

Syllabification in Southern Sámi*

Mikael Vinka**

1. Introduction

This paper will present an analysis of syllabification in Southern Sámi. We will consider the change in quantity in stem vowels, indicated by bold type in (1), which is determined by the phonological shape of the suffix, and the allomorphic variation observed in the suffixes, indicated by italics¹:

(1)	<i>singular</i>	'meat'	'shoe'
	nom.	berko:	ga:megə
	acc	berko:m	ga:megəm
	gen	berko:n	ga:megən
	ill	berko:se	ga:me:gasse
	iness	berko:sne	ga:me:gisnie
	elat	berko:ste	ga:me:gistie
	com	berko:jne	ga:me:ginnie

I will propose an analysis of these phenomena within moraic theory, as outlined in Hayes 1989 and related work. The general principles of moraic theory in combination with a few well-defined language particular constraints will be shown to handle the data in (1). It will be argued that the crucial factor in determining the surface forms hinges on whether a stem is vowel or consonant final.

*The main sources used for this paper are Knut Bergsland's traditional grammar *Sydsamisk grammatikk* and the Sámi-Norwegian dictionary *Aarjelsaemien-Daaroen Baakogærja*, edited by Knut Bergsland and Lajla Mattson-Magga, and the Swedish-Sámi dictionary *Vilhelmina Vefsen*, compiled by Erik Nilsson-Mankok. The phenomena treated in this paper are briefly touched upon in Bergsland's grammar, however in a sweeping theory neutral framework. Moreover, I have gathered additional information from recordings of Southern Sámi speech, mainly from educational material (Jaahkenelkien 1993).

**Lund University and McGill University. I would like to thank Professor Heather Goad for helpful discussions, comments and criticism of this paper. Thanks are also due to Professor Glyne Piggot, Hiro Hosoi, Walcir Cardoso and Steve Jimerson. A special thanks to Professor Jan-Olof Svantesson for making this paper fit into *Working Papers*! All errors are of course mine.

¹The following abbreviations will be used systematically throughout the paper: nom=nominative, acc=accusative, gen=genitive, ill=illative, iness=inessive, elat=relative,

2. Moraic theory

In this section I will outline the relevant parts of the basic theory adopted in the paper, namely moraic theory as outlined in Hayes 1989, 1995. We will particularly consider the following two issues: (i) syllabification algorithm and (ii) quantity-sensitivity.

Moraic theory provides us with three things of importance for approaching phonological phenomena related to the syllable. First of all, it imposes a syllabification algorithm which determines how individual segments are to be parsed into syllables. Secondly, it provides a straightforward account for the issue of how the quantity of segments such that 'long' vs. 'short' are assigned different structural configurations. Finally, the sum of the quantity of the segments within the domain of the quantity bearing units in the syllable determine the quantity of the entire syllable, hence the notion of quantity is relevant at dual levels, as it were (see Hayes 1989, 1995; McCarthy & Prince 1995). In what follows we will consider each of these aspects of moraic theory as they pertain to the issues to be treated in this paper.

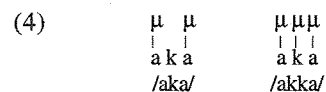
In moraic theory syllables are considered to be built up from weight bearing constituents, moras. Moras are typically present in the underlying representations of lexical items, and since the vocalic moras constitute the nucleus of the syllable, it follows that it is vowels that are dominated by moras underlyingly. Thus, assume that the underlying representation of the string /ka/ is as shown in (2), where μ =mora. Furthermore, in accordance with Hayes 1989:256, a string /ka:/ in which the vowel is long has the structure shown in (3), where the segment /a/ is associated with two moras. Thus, in a language with phonemic vowel length, potential monosyllabic words like [ka] and [ka:] are distinguished by the representations (2) and (3) respectively.



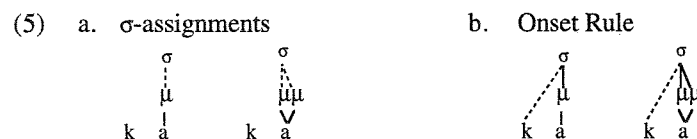
However, it is a well-known fact that not only vowel length but also consonant length may contrast. Japanese, for instance, provides many examples of this. Thus, the words /aka/ 'red' and /akka/ 'degeneration' are distinguished by whether /k/ is short or long. In moraic theory, consonants are generally not assigned moraic content underlyingly, except for gemi-

com=comitative, sg=singular, du=dual, pl=plural, prs=present tense, pst=past tense, ptc=participle, prog=progressive. Also, \hat{a} =[o], o=[u].

nates. Therefore, in /aka/ only the vowels are dominated by moras, whereas in /akka/ the consonant is as well. In (4), then, we have an illustration of the underlying difference between geminate consonants and plain ones:

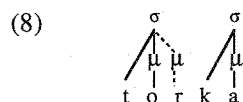
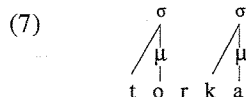
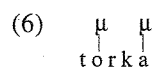


Next we will consider how syllables are constructed from the underlying representation. I will adopt the syllabification algorithm proposed in Hayes 1989, according to which syllabification proceeds in the following bottom-to-top fashion. Consider the underlying representation (2), /ka/. The mora is selected for domination by a syllable node, σ , which yields (5a). The next step is adjunction of the prevocalic consonant /k/ to the syllable, which henceforth will be referred to as Onset Rule, as is shown in (5b). The syllabification of (3), /ka:/ is given in parallel in (5).



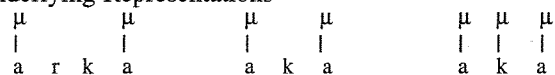
It is, however, not necessarily the case that a segment is underlyingly associated with a mora in order for it to be dominated by a mora in the syllable. The case in mind is that of closed syllables. For example a string like /torka/ has the underlying representation (6). Clearly, since /r/ is not geminate it cannot be underlyingly dominated by μ . When the syllabification rules introduced above are applied to the underlying representation (6), we obtain (7). As seen in (7), all segments but /r/ are licensed by some position in the syllable. Moreover, /rk/ cannot be parsed as a complex onset, since this would violate the sonority sequencing principle, which requires equal or preferably rising sonority in the onset (see e.g. Rice 1992). This means that /r/ must belong to the initial syllable. Hayes 1989:258 suggests that /r/ is assigned a mora by a rule Weight by Position (WBP), applying to consonants which are postvocalic but not prevocalic². This rule is ordered after the Onset Rule, illustrated in (5b). Thus we assume that WBP applies on (7), which results in the structure in (8).

²As pointed out in Hayes 1989:258, the rule Weight by Position is specific to languages where CVC syllables count as heavy, on a par with CVV.

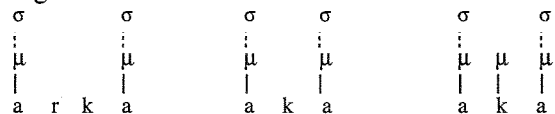


The reason why WBP is ordered after the Onset Rule rather than before it, is straightforward. If WBP were to apply before the Onset Rule, we would not be able to capture the fact that /aka/ and /akka/ in (4) constitute a minimal pair. Take /aka/ for example. Assuming that WBP applied before the Onset Rule, /k/ would be moraic by the time it is picked up as the Onset of the second syllable, which would incorrectly entail the surface form [akka], indistinguishable from the surface form of /akka/. Thus, WBP applies on unparsed non-moraic consonants in quantity sensitive languages. This is illustrated in (9), where the derivations of forms like /arka/, /aka/ and /akka/ are presented in detail.

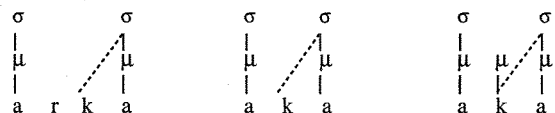
(9) a. Underlying Representations



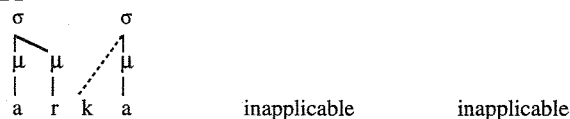
b. σ -assignments



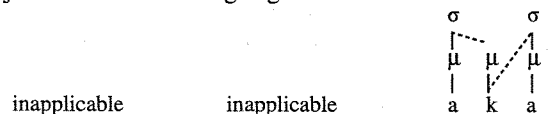
c. Onset Rule



d. WBP

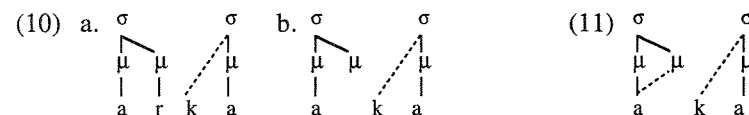


e. Adjunction of remaining segments



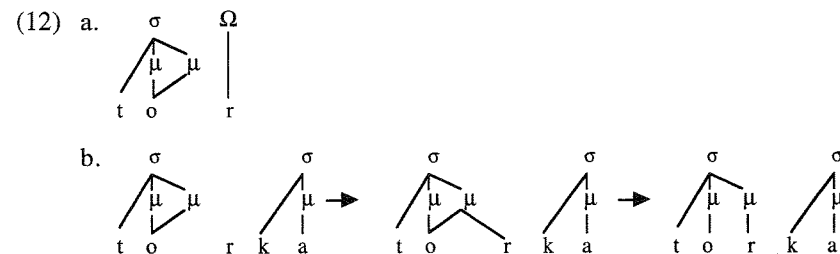
As seen in (9e), the final rule in the algorithm basically adjoins remaining moras to a syllable node. Thus, the lexical mora of /akka/ is adjoined to the initial syllable, which then means that /k/ spans over two syllables.

We have now introduced the syllable building machinery, and we have seen how it operates on underlying representations. Now we will turn to another aspect of moras, namely their property of being weight bearing. Hayes 1989 showed how moraic theory can account for phenomena such as compensatory lengthening, in a way that surpasses the treatments provided by CV and X-slot theory. Consider our example-word /arka/ from (9), where it was shown that /r/ is assigned a mora by WBP. Compensatory lengthening (CL) is characterized by the following: the deletion of a post-vocalic consonant renders lengthening of the vowel. Assume, for example, that /r/ of /arka/ were deleted for one reason or another, yielding (10b). This results in a 'dangling' mora. Vowel lengthening can now be analyzed as spreading of the vowel into the vacated mora, yielding (11).



As pointed out in Hayes 1989 and McCarthy & Prince 1995, moraic theory correctly predicts that compensatory lengthening will not occur if an onset consonant is deleted, since onsets are not weight bearing, or moraic.

Moraic theory also provides a straightforward account for closed syllable shortening, as pointed out by Kenstowicz 1994:437: "closed syllable shortening can be seen as a repair strategy to maintain a restriction against mora sharing." Thus, assume that we have a sequence /to:r/, where the final /r/ is extrasyllabic. Let us for the sake of convenience assume that extrasyllabic consonants are licensed in an Appendix, labeled Ω in (12a) (cf. Goldsmith 1990:127). If a suffix, for instance, attaches to /to:r/, /r/ is no longer at an edge and hence it can no longer be licensed as extrasyllabic. In order to escape stray erasure, /r/ adjoins to a preceding mora, which is the second of two which dominates /o/. Subsequently the vowel delinks, and hence shortens, as can be seen in (12b):

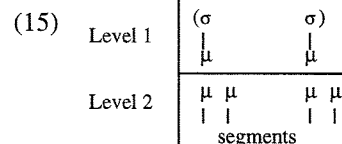
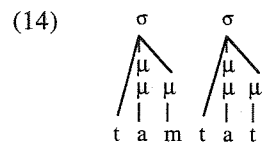


Moraic theory makes the prediction that neither compensatory lengthening nor closed syllable shortening will occur in quantity insensitive languages, since 'coda' consonants in such languages are not moraic.

Both CL and closed syllable shortening are examples of the relevance of weight as it pertains to segments, i.e. the relation mora-segment. It also a well-known fact that the relation syllable-mora is of great importance. The distinction between heavy and light syllables is a consequence of whether a syllable is bimoraic (heavy) or monomoraic (light). Syllable weight is crucial for the organization of syllables into feet. Feet are typically bimoraic. Thus in languages with trochaic feet, which we are concerned with in this paper, a foot consists of one heavy syllable or two light:

- (13) Possible trochees: $(\sigma_{\mu\mu})$ $(\sigma_{\mu} \sigma_{\mu})$

However, there are languages which do not conform easily to the standard foot typology in (13), Southern Sámi being one of them. In this language we find several indications that the language is quantity sensitive, as we shall see later. However, with respect to stress, it appears that syllables are parsed into feet in a left-to-right direction paying no attention to whether the syllables are heavy or light. Thus we have an asymmetry between the relation mora-segment on the one hand, and the relation syllable-mora, on the other; in the former case the language appears to be quantity sensitive, and with respect to the latter relation we seem to deal with a quantity insensitive language. In order to resolve this paradox, I will adopt the possibility of a two-layered moraic representation, as suggested by Hayes 1994:302 for Yupik languages. Consider the following abstract example (14). With respect to feet, the relation σ -level 1 μ counts, hence parsing the two syllables in (14) into one trochaic foot, which is illustrated in (15). However, the weight-depending properties of say closed syllable shortening, licensing of geminate consonants, operates between level 2 μ and the segments. This approach captures, as we shall see later, the fact that stress in Sámi consistently falls on every odd syllable, irrespective of apparent syllable weight.



3. Stems, inflectional morphology and allomorphy

We can observe that the major lexical categories noun, verb and adjective in Southern Sámi behave phonologically similarly with respect to inflectional suffixation. (The language lacks productive prefixes.) Two patterns are discernible, involving the length of the vowel in the syllable immediately preceding the suffix. The phonological shape of the suffix is crucial in determining vowel length. The two patterns are in complementary distribution. For a suffix of a particular shape, the relevant vowel is long in one type of stem, whereas it is short in the other type of stem. Consider the examples in (16), where a single consonant suffix has been added to the stems³. In the left column the vowel of the syllable preceding the suffix is long, whereas the vowel is short in the column to the right. In (17), the situation is reversed. Here the words in the left column have short vowels, whereas the right hand column displays long vowels. The bare forms of the stems, which are obtained by zero-suffixation (18), resemble (16)⁴:

- (16)
- | | | | |
|---------|----------------|---------|------------------|
| berko:m | 'meat.sg.acc' | ga:megm | 'shoe.sg.acc' |
| båata:h | 'come.prs.2sg' | dåeredh | 'follow.prs.2sg' |
- (17)
- | | | | |
|----------|-----------------|--------------|------------------|
| berkose | 'meat.sg.ill' | ga:me:gasse | 'shoe.sg.ill' |
| berkosne | 'meat.sg.iness' | ga:me:gisne | 'shoe.sg.iness' |
| båteden | 'come.prs.2du' | dåere:didden | 'follow.prs.2du' |
- (18)
- | | | | |
|--------|----------------|----------------|------------------|
| berko: | 'meat.sg.nom' | ga:meg/ga:megə | 'shoe.sg.nom' |
| båata: | 'come.prs.3sg' | dåered/dåeredə | 'follow.prs.3sg' |

Now, as we can see in (18), the left side stems appear to be vowel final, whereas the right side stems are consonant final. There is convincing evidence that this difference is the underlying reason for the asymmetrical behavior of the two types of stems. For example, if we causativize the verb *båete*: by adding the causative formative /t/, the derived verb *båetet* behaves on par with *dåered*. This is seen in (19):

- (19)
- | | | |
|---------|--------------|--------------|
| båata:h | båeteth | dåeredh |
| båteden | båete:tidden | dåere:didden |
| båata: | båetet | dåered |

We can thus make the following observational generalization. In a vowel-final stem, the vowel is long if the affix consists of a single

³In this paper I ignore the issue of ablaut and umlaut in Southern Sámi. See Vinka 1997 and Bergsland 1994.

⁴Notice the variation *ga:meg/ga:megə*. See section 5 below for some discussion concerning epenthetic vowels.

consonant. In a consonant-final stem, the vowel is short if the affix is null or a consonant. On the other hand, the vowel is short in a vowel-final stem if the suffix consists of one or more syllables, in contrast to consonant-final stems in which case the vowel is long. This is summarized in (20):

(20)

shape of stem	shape of suffix	
	C	σ
C-final	vc-c	vvc- σ
V-final	vv-c	v- σ

So far we have concentrated on what happens to the stem to which a suffix attaches. Let us now go back to (17) where we can observe allomorphic variation in the inflectional affixes. In (17) we find the singular illative suffix *-asse/-se*, inessive *-isne/-sne*, and second person dual present *-idden/-den*. As is clear, the suffixes are syllabic. We will assume in this paper that the bisyllabic form is underlying and that the short form is derived. The distribution of the allomorphs is such that the bisyllabic (vowel-initial) variant attaches to consonant-final stems, whereas the short (monosyllabic) allomorph attaches to vowel-final stems, as made explicit in (21).

(21)

berko-se	ga:me:g-asse
berko-sne	ga:me:g-isne
båete-den	dåere:d-idden

This subsection has outlined the general properties of variation in vowel quantity in Southern Sámi, and the basic pattern of allomorphy in the inflectional system. It will be shown that these phenomena can be accounted for in terms of syllabification and phonotactic constraints that hold in the language. However, before we embark on our investigation we must make clear some general phonotactic patterns, since they will turn out to be crucial for our understanding of the above.

3.1. General phonotactics of Southern Sámi

Sámi allows both long vowels and geminate consonants. However, the distribution of geminate consonants is rigidly restricted to occur only between odd and even numbered syllables:

(22)

a. aajjove	c. *aajovve
b. gaameegasse	d. *berkosse



In this respect it is also interesting to notice that the set of possible coda-onset clusters between odd and even numbered syllables is a superset of the

ones possible between even and odd syllables. In (23) we find a sample of possible coda-onset sequences between odd and even syllables. In (24) we can see that in non-derived contexts the only occurring consonant combinations between even and odd syllables are /rd/, /rg/, /ld/ /lg/:

(23) a. ber.koo	(24) a. jaa.mel.ge
b. lop.me	b. æv.ler.ge
c. gáp.čedh	c. jemh.kel.de
d. gag.ne.de	d. goek.ker.de
e. gieb.nie	e. *jaa.meg.nie
f. ñåam.ka.fid	f. *ñåa.kam.ka.fid

These facts can be understood as restrictions holding between syllables within the same foot, and as restrictions between adjacent syllables belonging to different feet. Thus, geminates can only occur within a foot, and they are prohibited in the context of (25a), being restricted to (25b):

(25)

a. *	b. $\sqrt{\quad}$
	

Furthermore, with the exception of the word initial syllable, all syllables must have an onset, a property which is fairly common across languages. Thus we do not find sequences like ...CV.VC... where each V is the nucleus of a syllable. This is seen in (26):

(26)

a. $\sqrt{\quad}$ aaj.jo.we	*aj.ow.e
b. $\sqrt{\quad}$ råw.na	*rå.u.na
c. $\sqrt{\quad}$ bar.ke.me	*bar.ke.am.me

We can thus formulate a constraint banning onsetless syllables:

(27) * $_{\sigma}$ [V (unless word initial)

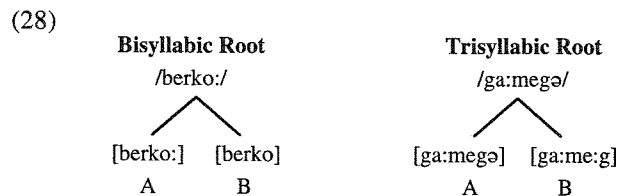
In this paper, it will be shown that (25) and (27) play a crucial role in determining the allomorphy observed in the previous subsection.

To sum up, we have seen that consonant clusters and geminates may occur more freely between odd and even syllables, than between even and odd ones. Throughout the paper we will assume that this is a reflection of the existence of feet. Finally, Sámi syllables require onsets, unless the syllable is word initial.

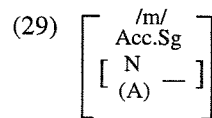
4. Bergsland's account for allomorphy

There exist no generative accounts of the alternation patterns observed in the previous section. In fact, the literature on Southern Sámi is overall extremely limited. The only work seriously treating the language is Bergsland 1994, who in 139 pages covers all aspect of Southern Sámi grammar. In spite of its unpretentious size, the work is considered the authority on Southern Sámi. Being a traditional grammar, Bergsland's analyses, to the extent that they exist, naturally differ from what is being done here. However, in this section we will take a look at how Bergsland approaches the present issues, and in doing so, we will view his approach from a generative perspective.

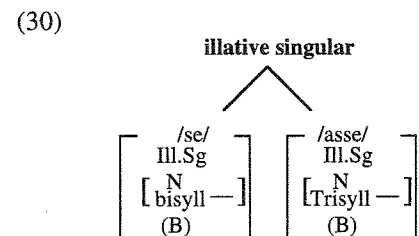
Bergsland accounts for these variations in the following way. He takes the nominative forms of nouns to equal their underlying, or basic, representation. The nominative form of 'shoe' is [ga:megə] and the nominative of 'meat' is [berko:]. Hence, the former is trisyllabic whereas the latter is bisyllabic. The accusative morpheme *-m* simply attaches to any nominal root or stem, yielding the forms [berko:m] and [ga:megəm]. The situation in the examples [berkose] and [ga:me:gasse] is somewhat trickier. Bergsland observes that each root comes in two shapes, so that trisyllabics like /ga:megə/ have one form where the second vowel is short, and another in which the second vowel is long. In (28) we equip the basic form with a tentative diacritic A, and the other form with a diacritic B:



Next, morphemes like the accusative singular *-m* simply attach to any root marked by A, which can be obtained by providing a specification in the lexical entry of the suffix, stating that the accusative singular morpheme attaches to any category N with the diacritic feature (A):



As for the illative singular suffix the situation becomes somewhat more complicated. Bergsland does not take a stand on which allomorph of *-se* and *-asse* is the underlying of the two, but rather stipulates something like (30):



Hence, /se/ selects bisyllabic roots of type (B), i.e. the variant with a short vowel, and /asse/ selects trisyllabic roots of type (B), yielding [berkose] and [ga:me:gasse] respectively.

As a descriptive account, Bergsland's solution certainly makes the right predictions. However, it has a large degree of redundancy, and it is also fairly contradictory. Consider for example /asse/ in (30). /asse/ sub-categorizes for a trisyllabic root, which means that it will attach to a noun like /ga:megə/ 'shoe', in (28). Nevertheless, the variant of the root to which it suffixes, namely [ga:me:g], is hardly trisyllabic. In fact, the only way for it to be trisyllabic would be to equip this form with a feature [+trisyllabic], which would only be relevant for morphological purposes. It is thus clear that a traditional account like Bergsland's does not meet the standards of a rigidly constrained theory like Principles and Parameters.

5. Syllabification

5.1. Consonant final stems

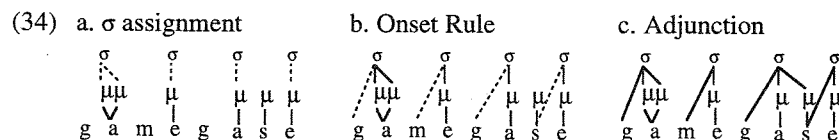
This section will provide an analysis of syllabification in Southern Sámi. By explicating the properties of syllabification and the constraints (25) and (27) particular to Southern Sámi, I intend to show that the inflectional pattern described previously follows straightforwardly; indeed it could not be in any other way than it is. To recapitulate, recall the following data:

(31)	berko:m	'meat.sg.acc'	ga:megm	'shoe.sg.acc'
	båata:h	'come.prs.2sg'	dåeredh	'follow.prs.2sg'
	berkose	'meat.sg.ill'	ga:me:gasse	'shoe.sg.ill'
	berkosne	'meat.sg.iness'	ga:me:gisne	'shoe.sg.iness'
	båateden	'come.prs.2du'	dåere:didden	'follow.prs.2du'

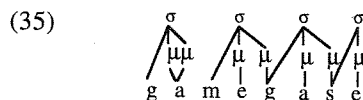
We will commence our investigation by approaching the consonant-final stems, i.e. the right column in (31). The question to be raised here is concerned with the vowel of the second syllable in V- and C-stems: Is it underlyingly long or short? Let us start by considering C-stems. Two possibilities are at hand: The shape is (i) CVX.CVVC, or (ii) CVX.CVC. I will argue that (i) is the correct choice, and that this follows from the syllabification algorithm provided by moraic theory. Let us begin with establishing that (ii) is not tenable, and in doing so we will consider the illative form /ga:me:gasse/. According to (ii), the underlying representations of the stem and the illative suffix are as shown in (32) and (33). Since the consonant in the suffix is a geminate, it must be dominated by a mora in its underlying representation, according to the assumptions spelled out in section 2.



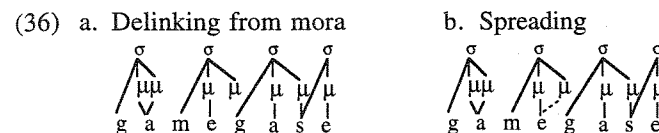
When /asse/ attaches to /ga:meg/, syllabification proceeds as follows. First, as seen in (34a), syllable nodes are assigned to vowels. Next, the onset rule applies, i.e. adjunction of prevocalic consonants, which yields (34b). Since there are no non-moraic postvocalic consonants in (34b), the weight by position rule will not apply. Instead remaining segments are adjoined to the syllable node, which is shown in (34c):



The only way to obtain the right surface form under assumption (ii), would be to add a stipulation saying that the suffix /asse/ attaches to a syllabified stem. In such an instance the final consonant of the stem would be assigned moraic status by the rule of weight by position. Otherwise we will have to invoke a blatant violation of the necessary conditions for the application of WBP, since there are no unparsed postvocalic consonants in (34). Nevertheless, if we were to maintain that the final vowel of the stem indeed is short, we would have to live with this. We would then arrive at an intermediate stage with the following representation:



We could then invoke constraint (25) banning geminates across a foot-boundary, which would force /g/ to delink from its dominating mora, and hence enable the preceding vowel to spread into the vacated mora, similarly to what we find in compensatory lengthening (36).

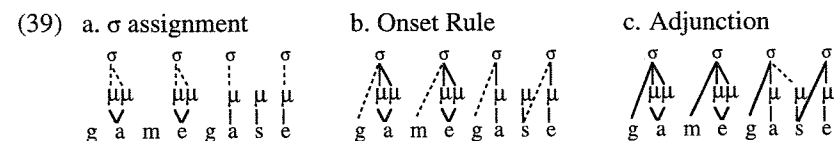


There are at least two points of concern with analyses like the ones just outlined. What makes the derivation implausible in one case is the fact that WBP would have to apply to an already parsed consonant. This is the immediate reason for abandoning the view that the vowel of the final syllable is short. Moreover, as we shall see below, the ban on geminates across a foot boundary in Southern Sámi is generally not repaired by delinking only the violating segment from its mora, but rather the whole mora is deleted. Another concern has to do with the parallel to compensatory lengthening. CL prototypically arises when a coda consonant is deleted, as in Hayes' (1989:260) Latin example *kasnus → ka:nus.

In (36), however, there is no deletion of a coda consonant which could trigger CL; at least not in the conventional way. In conclusion, there is good support for assuming that the vowel of the second syllable in consonant final stems is bimoraic, rather than monomoraic. I therefore claim that the underlying representation is not (32), but rather (37). We will now consider the syllabification process on the basis of (37). When /asse/ (33) attaches to (37), we obtain (38):



The syllabification algorithm applies in this way. First of all syllable assignment applies, as in (39a). Next the Onset Rule applies (39b). After this, there are no unparsed postvocalic segments left of the melodic tier. Hence WBP cannot apply. In (39c) the final rule in the algorithm applies, which adjoins the mora of the lexically moraic consonant to a syllable node.



As seen in (39), the derivation proceeds straightforwardly without appealing to auxiliary assumptions, as is the case if the underlying representation of the stem is as in (37). However, the position taken here requires us to address the question why we have the possible surface forms in (40):

- (40) Nom.sg ga:meg-ə
 Acc.sg ga:meg-ə-m / ga:meg-m
 Gen. sg ga:meg-ə-n / ga:meg-n

The fact that /ə/ in (40) is optional suggests that it is not part of the underlying forms. An even more compelling argument for this is found within the stress system of Southern Sámi, which is based on syllabic trochees, constructed from left to right, with the end rule applying on the leftmost foot. This means that we find stresses on every odd numbered syllable, while even numbered syllables have no stress. Furthermore, degenerate feet are not available:

- (41)
- | | |
|------------------------|-------------------------|
| *
(gaa.mee)(gas.se) | *
(bear.ko) se |
| *
(dåe.rie)(dam.me) | *
(båe.te)me |
| *
(nyj.se)(næj.ja:) | *
(nyj.se)(næj.je)se |

There are two exceptions to this general pattern. The first is the fact that a final odd syllable attracts stress if it is heavy and not part of a suffix which is subject to augmentation:

- (42) *
 (dåe.rie)(diéh)

The second exception arises when certain types of agglutinative morphology is added to a stem. For example, the causative formative is /t/, which attaches to a verbal stem. The progressive formative is /minnie/. Thus, adding /t/ to the consonant stem /dåeried/ gives [dåeredət], and suffixing /minnie/ gives [dåeredətəminnie], 'making x follow'. We thus have a six syllable word. Stress does not fall on every odd syllable, however, but only on the first and the fifth:

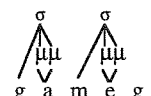
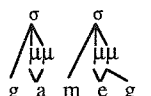
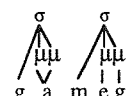
- (43) * (dåe.re)(də.tə)(min.nie) √ (dåe.red)ə.tə(min.nie)

Clearly, stress applies before epenthesis has a chance to construct anaptyctic syllables. The important point here is the fact that construction of anaptyctic syllables does not affect the quality of segments, which rather is determined by the syllabification relevant to stress⁵.

With this brief excursus into the Sámi stress system, let us go back to (40). Since epenthesis applies at a very late stage, we can argue that the relevant shape of the nominative form [ga:megə] 'shoe' is /ga:meg/. This is supported by the fact that [ga:meg] is also a possible surface form. Recall also the underlying form of 'shoe' from (37).

What we have in (40) are instances of closed syllable shortening (CSS) described in section 2. When a CV suffix is added to a string of the shape CVVC, the vowel shortens in order to accommodate the previously extrasyllabic consonant. Extrasyllabicity occurs only at a relevant prosodic edge, and the addition of a suffix makes the stem-final C lose its edgehood.

Forms like the nominative singular [ga:megə] can be analyzed in the following manner, following Blevins' (1995:224) analysis of Afar, where CSS occurs without the addition of a suffix. The underlying representation (37) syllabifies along the lines made clear above, which gives us (44a). In order to escape stray erasure (cf. example (52) below), /g/ adjoins to a preceding mora, which is the second of two which dominate /e/, yielding (44b). Subsequently the vowel delinks, and hence shortens, producing (44c).

- (44) a.  b.  c. 

The possible surface form [ga:megə] arises when a final schwa is inserted, which takes place fairly late. The addition does not trigger vowel lengthening. Moreover, [g] is rendered ambisyllabic.

It is thus clear that CSS also applies in accusative and genitive singular forms of nouns. These forms are derived in the same manner as the nominative singular, with the addition of overt suffixes.

The derivational procedures and stages presented so far have concentrated on nouns. However, the analysis generalizes to the other major lexical categories as well. Here I will present a few examples, which follow directly from the account given, as the reader easily can verify. In (45) we find a few samples from the verbal system, and (46) presents some adjectives.

⁵I assume that schwas head weightless syllables.

- | | | | |
|------|-------------|--------------------|-----------------------------|
| (45) | UR | /dâere:d/ 'follow' | /bâajto:d/ 'become fat' |
| | 1sg.prs | dâeredəm | bâajtodəm |
| | 3sg.prs | dâeredə | bâajtodə |
| | 2du.prs | dâere:didde:n | bâajto:didde:n |
| (46) | UR | /bijwe:l/ 'warm' | /juhte:l/ 'active, healthy' |
| | predicative | bijwelə | juhtelə |
| | comparative | bijwe:lâppo: | juhte:lâppo: |
| | superlative | bijwe:lommeš | juhte:lommeš |

I have shown that consonant final stems have a bimoraic vowel underlyingly in the second syllable. We saw that this assumption yields the right results with syllabic suffixes such as the illative and that it does so without having to resort to further assumptions and language specific stipulations. Moreover, we were able to give a natural account of nominative and accusative forms of nouns in terms of closed syllable shortening.

5.2. Vowel final stems

In this subsection we will consider vowel final stems. The first claim to be made is that vowel final stems end in a lexical long vowel. We will then consider derivations involving suffixation of monosegmental inflectional morphemes such as the accusative and genitive singular endings, and syllabic suffixes like the singular relative and plural comitative. I will show that the surface forms follow naturally under the general syllabification algorithm of moraic theory, in combination with two language particular constraints. The first of these involves a ban on long vowels in unfooted syllables, and the second constraint states that material which belongs to a lexical stem is exempt from the possibility of being parsed as extrasyllabic.

The relevant characteristics are seen in the following paradigm. As is familiar by now, monosegmental suffixes do not trigger closed syllable shortening, whereas syllabic suffixes appear to do so. As seen in (48) the same holds in the verbal system:

- | | | | | |
|------|-----------------|-----------------|-------------|---------------|
| (47) | <i>Singular</i> | | | |
| | nom | berko: | iness | berkosne |
| | acc | berko:m | elat | berkoste |
| | gen | berko:m | com | berkojne |
| | ill | berkose | | |
| (48) | | <i>Singular</i> | <i>Dual</i> | <i>Plural</i> |
| | 1.prs | čearo:m | čearo:n | čearobe |
| | 2.prs | čearo:h | čearoden | čearode |

As mentioned above, the claim is that vowel final stems in Southern Sámi involve a long vowel in the final syllable of the stem. Some evidence in support of such an analysis comes from the allomorphy of the 3rd person dual suffix. It may surface either as /jægan/ or /ægan/ (49). As expected, /ægan/ appears if the stem is consonant final, hence forcing the initial consonant of the suffix to delete. When the stem is vowel final, the initial consonant of the suffix escapes deletion. What is important here is the fact that the vowel preceding the suffix is long. If it were not underlyingly long, we would have to posit a rule that lengthens a vowel when followed by a suffix. But it is hard to see the overall motivation for such a rule in Sámi, as will become clear.

- | | | |
|------|-------------------------------------|-------------------------------|
| (49) | <i>C-stem</i> | <i>V-stem</i> |
| | dâeriedægan 'follow.3sg.du.prs' | bâetiejægan 'come.3sg.du.prs' |
| | bâetietægan 'come-cause.3sg.du.prs' | čearoejægan 'cry.3sg.du.prs' |

The long vowel analysis also enables us to straightforwardly approach allomorphy among case suffixes, as well as variations in vowel length. The initial consonant of the suffix deletes if the final segment of the stem is a consonant, whereas it remains if the stem ends in a vowel.

As a preliminary to the discussion to follow, assume that there is a rule in Sámi which shortens a long vowel if it occurs in a non-footed syllable, (50). That is, the language does not allow degenerate feet, but observes strict foot binarity. This scenario arises for word final odd syllables⁶.

- (50) (_{Foot} σσ) CVV → CV/_#

In the light of this, consider the following plural nouns in the nominative, accusative and inessive:

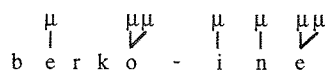
- | | | | |
|------|----------|-----------------|-----------------|
| (51) | | 'shoe' (C-stem) | 'meat' (V-stem) |
| | Nom.sg | ga:megə | berko: |
| | Nom.pl | ga:megh | berko:h |
| | Acc.pl | ga:me:giddie | berkojde |
| | Iness.pl | ga:me:ginnie | berkojne |

To begin with, let us assume that the underlying forms of the accusative and inessive suffixes are *-iddie* and *-innie*, respectively. When *-iddie* attaches to a C-stem, the final C becomes the onset of the first syllable of the affix, which entails lengthening the final vowel of the stem. This implies that the vowel is short in the underlying representation. Moreover, since we

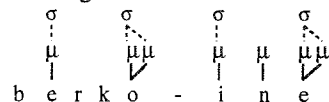
⁶Heavy syllables which are members of non-augmented suffixes are exempt from (50).

have stresses on the first and third syllables in /ga:me:giddie/, we conclude that the final long vowel is not subject to rule (50), since the footing would be (gaa.mee)(gin.nie). However, when *-iddie* suffixes to /berko:/, we see in (51) that the final vowel of the stem shortens and that the /i/ of the suffix becomes the coda of the second syllable⁷. This is a result of constraint (27), which bans onsetless syllables. Hence the vowel of the stem shortens in order to give room to the glide. This would give rise to an intermediate representation like /berkojnje/. Application of (50) being possible, the long vowel is shortened. This is illustrated in (52).

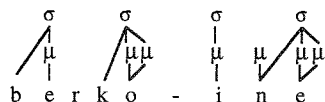
(52) a. Underlying Representation



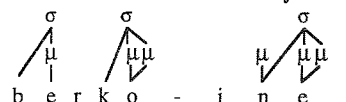
b. σ assignment



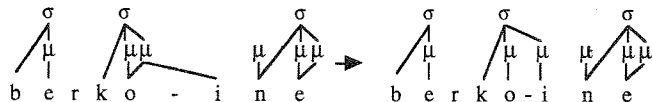
c. Onset Rule



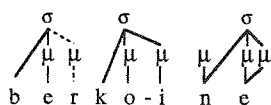
d. Delete Non-Onsetted Syllable



e. Closed Syllable Shortening



f. Weight by Position



h. Final Shortening:



The other route to take would be to claim that the final vowel of a V-stem is short. But this approach would be problematic for the nominative

⁷Here the suffix is augmented and hence it does not fulfill the minimal requirements for attracting stress.

plural form in (51), for example. We would have to account for why the addition of the suffix *-h* triggers lengthening in this case. Thus the assumption that the underlying vowel is short would give rise to a great number of complications that are avoided if it is analyzed as long.

We must consider the question why the monosegmental suffixes *-n* 'gen.sg', *-m* 'acc.sg', etc. do not trigger closed syllable shortening.

(53)	nom.	berko:
	acc	berko:m
	gen	berko:n

Recall from the previous subsection that we get closed syllable shortening in uninflected consonant final stems. These facts call for positing a descriptive constraint like the following:

(54)	*	$\begin{array}{c} \sigma \\ \swarrow \quad \searrow \\ \text{C} \quad \text{V} \quad \text{V} \quad \Omega \\ \text{[Stem]} \quad \text{<C>} \quad \text{]}\# \end{array}$
------	---	--

What (54) says is simply that no segment which is a member of a stem may be extrasyllabic. It is fairly clear that Sámi allows for extrasyllabicity, so closed syllable shortening must be triggered by a constraint like (54); if not we would expect the forms in (55a) to be well-formed like those in (55b). Since the final consonants in (55) are not present in the underlying forms of the stems, (54) does not apply, and hence they are extrasyllabic.

(55)	a. nom.	√berko:	b. nom.	√ga:meg
	acc	*berkom	acc	√ga:megm̩
	gen	*berkon	gen	√ga:megn̩

In this subsection I have shown that vowel final stems end in an underlyingly bimoraic vowel. The instances where this particular vowel surfaces as short were accounted for by invoking the ban on onsetless syllables, which in turn has consequences for the overall syllabification. The claim that the final vowel is long rather than short enables a more straightforward account of the data than if the contrary is claimed.

6. Final remarks

In this paper it has been shown that allomorphy and length variation in stem vowels in Southern Sámi can be accounted for by claiming a distinction between vowel final and consonant final stems. Once this distinction is made, the syllabification algorithm provided by moraic theory does the bulk

of the remaining job. We have also explicated some constraints particular to Southern Sámi which fill in the gaps in the overall theory. For example, we have seen that there is a constraint which prohibits material of a stem to be extrasyllabic, and how this constraint forces closed syllable shortening in bare consonant final stems. Also, a ban on geminates across foot boundaries and a requirement that syllables have onsets account for the patterns observable with respect to vowel final stems.

References

- Bergsland, Knut. 1994. *Sydsamisk grammatikk*. Karasjok, Norway: Davvi Girji.
- Bergsland, Knut & Lajla Mattson-Magga. 1993. *Áarjelsaemien-daaroen baakoegærja*. Alta, Norway: Idut.
- Blevins, Juliette. 1995. 'The syllable in phonological theory'. In J. Goldsmith (ed.), *The handbook of phonological theory*, 206-44. Oxford: Blackwell.
- Goldsmith, John. 1990. *Autosegmental and metrical phonology*. Oxford: Blackwell.
- Hayes, Bruce. 1989. 'Compensatory lengthening in moraic phonology'. *Linguistic Inquiry* 20, 253-306.
- Hayes, Bruce. 1995. *Metrical stress theory: principles and case studies*. Chicago: Chicago University Press.
- Jaahkenelkien, A. 1993. *Goltelidh jih soptsestidh*. Aarbortesne, Norway: Daasta berteme.
- Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Oxford: Blackwell.
- McCarthy, John & Alan Prince. 1995. 'Prosodic morphology'. In J. Goldsmith (ed.), *The handbook of phonological theory*, 318-66. Oxford: Blackwell.
- Nilsson-Mankok, Erik. 1974. *Svensk-sydsamisk ordlista: Vilhelmina vefsen*. Vilhelmina, Sweden: Independent publisher.
- Rice, Keren. 1992. 'On deriving sonority: A structural account of sonority relationships'. *Phonology* 9, 61-99.
- Vinka, Mikael. 1997. 'Southern Sámi ablaut and the role of the harmony foot', ms. McGill University, Montreal.

Implementing WordNet for Swedish adjectives

Caroline Willners

Introduction

A Swedish version of WordNet was created and around 300 Swedish adjectives, mainly from the semantic field of *strength*, were implemented. This paper is a documentation of the implementation. The purpose of the study was to investigate the possibilities of applying WordNet to Swedish and to illuminate general problems with WordNet as well as specific problems in the handling of adjectives. First, a short overview of WordNet is given, and then the WordNet categorisation of adjectives is reviewed. The section about the implementation gives hands-on knowledge of how to add a new adjectival lexical entry in WordNet. Then a description of the problems encountered and some general remarks follow.

WordNet

WordNet is a lexical reference system designed to reflect the organisation of human memory as well as to be a useful on-line dictionary. It contains approximately 95,600 English nouns, verbs, adjectives, and adverbs which are linked semantically. The system also contains syntactic and morphological information. Apart from the original English version developed at Princeton University, WordNets are being implemented for German, Spanish, and Dutch in the EuroWordNet project (Climent et al. 1996).

The most significant feature of WordNet is the semantic organisation. It supports synonymic, antonymic, hyponymic-hypernymic, and meronymic-holonymic relations. Because of this semantic approach the parts-of-speech categories can naturally be separated.

Data in WordNet are entered in 'lexicographers' source files', with somewhat different formats for the different parts of speech. The source files are then compiled by the program *Grinder* which generates the database that can be accessed by the window-based retrieval software. The