

Dialectal gemination in Finnish:  
Phonetics/phonology interplay in moraic theory

Christopher Spahr

Forum Paper  
Submitted in partial fulfillment of the degree of Master of Arts  
University of Toronto  
© 2011 Christopher Spahr



# Contents

<b>Contents</b>	<b>i</b>
<b>Acknowledgments</b>	<b>iii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Finnish Phonology . . . . .	3
1.2 Gemination and Degemination . . . . .	7
<b>2 Gemination: A Dialectal Phenomenon</b>	<b>11</b>
2.1 “Standard Finnish” and Finnish Dialects . . . . .	12
2.2 An Informal Typology of Dialectal Gemination . . . . .	15
2.2.1 Phonetic Gemination . . . . .	16
2.2.2 Phonemic Gemination . . . . .	17
2.2.3 Morphological Gemination . . . . .	21
2.2.4 Summary . . . . .	22
<b>3 Moraic Structure</b>	<b>23</b>
3.1 Mora Count and Moraic Theory . . . . .	24

3.2	Second Mora Lengthening . . . . .	30
3.3	Defining the Mora: What does it do? . . . . .	33
3.3.1	Length and Duration . . . . .	33
3.3.2	Weight and Timing . . . . .	34
<b>4</b>	<b>Second Mora Lengthening and Primary Gemination</b>	<b>37</b>
4.1	Evidence for a Correlation . . . . .	38
4.2	Modeling Primary Gemination . . . . .	40
4.2.1	A Previous Account . . . . .	42
4.2.2	The Synchronic Connection . . . . .	45
4.2.3	Gemination Before Diphthongs . . . . .	49
4.2.4	Gradient Effects: A Model for Change . . . . .	57
4.2.5	Summary . . . . .	59
<b>5</b>	<b>Additional Issues and Closing Remarks</b>	<b>61</b>
5.1	Further Theoretical Issues . . . . .	61
5.1.1	Underlying Representations . . . . .	62
5.1.2	Segmental Phonology and Syllabification . . . . .	63
5.1.3	Phonetic Interpretation . . . . .	65
5.2	Conclusions . . . . .	66
5.3	Directions for Future Research . . . . .	67
	<b>References</b>	<b>69</b>

# Acknowledgments

Writing a forum paper is not an easy task, or at least so the cliché goes. Actually I found it quite manageable, though by no means trivial; the real challenge was making it to the forum meetings at 9 o'clock every Tuesday morning. Nonetheless, there are some people I should mention who made the process easier. Keep in mind that none of the errors or bad calls in this paper are the fault of anyone mentioned below. They are entirely my own.

First and foremost, because misery loves company, I'd like to thank my fellow members of the 2010–2011 MA forum group. I give the names of all thirteen of them in alphabetical order, as not to show favourites:<sup>1</sup> Andrei Anghelescu, Elise Benallick, Marisa Brook, James Byrnes, Erin Chen, Radu Craioveanu, Ross Godfrey, Isobel Marr, Christian Mutikainen, Madeline “Maddie” Shellgren, Michelle Stella, Brigid Thurgood, and Holly Young. It is a great shame that I may not have succeeded in my goal of earning myself a footnote for my invaluable assistance in each and every one of your own forum papers, but I hope that I provided at least some moral support in the process, as you all certainly did for me. I wish you all the greatest success in your future endeavours, academic or otherwise.

I'd also like to thank the graduate coördinator Sali Tagliamonte for giving us a hard time when it was necessary, and for acting impressed when I pointed out that I managed to finish my full two semesters of MA coursework without missing a single class, including those 9a.m. forum “get-togethers”.

This paper would not have been possible without academic guidance, and so I owe my sincerest thanks to my advisor, Elan Dresher. His knowledge, thoroughness, humour, and tolerance for me when my opinions on

---

<sup>1</sup>This is not to say that I don't have favourites, as I do of course; you know who you are. Extra credit if you've ever pushed me to/helped me use L<sup>A</sup>T<sub>E</sub>X or have spent time with me in an Ethiopian restaurant.

some specific representational issues differed has helped me both to proceed in my never-ending quest to become a real phonologist, and to “be myself”, as it were (after all, what good is linguistics without a little style?) Furthermore my thanks go out to my second reader, Keren Rice, for having a fresh look at the whole paper and suggesting a number of revisions, fortunately none of them too major!

I would also like to thank Elizabeth Cowper. Although I only know her academically as a course instructor in Intermediate Syntax (which I took to calling “remedial syntax”, due to my own shortcomings in the subject area), it was in this class that I came to really internalize the importance of making clear one’s assumptions in linguistic writing. Unfortunately, this has made me a bit hyper-aware, and has sort of spilled over in to my non-linguistic life (as if I have one of those!). My apologies if the expected quality of argumentation is not readily apparent in the present work.

For giving me some native speaker intuitions, and reaffirming my suspicions about where primary gemination can apply (it really can to *any* CV́CVV word, it seems!), I thank Jari Similä.

I also thank my parents, Linda Spahr and Stephen Spahr, for supporting my decision to pursue academia instead of getting a real job.<sup>2</sup> An MA is just the first step down an exceedingly long path, but I trust that they’ll continue to cheer me on along the way.

Finally, I must make mention of the bidding which took place during the summer meeting of the MA cohort, after noticing how often a paper on this topic makes use of the word “gemination”. The number of instances of this string<sup>3</sup> was compared to the list of bids from the meeting several months earlier, according to “Price Is Right” rules (closest without going over). The winner was Madeline Shellgren. Congratulations, Maddie!

And with that, I give you my year’s work.

Christopher Spahr  
September 17, 2011

---

<sup>2</sup>Don’t think that this would have been impossible without their support, though, as I would have gone ahead and done it anyway.

<sup>3</sup>The version that went out to my second reader contained 139, for the record, and that’s not counting headers or the table of contents. Subsequent revisions seem to have added several more!

# Chapter 1

## Introduction

The Finnish language possesses a quantity distinction between short and long vowels and singleton and geminate consonants, regardless of primary stress placement (which always occurs on the initial syllable), as well as agglutinative morphology. Because of this combination, Finnish has proved an interesting testing ground for both phonological and morphological studies in the theoretical linguistic literature.

The vast majority of work in generative phonology has been on “Standard Finnish”, often without defining the term. Because spoken Finnish is really a collection of dialects varying mainly in their phonology, it is a wonder that so much theoretical work has gone into modeling the system of an artificially constructed standard, rather than looking at dialects as the individual varieties that they are.

The focus here is on a phonological process present in a number of

dialects of Finnish known as primary gemination. In primary gemination, a singleton consonant is geminated between a stressed short vowel and an unstressed long vowel:

$$(1.1) \quad C\acute{V}C_{\alpha}VV \rightarrow C\acute{V}C_{\alpha}C_{\alpha}VV$$

Although primary gemination is well known in the Finnish literature, most publications have been either descriptive (see Paunonen 1973 for an example of an article in English) or quantitative (such as Nahkola 1987); there has been almost no research done within generative phonology looking at primary gemination. The only example of generative work along these lines of which I am currently aware is Harrikari (2003), which focuses on two related but distinct types of dialectal gemination.<sup>1</sup>

The goals of the present study are twofold. The first is to expand the generative phonological literature on Finnish dialectal phonology by examining a dialectal phenomenon (primary gemination), placing it within the context of both the typology of related dialectal gemination processes and the diachronic rise of those processes. The second is to provide an original contribution to moraic theory with regards to the interplay between phonetics and phonology. The analysis presented in chapter 4 is based on evidence of this interplay: the additional weight attributed by a phonetic process (second mora lengthening) causes phonological gemination under the right prosodic conditions (in words of the shape CVCVV).

---

<sup>1</sup>These are, namely, “special” gemination and South-Western gemination, both of which are discussed in chapter 2.



The remainder of this chapter gives an overview of Finnish phonology. Chapter 2 covers the sociolinguistic status of Finnish dialects and provides a typology of dialectal gemination processes. Chapter 3 discusses moraic structure in Finnish and the system of representations used in this study with regards to syllable structure. Chapter 4 lays out the analysis at the core of this study, identifying second mora lengthening as the root cause of primary gemination. Chapter five touches on some additional theoretical issues and concludes the paper.

## 1.1 Finnish Phonology

The spoken Finnish language exists as something of a dialect continuum, and the greatest source of variation between Finnish dialects is in phonology. This section gives a basic overview of Finnish phonology in order to give the reader a sense of what all dialects of Finnish have in common.

Finnish has eight vowel phonemes, which are given in Figure 1.1. The vowels are given using Finnish orthography; approximate IPA values are given in brackets where they differ from the orthography. Suomi *et al.* (2008:23) note that all dialects of Finnish have these same eight vowel phonemes, although the exact phonetic realization of these phonemes can vary slightly between dialects.

Finnish has a system of front-back vowel harmony. In this system, vowels in suffixes added to words must agree in backness with vowels of

Figure 1.1: Vowel Phonemes of Finnish

	Front	Back
High	i y	u
Mid	e ö [ø]	o
Low	ä [æ]	a [ɑ]

the stem. Front and back vowels cannot co-occur within a non-compound word. The vowels /i/ and /e/ are neutral, and can co-occur with either front or back vowels, as in (1.2c, d) (examples from Karlsson 1999:17):

- (1.2) a. *talo-ssa* ‘in the house’    b. *kylä-ssä* ‘in the village’  
       c. *Pori-ssa* ‘in Pori’            d. *käde-ssä* ‘in the hand’

When stems contain only neutral vowels, attached suffixes harmonize to front vowels:

- (1.3) a. *venee-ssä* ‘in the boat’    b. *miehe-llä* ‘on the man’

All eight vowels can appear as either long or short. Long vowels are indicated in the orthography and in the present work by doubling the vowel. Vowels are contrastive for length in all positions – the placement of stress is irrelevant to the realization of the phonological length of a vowel. Furthermore, neither the placement of stress nor the length of the vowel have any significant effect on vowel quality. Some examples of the independence of vowel quantity and stress are given in (1.4).

Figure 1.2: Diphthongs of Standard Finnish

i-final	u-final	y-final	mid-final
äi	au	äy	ie
ui	ou	öy	yö
ai	eu		uo
oi	iu		
öi			
yi			

- (1.4) a. *túle* ‘come!’      b. *túule* ‘blow! (of the wind)’  
 c. *túlee* ‘it comes’    d. *túulee* ‘it blows’

Finnish also has a number of diphthongs. Karlsson (1999:14) lists 16 common diphthongs, shown in Figure 1.1, divided according to their second vowel melody. While this represents the diphthongs of Standard Finnish, the exact diphthong inventories of different dialects show some variation.

Describing the phonemic status of consonants in Finnish is a somewhat more complicated matter. The reason for this is that the “native” consonant inventory is augmented to varying degrees by loan words depending on several factors, including the amount of exposure to other languages the dialect has had, and the level of knowledge an individual speaker has of other languages. Figure 1.3, reproduced from Suomi *et al.* (2008:38) lists the main allophones of the primary phonemes of Finnish.<sup>2</sup>

---

<sup>2</sup> There are several differences between the notation used in Figure 1.3 and the

Figure 1.3: Consonant Phonemes of Finnish

	Bilabial	Labiodental	Laminal dentalveolar	Apical alveolar	Laminal alveolar	Palatoalveolar	Palatal	Velar	Glottal
Plosive	p (b)		t̪					k (g)	
Semiplosive				d					
Fricative		f			s	(ʃ)			
Glottal continuant									h
Nasal	m			n				ŋ	
Trill				r					
Lateral approximant				l					
Central approximant		v					j		

Phonemes given in parentheses have marginal status, depending on speaker, but are not considered to be native phonemes, and there is a great deal of variation with regards to how certain phonemes are produced. For example, /s/ can often be pronounced closer to [ʃ]. However, Suomi *et al.* speculate that /ʃ/ has a less marginal status (it is considered less foreign) than /b/ or /g/, despite the presence of (near) minimal pairs such as *pussi* ‘sack’ and *bussi* ‘bus’. Similarly, although /f/ is not originally a native

---

orthography: /ŋ/ is written as either <n> before <k>, or as <ng>, representing a geminate velar nasal [ŋŋ]. The phoneme /t̪/ is written simply as <t>, and /v/ is written <v>. The (marginal) phoneme /ʃ/ can be written variably as <sh>, <š>, or simply <s>. Finnish orthography is used in transcriptions in the present work.

phoneme, Suomi *et al.* consider it to be much more thoroughly integrated into the inventory.

The status of /d/ requires special discussion. Suomi *et al.* classify it as a “semiplosive”, but it unclear exactly what is meant by this. What is clear, however, is that /d/ does not pattern with the other voiced plosives, which have entered the language only through loans. /d/ only appears as the weak grade of /t/, resulting from consonant gradation. Its exact phonetic realization differs from dialect to dialect, ranging from [d], in Standard Finnish, to a flap homophonous with singleton /r/, to the creation of vowel hiatus.

There is also a significant amount of allophonic variation in consonants resulting from co-articulation. However, this will not be covered further because it is of little relevance to the present study.

## 1.2 Gemination and Degemination

Much like vowels, consonants in Finnish show a two-way quantity opposition between long and short, that is, consonants may appear as geminates. Geminate consonants can only appear word-internally, and there are several restrictions on what consonants can be geminated. /v/, /j/, /h/, and /d/ cannot appear as geminates in underlying forms,<sup>3</sup> and only /p/, /t/, /k/, and /s/ can appear as geminates within clusters following sonorants:

---

<sup>3</sup>There are a few exceptions to this. /d/ can appear as a geminate in a few recent loans, and /h/ appears as a geminate in the word *hihhuli* ‘religious fanatic’.

- (1.5) a. *korpi* ‘wilderness’    b. *korppi* ‘crow’  
       c. *kansa* ‘people’        d. *kanssa* ‘with’

There are several phenomena in Finnish which can alter the phonological quantity of consonants. The first of these is consonant gradation. Consonant gradation is a morpho-phonological process which causes the weakening of consonants and consonant clusters containing /p/, /t/, and /k/ before closed syllables in certain morphological environments.

When an applicable suffix creates a closed syllable before an underlying geminate stop and consonant gradation applies, as in (1.6), the result is degemination; the geminate becomes the corresponding singleton consonant (examples from Karlsson 1999:28-29):

- (1.6) a. *ka**ap**pi* ‘cupboard’    b. *ka**a**pi-ssa* ‘in the cupboard’  
       c. *ma**tt**o* ‘mat’            d. *ma**t**o-lla* ‘on the mat’  
       e. *ku**k**ka* ‘flower’        f. *ku**k**a-n* ‘of the flower’

However, degemination is not the only way that consonant gradation manifests itself. There also exists qualitative gradation in which a singleton intervocalic consonant can change to a voiced consonant:

- (1.7) a. *tu**p**a* ‘hut’        b. *tu**v**a-ssa* ‘in the hut’  
       c. *ka**t**u* ‘street’    d. *ka**d**u-lla* ‘on the street’

/k/ can undergo one of several changes, depending on phonological en-

vironment. It can either delete entirely, leaving hiatus, as in (1.8b), become /v/ in the environment of /u/, as in (1.8d), or become /j/ in the environment of a following /e/, as in (1.8f):

- (1.8) a. *mä**k**e*- ‘hill’      b. *mäe-llä* ‘on the hill’  
       c. *puk**k**u* ‘dress’      d. *pu**v**un* ‘of the dress’  
       e. *pol**k**-e* ‘trample’    f. *pol**j**e-n* ‘I trample’

Finally, /p/, /t/, and /k/ can fully assimilate to a preceding homorganic sonorant in the same cluster, resulting in a geminate sonorant:

- (1.9) a. *amp**p**u*- ‘shoot’    b. *amm**m**u-mme* ‘we shoot’  
       c. *ranta* ‘shore’      d. *rann**n**a-lla* ‘on the shore’  
       e. *ke[ŋ]**k**ä* ‘shoe’    f. *ke[ŋ]**ŋ**ä-n* ‘of the shoe’  
       g. *kult**a*** ‘gold’      h. *kull**a**-n* ‘of the gold’  
       i. *part**a*** ‘beard’     j. *par**r**a-ssa* ‘in the beard’

There are also several processes which can *cause* gemination. The first of these is a sandhi phenomenon which Suomi *et al.* (2008:44) term “boundary lengthening”. Boundary lengthening describes the creation of a geminate consonant across a word boundary following certain morphemes. Words which trigger boundary lengthening include, for example, the second person imperative forms of verbs. Vowel-final boundary-lengthening words cause a geminate of the initial consonant of the following word:

- (1.10) a. *mene* [mene] ‘go!’ b. *mene pois* [mene**pp**ois] ‘go away!’

Interestingly, this process allows the creation of geminates of all consonants which can appear word-initially. Thus, geminate /h/, /v/, and /j/ can be created (/d/ does not appear word-initially in native words, example from Karlsson 1999:12):

- (1.11) a. *ole* [ole] ‘be!’ b. *ole hiljaa* [ole**hh**iljaa] ‘be quiet!’

Even more interestingly, boundary lengthening can create geminate glottal stops across a word boundary when the second word begins with a vowel, as in (1.12). This occurs despite the fact that the glottal stop is apparently non-phonemic in Finnish, appearing only word-initially and intervocalically to break up hiatus.

- (1.12) a. *mene* [mene] ‘go!’ b. *mene ulos* [meneʔ:ulos] ‘go up!’

Finally, there are dialectal processes which cause gemination before long vowels. The details of how these work vary from dialect to dialect, but I refer to them together with the umbrella term “dialectal gemination”. These processes are the focus of the present study, and their nature is covered in detail in the next chapter.



## Chapter 2

# Gemination: A Dialectal Phenomenon

As the purpose of this work is to provide an account of a dialectal phonological process, it is necessary to describe this and related phenomena in the context of the Finnish speech community as a whole. This chapter serves to give background on the situation of “Standard Finnish” and Finnish dialects as they relate to each other.

I begin by describing the sociolinguistic situation with regards to different varieties of Finnish, followed by giving a basic typology of the continuum of different processes which can fall under the label “dialectal gemination”.

## 2.1 “Standard Finnish” and Finnish Dialects

In some sense, the concept of “Standard Finnish” is a misnomer. Even in an idealized situation, the Finnish language does not exhibit a great deal of homogeneity from a phonological perspective. Instead, Standard Spoken Finnish (SSF) can be described as a set of prescriptive morpho-syntactic rules which are used in more formal registers, but with a given speaker’s phonology.

Suomi *et al.* (2008: 7) explain that for the majority of Finnish speakers, two varieties are used, both a local variety and SSF.<sup>1</sup> However, SSF is not as prominent in society as one may think, at least as far as the frequency with which a speaker may produce it, rather than hearing it:

“Usually, the local variety is used in informal speaking situations, and SSF in formal ones; however, some speakers, especially elderly ones, do not necessarily speak SSF on any occasion – and even many younger people never have the chance or duty to speak in formal situations.”

---

<sup>1</sup> It is my personal opinion that the differences between Standard Finnish and many spoken varieties, even those considered less dialectal, are great enough that they border on a situation of *diglossia*. For example, inflectional paradigms are often radically different in spoken varieties (such as loss of number marking in the third person plural on verbs, the use of passive verb forms for the first person plural, and the use of different noun cases in certain constructions), and numerous common lexical items appear in shorter forms which do not seem to be derivable merely through synchronic phonological reduction. Any in-depth discussion on how many grammars a given speaker commands, and how unrelated they are, however, is clearly beyond the scope of the present study.

Even when SSF is used, there are almost invariably “local colourings” involved, especially prosodic ones (most importantly, as will be seen for our purposes, durational ones). Unlike in many other standard languages of the world, local features are not a source of social stigma, and in fact increasingly informal varieties of SSF are becoming the norm even in more formal situations (Suomi *et al.* 2008: 7). Certainly it seems that local varieties of Finnish are not on the way out, but perhaps even on the way in, and this in spite of the fact that Finnish is spoken in a modern country with national media, not to mention massive influence of English.<sup>2</sup>

In fact it can be said that there is something rather unique about the creation of Standard Spoken Finnish in comparison with other standard languages.

As is often the case with standard languages, there is no particular variety upon which SSF is based. Because Swedish was the language of power at the time that SSF was created, the result was a standard language built upon neither anyone’s native language nor the prestige language spoken at the time. For this reason, SSF is truly an artificial language; it was based on Standard Written Finnish, which was in the 19th century itself “consciously created ... as a compromise between the various dialects.” (Suomi *et al.* 2008: 7)

---

<sup>2</sup> This is not to say that the language isn’t changing, of course, as such a situation would be impossible. I merely mean to point out that there is a great deal of variation in spoken Finnish, even in what are considered formal and “standard” varieties thereof; there is little value in discussing data from “Standard Finnish” without elaborating on what is meant by the term, as there is no homogenous source from which to draw such data.

Today, the majority of speakers begin by learning to speak a dialect at home as their native variety, and only later learn the rules of SSF. Nonetheless, the local variation, particularly with regards to phonetics, is still present, even after learning the standard.

This local variation has been exploited by phonetic studies of segmental duration such as Wiik and Lehiste (1968), Lehtonen (1970), and more recently Ylitalo (2004, 2009). Ylitalo (2009) in particular has shown there to be systematic differences in the prosodic features of different local varieties of Standard Spoken Finnish, and their prevalence is further reinforced by the fact that the differences do show up systematically despite the fact that the data on which they are based were collected in a formal laboratory setting.

However, there does exist some division between SSF and local dialects in regards to phonetic variation versus phonemic variation. For example, the gemination processes which are the focus of the current study are in their pure form phonemic; they are the result of phonological processes rather than phonetic variation. Thus, although a local dialect might have dialectal gemination, the corresponding local variety of SSF would not.

Ylitalo (2004), a study of vowel duration in the so-called “half-long vowel” of the second syllable in the Oulu region dialect, notes that her elicited words include those of the form CV.CVV (in speakers’ recorded SSF speech), despite the fact that words of this shape do not exist in the local dialect, due to dialectal gemination which would turn them into CVC.CVV.

She explains that this is because salient features of a local dialect (such as dialectal gemination) can be consciously removed from one's speech, while non-contrastive features, such as the half-long vowel, remain. Thus, although her study looks at Standard Spoken Finnish speech, which lacks dialectal gemination, the half-long vowel stills appears consistently.

These dialectal gemination processes are a key example of the sort of phonemic process which would not appear in an individual's SSF speech. However, because they are present in informal speech of local varieties, and local features are not, in general, stigmatized in Finnish, they are undoubtedly part of the grammar of a speaker of such a local variety.

## 2.2 An Informal Typology of Dialectal Gemination

The use of the term “dialectal gemination” when referring to Finnish is a potentially confusing one. All dialects of Finnish use underlying geminates contrastively, and so questioning the existence of phonological geminates in the language is not the matter at hand. However, the presence and exact properties of processes causing gemination to occur as a result of phonological structure varies a great deal between different dialects of Finnish.

Thus, the purpose of this section is to provide a typology of these gemination processes, although this typology is not necessarily intended to be exhaustive. Instead, I mean to convey the amount of diversity present

among the major types of dialectal gemination processes discussed in the literature, as well as to provide sufficient background on the kinds of general trends in the changes of prosodic structure of Finnish as a whole.

The processes are described in the order that they are said to arise diachronically, but it is necessary to note that not all varieties have yet gone through all stages. The description in this section follows mainly that of Nahkola and Palander (1981) and Paunonen (1973). In accord with Nahkola and Palander, I thus subdivide the discussion into three major types of gemination: phonetic, phonemic, and morphological.

### 2.2.1 Phonetic Gemination

Dialectal gemination begins as a slight phonetic lengthening of a singleton consonant after a short stressed vowel and before a long vowel, when that long vowel is the result of the elision of an intervocalic consonant. Such phonetic lengthening is merely allophonic; the length is not contrastive. This change is illustrated in the following example, in which the partitive case suffix, originally /*ða*/, loses its onset, resulting in a long /*a*/. The consonant preceding this long vowel then shows phonetic lengthening.

(2.1) Change of historical \**vika-ða* ‘fault (partitive case)’

*vi.ka.-ða* > *vi.kaa* > *vi.k:aa*

This phonetic lengthening, however, is apparently not very stable. It is allophonic in that it is not contrastive, but likely due to its perceptual

proximity to the full phonological geminates in the language, it serves as an intermediate phase on the way to phonologization.

## 2.2.2 Phonemic Geminata

### Primary Geminata

Paunonen (1973) speculates that within just a few generations, phonetic geminates become reanalyzed as full phonemic geminates, and thus become equivalent to the contrastive phonological geminates of the language. At this stage, a variety has acquired a phonological process in which “the initial consonant of a syllable following a short stressed syllable is geminated if the second syllable contains a long vowel or diphthong comparable to a long vowel.”<sup>3</sup> (p.146) This is the so-called “primary gemination” (*yleisgeminatio*), which can be descriptively formalized as below:

(2.2) *Primary Geminata*

$$CVC_{\alpha}VV \rightarrow CVC_{\alpha}C_{\alpha}VV$$

A variety undergoing a change to primary gemination therefore ends up with stem alternations in certain forms, such that certain lexical items may have a singleton consonant in the nominative, but a geminate in, for example, the partitive:

---

<sup>3</sup>The status of diphthongs will be addressed below.

- (2.3) a. /vika/ ‘fault (nominative)’ → *vika*  
 b. /vika-a/ ‘fault (partitive)’ → *vikkaa*

Because geminates are already phonologically contrastive, gaining gemination before long vowels can create homophony in certain case forms for certain words, compared to a “standard” variety without the gemination process. As a result, Nahkola and Palander (1981: 12) note that some varieties may end up with “overlong geminates” on the surface forms of words with underlying geminates, as a means of preserving contrasts. For example, the following table shows the illative case forms for *laki* ‘law’ and *lakki* ‘cap’ as they would appear in three types of varieties: non-geminating, homophony-producing, and overlong-geminating.

(2.4) The Potential for Homophony

Underlying Form	Non-geminating	Homophony-producing	Overlong
/laki-in/ ‘law (illative)’	la.kiin	lak.kiin	lak.kiin
/lakki-in/ ‘cap (illative)’	lak.kiin	lak.kiin	lak.k:iin

Dialects with primary gemination also show some variety in terms of how they treat diphthongs as possible triggers for gemination. There are, firstly, varieties in which diphthongs do not cause gemination at all.

In varieties in which diphthongs can cause gemination, there is a further



split: either all diphthongs cause gemination, or only a subset of diphthongs cause gemination. Diphthongs which cause gemination in these “selective dialects” are the ones which Paunonen describes in the above quote as “comparable to a long vowel.” How is the difference between the two types of diphthongs determined?

In “selective” dialects, only surface diphthongs which are the result of some contraction from the underlying forms trigger gemination. That is, the diphthongs are created when some material is deleted due to other phonological processes. The result is that “selective” dialects create contrasts which do not exist in varieties which do not geminate before diphthongs and those which geminate before all diphthongs. Consider the surface forms of /sano-i-n/ ‘I said’ and /sano-ta-i-n/ ‘to say (instructive plural)’:

(2.5) Variability of Gemination Before Diphthongs

Gemination Before Diphthongs	/sano-i-n/	/sano-ta-i-n/
None	<i>sanoïn</i>	<i>sanoïn</i>
All	<i>sannoïn</i>	<i>sannoïn</i>
“Selective”	<i>sanoïn</i>	<i>sannoïn</i>

Indeed, Paunonen mentions reports that there exist varieties spoken in Savo in which gemination-causing diphthongs are perceptibly longer than non-gemination-causing diphthongs, and affirms based on his own field work that the difference is in fact quite salient. In these dialects, the diphthongs causing gemination are in fact directly comparable to long vowels, as they

are equivalent in length to long vowels, whereas non-geminating diphthongs would be equivalent in length to short vowels.

In those dialects with a length contrast on diphthongs, then, long diphthongs pattern with long vowels in causing gemination, while short diphthongs pattern with short vowels in that they do not. In varieties without such a length contrast, however, it is clear that the morphophonology must have some effect on the creation of geminates, rather than gemination being caused simply with reference to length. Some possible analyses of this are discussed in section 4.2.3.

### “Special” Gemination

There is another kind of phonemic gemination, the so-called “special gemination” (*erikoisgemmaatio*). In special gemination, intervocalic singleton consonants are geminated before all long vowels, without needing to be positioned after a stressed short vowel. Thus special gemination can be formulated in the same way as primary gemination, but with a simpler environment:

#### (2.6) *Special Gemination*

$$C_{\alpha}VV \rightarrow C_{\alpha}C_{\alpha}VV$$

Nahkola and Palander (1981: 12) make a point of noting that “special gemination appears in those dialects in which primary gemination has al-

ready been established.”<sup>4</sup> Thus, special gemination seems to be the result of the full analogical extension of primary gemination to occur before all long vowels, regardless of the context of position within a prosodic word.

### 2.2.3 Morphological Gemination

The final stage of dialectal gemination is “morphological gemination”, characteristic of the South-Western dialects of Finnish. In these dialects, there are no longer long vowels or diphthongs in non-initial syllables. However, the gemination once triggered by the presence of a long vowel or diphthong remains as the sole indicator of the morphological alternations with which it originally occurred:

(2.7) a. *vika* ‘fault (nominative)’    b. *vikka* ‘fault (partitive)’

Harrikari (2003) notes, however, that this South-Western morphological gemination has certain other restrictions: gemination only occurs with the phonemes /p/, /t/, /k/, and /s/. Morphological gemination can thus be described as the fossilized morpho-phonological remains of a once fully phonologically and prosodically active system. Such degradation is in line with the direction of change seen in dialectal gemination in general: the process proceeds from allophonic alternation to a full phonological process, and then remains as a morphological alternation when the phonological environment is lost.

---

<sup>4</sup>“erikoisgeminaatio esiintyy niissä murteissa, joissa yleisgeminaatio on jo vakiintunut.” Translation my own.

Although it is beyond the scope of the current work to examine in any detail, it should be noted that the South-Western dialects of Finnish are not alone in showing morphological gemination. The Estonian language shows some similar characteristics in terms of how quantity is used in bisyllabic words to convey morphological information, as well as a similarity in that it does not possess long vowels or diphthongs in non-initial syllables.

#### **2.2.4 Summary**

The focus of this paper is on phonological gemination, especially primary gemination. How can primary gemination be represented, and how does it develop from phonetic gemination? Because phonological gemination does arise from phonetic gemination, there must be a point at which the process goes from phonetic to phonological. The next chapter lays out the moraic structure of Finnish, in order to provide a context in which to discuss gemination processes.

# Chapter 3

## Moraic Structure

Finnish is a “full-fledged quantity language” (Suomi *et al.* 2008: 39), that is, both vowels and consonants are contrastive for length, regardless of stress placement. Because of this, the system that we use to represent quantity phonologically is extremely important, and must reflect other aspects of the phonological system.

Furthermore, ideally the system of representations should to some extent reflect the phonetic properties of the surface forms derived from those representations, if there is no reason otherwise not to. The purpose of this chapter is to make explicit the way in which my model represents the quantity oppositions in the phonology of Finnish, while attempting to consider the phonetics as well.

### 3.1 Mora Count and Moraic Theory

It is the goal of a system of the representation of quantity to reflect phonological weight and length, and ideally, at least, to approximate phonetic length. At the same time, such a system should attempt to adhere to linguistic tendencies<sup>1</sup> revealed by previous phonological research, deviating only when there is evidence justifying such a deviation.

I adopt a system of representations within moraic theory more or less in line with that of Hayes (1989). However, in order to apply the model, we must first make a decision as to the role that mora count and syllabification, not to mention syllable weight,<sup>2</sup> play in the phonological system of the language. Consider the table in Figure 3.1, reproduced from Suomi *et al.* (2008), representing the most common syllable types in Finnish. The column labeled “Proportion” indicates the percentage of syllables in the language of that type, based on a corpus study.

Note that several syllable types are said to be trimoraic under this model.<sup>3</sup> Unfortunately, at no point is it made clear what exactly is meant by “mora” in the discussion of the data, and in fact no phonological criteria are mentioned other than a requirement of open class words to be mini-

---

<sup>1</sup>I hesitate to use the potentially controversial term “universals”, as there exist counter-examples for many supposed linguistic universals, including with regards to quantity.

<sup>2</sup>The role of syllable weight in Finnish is significant, at least, for determining the placement of secondary stress.

<sup>3</sup>Suomi *et al.* (2008) also make mention of less common syllable types, including quadrimoraic syllables. However, the only three examples provided are **Kuortti**, **Suortti**, and **Jotaarkka**. Since these are all proper names, and since two of the three are nearly identical, such forms will be disregarded for this study.

Figure 3.1: Basic Finnish Syllable Types (Suomi *et al.* 2008:65, glosses added are my own)

Syll. Type	Proportion	Example	Weight	N of Mora	Rhyme
CV	40.4	<b>ta</b> .lo ‘house’	light	1	V
CVC	27.5	<b>tas</b> .ku ‘pocket’	heavy	2	VC
CVV	12.7	<b>saa</b> .ri ‘island’	heavy	2	VV
CVVC	9.6	<b>viet</b> .to ‘celebration’	heavy	3	VVC
VC	3.9	<b>es</b> .te ‘barrier’	heavy	2	VC
V	3.9	<b>o</b> .sa ‘part’	light	1	V
VV	1.2	<b>au</b> .to ‘car’	heavy	2	VV
CVCC	0.6	<b>kilt</b> .ti ‘nice’	heavy	(3)	VCC
VVC	0.3	<b>aal</b> .to ‘wave’	heavy	3	VVC
VCC	0.1	<b>ark</b> .ku ‘coffin’	heavy	(3)	VCC

mally bimoraic. Thus, it seems that the mora count is equivalent to the number of segments in the rhyme. The problem with this definition is that it posits trimoraic (or longer) syllables, a situation not widely discussed in the theoretical literature.

An additional question raised by a moras-as-segments analysis concerns the segmental representation of geminates. Harrikari (1999) points out that the literature on Finnish has traditionally analyzed geminates as a sequence of two identical segments, rather than as a single long segment, because they pattern as consonant clusters with regards to syllable structure.

However, she argues that they should be considered single segments, because they do not pattern as clusters with regards to epenthesis, as well as because such sequences would trigger violations of the Obligatory Contour Principle. She argues further that Optimality Theoretic lexicon optimiza-

tion results in underlying representations and outputs as in (3.1b), with a single moraic segment syllabified as an onset, rather than as in (3.1a), as two segments spanning the syllable boundary.

- (3.1) a. /palkka/ → pa.lak.ka  
 b. /palk<sub>μ</sub>a/ → pa.la.k<sub>μ</sub>a

However, this analysis, too, is problematic: Here, geminates are considered to belong only to the syllable for which they serve as an onset, apparently serving as moraic onsets. The representations used do not take into account the autosegmental nature of syllable weight. This is undesirable for two reasons: first, moraic onsets are typologically rare at best,<sup>4</sup> and second, syllables preceding geminates seem to be closed by those geminates; they pattern as heavy syllables in Finnish. The major problem seems to be the system of representations used: Harrikari's system does not consider moras to be part of hierarchical syllable structure.<sup>5</sup>

So how then should quantity be represented in Finnish? Hayes (1989), following earlier work using moras to represent internal syllable structure, uses an autosegmental system with three tiers. The lowest represents the segmental melody, and is tied to two prosodic tiers. Weight-bearing (moraic) melodies are tied to moras, represented with the Greek letter mu ( $\mu$ ). These are then associated with the higher syllabic tier, represented

---

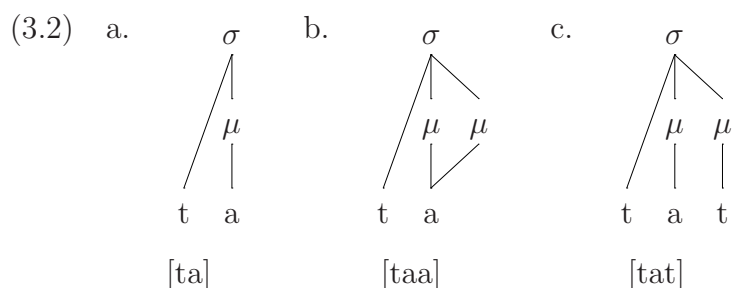
<sup>4</sup>See Topintzi (2010) for extensive discussion.

<sup>5</sup>This is apparently not the only problem with Harrikari's paper; Kari Suomi's (2000) response to it claims that some of the epenthesis data itself on which the paper is based is in fact entirely unattested, having been mis-cited numerous times by previous authors.

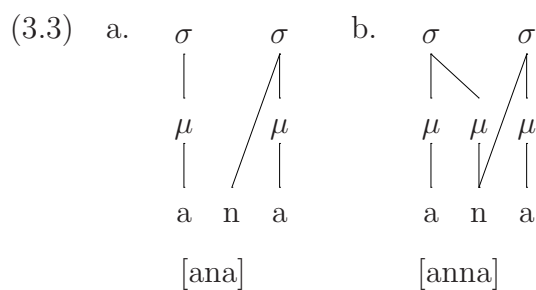


with the Greek letter sigma ( $\sigma$ ).

Syllable onsets, which do not contribute to syllable weight, are tied directly to the syllabic tier, not being associated with any moras, whereas syllable codas, in a quantity-sensitive language, are moraic. A short vowel is associated with one mora, whereas a long vowel is associated with two. Thus, Hayes arrives at the representations in (3.2):



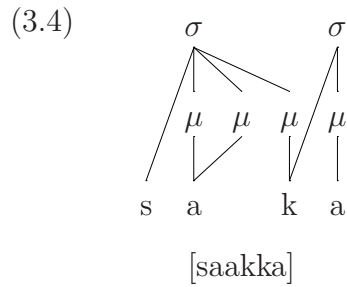
Under this theory, geminates are moraic intervocalic consonants. They are syllabified such that they contribute this weight to the preceding syllable, while at the same time serving as the onset of the following syllable, in the so-called “flopped” representation, as shown in (3.3b):



In Hayes’s theory, syllable weight is a binary distinction of light versus heavy in the absence of phonologically “ultraheavy” syllables. A light

syllable is a syllable node dominating one mora, and a heavy syllable is a syllable node dominating two moras. Because Finnish does not have more than a binary weight distinction, it is desirable to allow a maximum of two moras in our representations.

Using Hayes's theory as outlined above, maximum bimoraicity becomes problematic when analyzing syllables consisting of a long vowel followed by a geminate consonant, in which we would expect a trimoraic syllable. Consider (3.4), which shows the would-be moraic structure for *saakka* 'until':



As an alternative, I argue for a representation in which the long vowel of the complex nucleus shares its second mora with the following geminate, thus allowing a phonological distinction while retaining a binary distinction in mora count. Such a representation for *saakka* is given in (3.5) below:

(3.5) “Mora sharing” with complex nuclei

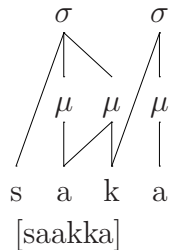


Figure 3.2: Mean segment durations of CVVCV and CVVCCV words (Lehtonen 1970, all measurements in milliseconds)

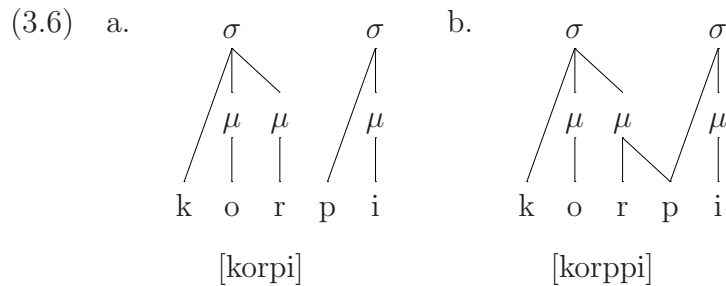
	CVVCV	CVVCCV
C <sub>1</sub>	71	80
<b>V<sub>1</sub></b>	<b>150</b>	<b>134</b>
C <sub>2</sub>	69	142
V <sub>2</sub>	48	43

Choosing such a representation makes the prediction that the long vowel in the first syllable should not be as long as a fully bimoraic vowel, as it does not have to itself the weight of two full moras. Indeed, the phonetic data seem to support this prediction. Figure 3.2 gives Lehtonen’s (1970:120-121) data on segment length in words of the shape CVVCV and CVVCCV. The mean duration of the long vowel in words of the shape CVVCV is 150 milliseconds, whereas that of CVVCCV words is 134 milliseconds, somewhat shorter. Thus, while both are phonemically long, there is a phonetic difference reflected in the representational system: the phonetically shorter vowel is forced to share its second mora with the following geminate.

“Mora sharing” is used similarly for representing complex codas. Consider the contrast between *korpi* ‘wilderness’ and *korppi* ‘crow’. The only difference phonologically concerns syllabification. In *korppi*, the consonantal melody /p/ serves as both the onset of the second syllable and part of the coda of the first syllable, whereas in *korpi*, it serves only as the onset

of the second syllable.

This distinction is shown in (3.6). In (a), /p/ is not associated with the first syllable at all. In (b), however, /p/ shares the second mora of the first syllable, serving as part of its coda, while at the same time serving as the onset of the second syllable, again in a “flopped” representation.



Much as (3.5) predicts a shorter vowel in the first syllable, because the second mora does not contribute to it entirely, (3.6) predicts that the /r/ in *korppi* should be shorter than that in *korpi*, because it must share its mora with another segment. This does in fact seem to be the case. Karlsson (1999:13) notes that sonorants in such positions are pronounced longer: “Before a short **p, t, k** or **s** the consonants **l, r, m, n** and **ŋ** are fairly long... Thus **kanssa** ‘together with’ is pronounced [kans:a] while **kansa** ‘people’ is pronounced either [kan:sa] or [kansa].”

## 3.2 Second Mora Lengthening

Suomi and Ylitalo (2004) refer to a regular phonetic process of second mora lengthening (which I henceforth term SML) in Northern Finnish, whereby

the second mora of a word is systematically lengthened phonetically provided it is voiced. One indicator of the presence of SML in a given variety of Finnish is the feature traditionally called the half-long vowel (*puolipitkä vokaali*).

The half-long vowel appears in words of the shape CVCV. In such words, the second, unstressed short vowel is actually pronounced phonetically longer than the first, stressed, short vowel. This is supported by numerous phonetic studies, including Suomi and Ylitalo (2004), Ylitalo (2004, 2009), Lehtonen (1970), and Wiik and Lehiste (1968), among others. The dialectal nature of SML and the half-long vowel can be seen in Figure 3.3, reproduced from Ylitalo (2009: 52), which shows the durations of segments in two, three, and four syllable-long words, in three different varieties, when pronounced without being accented in the carrier phrase. The Oulu and Turku varieties show the half-long vowel, while the Tampere variety does not. Thus in Oulu and Turku,  $V_2$  is consistently longer than  $V_1$ , whereas the opposite is the case in Tampere.

I indicate this phonetic lengthening in (3.7) with a length mark (:). This is in opposition to fully phonemic long vowels, which are represented as VV.

(3.7) a. *sána:* ‘word’

b. *víka:* ‘fault’

In the system of representations used here, a lengthened second mora is represented by a superscript plus sign ( $\mu^+$ ). The application of SML creating a half-long vowel can be seen in (3.8):

Figure 3.3: Segmental Duration in Three Varieties (milliseconds)

	WL	C <sub>1</sub>	V <sub>1</sub>	C <sub>2</sub>	V <sub>2</sub>	C <sub>3</sub>	V <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	V <sub>4</sub>	TWD
Oulu	2	65	<b>57</b>	71	<b>87</b>						280
	3	75	<b>58</b>	74	<b>63</b>	42	60				372
	4	73	<b>62</b>	74	<b>70</b>	46	47	50	57	44	523
Turku	2	71	<b>56</b>	85	<b>65</b>						277
	3	83	<b>60</b>	91	<b>56</b>	44	57				391
	4	81	<b>63</b>	87	<b>60</b>	48	47	53	70	31	540
Tampere	2	73	<b>58</b>	72	<b>53</b>						256
	3	81	<b>63</b>	76	<b>48</b>	44	56				368
	4	80	<b>65</b>	75	<b>50</b>	46	48	53	61	36	514

WL: Word Length (in syllables)

TWD: Total Word Duration

$$(3.8) \quad \begin{array}{c} \sigma \quad \sigma \\ \diagup \quad \diagup \\ \mu \quad \mu \\ \diagdown \quad \diagdown \\ s \quad a \quad n \quad a \\ \text{[sana]} \end{array} \rightarrow \begin{array}{c} \sigma \quad \sigma \\ \diagup \quad \diagup \\ \mu \quad \mu^+ \\ \diagdown \quad \diagdown \\ s \quad a \quad n \quad a \\ \text{[sana:]} \end{array}$$

Because the result of SML is sub-phonemic, the plus sign does not indicate a prosodic autosegment which can be present in underlying representations, but rather one which is applied by the process of SML. The significance of this to the analysis adopted in the present study will be outlined in detail in the next chapter.

### 3.3 Defining the Mora: What does it do?

Although its traditional use has been as a unit of syllable weight, the more recent use of the mora within moraic theory (e.g. Hayes 1989) has included its use in the representation of quantity.<sup>6</sup> “Long” phonemes, that is, long vowels and geminate consonants, are specified underlyingly for their moraic values in order to convey phonemic length.

I propose, however, that through the interpretation of moraic structure, it is also possible for this same structural information to regulate timing as well, without a separate skeletal tier, but rather with the “segments” on the melodic tier representing feature matrices directly tied into the syllable structure.

#### 3.3.1 Length and Duration

As discussed in section 3.1, phonological length is specified underlyingly using moras: a short vowel bears one mora, while a long vowel bears two moras. Short consonants bear no moras, while long (geminate) consonants bear one mora. Length is phonological in that it is used to convey meaning; a difference in length is as important in distinguishing meaning in the phonological system of Finnish as any other phonemic distinction.

Duration, on the other hand, is phonetic. It is possible (as many studies have) to measure the phonetic duration of segments, but without knowledge

---

<sup>6</sup>For a general overview of the use of moras and the CV tier, see Broselow (1996) and Szigetvári (2011).

of the phonological system of a language, duration alone is not enough to determine phonological length. Duration can vary depending on position within prosodic structure, and depending on stress and word accent.

Furthermore, and most important to the present study, duration can be modified by a sub-phonological process such as SML. The additional moraic material attributed to a lengthened mora translates into phonetic lengthening of the melody dominated by that mora. That is, the targeting of a phonological unit such as a mora by a phonetic process can cause phonetic alterations to the segment associated with that unit.

### 3.3.2 Weight and Timing

The traditional use of the term “mora” has been in reference to syllable weight, and this is a distinction which I retain here in addition to the representation of phonological length. This is because syllable weight is important in Finnish phonology at least for determining secondary stress placement.<sup>7</sup> On the assumption that we wish to retain this binary weight distinction, we can arrive at the following definitions of the representation of syllable weight:

- (3.9) **Light syllable:** any syllable node dominating less than two moras  
**Heavy syllable:** any syllable node dominating at least two moras

---

<sup>7</sup>See Suomi *et al.* chapter 9 for a general overview of secondary stress and syllable weight in Finnish.



How can we account for the notion of timing? By timing, I refer to the actual time it takes to pronounce a given “segment”, regardless of phonological length or weight. That is, by having timing associated with a melody, there is an actual phonological segment of which there can be a duration. In CV phonology (Clements and Keyser 1983), timing is represented with a CV tier. The elements C and V “define functional positions (peak versus non-peak) within the syllable” (Clements and Keyser 1983:10). However, with the present model, because all syllable peaks, that is, nuclei, are underlyingly moraic, this function of the CV tier is redundant.

As for the timing of onset consonants, which bear no mora, the solution is quite intuitive: a consonant is a single articulatory event. The pronunciation of a consonant necessitates the amount of time it takes to pronounce, and so the timing of weightless onsets does not require phonological representation; any such specification would be inherently redundant. This is as opposed to vowels, which reflect peaks in sonority with virtually no articulatory constriction in the vocal tract. For vowels, specification of timing (along with length, both in the form of moras) is entirely necessary, as there are no physiological articulatory constraints limiting their timing.

If we conceive of an “ideal” flow of speech as a repeating string of CV syllables, with all segments equal in duration, then all Vs (peaks) are moraic while all Cs (non-peaks) are non-moraic. This would represent a completely even speech rhythm, consisting entirely of light syllables. Within such an ideal, a heavy syllable, consisting of two moras, due either to a long vowel,

a coda consonant, or a split between both, would result in an interruption of this otherwise perfect “flow”, causing the speech rhythm to “hang up” during the more weighty syllable. Such a syllable causes this actual “hang-up” because of its greater duration; but because it is an interruption of “flow”, its actual effect is to cause that syllable to be prominent.

Although this is a somewhat abstract explanation of what constitutes syllable prominence, it seems a sure thing that greater quantity in a full-fledged quantity language such as Finnish should be connected to syllable weight. The use of the mora to represent both is thus quite justified for a language such as Finnish, especially considering the presence of quantity sensitivity.

At any rate, it is clear that there are connections between moraic structure and phonetic length, and a connection should exist between the two in the phonetics/phonology interface. At the very least, the mora does more work in the phonology than simply reflecting syllable weight. An example of the connection between phonetics and phonology is demonstrated in greater detail in the next chapter.

## Chapter 4

# Second Mora Lengthening and Primary Gemination

In this chapter I present an analysis of the process of primary gemination. This analysis involves a connection between the phonetic process of second mora lengthening (SML) discussed in section 3.2 and primary gemination itself.

Because an analysis of primary gemination which is based on SML presupposes that varieties which have primary gemination also have SML, I begin by giving evidence for the connection between these two processes across different dialects. I then present the technical details of the analysis itself, followed by implications for the applicability the model to the diachronic rise of gemination processes.

## 4.1 Evidence for a Correlation

The analysis proposed in the present work involves a unified account of the processes of second mora lengthening (SML) and primary gemination, and thus it posits a correlation between the two phenomena. That is, because SML is assumed to be the cause of primary gemination, the prediction is made that varieties with primary gemination should represent a subset of those with SML. If we assume that the presence of the half-long vowel is the result of SML, then finding the presence of the half-long vowel in a given variety should be sufficient to indicate that that variety has SML.

One source of such data comes from Kettunen (1940). This dialect atlas is the culmination of two decades of field work conducted by the author, Lauri Kettunen, throughout almost all of Finland, and contains 213 maps of various features. Map number 2 shows the presence of different types of gemination processes, namely primary gemination, special gemination, and South-Western (morphological) gemination.

Map 198 shows the presence of the half-long vowel. Based on this map, there appears to be almost no correlation between the half-long vowel and primary gemination; the half-long vowel is only shown as being present in the South-West—a tiny subset of the dialects with gemination. How do we account for this discrepancy?

The problem seems to revolve around two different definitions of the half-long vowel. The half-long vowel (and thus SML) is a phonetic characteristic

of an individual's speech. Lauri Kettunen's (1940) field work was apparently conducted by ear and based on hand-written field notes; no measurements were involved at all. Kettunen may well have been incapable of hearing the difference, even if he was listening for it.

On the other hand, studies have been conducted measuring the *phonetic* duration of the half-long vowel, such as that by Wiik and Lehiste (1968). They shed some light on the two different meanings of the half-long vowel (1968: 572):

“The half-long vowel is easily *heard* in the Standard Finnish speech of those from the South-West, but not in the speech of, for example, those from Savo. This may be due to the fact that the term “half-long vowel” as generally used in Finland refers not only to the duration of the vowels, but also to other characteristics, such as fundamental frequency and intensity contours of the words in which the half-long vowel appears. It is possible that these other characteristics differ in the two dialect areas in which the half-long vowel was found to occur.”

So it seems that the half-long vowel, and thus, SML<sup>1</sup>, may be far more wide-spread than is reported in early surveys such as Kettunen (1940). Clearly it is no coincidence that the only area in which Kettunen shows the half-long vowel is in the South-West—the very same area that Wiik and

---

<sup>1</sup> It should be borne in mind that the primary concern is in fact with the *lengthening*, and thus phonetic duration, and not with other prosodic characteristics.

Lehiste say it is easy to hear! Wiik and Lehiste's map of the home towns of the speakers in their study indicates that the half-long vowel is indeed far more widespread.

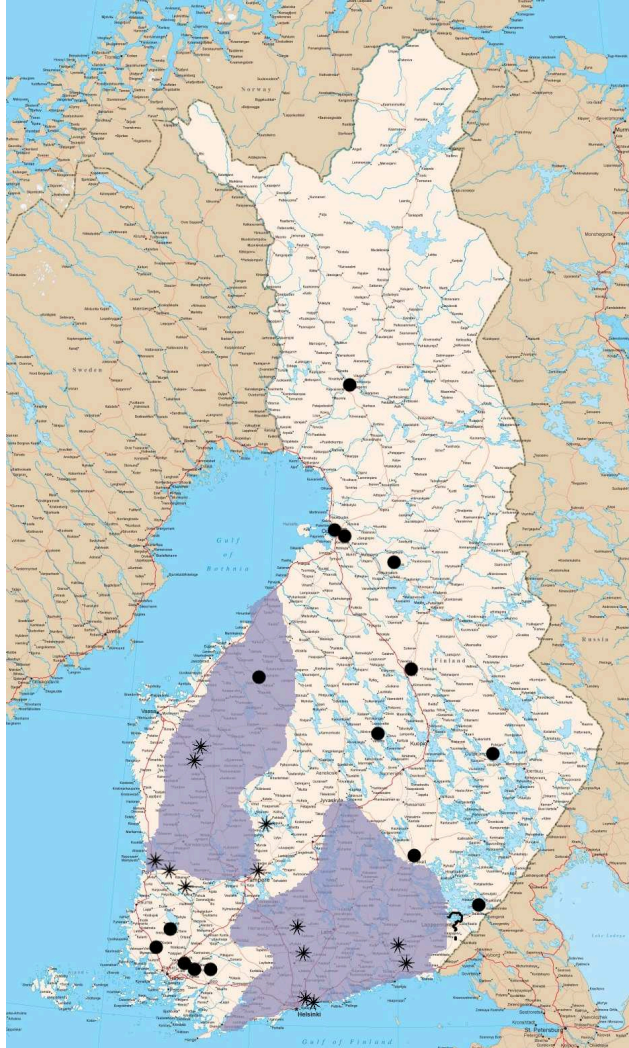
In fact, when the phonetic data in Wiik and Lehiste (1968) is overlaid on Kettunen's (1940) map of dialectal gemination, the correlation is very strong, and thus, we can conclude that there is in fact sufficient evidence for the two dialectal features occurring together. This overlay is shown in figure 4.1.

It should also be noted that the conclusion that SML and primary gemination are correlated has been reached independently by Nahkola (1987: 25). On the whole, however, this seems to be largely under-discussed in the literature, and so we can take the fact that two researchers (Nahkola and myself) have come to the same conclusion independently to provide some additional strength to the arguments based on this correlation. Furthermore, Nahkola's analysis is somewhat less technical from a theoretical point of view—this will be discussed in more detail below.

## 4.2 Modeling Primary Gemination

The analysis of primary gemination which I propose in this section involves an interaction of the phonetic process of SML targeting a vowel in certain phonological structures to create a phonemic alternation in consonantal quantity. That is to say, a phonetic lengthening triggers the creation of

Figure 4.1: Combination of Kettunen (1940) map 2 and Wiik and Lehiste’s map of speaker home towns (1968).



- Circles indicate home towns of speakers *with* the half-long vowel.
- Stars indicate home towns of speakers *without* the half-long vowel.
- The question mark indicates the home town of a speaker for whom the presence/lack of the half-long vowel was an “intermediate case”.
- Shaded areas *lack* gemination, according to Kettunen (1940).

phonological length.

The significance of this model is twofold. First, it presents a unified account of several prosodic features which co-occur in several dialects: the half-long vowel (as SML) is the factor causing primary gemination. Second, it provides evidence for gradient phonetic phenomena interacting with and affecting phonological classes, in that SML can apply gradiently, causing phonological gemination only when it reaches a certain threshold.

### 4.2.1 A Previous Account

As mentioned in section 4.1, Nahkola (1987) posits a connection between SML and primary gemination.<sup>2</sup> This connection is made with regards to historical change: the long vowel in the second (unstressed) syllable is said to have occurred as a result of the elision of an intervocalic consonant:

$$(4.1) \quad *sána-\check{d}a > sánaa$$

However, in the case that the second vowel in the earlier form in (4.1) has already become lengthened to a half-long vowel, then the result of elision is actually an overlong (“*ylipitkä*”) vowel, giving *sana:a*. The resulting structure of a stressed short vowel followed by an unstressed overlong vowel creates a severe imbalance in weight. Gemination then occurs to remove this imbalance. This explanation is rather brief, however, and no elaboration

---

<sup>2</sup>It should be noted that Nahkola thanks Heikki Paunonen (the author of Paunonen 1973) for suggesting this explanation.



is given as to the nature of this compensation. Is it assumed to be some sort of compensatory lengthening, or perhaps some sort of metrical process? My analysis serves not to oppose the explanation of overlength motivating gemination, but rather to spell out in more detail exactly what is going on in terms of the resulting alternation in quantity within a more modern generativist framework.

The steps to the change as explained by Nahkola are illustrated as a historical “derivation” in (4.2):

(4.2) Diachronically Derived Gemination

	<i>*sána-ǎa</i>
Half-Long Vowel	<i>sána:ǎa</i>
Elision	<i>sána:a</i>
Primary Gemination	<i>sánnaa</i>

There are several problems with such an historical analysis. First, Nahkola himself admits that this only works provided that the half-long vowel occurred before elision took place. However, elision is a far more wide-spread phenomenon. To the best of my knowledge, all varieties of Finnish have elision, yet the half-long vowel is only present in a subset of varieties which have elision.

At any rate, the order in which the two changes occurred historically is not important to my analysis, as the key fact is that their co-existence gives rise to long vowels targeted by second mora lengthening. Nahkola’s explanation requires an additional assumption about the order of historical

change which need not be made in order to provide a satisfactory explanation.

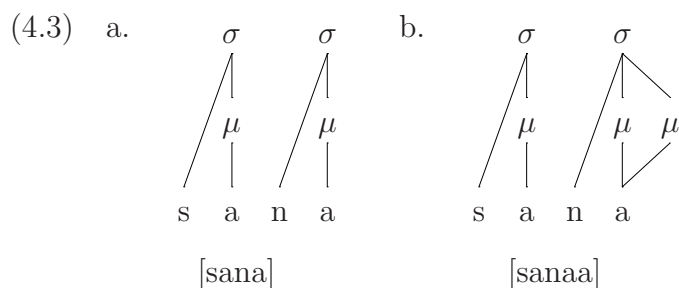
Second, this explanation is somewhat against the modern generativist approach to phonology assumed in the present study, whereby different morphemes are combined to derive surface forms. This explanation assumes that it is only forms of words which change historically, rather than their component parts (such as morphemes and phonemes) or the processes which relate those component parts. Whole forms are stored in the lexicon, in our terms, rather than individual pieces being stored, which are then combined by a number of syntactic, morphological, and phonological processes.

Furthermore, although this explanation appeals to historical factors, it fails to explain adequately the issue of the gradient rise of primary gemination: If gemination is assumed to arise as a phonetic (non-phonemic) lengthening of the intervocalic consonant, then the consonant would have to be phonologically weightless at the point before it is re-analyzed as being a phonological geminate. However, if gemination occurred in order to compensate for the imbalance in weight caused by having an overlong vowel in the second syllable, then there is no way to explain a gradient weightless lengthening as an intermediate stage. My proposal offers the explanation that this weight and length originates in the overlong vowel and gradiently “spills over” onto the preceding segment, the syllable’s weightless onset.

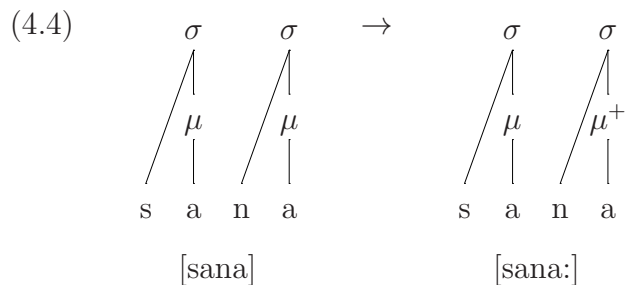
### 4.2.2 The Synchronic Connection

Instead of a purely historically-based explanation of primary gemination, I propose to develop and incorporate a synchronic model in which SML is the root cause of gemination, in addition to being the cause of the half-long vowel.

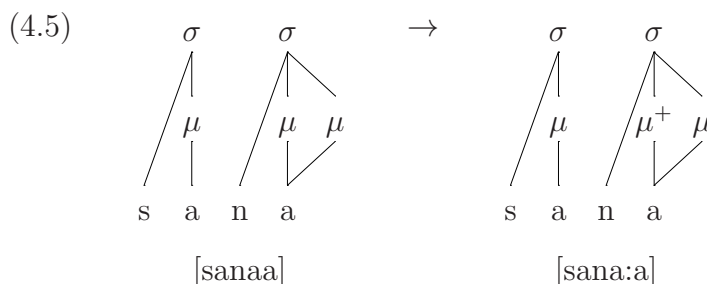
Recall that within the model of moraic theory adopted here, short vowels are represented as vowels associated with a single mora, while long vowels are associated with two moras. This is illustrated for the words *sana* and *sanaa* below:



If SML is applied to (4.3a), then the result is a half-long vowel in the second syllable. This process occurs as shown in (3.8) on page 32, repeated below as (4.4) for convenience:



The half-long vowel resulting from (4.4) has a phonetic length perceptually approaching that of a phonologically long vowel. However, it is still within the bounds of what is interpreted as a phonologically short vowel. When SML is applied to (4.3b), however, the result is an overlong vowel, as shown in (4.5):



The question is now, what becomes of the additional weight and length attributed by SML in (4.5)? Because the length of the vowel, after being affected by SML, is greater than what is generally within the confines of a long vowel in the language, there is some pressure to create a third phonological category of quantity, “overlong”, in addition to the already present long and short vowels.

However, there is a singleton consonant directly to the left of the newly-created overlong vowel, and the language already possesses a fully-established binary distinction in phonological length for consonants. That is to say, there is room phonologically on the adjacent short consonant to host the additional length attributed by SML.

Furthermore, it is not unreasonable to assume a restriction on the amount of length which can be borne by a single unstressed vowel melody, particu-

larly when, as is the case here, the result is a foot consisting of a stressed light syllable and an unstressed heavy syllable. The application of primary gemination serves to level the imbalance, creating instead a foot of two heavy syllables:

$$(4.6) \quad [\acute{L}.H]_{foot} \rightarrow [\acute{H}.H]_{foot}$$

Hyde (2011: 1061-1062) provides a general overview of foot structure. Typologically, there is a preference for trochees (binary feet in which the left branch is stressed) to have both syllables be equal in weight.

Thus, because of the phonological availability of gemination on the consonant, and the desire to correct the imbalance, there is cause for the additional weight of the overlong vowel to “spill over” to the onset of the second syllable. This additional weight is then re-syllabified as a coda of the first syllable, causing gemination:

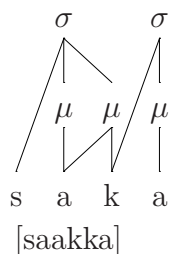
$$(4.7) \quad \begin{array}{ccc} \begin{array}{c} \sigma \qquad \sigma \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \mu \quad \mu^+ \quad \mu \\ | \quad | \quad | \\ s \quad a \quad n \quad a \\ [sana:a] \end{array} & \rightarrow & \begin{array}{c} \sigma \qquad \sigma \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \mu \quad + \quad \mu \quad \mu \\ | \quad | \quad | \quad | \\ s \quad a \quad n \quad a \end{array} & \rightarrow & \begin{array}{c} \sigma \qquad \sigma \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \mu \quad \mu \quad \mu \quad \mu \\ | \quad | \quad | \quad | \\ s \quad a \quad n \quad a \\ [sanna:a] \end{array} \end{array}$$

It is worth acknowledging that in this model, the “half-mora” or so of extra weight attributed by SML (represented by the “+”) is automatically reinterpreted as a full mora when it is applied to a consonant. This is

because consonants in Finnish appear to have less flexibility in terms of their phonetic duration within a given phonological category.

Let us for a moment return to the concept of “mora sharing” in syllables of the shape CVVC. Consider the representation of *saakka* from (3.5), repeated below as (4.8) for convenience:

(4.8) “Mora sharing” with complex nuclei



It was noted in section 3.1 that the long vowel in words of the shape shown in (4.8) (CVVCCV) is phonetically shorter than that in words of the shape CVVCV. This was used as justification for the splitting of the quantity attributed by this mora among two segmental melodies. The same correlation is found for geminate length. Figure 4.2 shows the segmental durations of the intervocalic geminates in words of the shape CVCCV (in which the geminate bears its own full mora) and CVVCCV (in which the geminate’s mora is shared with the second mora of the long vowel). The geminate in CVVCCV is seen to have a shorter mean duration.

Figure 4.2: Geminate Duration (Lehtonen 1970: 117)

	CVCCV	CVVCCV
C <sub>1</sub>	86	96
V <sub>1</sub>	66	132
C <sub>2</sub>	<b>165</b>	<b>143</b>
V <sub>2</sub>	43	42

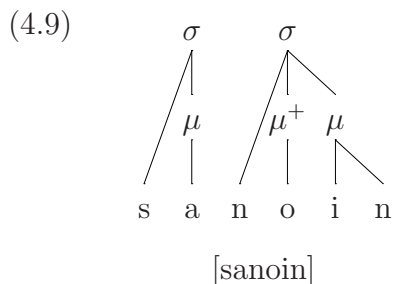
### 4.2.3 Gemination Before Diphthongs

There are several representational issues which must be addressed with regards to gemination before diphthongs. As discussed in section 2.2.2, different varieties treat diphthongs in different ways for the purposes of considering diphthongs to be triggers for primary gemination. The analysis outlined above relies on a restriction on the amount of moraic material that can be associated with a single vowel melody. Whether diphthongs are represented as single vowel melodies, and whether they are mono- or bimoraic is thus vital to accounting for the variation in diphthong behavior.

For varieties in which diphthongs never trigger gemination, diphthongs can simply be represented as two adjacent vowel melodies, each tied to its own mora under a single syllable node, as in (4.9). Thus, even if SML is applied, there is no single vowel melody bearing more than two full moras, and so there is no motivation for the additional weight to “spill over” and cause gemination.<sup>3</sup>

---

<sup>3</sup>Note that in (4.9), “mora sharing”, as outlined in section 3.1, still applies represen-



For varieties which geminate before all diphthongs, on the other hand, diphthongs must be analyzed as comprising a single complex melody bearing two moras. Such an analysis requires some considerations as to what exactly is being represented on the melodic tier. Because phonetic and phonological length of rhyme constituents are both taken to be represented by moras in the present study, there is no need for the melodic tier to play a role in timing directly. I thus propose that it is possible for some varieties to analyze diphthongs as single nodes on the melodic tier.

Let us first consider for a moment the concept of complex segments.

Broselow (1996) notes that articulatorily complex consonants, such as the

tationally, since the second syllable in *sanoin* is of the shape CVVC. However, phonetic data from Lehtonen (1970) show that the long vowel in CVCVVC words is almost identical to (and in fact, slightly longer than) that of CVCVV words, as in (i). Thus, the vowel nucleus is assumed to bear the equivalent of two full moras.

(i)

	CVCVV	CVCVVC
C <sub>1</sub>	72	63
V <sub>1</sub>	67	68
C <sub>2</sub>	81	82
V <sub>2</sub>	<b>145</b>	<b>155</b>
C <sub>3</sub>		58

It is not entirely clear why these long vowels are the same length, however it seems that the variability in phonetic length caused by mora sharing is far more robust in stressed syllables. For this reason, I set aside these data as beyond the scope of the present discussion.



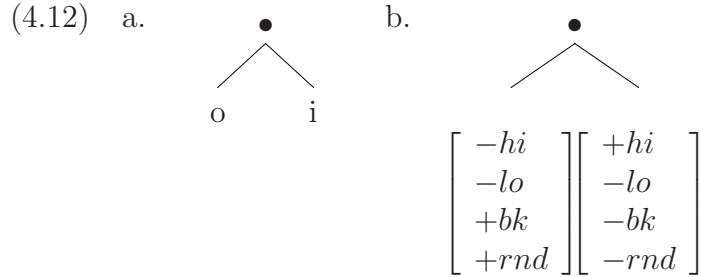
affricate  $[t^s]$  are often represented as the melodies  $[t]$  and  $[s]$  associated to a single root node:

$$(4.10) \quad [t^s] = \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ t \quad \quad s \end{array}$$

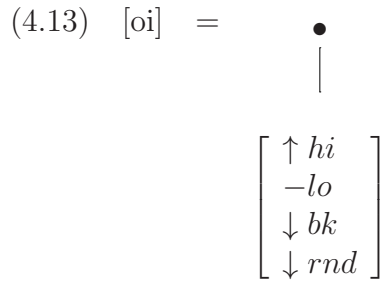
Similarly, a simple segment can be represented as a single melody attached to a root node; consider (4.11), which shows the segmental representation for the monophthongal melody  $[o]$ . In (4.11b), the feature specifications for the vowel melody itself replace the shorthand notation of  $[o]$ .

$$(4.11) \quad \text{a.} \quad \begin{array}{c} \bullet \\ | \\ o \end{array} \quad \text{b.} \quad \begin{array}{c} \bullet \\ | \\ \left[ \begin{array}{c} -hi \\ -lo \\ +bk \\ +rnd \end{array} \right] \end{array}$$

A parallel can be drawn to the complex melody of the affricate in (4.10) in order to represent a diphthong. In (4.12a), the complex melody of the diphthong  $[oi]$  is shown as  $[o]$  and  $[i]$  linked to a single root node. (4.12b) gives the feature matrices for both vowel melodies, again, both linked to the root node.

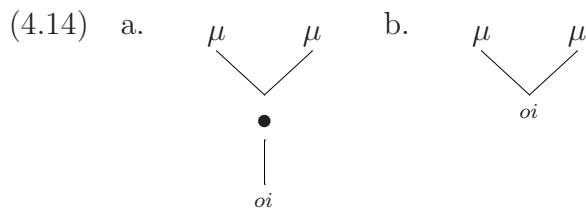


Because dialects with representations such as that in (4.12) do effectively consider diphthongs to be single “segments”, their feature matrices can be further simplified. In (4.13), an up arrow ( $\uparrow$ ) indicates a global change from a negative to positive value of a feature, and a down arrow ( $\downarrow$ ) represents a global change from a positive to a negative value of a feature. Under this system of notation, plus (+) and minus (−) represent features which remain the same throughout the duration of the segment.



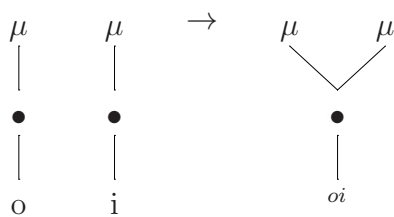
Since [oi] can be understood to be a single segment, it can be further simplified notationally as [ $oi$ ]. The moraic material associated with a segment is then linked above the root node, as shown in (4.14a). However, since it is understood that a single root node in this system represents a

single segment, and not that segment's phonological length, we can omit the association line between the root node and the melodic tier, and instead leave just the melody attached the relevant moraic material, as in (4.14b).



Diphthongs can be both derived and underlying in Finnish. In the case of an underlying /oi/, the representation in a speaker's lexicon would simply be as in (4.13). In the case of derived diphthongs, however, the surface diphthong [oi] must have originated as two separate melodies, /o/ and /i/, linked to two adjacent root nodes. For varieties which analyze diphthongs as single complex melodies, I posit a process whereby two adjacent nuclear melodies within the same syllable fuse together, preserving any moras associated with them:

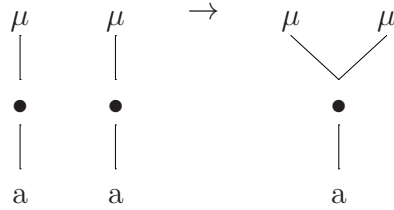
(4.15) “Nuclear Fusion”



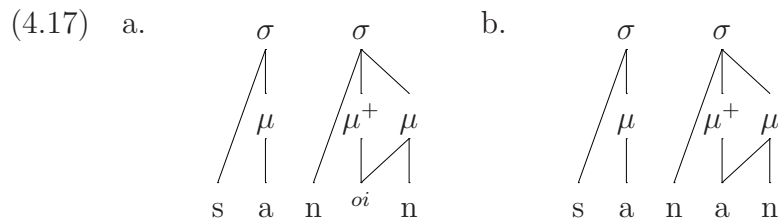
This is by no means unprecedented based on language-internal evidence from Finnish, as derived long vowels behave in the same way: two adjacent

identical vowels become long vowels on the surface, and behave as a single bimoraic vowel melody for the purposes of phonological computation with regards to the creation of geminates:

(4.16) “Fusion of Derived Long Vowels”



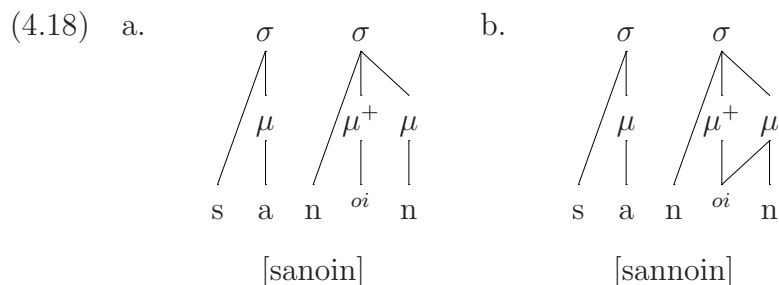
With this understanding of complex nuclear melodies, it is a simple matter to account for the different ways in which varieties may treat diphthongs for the purpose of triggering primary gemination. In varieties in which gemination occurs before all diphthongs, the nuclear melody is always bimoraic, as in (4.17a). This is completely analogous to a geminating word with a long vowel in the same position, such as *sanaan* ‘word (illative case)’, given in (4.17b).



In both cases in (4.17), the vowel in the second syllable is directly dominated by two moras, one of which is lengthened by SML, creating overlength. Thus, as in earlier examples, the conditions for primary gemination

are met and “spill over” occurs, yielding *sannoin* for (4.17a) and *sannaan* for (4.17b).

There are several solutions for “selective” varieties, that is, those in which gemination only occurs before some diphthongs. In the case of Savo dialects which Paunonen (1973) has reported as having two distinct diphthong lengths, gemination occurs before long diphthongs, but not before short diphthongs. In such a variety, long diphthongs would be bimoraic, like long vowels, and short diphthongs would be monomoraic, like short vowels, as in (4.18).<sup>4</sup>



The final issue deals with the analysis of diphthongs in “selective” varieties in which there is apparently no surface distinction in diphthong length between those that cause gemination and those that do not. There are several possible explanations for this:

First, it is possible that the two types of diphthongs are still interpreted as either mono- or bimoraic, as with those in (4.18). In this case, there

---

<sup>4</sup>Note that the representations in (4.18) and (4.20) reflect the moraic structure *before* primary gemination is applied in the derivation; the purpose of these examples is to show the moraic structure of the second syllable’s nucleus.

would still be a phonological distinction in weight for the purposes of computation. This would then require us to posit some principle requiring diphthongs to surface as phonetically long regardless of their underlying phonological status. However, this explanation is undesirable given the spirit of the present study, which aims to use phonological representations of quantity which represent the phonetic reality as closely as possible, in order to account for interplay between the two.

Second, the issue could be considered more of a morphophonological one than one dealing purely with phonological structure. This is because the diphthongs that cause gemination are derived from different underlying morphology than those that do not, for example as in (4.19).<sup>5</sup> The phonological alternation of primary gemination could then be referenced in a morphological rule or constraint.

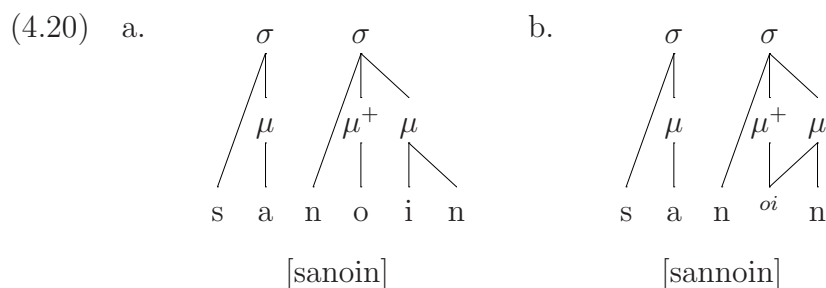
(4.19)	/sano-i-n/ 'I say'	Non-geminating	"Selective"
		<i>sanoïn</i>	<i>sanoïn</i>
	/sano-ta-i-n/ 'to say (instr. plural)'	<i>sanoïn</i>	<i>sannoïn</i>

Third, it could be that the two types of diphthongs are actually identical in length, but have different (morpho-)phonological structure. This requires that a "selective" variety in which diphthongs are the same length

---

<sup>5</sup>The exact details of the phonological processes resulting in the surface structures in (4.19) are not important, as I do not mean to take a stance on all details of the segmental phonology of Finnish. However, it is clear that there is different underlying morphological structure in play.

only optionally applies “nuclear fusion”, depending on the morphological conditions which gave rise to the diphthong. Thus we allow for, within the same grammar, the representations used in both (4.9) and (4.18b), shown in contrast in (4.20):



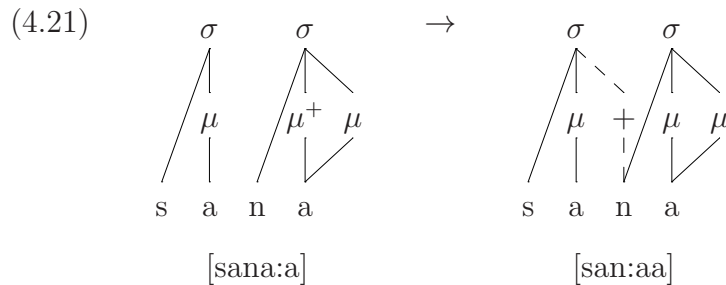
Finally, it is possible that there is in fact a phonetic difference in the quantity of the gemination-causing and non-gemination-causing diphthongs in this variety, and it has merely gone unreported due to an insufficient amount of data having been collected. As far as I am aware, there is no phonetic study examining this. At any rate, choosing between these possible solutions is left for future research, as further investigation is not possible at this time.

#### 4.2.4 Gradient Effects: A Model for Change

In addition to providing an analysis of the synchronic process of primary gemination which unifies several prosodic features of Finnish dialects, the model I have presented above is advantageous in that it can account for the diachronic rise of the process of primary gemination.

The additional length contributed by the lengthened second mora (represented as “+”) has at no point been explicitly defined. I have referred to it as being equivalent to something like a “half mora”, but have made no attempt to formalize what this actually means. The reason for this is that the amount of actual length which it indicates can differ from speaker to speaker and dialect to dialect, and also depending on the word’s position within a prosodic phrase.

The inherently gradient nature of this additional length allows for a model of gemination in line with the assumed rise of primary gemination (see section 2.2.1). Gemination is said to originate as phonetic lengthening of the intervocalic consonant, lasting for perhaps a few generations before being re-analyzed as a phonemic geminate. This could be represented as the intermediate stage of the derivation in (4.7), instead acting as the final surface representation in (4.21). Because the length attributed to the consonant does not get interpreted as a full mora, phonemic gemination does not occur; instead, the lengthening is merely phonetic.



This is also in line with the proposal that consonants disprefer gradient lengthening, preferring to be either long or short, while vowels allow greater



variability in their length within a given phonemic class of quantity. The situation in (4.21) is indeed an unstable and dispreferred one, leading to a rapid re-analysis by subsequent generations of learners. Half-length does not sit as well with consonants as it does with vowels. Once the gradient length attributed by SML reaches a certain threshold, full phonemic primary gemination is born.

#### **4.2.5 Summary**

I have presented a model for primary gemination which unifies a phonetic process of second mora lengthening (SML) with a phonological alternation in consonant length, providing evidence for interplay of phonetic and phonological quantity.

A further advantage of this model is that it incorporates the gradient nature of the length attributed by SML, and thus makes a prediction in line with the assumed diachronic rise of gemination as phonetic lengthening. Thus it is compatible with historic change as well as with a synchronic system.



## Chapter 5

# Additional Issues and Closing Remarks

### 5.1 Further Theoretical Issues

As it stands, there are some further theoretical issues which should be addressed with regards to how the analysis in the preceding chapter is presented, in terms of what exactly the assumptions are about how syllabification is accomplished based on moraic and segmental structure.

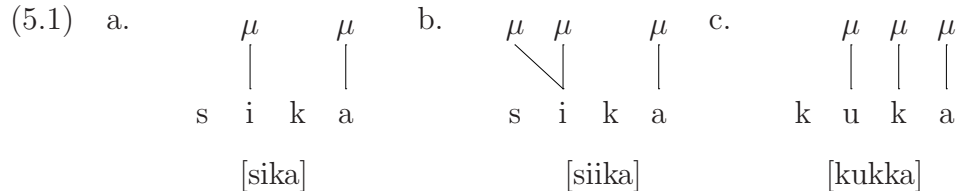
There is some question as to how the structure represented by syllable trees is built and interpreted. This section seeks to make clear what must be present underlyingly, how that information is processed, and at what point in the derivation primary gemination fits in.

The overarching assumption is that at the end of a derivation, struc-

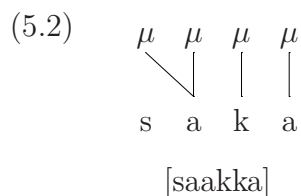
ture should represent syllabification with a binary distinction in phonological quantity. That is, syllables are either monomoraic or bimoraic. The phonological structure built during syllabification is then interpreted at the phonetics/phonology interface, and produces an output with the correct phonetic quantity.

### 5.1.1 Underlying Representations

Underlying representations consist primarily of a string of segmental melodies. Following Hayes (1989), I assume that vowels are underlyingly moraic. Short vowels have one mora, while long vowels have two moras. Underlying geminates also have a mora, while all other consonants are non-moraic. Examples of the underlying forms of *sika* ‘pig’, *siika* ‘whitefish’, and *kukka* ‘flower’ are given in (5.1).



In words of the shape CVVCCV, the long vowel is assumed to be bimoraic, and the geminate moraic, at the level of underlying representations; “mora sharing” is not assumed to be a property of underlying representations. The underlying form of *saakka* ‘until’ is given in (5.2).



It is this information which is used as the input to the phonological processes of the language, which can alter both segments and prosodic structure.

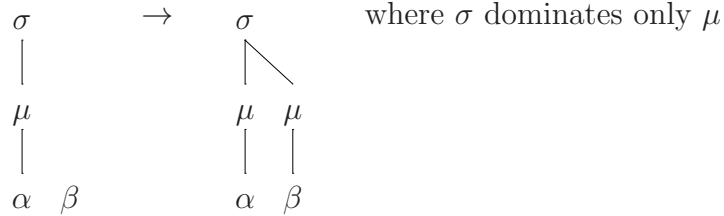
### 5.1.2 Segmental Phonology and Syllabification

The first stage of phonological processing involves the application of segmental processes and the building of prosodic structure, including syllabification. This stage includes morphophonological processes such as consonant gradation, which can alter moraic structure, for example, by reducing geminate stops to singleton stops when they constitute the onsets of closed syllables.

Coda consonants which do not constitute geminates are moraic in Finnish, as they contribute to syllable weight.<sup>1</sup> However, because they are not part of geminates, they are non-moraic underlyingly. They receive their weight only with reference to the syllabic structure which is built up later during the derivation. In order to assign this mora, Hayes (1989:258) posits a process of “weight by position”, reproduced in (5.4).

---

<sup>1</sup>This is the case at least insofar as the effect of syllable weight on secondary stress is concerned.

(5.3) *Weight by Position*

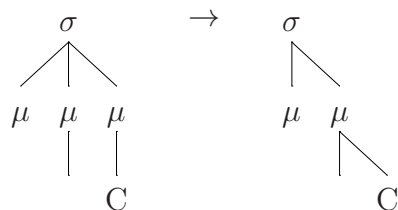
However, because Finnish does not allow complex onsets in native words,<sup>2</sup> “weight by position” can actually be further simplified:

(5.4) **Weight by Position:** Assign a mora to any consonant which is not immediately followed by a vowel.

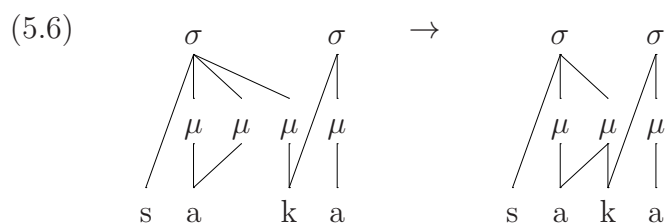
The rule in (5.4) will tend to generate trimoraic syllables, and so another thing which must be accounted for by syllabification is “mora sharing” (see page 28). Mora sharing is necessary in order for syllables to retain a binary weight opposition between light (monomoraic) and heavy (bimoraic). This is accomplished by the application of a process which I term “mora compression”.

Mora compression applies to any trimoraic syllable. All segmental melodies dominated by the last two moras of a trimoraic syllable are “remoraified” so that they are dominated by a single mora. This is formalized in (5.5).

<sup>2</sup>Indeed, speakers with little to no knowledge of foreign languages are well-known to simplify complex onsets even in loan words, such that *presidentti* ‘president’, for example, is pronounced [resitentti].

(5.5) *Mora Compression*

The application of this to *saakka* ‘until’ is shown in (5.6) below; the trimoraic syllable is modified to be bimoraic.



This stage of phonological processing is also where prosodic structure is assigned to diphthongs. As discussed in section 4.2.3, there is a great deal of variation in terms of how different dialects analyze diphthongs, both in terms of their weight and their status as individual nuclei. For this reason, the particular formalization necessary to adequately describe the processes responsible for assigning this structure is equally diverse.

### 5.1.3 Phonetic Interpretation

After phonological processing has been completed and prosodic structure assigned, forms are passed off to the phonetics/phonology interface for phonetic interpretation, in order to have both duration and other prosodic fac-

tors, such as stress and intonation, applied.

The exact correlates of the duration of different phonological quantities in different areas of word structure vary greatly from dialect to dialect, and depending on the details of word structure. Quantitative studies such as Lehtonen (1970) and Ylitalo (2009) provide excellent discussion illustrating the complications and variation present, and mapping every detail of the interface which interprets this is a task beyond the present study.

One process which takes place at this interface, however, is second mora lengthening (SML). Because of the position in the derivation that SML occupies, when it applies in the right prosodic context, namely, in words of the shape  $C\acute{V}CVV$ , the result is a sort of “last resort” resyllabification. This occurs as a result of the “spill over” of weight which it contributes to the unstressed long vowel, pushing its perceptual bounds. This resulting primary gemination is the only place that the model predicts that a phonological alternation can occur after syllabification has been completed, since, to the best of my knowledge, there are no other phonetic processes in Finnish which can trigger such a change in phonological prosodic structure.

## 5.2 Conclusions

In this study, I have argued for a connection between second mora lengthening and primary gemination, whereby the extra weight attributed by second mora lengthening creates a half-long vowel in the second syllable of  $CVCVV$



words. This additional phonetic length and weight “spills over” onto the onset of the second syllable, and is reinterpreted as a phonological geminate.

The implication for moraic theory is that a phonetic process which spells out the duration of a segment, applying after syllabification and moraic structure has been established for a given form, can still result in a phonological alternation which must alter syllable structure at the end of a derivation.

Furthermore, I have helped to develop a deeper understanding of the connection between different prosodic phenomena in Finnish dialects by connecting two of them in a model which can account for both the synchronic system and diachronic change from phonetic allophony to a phonological alternation.

### 5.3 Directions for Future Research

There are several topics which should be examined in future research related to the present study. Firstly, secondary stress is a matter which has been dealt with somewhat vaguely. This is because there seems to be a great deal of disagreement as to what exactly constitutes secondary stress in Finnish, and anecdotal investigation thus far on my part has suggested that the placement of secondary stress differs between varieties depending on other prosodic factors, such as the half-long vowel. There is certainly much work to be done in this regard, in terms of both collecting and analyzing data

from different varieties.

The treatment of diphthongs in different geminating varieties has also been treated somewhat speculatively. This, again, is due to a lack of data on the matter. Field research and/or analysis of existing dialect recordings would be extremely worthwhile.

There is also a question raised in section 4.2.1 as per the order in which elision and the half-long vowel occurred historically. Although it is not necessarily vital to the analysis of dialectal gemination presented in this paper, it would certainly be useful to find out, and I suspect that this would be possible to do by looking at other dialectal features or dialectal variants of certain vocabulary items.

Finally, the specifics of the way that syllable structure is parsed should be revisited in more detail. The need for a derivation to create trimoraic syllables via weight-by-position and then change these to bimoraic via mora compression may be considered somewhat redundant by some. For this reason, it would be fruitful to analyze the data under different assumptions of phonological “processes”, for example within a non-derivational framework.

# References

- Broselow, Ellen. 1996. Skeletal Positions and Moras. In: John Goldsmith (ed.) *The Handbook of Phonological Theory*. Oxford: Blackwell.
- Clements, George N. and Samuel Jay Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: The MIT Press.
- Harrikari, Heli. 1999. Epenthesis, Geminates, and the OCP in Finnish. *Nordic Journal of Linguistics* 22: 3–26.
- Harrikari, Heli. 2003. From Phonology to Morphology: A Synchronic Account of Dialectal Gemination in Finnish. *MIT Working Papers in Linguistics Volume 45: Proceedings of the Eleventh Student Conference in Linguistics*.
- Hayes, Bruce. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry* 20: 253–306
- Hyde, Brett. 2011. The iambic-trochaic law. In: Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.) *The Blackwell Companion to Phonology*. Wiley-Blackwell.

- Karlsson, Fred. 1999. *Finnish: An Essential Grammar*. London: Routledge.
- Kettunen, Lauri. 1940. *Suomen Murteet III A. Murrekartasto*. Helsinki.
- Lehtonen, Jaakko. 1970. *Aspects of Quantity in Standard Finnish*. Jyväskylä: Jyväskylä University.
- Nahkola, Kari. 1987. *Yleisgeminaatio: Ääntenmuutoksen synty ja vaihteet kielisysteemissä erityisesti Tampereen seudun hämäläismurteiden kannalta*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Nahkola, Kari and Marjatta Palander (eds.). 1981. *Huomioita Suomen Murteiden Geminaatioilmiöistä*. Folia Fennistica & Linguistica. Tampere.
- Paunonen, Heikki. 1973. On the Primary Geminization of Finnish Dialects. *Finnisch-Ugrische Forschungen* 40: 146–164.
- Suomi, Kari. 2000. Nearly a Million Ostrobothnians Cannot Be Wrong: Comments on an Optimality Theoretic Analysis of “Ostrobothnian”. *Nordic Journal of Linguistics* 23: 89–93.
- Suomi, Kari, Juhani Toivonen and Riikka Ylitalo. 2008. *Finnish Sound Structure: Phonetics, Phonology, Phonotactics, and Prosody*. Sudia Humoniora Ouluensia 9. Oulu: University of Oulu.
- Suomi, Kari and Riikka Ylitalo. 2004. Durational and Tonal Correlates of Accent in Finnish. *Journal of Phonetics* 32: 35–63.

- Szigetvári, Péter. 2011. The Skeleton. In: Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume and Keren Rice (eds.) *The Blackwell Companion to Phonology*. Wiley-Blackwell.
- Topintzi, Nina. 2010. *Onsets: Suprasegmental and Prosodic Behavior*. Cambridge: Cambridge University Press.
- Wiik, Kalevi and Ilse Lehist. 1968. Vowel Quantity in Finnish Disyllabic Words. *Congressus Secundus Internationalis Fenno-Ugristarum*. Helsinki.
- Ylitalo, Riikka. 2004. Toisen tavun vokaalin puolipidennyksestä oulunseutulaisten puheessa. *Virittäjä* 108: 414–422.
- Ylitalo, Riikka. 2009. The Realization of Prominence in Three Varieties of Standard Spoken Finnish. Doctoral Thesis. University of Oulu.