some weak preterite subjunctives of otherwise strong verbs in Modern Icelandic« in Arkiv för nordisk filologi 86:139–78. Lund, 1971. (Referred to as Orešnik 1971 a.)
Orešnik, Janez. »Morphophonemic notes on the Modern Icelandic imperative singular.« Forthcoming. (Referred to as Orešnik 1971 b.)
Orešnik, Janez. »On the epenthesis rule in Modern Icelandic.« Forthcoming. (Referred to as Orešnik 1972.)

POVZETEK

V duhu generativne fonologije se raziskuje vprašanje, ali je med sestavnimi deli novoislandskih sestavljenih besed kaka fonološka meja. Dveh pravil novoislandske slovnice — tistega o mehčanju zadnjenebnih soglasnikov in tistega o fonetični dolžini poudarjenih samoglasnikov — sploh ni mogoče pravilno izreči, če se na omenjenem mestu ne postulira obstoj kake meje, morfemske ali besedne ali katere druge. Odprto pa ostaja vprašanje o naravi te meje.