Arto Anttila

VARIATION AND OPACITY

ABSTRACT: Phonological variation and phonological opacity have been extensively studied independently of each other. This paper examines two phonological processes that simultaneously exhibit both phenomena: Assibilation and Apocope in Finnish. The evidence converges on two main conclusions. First, variation results from the presence of MULTIPLE METRICAL SYSTEMS within Finnish. Assibilation and Apocope are metrically conditioned alternations and the segmental variation reflects metrical variation. The metrical analysis explains a number of apparently unrelated phenomena, including typological asymmetries across dialects, quantitative asymmetries within dialects, differences between nouns and verbs, differences among noun classes, and the loci of lexical frequency effects. Second, phonological opacity arises from MORPHOLOGICAL LEVEL ORDERING. By interleaving transparent phonologies with independently motivated morphosyntactic constituents (stems, words, phrases) we derive the transparent and opaque interactions of four phonological processes, including Assibilation and Apocope.

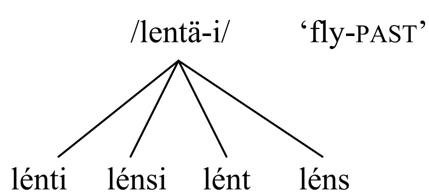
ABBREVIATIONS: 1P/2P/3P – first/second/third person; ALL – allative; CAUS – causative; CLIT – clitic; COND – conditional; ESS – essive; GEN – genitive; ILL – illative; INE – inessive; PAR – partitive; PAST – past tense; PL – plural; PX – possessive suffix; SG – singular; SUP – superlative; TRA – translative; QUE – question

1. Introduction

Phonological variation and phonological opacity have been extensively studied independently of each other. Variation arises when one input yields multiple outputs. Opacity arises when a phonological process applies even if its conditioning environment is not met on the surface (overapplication), or conversely, fails to apply even if its conditioning environment is met on the surface (underapplication).

As an illustration, consider two phonological processes in Finnish: ASSIBILATION which fricativizes a short nongeminate *t* before *i* (*t* → *s* /_ *i*) and APOCOPE which deletes a short *i* in an unstressed syllable (*i* → ∅).

(1) Assibilation and Apocope (Southeastern Finnish, Laalo 1988)



The example in (1) illustrates both variation and opacity. Both Assibilation and Apocope are variable in this dialect, which results in four distinct surface forms: *lenti ~ lensi ~ lent ~ lens*. The form *lens* is an example of overapplication opacity: Apocope has removed the vowel *i* that triggers Assibilation, but Assibilation nevertheless applies. As we will see in a moment, patterns like (1) are not particularly rare or marginal, suggesting that variation and opacity are part of the core phonology of the language.

This paper examines the interactions of Assibilation and Apocope. We will work out the empirical generalizations, derive them from general principles, and follow up on their implications for phonological theory. The evidence converges on two main conclusions. First, variation results from the presence of multiple metrical systems within Finnish. Assibilation and Apocope are metrically conditioned alternations and the segmental variation reflects metrical variation. Second, phonological opacity arises from morphological level ordering. By interleaving transparent phonologies with independently motivated morphosyntactic constituents (stems, words, phrases) we derive the transparent and opaque interactions of four phonological processes, including Assibilation and Apocope.

Our data come from a number of different sources. Most of the dialect data come from Klaus Laalo's monograph on Finnish Assibilation (Laalo 1988) which contains approximately 3,200 examples of Assibilation and Apocope from 25 regional dialects, given in phonemic transcription in their original sentence context. For the purposes of this study, Laalo's corpus was annotated phonologically by the present author. The annotated corpus is easy to manipulate on a computer, which greatly helps in exploring the relationships among the alternations. We have also made some use of the recently launched *Electronic Morphology Archives for Finnish Dialects*, available at the Finnish IT Center for Sciences http://www.csc.fi/, which overlaps with Laalo's corpus. Our data on spoken Helsinki Finnish come from Heikki Paunonen's *Spoken language in the Helsinki area 1972-1974* corpus, available at the University of Helsinki Language Corpus Server http://www.ling.helsinki.fi/uhlcs/. The corpus consists of interviews with 126 speakers of Helsinki Finnish, collected and transcribed by Paunonen and his associates in the early 1970's and documented in Paunonen 1995, approximately 500,000 word forms in all. Our data on Written Standard Finnish come mainly from three newspaper corpora: the *Aamulehti 1999* corpus (16,608,843 word forms), the *Turun Sanomat 1999* corpus (11,821,904 word forms) and the *Kaleva 1998-1999* corpus (9,758,628 word forms), compiled by the Research Institute for the Languages of Finland, the Department of General Linguistics at the University of Helsinki, and the Finnish IT Center for Sciences, available at http://www.csc.fi/.

The paper is structured as follows. Section 2 situates Assibilation and Apocope within the general context of Finnish phonology and morphology and motivates level ordering. Assibilation is a lexical process that applies to stems, Apocope is a postlexical process that applies across the board, and lexical processes feed postlexical processes. Section 3 puts forward a metrical analysis of Assibilation that explains a number of apparently unrelated phenomena, including typological asymmetries across dialects, quantitative asymmetries within dialects, differences between nouns and verbs, differences among noun classes, and the loci of lexical frequency effects. Section 4 shows that the same metrical analysis also accounts for Apocope. Section 5 examines the consequences of the Finnish evidence for the theory of Comparative Markedness (McCarthy 2003). Section 6 concludes the paper.

2. THE INTERACTION OF ASSIBILATION AND APOCOPE

We start by situating Assibilation and Apocope within the general context of Finnish phonology. Two additional processes will be helpful: Vowel Deletion which deletes a stem-final low vowel before *i* (*a, ä* → ∅ / _ *i*) and Degemination which shortens a geminate stop in the onset of a closed syllable (*pp, tt, kk* → *p, t, k* / _VC.). All four processes are illustrated in (2).

(2)	Vowel Deletion	a → ∅ / _i	huuta-i → huuti	'shout-PAST'
	Assibilation (optional)	t → s / _i	huut-i ~ huusi	'shout-PAST'
	Degemination	tt → t / _VC.	ott-i-n → otin	'take-PAST'
	Apocope (optional)	i → ∅	huus-i ~ huus	'shout-PAST'

In a derivational analysis (Chomsky and Halle 1968), Assibilation must precede Apocope. This is demonstrated by the interactions in (3).

(3) a. Vowel Deletion feeds Assibilation:
/huuta-i/ → huuti → huusi 'shout-PAST'
b. Degemination counterfeeds Assibilation:
/otta-i-n/ → ottin → otin (→ *osin) 'take-PAST-1P.SG'
c. Apocope counterfeeds Degemination:
/hakkat-i/ → hakkasi → hakkas (→ *hakas) 'beat-PAST'

The interactions in (3) entail the rule ordering Vowel Deletion < Assibilation < Degemination < Apocope. This ordering is also supported by other interactions. For example, Vowel Deletion precedes and therefore feeds Degemination (/otta-i-n/ → *ottin* → *otin*) and Apocope follows and therefore counterbleeds Assibilation (/huuta-i/ → *huuti* → *huusi* → *huus*). The ordering is also consistent with dialect typology. In dialects where both Assibilation and Apocope are optional, we expect four-way variation, illustrated in (4) for the verb / lentä-i/ 'fly-PAST'.

(4)

	/lentä-i/		/lentä-i/		/lentä-i/	/lentä-i/
Vowel Deletion	lenti		lenti		lenti	lenti
Assibilation (optional)	--		lensi		--	lensi
Degemination	--		--		--	--
Apocope (optional)	--		--		lent	lens
	[lenti]	~	[lensi]	~	[lent] ~	[lens]

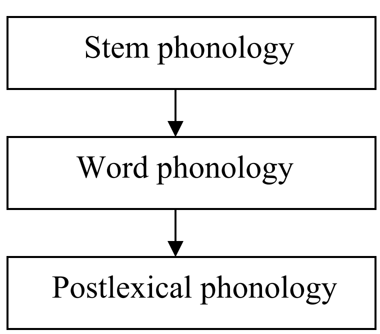
An example of such a dialect is Southeastern Finnish:

(5) a. lent-i sii-he kuva-m peä-le (Jaakkima, Southeast, *ti*-variant)
fly-PAST it-ILL picture-GEN top-ALL
'flew there, onto the picture'
b. mulla-t lens-i-it (Hiitola, Southeast, *si*-variant)
soil-PL fly-PAST-3P.PL
'soil flew (was scattered) around'
c. se lent ko lintu (Hiitola, Southeast, *t'*-variant)
it fly(-PAST) like bird
'it flew like a bird'
d. se ku lens moa-ha (Kurkijoki, Southeast, *s'*-variant)
it EMPH fly(-PAST) ground-ILL
'it flew to the ground'

The dialect pattern in (5) implies that Assibilation crucially precedes Apocope. If we reverse the ordering, we get three-way variation: *lenti~lensi~lent*, but no **lens*. Such dialects are systematically absent. We conclude that the ordering theory is well supported by the facts, suggesting that something about it is right and worthy of further exploration.

A particular interpretation of phonological ordering is put forward in Stratal Optimality Theory (Kiparsky 2000, 2003; see also Bermúdez-Otero 1999, Cohn and McCarthy 1994/1998, Hale, Kissock, and Reiss 1998, Itô and Mester 2002, Kenstowicz 1995, McCarthy and Prince 1993, Orgun 1996, Rubach 2000). The central hypothesis of Stratal Optimality Theory is that phonological ordering reflects morphological ordering. Morphology is divided into three levels called STEM LEVEL, WORD LEVEL, and POSTLEXICAL LEVEL, and morphological and phonological operations apply in tandem within each level. For precedents in Lexical Phonology and Morphology, see e.g. Kiparsky 1982, Mohanan 1986, cf. Goldsmith 1993. The main difference between Lexical Phonology and Stratal Optimality Theory is that the latter has no ordered phonological rules. Instead, each morphological level is associated with an optimality-theoretic phonological grammar (Prince and Smolensky 1993/2004). This predicts that phonological processes should interact transparently within a level, but not necessarily across levels because the levels are serially ordered: the output of stem-level phonology is the input to word-level phonology, and the output of word-level phonology is the input to postlexical phonology. Opacity arises because word-level processes can mask stem-level processes and postlexical processes can mask both stem-level and word-level processes.

(6) Stratal Optimality Theory (Kiparsky 2003)



Stratal Optimality Theory predicts that phonological ordering should reduce to morphological ordering. The Finnish case is interesting because we have four ordered rules, but only three ordered levels. This predicts that some two adjacent rules must belong to the same level and interact transparently. There are exactly three options:

(7) a. Vowel Deletion and Assibilation interact transparently at the stem level.
b. Assibilation and Degemination interact transparently at the word level.
c. Degemination and Apocope interact transparently at the postlexical level.

The choice is uniquely dictated by the facts: only Vowel Deletion and Assibilation interact transparently. The rule ordering must thus be translated into level ordering as shown in (8).[1]

(8)

RULE ORDERING		STRATAL OPTIMALITY THEORY
1. Vowel Deletion	→	1. Stem Level
2. Assibilation	→	1. Stem Level
3. Degemination	→	2. Word Level
4. Apocope	→	3. Postlexical Level

Based on purely phonological evidence, we have now located the four phonological processes at specific morphological levels. Three morphological predictions now follow:

(9) a. Vowel Deletion and Assibilation should interact with stem-level morphology.
b. Degemination should interact with word-level morphology.
c. Apocope should have no morphological or lexical conditions at all.

Since Finnish is a suffixing language, stem-level suffixes precede word-level suffixes which precede clitics, i.e. [[[[*root*] *stem-level suffixes*] *word-level suffixes*] *clitics*]. The theory thus predicts that Vowel Deletion and Assibilation should apply in the smallest domain which contains the root and some number of suffixes immediately on its right (the stem domain); Degemination should apply in the intermediate domain which contains the stem domain and the suffixes immediately on its right (the word domain); and Apocope should apply in the largest domain which contains the word domain and the clitics immediately on its right (the postlexical domain). These predictions could easily be false: there is no necessary connection between the opacity of a phonological process and its morphological environment. Remarkably, all three predictions are confirmed.

---

[1] Based on a complementary set of evidence, Kiparsky (2003) motivates the existence of two levels in Finnish. Kiparsky distinguishes between stem-level lexical stress which interacts with morphophonemics and word-level rhythmic stress which does not. The analysis differs from the present one in that it locates Degemination at the stem level. It seems possible to reconcile the two analyses by assuming that Degemination is a word-level process (as argued here and also in Kiparsky 1993a) and that rhythmic stress is postlexical. A detailed demonstration must be left for future work.

Finnish bound morphemes are traditionally divided into four basic classes by linear position (Karlsson 1982, p. 231): DERIVATIONAL SUFFIXES (position 1), SIGNATURES (position 2), ENDINGS (position 3), and CLITICS (position 4). We propose that the four traditional position classes are mapped onto the three morphological levels of Stratal Optimality Theory as shown in (10).[2]

(10) Finnish morphotactics and level ordering

<table>
<tr><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td></tr>
<tr><td>root</td><td>derivational suffixes</td><td>signatures</td><td>endings</td><td>clitics</td></tr>
<tr><td colspan="3">Stem Level</td><td></td><td></td></tr>
<tr><td colspan="4">Word Level</td><td></td></tr>
<tr><td colspan="5">Postlexical Level</td></tr>
</table>

(11) Examples of the content of each position class:
a. Derivational suffixes: /-impA/ ‘superlative’, /-ime/ ‘instrument’
b. Signatures: number in nominals; tense and mood in finite verbs
c. Endings: case in nominals; person in finite verbs
d. Clitics: /-kO/ ‘question particle’; /-kin/, /-kAAn/ ‘even’

(12) A morphologically analyzed sample word:
/uute-impA-i-ssA-kin/ → *uusimmissakin*
‘new’ (0) ‘superlative’ (1) ‘plural’ (2) ‘inessive’ (3) ‘even’ (4)
‘even in the newest ones’

First, consider the morphological context of Assibilation. According to Karlsson (1982, p. 343), Assibilation occurs in the following morphological contexts: plural /-i/, e.g. /vuote-i-nA/ → *vuosina* ‘year-PL-ESS’, past tense /-i/, e.g. /huuta-i-vAt-kO/ → *huusivatko* ‘shout-PAST-3P.PL-QUE’, and superlative /-impA/, e.g. /uute-impA-nA/ → *uusimpana* ‘new-SUP-ESS’. In addition, Karlsson lists a number of “lexicalized” contexts, namely *vesitse* ‘by water’ (prolative), *kynsiä* ‘to claw’ (*i*-continuative), *kansittaa* ‘to furnish with a cover’ (*ittA*-causative), *tosikko* ‘a person with no sense of humor’ (*ikkO*-derivative), *vesistö* ‘lake and river system’ (*istO*-derivative), and *virsikäs* ‘verseful’ (*ikAs*-noun). Finally, Karlsson mentions the productive derivational suffix /-inen/ which derives adjectives from nouns and triggers Assibilation optionally, e.g. *vesinen* ~ *vetinen* ‘watery’. Crucially, all these suffixes are either signatures (plural, past tense) or can be classified as derivational (all the rest). We therefore conclude that derivational suffixes and signatures belong to the stem level. In addition, Karlsson (1982, p. 343) notes that “Assibilation occurs in largely the same morphological contexts as vowel changes”, confirming the prediction that both Assibilation and Vowel Deletion belong to the stem-level phonology.

Not all /i/-initial stem-level suffixes trigger Assibilation. Examples include the derivational suffix /-ime/ ‘instrument’, e.g. /lentä-ime-n/ → *lentimen* / **lensimen* ‘fly-instrument-GEN’, and the signature /-isi/ ‘conditional’, e.g. /tunte-isi/ → *tuntisi* / **tunsisi*

[2] The Finnish terms are are *kanta* ‘root’, *johdin* ‘derivational suffix’, *tunnus* ‘signature’, *pääte* ‘ending’, and *liite* ‘clitic’.

‘feel-COND’. Also note that the derivational suffix /-inen/ only triggers Assibilation optionally, e.g. /vete-inen/ → *vesinen* ~ *vetinen* ‘watery’. This is not surprising: different stem-level suffixes often participate in slightly different phonological subregularities; see Anttila 2002a for Finnish and Raffelsiefen 1999 and Zamma 2005 for parallel cases in English. Crucially, if Assibilation does occur, it occurs at the stem level.

Next, consider the morphological context of Degemination. Karlsson's (1982, p. 325) statement is best cited verbatim:

> Normal Grade Alternation [which includes Degemination, AA] is triggered mainly by ***case suffixes*** in nominals and ***person suffixes*** in verbs. Derivational suffixes that meet the relevant structural requirements behave similarly, e.g. *heikko* ‘weak’ : *heikottaa* ‘to feel weak’, *tarkka* ‘accurate’ : *tarkempi* ‘more accurate’, *purkki* ‘can’ : *purkisto* ‘set of cans’, *keikku* ‘swing’ : *keikahtaa* ‘topple’. But suffixes that are morphotactically to the right of case and person endings do not trigger Grade Alternation, even when they close the syllable. This applies above all to possessive suffixes, e.g. *takki* ‘coat’ : *takkinsa* ‘coat-3P.PX’. Clitics never trigger Grade Alternation. (Karlsson 1982, p. 325) [Translation mine, AA]

The principal triggers of Degemination are thus case and person suffixes. Both belong to the class of endings. We therefore conclude that endings belong to the word level. The fact that Degemination can also apply to material located to the left of endings is expected, given the assumption that the output of the stem level is the input to the word level.

Finally, Apocope contrasts with both Assibilation and Degemination in applying across the board. As shown in (13), Apocope applies at least to the past tense /-i/, the conditional /-isi/, the possessive suffix /-si/, the translative case suffix /-ksi/, and even to plain roots like /yksi/ (Karlsson 1982, p. 147, Rapola 1965, pp. 328-30, Rapola 1966, p. 493). This is consistent with the prediction that Apocope is a postlexical process.

(13)
a. makas-i ~ makas ‘lie-PAST’
b. olta-isi ~ oltais ‘be-COND’
c. talo-si ~ talos ‘house-2P.PX’
d. puhtaa-ksi ~ puhtaaks ‘clean-TRA’
e. yksi ~ yks ‘one’

Independent evidence for placing Assibilation and Apocope at different levels comes from a constraint that bans adjacent fricatives, a special case of the Obligatory Contour Principle (OCP) which bans adjacent identical elements (Goldsmith 1976, Leben 1973). This constraint blocks Assibilation in the sequences /sti/ and /hti/. For parallel effects in other languages, see e.g. Gussenhoven and Jacobs 1998, p. 47 (English), and Hall 2004, pp. 1056-60 (German).

(14) The OCP blocks Assibilation
a. /hiihtä-i/ → hiihti (*hiihsi) ‘ski-PAST’
b. /varasta-i/ → varasti (*varassi) ‘steal-PAST’

This pattern is exceptionless, not only in all dialects of Finnish, but in all Finnic languages (Posti 1953, pp. 50-51). Laalo briefly mentions it at the beginning of his monograph before moving on to environments where there is dialectal variation to report. Strikingly, Apocope is immune to this constraint and freely creates sequences of adjacent fricatives irrespective of morphosyntactic environment. Examples of adjacent fricatives are shown in (15).

(15) The OCP does not block Apocope

a. /imeltä-i-vät/ → imelsvät 'sweeten-PAST-3P.PL' (Jämsä)

b. /piirtä-i-hän/ → piirshän 'draw-PAST-CLIT' (Siilinjärvi)

c. /kuumenta-i si-tä/ → kuumens sitä 'heat-PAST it-PAR' (Ruokolahti)

These examples show that sequences of adjacent fricatives created by Apocope are found across stem-affix boundaries ((15a)), across word-clitic boundaries ((15b)), and across words ((15c)). Adjacent fricatives can also result from combining words in syntax, e.g. *miäs sitä hointi* 'a man was caring for it [a horse]' (Parkano). These facts suggest that the OCP is active in stem phonology, but inactive in postlexical phonology.

The fundamental difference between Assibilation and Apocope has not gone unnoticed in earlier work. In his comprehensive description of Finnish phonology and morphology, Karlsson (1982) places Vowel Deletion, Assibilation, and Degemination under morphophonological rules "that operate in partly phonological, partly morphological contexts" (p. 312), but Apocope under phonological rules, more specifically articulatory reduction rules "that largely correspond to the allophonic rules of structuralist phonology" (p. 20). In terms of the theory adopted here, this is the distinction between lexical and postlexical phonology.

In sum, the evidence suggests that Assibilation and Apocope are located at different morphological levels. This immediately explains all the opacity facts discussed in this section. What remains to be shown is how the phonology itself works. This topic will be addressed next.

3. STEM-LEVEL PHONOLOGY: ASSIBILATION

3.1. *Verbs*

The central phonological generalization about Assibilation is that the process is sensitive to stem length (Laalo 1988, pp. 10-11). We start by illustrating the length effect from Standard Finnish verbs (Anttila 2003). The descriptive generalizations are stated in (16); examples are given in (17).

(16) In Standard Finnish verbs, Assibilation is

a. blocked after a monomoraic first syllable (μ.*ti*)
b. variable and/or lexically conditioned after a bimoraic first syllable (μμ.*ti*)
c. obligatory after a trimoraic first syllable (μμμ.*ti*)
d. obligatory after two or more syllables (σ.σ.*ti*)

(17)

a.	/vetä-i/	ve.ti		*ve.si	'pull- PAST'
b.	/hoita-i/	hoi.ti		*hoi.si	'take.care- PAST'
	/vuota-i/	vuo.ti	~	vuo.si	'seep- PAST'
	/pyytä-i/	*pyy.ti		pyy.si	'ask-PAST'
c.	/kaarta-i/	*kaar.ti		kaar.si	'veer- PAST'
d.	/tilat-i/	*ti.la.ti		ti.la.si	'order-PAST'

In regional dialects, the length generalization emerges as a typological implication: if Assibilation applies to a stem in a given dialect, it applies to any longer stem in the same dialect. As shown in (18), this generalization holds in 24 dialects out of 25 in Laalo's corpus. To minimize the possible interference of other phonological and morphological factors, we have limited our attention to 3rd person /a, ä/-final verbs only. This subcorpus contains 1,860 word forms representing 194 distinct lexemes. Laalo's corpus also contains three /e/-final verbs: /pote-/ 'to be ill', /lähte-/ 'leave', and /tunte-/ 'feel'. These verbs are excluded for the moment, but will be brought into the discussion later. We have also excluded two /a, ä/-final verbs where the stem-initial consonant is *t*: /tietä-/ 'know' and /taita-/ 'be.able.to-PAST'. These verbs undergo virtually categorical Assibilation (*tiesi/*tieti*, *taisi/*taiti*) due to an independent dissimilatory constraint that prohibits identical onsets in adjacent syllables (Laalo 1988, pp. 14-15, 170-176), especially if the onset is *t.*

(18) The typological length effect in Laalo's corpus

	μ	μμ	μμμ, σσ	ATTESTED	SAMPLE DIALECT
a.	ti --	ti~si	ti~si	18 dialects	Southeast
b.	ti --	ti~si	-- si	3 dialects	Vermlanti
c.	ti --	ti --	ti~si	1 dialect	Northern Ostrobothnia
d.	ti --	ti --	ti --	1 dialect	Western Savo
e.	ti --	-- si	-- si	1 dialect	Ingria
f.	ti --	ti~si	ti --	1 dialect	Southern Ostrobothnia [*si*, N=1]

The only counterexample to the length generalization is the Southern Ostrobothnia dialect where the *si*-form appears in bimoraic stems—exactly once in Laalo's corpus—but never in longer stems. The problematic token is *pyysi* 'ask-PAST', recorded from the parish of Isojoki.[3]

[3] The recently launched *Electronic Morphology Archives for Finnish Dialects* turns up a number of additional counterexamples from Isojoki: *löysi* 'find-PAST', *huusimme* 'shout-PAST-1P.PL', *hoisin* 'care-PAST-1P.SG'. More interestingly, we also find Assibilation in one trimoraic stem: *kiersi* 'turn-PAST'. This suggests that the Isojoki dialect is in fact quite well-behaved and belongs to type (18a). The special status

The length effect also emerges quantitatively: the longer the stem, the more common Assibilation. As shown in (19), this generalization holds in 20 out of the 25 dialects in Laalo's corpus. In five dialects, we find what appears to be the opposite bias, but in each case the difference is too small to be statistically significant. Some caution in interpreting the numbers is warranted because Laalo's corpus is not a collection of naturally occurring spoken texts, but a dialectological database compiled by hand from archival material only some of which is running text (Laalo 1988, p. 50). Nevertheless, Laalo clearly intended his data to be statistically representative: "I have condensed the data in such a way that the examples reflect proportional frequencies" (Laalo 1988, p. 54). Given this, it seems reasonable to assume that the quantitative patterns in his corpus are genuine. Data from other sources also support the length generalization, as we will see in a moment.[4]

of the Isojoki dialect may have to do with its geographical location: Isojoki is the southernmost parish of Southern Ostrobothnia, bordering Pori Region (Merikarvia, Siikainen) and Upper Satakunta (Honkajoki) where Assibilation is common.

[4] In the aggregate data, the Assibilation rate is 47.1% for bimoraic stems and 52.5% for longer stems which comes out statistically significant (p = 0.0345, Fisher's exact test). The corpus is too small for demonstrating significance for each individual dialect, but the five dialects where we do reach significance (Central Häme, South-Eastern Häme, Southwest-Northern, Northern Finland, Päijät-Häme) the pairs go in the right direction. In the five dialects where the pairs go in the wrong direction the distribution is not significant: Southeast (p = 0.8888), Central Savo (p = 0.2784), Central Ostrobothnia (p = 0.4089), Southern Ostrobothnia (p = 0.3571), Pori region (p = 0.6486).

(19) The quantitative length effect in individual dialects. N = number of tokens.

REGIONAL DIALECT	OBSERVED ASSIBILATION %			DIFFERENCE%	N
	μ	μμ	μμμ/σσ		
Southeast	0	64.0	62.6	−1.4	273
Central Savo	0	39.1	31.3	−7.8	230
Central Ostrobothnia	0	48.3	40.8	−7.5	161
Central Häme	0	42.4	73.3	+30.9	134
South-Eastern Häme	0	58.0	86.8	+28.8	126
Central Finland	0	26.3	47.0	+20.7	85
Northern Carelia	0	9.1	10.9	+1.8	77
Kainuu	0	4.2	5.8	+1.6	76
Southwest-Northern	0	76.7	97.7	+21.0	73
Upper Satakunta	0	43.5	59.2	+15.7	72
Southern Ostrobothnia	0	4.0	0.0	−4.0	70
Northern Finland	0	28.6	72.9	+44.3	62
Western Uusimaa	0	72.0	75.9	+3.9	54
Northern Ostrobothnia	0	0.0	5.0	+5.0	52
Pori region	0	85.7	66.7	−19.0	37
Western Savo	0	0.0	0.0	0	36
Lower Satakunta	0	30.0	48.0	+18.0	35
Vermlanti	0	89.5	100.0	+10.5	33
Turun ylämaa	0	44.4	69.6	+25.2	32
Northwest	0	36.4	72.2	+35.8	29
Päijät-Häme	0	50.0	96.0	+46.0	29
Southwest-Southern	0	88.9	100.0	+11.1	25
Ingria	0	100.0	100.0	0	24
Southern Häme	0	60.0	77.8	+17.8	23
Somero and Somerniemi	0	60.0	100.0	+40.0	12

What is the nature of the variation within each regional dialect? An anonymous reviewer suggests various possible scenarios. For example, each variable dialect may consist of invariant speakers with different idiolects. Each speaker may control a number of different sociolects, with the result that the total output of a single individual would be an amalgamation of different grammars. Finally, closer inquiry into individual variation may reveal free variation unconditioned by contextual factors, at least to the extent that we can currently understand or study. Unfortunately, Laalo's corpus is not nearly large enough to decide among these competing scenarios and it is entirely possible that all of them are simultaneously true. Given the absence of sufficient data on individual speakers, we must leave the question open.

Throughout the preceding discussion, we have been assuming that stem length is measured in terms of metrical units (moras, syllables). Another obvious possibility is that stem length is measured in terms of segments. This is a plausible alternative, especially if we assume that the length effect is motivated by ease of word recognition (Laalo 1988, pp. 10-11): the more nonalternating segments there are in a word form, the easier it is to identify the lexeme. Thus, Assibilation in long stems would be less disruptive of word recognition than Assibilation in short stems. The segmental hypothesis will be evaluated in section 3.5. after we have seen the whole spectrum of data.

3.2. *The analysis*

The analysis we propose is metrical. Finnish word stress is assigned by laying down left-headed, left-aligned feet, with main stress on the leftmost foot. Adjacent syllables are never stressed and there are no degenerate feet. For the details of Finnish word stress, see e.g. Carlson 1978, Elenbaas and Kager 1999, Hanson and Kiparsky 1996, Kiparsky 2003, and Sadeniemi 1949. Our proposal is that Assibilation applies to EXTRAMETRICAL coronal stops. Examples are given in (20).

(20)	a. (vé.ti)	'pull-PAST'	syllabic trochee
	b. (húu)si ~ (húu.ti)	'shout-PAST'	moraic trochee~syllabic trochee
	c. (pá.ran)si ~ (pá.ran.ti)	'heal-PAST'	syllabic trochee~dactyl

In (20a), the first syllable is monomoraic and consequently /t/ falls within the foot. There are no other metrical parses available. In particular, *(*vé*)*si* is not an option because degenerate feet are disallowed. The situation is different in (20b). Since the first syllable is bimoraic, the speaker can choose between two types of binary feet: a moraic trochee with Assibilation, i.e. (*húu*)*si*, and a syllabic trochee without Assibilation, i.e. (*húu.ti*). In (20c), the speaker has a choice between a syllabic trochee with Assibilation, i.e. (*pá.ran*)*si*, and a dactyl without Assibilation (*pá.ran.ti*). Finnish phonology allows both options, but since ternary feet are universally dispreferred, one might expect that syllabic trochees, and hence Assibilation, would be preferred. This is indeed what we find.

The claim is that Assibilation is conditioned by metrical structure and that variation in Assibilation reflects variation in metrical structure. The crucial phonological constraints are listed in (21).

(21)	IDENT$_{\varphi}$	Do not alter the features of a segment within a foot.
	PARSE-σ	Syllables belong to feet.
	*TERNARY	Prosodic constituents are at most binary.
	*TI	No *ti*-sequences. (Kim 2001)

The tableaux in (22) show the constraint violations for the four stem types represented in Laalo's corpus. Harmonically bounded candidates are indicated by "#". No rankings are intended.

(22) Assibilation: The violation patterns

(a) Monomoraic stems

/vetä-i/ ‘pull-PAST’	IDENT$_{\varphi}$	*TI	PARSE-σ	*TERNARY
a. (vé.ti)		*		
b. (vé.si)	*			

(b) Bimoraic stems

/huuta-i/ ‘break-PAST’	IDENT$_{\varphi}$	*TI	PARSE-σ	*TERNARY
a. (húu)si			*	
b. (húu.ti)		*		
c. (húu.si)	*!			
d. #(húu)ti		*	*	

(c) Trimoraic stems

/kaarta-i/ ‘veer-PAST’	IDENT$_{\varphi}$	*TI	PARSE-σ	*TERNARY
a. (káa)r.si			*	
b. (káar.ti)		*		*
c. (káar.si)	*			*
d. #(káar)si			*	*
e. #(káar)ti		*	*	*
f. #(káa)rti		*	*	

(d) Disyllabic stems

/paranta-i/ ‘improve-PAST’	IDENT$_{\varphi}$	*TI	PARSE-σ	*TERNARY
a. (pá.ran)si			*	
b. (pá.ran.ti)		*		*
c. (pá.ran.si)	*			*
d. #(pá.ran)ti		*	*	

A note on trimoraic syllables is due. Following Bye’s (1997) analysis of Estonian, we assume that apparent trimoraic syllables are in fact sesquisyllables. This means that the /r/ in (*káa*)*r.si* is not included in the initial syllable, but is dominated by a freestanding mora.[5] We further assume that if /r/ is included in the first foot, as in (*kaar.ti*), (*kaar.si*), (*kaar*)*ti*, and (*kaar*)*si*, it must also be syllabified. Each of these four candidates would thus contain a trimoraic syllable, in violation of *TERNARITY. This analysis correctly predicts that trimoraic and disyllabic stems are metrically equivalent for the purposes of Assibilation.

[5] An alternative would be to assume that /r/ forms a minor syllable of its own, i.e. (*káa*).*r.si,* or perhaps an empty-headed syllable, as suggested independently by Glyne Piggott (p.c.) and Marc van Oostendorp (p.c.). The immediate problem with such an analysis is that it conflicts with native speaker intuitions about syllable count and the distribution of stress.

Next, we must address the issue of variation. Here we adopt a theory of variation that makes very few assumptions: the Multiple Grammars Theory (Anttila 2002b, to appear, Kiparsky 1993b, Kroch 1989). The central assumption is stated in (23):

(23) The Multiple Grammars Theory: Variation results from multiple invariant grammars within or across individuals.

The Multiple Grammars Theory is neutral with respect to variation within speakers and variation across speakers. In the context of Optimality Theory, the theory simply states that any combination of strict rankings is a possible grammar. This is far from saying that anything goes; the theory is constrained by the factorial typology. By ranking the 4 constraints in all the 4! = 24 possible ways, we only derive the six patterns in (24), computed by OTSoft (Hayes, Tesar, and Zuraw 2003). The gray cells indicate Assibilation.[6]

(24) The factorial typology

Verbs (regional dialects)

	1	2	3	4	5	6
μ	(ve.ti)	(ve.ti)	(ve.ti)	**(ve.si)**	**(ve.si)**	**(ve.si)**
μμ	(huu.ti)	(huu.ti)	**(huu)si**	**(huu.si)**	**(huu.si)**	**(huu)si**
μμμ	(kaar.ti)	**(kaa)r.si**	**(kaa)r.si**	**(kaar.si)**	**(kaa)r.si**	**(kaa)r.si**
σσ	(pa.ran.ti)	**(pa.ran)si**	**(pa.ran)si**	**(pa.ran.si)**	**(pa.ran)si**	**(pa.ran)si**

Under the Multiple Grammars Theory, each regional dialect can be represented as some combination of patterns 1, 2, and 3: Assibilation is possible everywhere except in monomoraic stems. By considering all the possible combinations of these three patterns, we obtain six possible dialect types, three of which are invariant, three variable. We will address patterns 4, 5, and 6 shortly.

(25) The predicted dialect typology. N = number of attested dialects.

	μ	μμ	μμμ/σσ	N	SAMPLE DIALECT
a	*ti*	*ti*	*ti*	1	Western Savo
b.	*ti*	*ti*	*si*	--	--
c.	*ti*	*si*	*si*	1	Ingria
d.	*ti*	*ti*	*ti~si*	1	N. Ostrobothnia
e.	*ti*	*ti~si*	*ti~si*	18	South-East
f.	*ti*	*ti~si*	*si*	3	Vermlanti

[6] One may wonder what happens when the verb gets longer through person agreement suffixation, e.g. /paranta-i-vat/ 'heal-PAST-3P.PL'. The stem-level grammar correctly predicts variation, e.g. (*pa.ran*)*si.vat*~(*pa.ran*)(*ti.vat*). However, person agreement suffixes only come in at the word level, witness their ability to trigger Degemination: /otta-i-n/ 'take-PAST-1P.SG' → *ot.tin* → *o.tin*. For this reason, they are irrelevant here.

The resulting typology covers 24 of the 25 dialects in Laalo's corpus. One dialect is predicted, but not observed: this is dialect (25b). One dialect is observed but not predicted: this is the Southern Ostrobothnia dialect (but see footnote 3).

More generally, the factorial typology entails the following prediction:

(26) The typological prediction: For all dialects, Assibilation in a shorter stem implies Assibilation in a longer stem where stem length is measured in metrical units: $\mu < \mu\mu < \mu\mu\mu, \sigma\sigma$.

The Multiple Grammars Theory straightforwardly generalizes to the quantitative length effect illustrated in (19). Here we assume the following quantitative interpretation of multiple grammars:

(27) A quantitative interpretation of Multiple Grammars (Anttila 1997): The probability of an output is proportional to the number of grammars predicting this output.

The key observation is that the factorial typology is asymmetrical. For every individual pattern in (24), Assibilation in a shorter stem implies Assibilation in a longer stem, but not vice versa. Since the asymmetry holds for all individual patterns, it will also hold for any combination of patterns. This implies that the Multiple Grammars Theory will preserve the asymmetry quantitatively. The following prediction is thus guaranteed:

(28) The quantitative prediction: When stem length increases, Assibilation rate will either increase or remain the same, but never decrease.

The typological prediction and the quantitative prediction follow directly from the factorial typology and are therefore independent of rankings. This means that they have very general validity: they follow under the Multiple Grammars Theory, even if we allow multiple copies of the same ranking; they follow under Partially Ordered Grammars (Anttila 1997, Anttila and Cho 1998) which is a special case of the Multiple Grammars Theory with additional constraints on the ranking relation; and they follow under Stochastic Optimality Theory (Boersma 1998, Boersma and Hayes 2001) where constraints have real-number weights. This is because in all these theories the factorial typology is the same.

3.3. *Nouns*

The length effect only emerges in verbs; it is strikingly absent in nouns. Instead, nouns exhibit two kinds of categorical behavior: Assibilation applies in a closed class of /e/-final nouns and is blocked everywhere else. Examples of /e/-final nouns are given in (29a); examples of regular nouns are given in (29b).

(29)	a. /vete-i-nä/	*vetinä	vesinä	‘water-PL-ESS’
	/vuote-i-na/	*vuotina	vuosina	‘year-PL-ESS’
	/tuhante-i-na/	*tuhantina	tuhansina	‘thousand-PL-ESS’
	b. /sota-i-na/	sotina	*sosina	‘war-PL-ESS’
	/vuota-i-na/	vuotina	*vuosina	‘skin-PL-ESS’
	/suunta-i-na/	suuntina	*suunsina	‘drection-PL-ESS’
	/egypti-i-en/	Egyptien	*Egypsien	‘Egypt-PL-GEN’
	/varastoi-nti-i-en/	varastointien	*varastoinsien	‘store-ing-PL-GEN’

These patterns would be easy enough to describe by marking the /e/-final nouns with the feature [+Assibilation] and the regular nouns with the feature [−Assibilation]. This would amount to recognizing two classes of lexical exceptions to the length generalization. While the exception feature analysis would straightforwardly account for Assibilation itself, it would not relate the pattern to anything else in the language. What one would like to understand is how exactly nouns are different from verbs, and from each other, and more ambitiously, why there should be such differences. Ideally, the explanation should generalize beyond Assibilation.

The following hypothesis now suggests itself. If Assibilation is metrically conditioned, as we have argued, then perhaps the differences among word classes reflect differences in metrical structures. This idea can be implemented by associating different word classes with different constraint rankings, or COPHONOLOGIES (see e.g. Anttila 2002a, Inkelas 1998, Orgun 1996, Raffelsiefen 1999, Zamma 2005), as in (30).

(30) The factorial typology divided among word classes[7]

Regular nouns → 1; Verbs (regional dialects) → 1, 2, 3; /e/-nouns → 4, 5, 6

	1	2	3	4	5	6
μ	(so.ti)	(ve.ti)	(ve.ti)	**(ve.si)**	**(ve.si)**	**(ve.si)**
μμ	(vuo.ti)	(huu.ti)	**(huu)si**	**(vuo.si)**	**(vuo.si)**	**(vuo)si**
μμμ	(suun.ti)	**(kaa)r.si**	**(kaa)r.si**	**--**	**--**	**--**
σσ	(e.gyp.ti)	**(pa.ran)si**	**(pa.ran)si**	**(tu.han.si)**	**(tu.han)si**	**(tu.han)si**

How is the metrical cophonology analysis better than the exception feature analysis? At first blush, the two may seem empirically indistinguishable: the exception feature analysis distinguishes the word classes by associating regular nouns with the feature [−Assibilation], /e/-final nouns with the feature [+Assibilation], and by leaving the verbs either unspecified (variable verbs) or by associating them individually with [+Assibilation] or [−Assibilation] as the facts may require. The metrical cophonology

[7] The noun forms are plural stems. One may wonder what happens when the nouns get longer through case suffixation, e.g. /tuhante-i-na/ ‘thousand-PL-ESS’, /varastoi-nti-i-en/ ‘store-ing-PL-GEN’, etc. The stem-level grammar makes the correct predictions, e.g. (*tu.han*)*si.na* and (*va.ras*)(*toin.ti.en*). However, case suffixes only come in at the word level, witness their ability to trigger Degemination: /lakko-n/ ‘strike-GEN’ → *lakon*. For this reason, they are irrelevant here.

analysis accomplishes the same by associating different word classes with different metrical grammars. Both analyses require lexical marking. How are they different empirically?

The metrical cophonology analysis has one major advantage: it makes predictions beyond Assibilation. The factorial typology in (30) reveals the following predictions:

(31) Predicted metrical differences among word classes:

a. Regular nouns are exhaustively footed.
b. Verbs allow optional extrametricality, except in monomoraic stems.
c. /e/-nouns allow optional extrametricality, except in monomoraic stems.

The difference between the analyses is now clear: the metrical cophonology analysis makes specific predictions about the metrical structures associated with particular word classes. These predictions can be independently confirmed or disconfirmed by using other metrically conditioned processes as diagnostics. The exception feature analysis makes no such predictions. As we will see in a moment, the metrical predictions are confirmed by evidence from Apocope.

The complete factorial typology which includes all six patterns and their possible combinations is shown in (32). All in all, ten distinct patterns are predicted.

(32) The predicted dialect typology

	μ	μμ	μμμ/σσ	SAMPLE DIALECT
a.	*ti*	*ti*	*ti*	verbs, W. Savo; regular nouns, all dialects
b.	*ti*	*ti*	*si*	--
c.	*ti*	*si*	*si*	verbs, Ingria
d.	*si*	*si*	*si*	/e/-final nouns, all dialects
e.	*ti*	*ti*	*ti~si*	verbs, Northern Ostrobothnia
f.	*ti*	*ti~si*	*ti~si*	verbs, South-East
g.	*ti*	*ti~si*	*si*	verbs, Vermlanti
h.	*ti~si*	*ti~si*	*ti~si*	/e/-final verbs, Central Savo
i.	*ti~si*	*ti~si*	*si*	--
j.	*ti~si*	*si*	*si*	--

/e/-final nouns instantiate pattern (32d). Marginal evidence for patterns (32h-j) comes from /e/-final verbs which were briefly mentioned earlier: /pote-/ 'to be ill', /lähte-/ 'leave', and /tunte-/ 'feel'. Of these verbs, /pote-/ is monomoraic. This verb exhibits variable Assibilation *poti ~ posi* in four eastern dialects (South-East, Central Savo, South-Eastern Häme, Northern Carelia) instantiating one of the patterns (32h-j) although it is impossible to tell which one. We have arbitrarily placed the /e/-final verbs under (32h). In all four dialects, Assibilation is also found in bimoraic /e/-final verbs, as predicted.

We have argued that different parts of the lexicon are associated with different grammars. This still leaves us with a deeper question: why should the lexicon be partitioned in this way? In particular, why should there be two classes of nouns with different metrical structures? In this case, the rationale appears to be historical: /e/-final nouns belong to the oldest stratum of Finnish vocabulary whereas regular nouns are typically younger (Itkonen 1984; for similar lexical stratification in Japanese, see Itô and Mester 1995, 1999). According to Itkonen (1984, p. 70), /e/-stems formed the great majority of disyllabic stems during the Uralic and Finno-Ugrian period (Itkonen lists 102 stems) and were approximately as many as /a/-stems and /ä/-stems taken together (Itkonen lists 83 stems), whereas in modern Finnish the number of /e/-stems has shrunk to approximately one eighth of the number of /a/ and /ä/-stems taken together due to a process that may have started already in the Finno-Permian period. One might speculate that /e/-final nouns represent an earlier metrical system that still persists in the old vocabulary, whereas the more recent vocabulary reflects a historical change towards a more permissive foot structure which favors exhaustive parsing even at the expense of ternary feet. The association of different parts of the lexicon with different phonological grammars would thus be synchronically arbitrary, but admit a historical explanation.

3.4 *Lexical frequency effects*

We have provided evidence that Assibilation is metrically conditioned. One important question remains: what determines the Assibilation behavior of an individual verb in environments where the metrical analysis predicts variation?

Paunonen (1974) studied the Assibilation behavior of 41 verbs by means of an elicitation test. 36 native speakers were asked to grade the *ti* and *si* forms of each verb on the scale 3 = "very good", 2 = "not very good, but possible", 1 = "impossible". The list contained 2 monomoraic, 23 bimoraic, and 16 trimoraic verbs in both *ti*- and *si*-forms, presented in random order, 82 items in all. By subtracting the average *ti*-score of each verb from the average *si*-score of the same verb, Paunonen derived an index between +2.0…−2.0 that reflects the relative naturalness of Assibilation for each verb: the higher the index, the better the *si*-form. Three examples are shown in (33). The results for all 41 verbs are shown in (34)-(35).

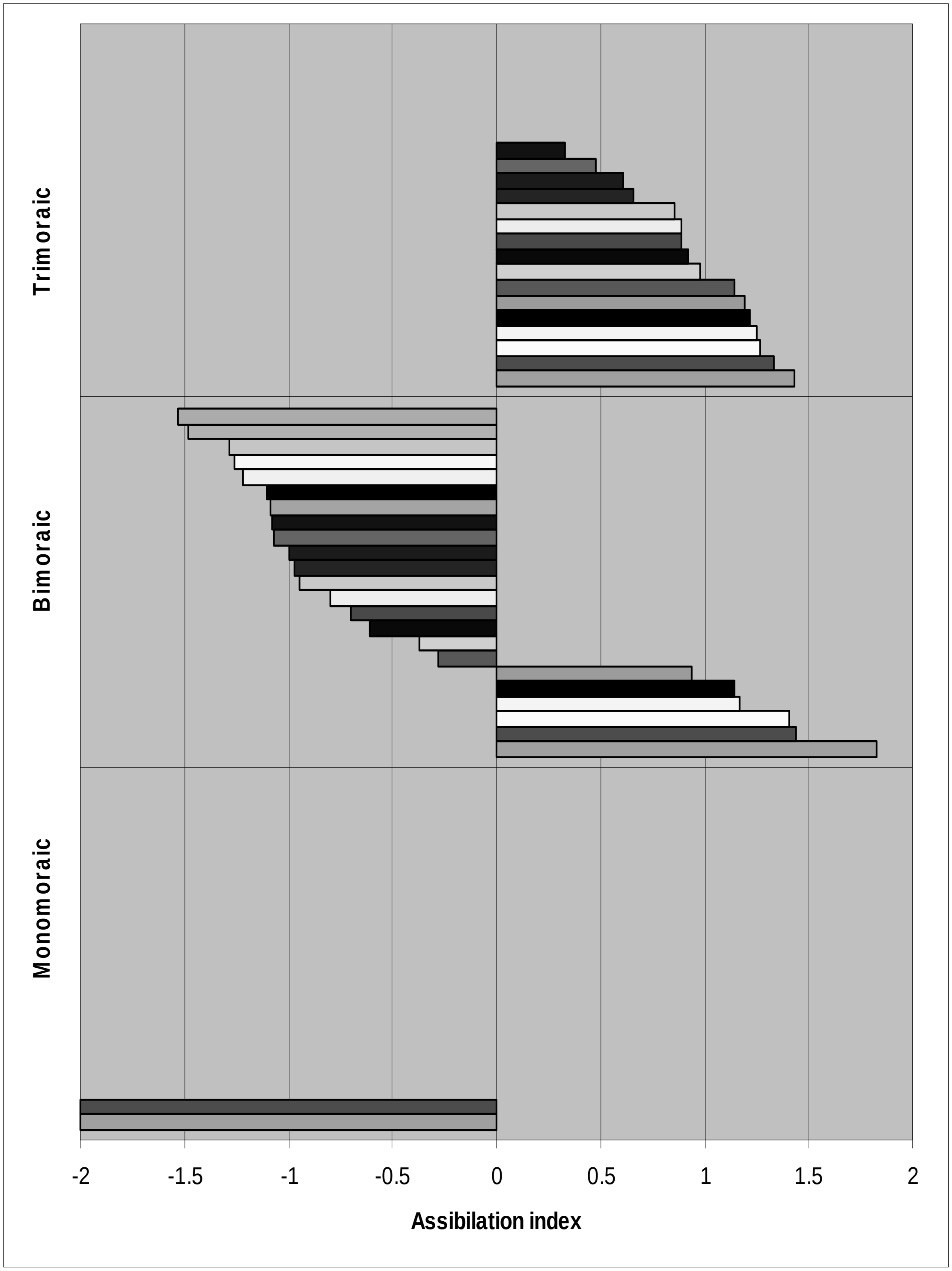
(33) Assibilation index (Paunonen 1974)

(a)	*tiesi*	3.00	*tieti*	1.17	*si* − *ti* = +1.83	'know-PAST'
(b)	*kiisi*	2.26	*kiiti*	2.63	*si* − *ti* = −0.37	'speed-PAST'
(c)	*siesi*	1.46	*sieti*	2.94	*si* − *ti* = −1.48	'tolerate-PAST'

(34) Assibilation by verb, average values across 36 subjects (Paunonen 1974). Monomoraic stems are shown in italics, bimoraic stems in boldface, trimoraic stems in regular font.

	Assibilation index (+)		Assibilation index (−)
tietää ‘know’	+1.83	**soutaa ‘row’**	−0.28
löytää ‘find’	+1.44	**kiitää ‘speed’**	−0.37
rientää ‘hurry’	+1.43	**hyytää ‘freeze’**	−0.61
huutaa ‘shout’	+1.41	**vuotaa ‘seep’**	−0.70
piirtää ‘draw’	+1.33	**häätää ‘evict’**	−0.80
kääntää ‘turn’	+1.27	**yltää ‘reach’**	−0.95
työntää ‘push’	+1.25	**liitää ‘glide’**	−0.97
myöntää ‘admit’	+1.22	**säätää ‘decree’**	−1.00
ääntää ‘utter’	+1.19	**jäätää ‘freeze’**	−1.07
pyytää ‘ask’	+1.17	**syytää ‘spew’**	−1.08
kieltää ‘deny’	+1.14	**jäytää ‘afflict’**	−1.09
lentää ‘fly’	+1.14	**hoitaa ‘care’**	−1.10
siirtää ‘move’	+0.98	**noutaa ‘fetch’**	−1.22
murtaa ‘break’	+0.94	**sortaa ‘oppress’**	−1.26
kiertää ‘veer’	+0.92	**joutaa ‘have.time’**	−1.28
kiiltää ‘glisten’	+0.89	**sietää ‘tolerate’**	−1.48
hiertää ‘rub’	+0.89	**kyntää ‘plow’**	−1.53
puoltaa ‘defend’	+0.86	*pitää* ‘hold’	−2.00
puurtaa ‘toil’	+0.66	*vetää* ‘pull’	−2.00
uurtaa ‘groove’	+0.61		
siintää ‘shimmer’	+0.48		
suoltaa ‘spew’	+0.33		

(35) Verb length and Assibilation index (data from Paunonen 1974)



The three groups of verbs emerge fairly clearly in Paunonen's study. Both monomoraic stems are at the negative extreme, all trimoraic stems are positive, and bimoraic stems vary widely depending on the lexeme: 6 are positive (*tietää, löytää, huutaa, pyytää, lentää, murtaa*), the rest are negative. The bimoraic verb with the highest Assibilation index (+1.83) is *tietää* 'know' where onset dissimilation plays a role. But what explains the great differences among the remaining bimoraic verbs?

Laalo (1988, pp. 18-22) suggests that one of the relevant factors is lexical frequency: Assibilation is favored in high-frequency verbs and disfavored in low-frequency verbs. Lexical frequency effects are familiar from a number of languages, see e.g. Bybee 2001, 2002, Fidelholtz 1975, Hay 2003, Hooper 1976, Jurafsky, Bell, Gregory, and Raymond 2001, Kang 2003, Myers 2003, Myers and Guy 1997, Phillips 1984, 2001, and Pierrehumbert 2001. Since Laalo's corpus is not suitable for studying lexical frequencies, we examined Paunonen's verbs in three large newspaper corpora: the *Aamulehti 1999* corpus (AL, 16,608,843 word forms), the *Turun Sanomat 1999* corpus (TS, 11,821,904 word forms) and the *Kaleva 1998-1999* corpus (KA, 9,758,628 word forms). In our verb list, we included all the 41 verbs listed by Paunonen (1974) plus one monomoraic stem: /itä-/ 'sprout'. The search was based on the 3rd person plural past tense forms, e.g. *pitivät* 'hold-PAST-3P.PL'.[8]

The corpus results converge with Laalo's dialect data as well as Paunonen's elicitation test. First, monomoraic stems never undergo Assibilation irrespective of frequency. The numbers in (36) show the frequency of the 3rd person plural past tense forms in each corpus (AL, TS, KA). The percentages indicate Assibilation rates.

(36) Monomoraic stems: No Assibilation.

		AL%	TS%	KA%	AL	TS	KA	Total frequency
ti	pitää 'hold'	0	0	0	578	454	333	1,365
	vetää 'pull'	0	0	0	145	107	69	321
	itää 'sprout'	--	--	0	--	--	2	2

Bimoraic stems fall into three groups: only *si* (6 verbs), variation (3 verbs), and only *ti* (13 verbs). Strikingly, the six categorical *si*-verbs are the very same verbs that received a positive Assibilation index in Paunonen's elicitation test 25 years earlier, namely *löytää, tietää, huutaa, pyytää, lentää*, and *murtaa*.

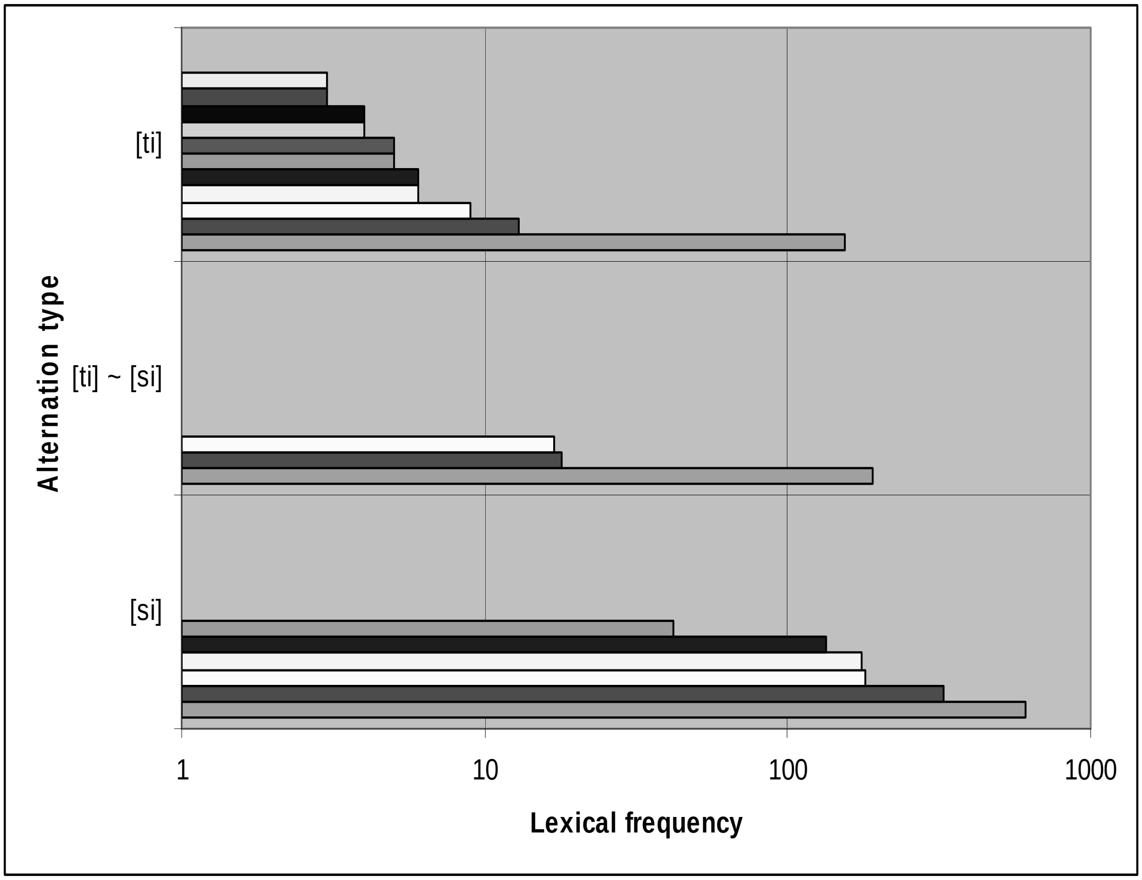
[8] We could have easily tripled the number of examples by searching for the 3rd person singular form. However, the longer plural form avoids the homonymy problem inherent in the singular, e.g. *vuosi* 'seep-PAST' or 'year', *nousi* 'fetch-PAST' or 'rise-PAST', *tiesi* 'know-PAST' or 'road-2P.PX', *kynsi* 'plow-PAST' or 'nail', *syysi* 'spew-PAST' or 'fault-2P.PX', *jousi* 'have.time-PAST' or 'spring'. There remain two potential homonymies: *nousivat* 'fetch-PAST-3P.PL' or 'rise-PAST-3P.PL' and *kynsivät* 'plow-PAST-3P.PL' or 'claw-PAST-3P.PL' which had to be weeded out manually.

(37) Bimoraic stems: Variation and lexical frequency

		AL%	TS%	KA%	AL	TS	KA	Total frequency
si	löytää ‘find’	100	100	100	257	217	139	613
	tietää ‘know’	100	100	100	133	121	75	329
	huutaa ‘shout’	100	100	100	79	42	59	180
	pyytää ‘ask’	100	100	100	81	49	45	175
	lentää ‘fly’	100	100	100	77	34	23	134
	murtaa ‘break’	100	100	100	24	8	10	42
ti ~ si	yltää ‘reach’	87.8	77.3	88.6	82	66	44	192
	vuotaa ‘seep’	14.3	20.0	50.0	7	5	6	18
	kiitää ‘speed’	50.0	50.0	55.6	2	6	9	17
ti	hoitaa ‘care’	0	0	0	48	62	44	154
	soutaa ‘row’	0	0	0	5	7	1	13
	noutaa ‘fetch’	0	0	0	3	5	1	9
	sortaa ‘oppress’	0	0	0	3	1	2	6
	sietää ‘tolerate’	0	--	0	4	--	2	6
	liitää ‘glide’	0	0	0	3	1	1	5
	joutaa ‘have.time’	0	0	0	2	1	2	5
	säätää ‘decree’	0	0	0	2	1	1	4
	syytää ‘spew’	0	0	--	3	1	--	4
	häätää ‘evict’	--	0	--	--	3	--	3
	jäytää ‘afflict’	0	0	--	2	1	--	3
	jäätää ‘freeze’	0	--	--	1	--	--	1
	kyntää ‘plow’	--	0	--	--	1	--	1
	hyytää ‘freeze’	--	--	--	--	--	--	0

The diagram in (38) visualizes the relationship between Assibilation and lexical frequency. Lexical frequencies are plotted on a logarithmic scale reflecting the common assumption that humans process frequency information logarithmically.

(38) Assibilation and lexical frequency: Bimoraic stems



The picture shows that high-frequency verbs tend to undergo categorical Assibilation. If we compare the log frequencies of *si*-verbs and the rest of the verbs, the difference between the means is highly significant ($t = 4.8967$, $df = 20$, $p < 0.0001$, two tailed). The difference between the variable verbs and *ti*-verbs is also significant ($t = 2.4550$, $df = 14$, $p = 0.0278$). This is consistent with the familiar generalization that high-frequency words undergo phonological reduction more than low-frequency words. There are two outliers: the *ti* ~ *si* class has one high-frequency verb *yltää* 'reach' and the *ti*-class has one high-frequency verb *hoitaa* 'care'.[9]

Finally, trimoraic stems exhibit virtually categorical Assibilation, irrespective of frequency. The only exception is the unassibilated form *kääntivät* 'turn-PAST-3P.PL' which occurs once in each corpus. This datum can be traced back to a single news item provided by the Finnish News Agency (STT) which was copied by all three newspapers.

[9] Why should *yltää* 'reach' and *hoitaa* 'care' be special? Both verbs assibilate less than one would expect based on their lexical frequency. In the case of *yltää*, Laalo (1988, p. 198) suggests that this may be because of the absence of an initial onset which makes the verb shorter in terms of segment count. This hypothesis will be discussed and rejected in section 3.5. In the case of *hoitaa*, one might suggest that the initial *h* exerts a dissimilatory influence on the onset of the following syllable, inhibiting Assibilation. The problem is that the verb *huutaa* 'shout' also has an initial *h*, but nevertheless assibilates categorically. These two verbs seem genuine lexical exceptions.

(39) Trimoraic stems. Nearly categorical Assibilation.

		AL%	TS%	KA%	AL	TS	KA	Total frequency
si	myöntää 'admit'	100	100	100	68	59	47	174
	kiertää 'turn'	100	100	100	57	45	31	133
	siirtää 'shift'	100	100	100	40	31	32	103
	kieltää 'deny'	100	100	100	32	18	22	72
	rientää 'hurry'	100	100	100	17	9	6	32
	työntää 'push'	100	100	100	14	10	7	31
	puoltaa 'defend'	100	100	100	10	8	7	25
	piirtää 'draw'	100	100	100	7	5	6	18
	hiertää 'rub'	100	100	100	3	3	3	9
	puurtaa 'toil'	100	100	100	1	6	1	8
	kiiltää 'glitter'	100	--	100	3	--	2	5
	suoltaa 'spew'	100	100	100	1	2	1	4
	siintää 'shimmer'	100	100	--	1	1	--	2
	ääntää 'utter'	--	--	--	--	--	--	--
	uurtaa 'groove'	--	--	--	--	--	--	--
ti~si	kääntää 'turn'	98.0	97.4	95.5	50	39	22	111

Similar patterns can be observed in other genres. The dialect of the first person narrator in Mika Waltari's novel *Sinuhe egyptiläinen* (*The Egyptian*, first published in 1945, translated into English by Naomi Walford) has no Assibilation in monomoraic verbs, variable Assibilation in bimoraic verbs, and categorical Assibilation in trimoraic verbs. Sinuhe's dialect also provides an example of free variation within an individual: the past tense of *murtaa* 'break' varies between *murti* (4 times) and *mursi* (6 times). In each case, the verb is used to denote the breaking of a concrete object, such as a gate, a door, or a rock. The fact that this variation managed to pass muster with the editors of a prestigious publishing house (Werner Söderström) shows that it is unremarkable. The lexical frequency pattern is also familiar, except that this time the verb *vuotaa* 'seep' is somewhat out of line.

(40) Assibilation rates and lexical frequencies of bimoraic 3rd person verbs in Mika Waltari's *Sinuhe egyptiläinen*, available at http://www.ling.helsinki.fi/uhlcs/

		Assibilation %	Lexical frequency
si	huutaa 'shout'	100	178
	tietää 'know'	100	151
	pyytää 'ask'	100	44
	löytää 'find'	100	21
	lentää 'fly'	100	19
ti ~ si	murtaa 'break'	60	10
ti	vuotaa 'seep'	0	35
	soutaa 'row'	0	7
	hoitaa 'care'	0	4
	noutaa 'fetch'	0	2
	sietää 'tolerate'	0	1

We conclude that verbs in Written Standard Finnish combine patterns 2-3 in the factorial typology, with variation only in bimoraic stems.

(41) Regular nouns → 1; Verbs (Written Standard) → 2, 3; /e/-nouns → 4, 5, 6

	1	2	3	4	5	6
μ	(so.ti)	(ve.ti)	(ve.ti)	**(ve.si)**	**(ve.si)**	**(ve.si)**
μμ	(vuo.ti)	(huu.ti)	**(huu)si**	**(vuo.si)**	**(vuo.si)**	**(vuo)si**
μμμ	(suun.ti)	**(kaa)r.si**	**(kaa)r.si**	**--**	**--**	**--**
σσ	(e.gyp.ti)	**(pa.ran)si**	**(pa.ran)si**	**(tu.han.si)**	**(tu.han)si**	**(tu.han)si**

Bimoraic verbs support the familiar generalization that high-frequency words undergo phonological reduction more than low-frequency words. Interestingly, the generalization fails to extend to the rest of the data. Monomoraic verbs never assibilate, irrespective of frequency, and trimoraic verbs virtually always assibilate, irrespective of frequency. This is particularly striking in verbs like *pitää* 'hold' which never assibilates despite its extremely high usage frequency (1,365 tokens in the newspaper sample) and verbs like *puurtaa* 'toil' which always assibilates despite its very low usage frequency (8 tokens in the newspaper sample). This is the opposite of what one would expect based on usage frequency alone. As a general statement about language, the generalization that high frequency leads to phonological reduction is thus clearly false.

The diagnosis is obvious: lexical frequency effects are embedded in phonological structure, which is independent of frequency and which can override frequency effects. In Finnish, lexical frequency effects only emerge in environments where phonology underdetermines metrical parsing. One possible interpretation is that individual lexical items exploit metrical variation to set up frequency-based lexical patterns. Thus, high-frequency verbs like *pyytää* 'ask' typically choose a moraic trochee with extrametricality, i.e. (*pyy*)*si* 'ask-PAST', whereas low-frequency verbs like *liitää* 'glide' typically choose a syllabic trochee with exhaustive parsing, i.e. (*liiti*) 'glide-PAST'. Since this metrical choice is only available in bimoraic verbs, it follows that lexical frequency effects are suppressed elsewhere. In sum: (i) lexical frequency effects only emerge in environments where phonology allows metrical choice; (ii) high-frequency words prefer incomplete metrical parsing, low-frequency words exhaustive parsing; (iii) incomplete parsing results in segmental reduction, exhaustive parsing results in segmental faithfulness. The relationship between lexical frequency and phonological reduction is thus indirect and mediated by metrical constraints.

Lexical effects can be described in terms of cophonologies (see e.g. Anttila 2002a). In this case, the mutual ranking of PARSE-σ and *TI is not fixed by the general phonology of verbs, but decided on a lexeme-by-lexeme basis: high-frequency verbs typically choose *TI >> PARSE-σ which results in extrametricality; low-frequency verbs typically choose PARSE-σ >> *TI which results in exhaustive parsing; and yet other verbs choose neither ranking and remain variable. The analysis raises but does not answer the question why high usage frequency should be connected with incomplete metrical parsing. Finally, the existence of outliers like *yltää* 'reach' and *hoitaa* 'care' suggests that frequency alone will not be enough to predict the behavior of individual verbs, but we have to leave open the possibility that the choice among cophonologies may be a lexeme-specific matter.[10]

3.5. *Is the length effect metrical or segmental?*

Until now, we have been assuming that Assibilation is sensitive to stem length measured in terms of metrical units (moras, syllables). Another possibility is that stem length is measured in terms of segments. In fact, this would seem to be the most plausible measure if we assume, with Laalo (1988, pp. 10-11), that the length effect is motivated by ease of word recognition: the more nonalternating segments there are in a word form, the easier it is to identify the lexeme. For example, Assibilation would fail to apply in /vetä-/ 'pull-PAST' (*veti / *vesi*) because there are only two nonalternating segments (*ve*), but Assibilation would apply optionally in /lentä-/ 'fly-PAST' (*lenti ~ lensi*) because there are three nonalternating segments (*len*). Clearly, to recognize a word form as a member of a particular lexeme we must consider its melodic properties, such as segments or features, not just its metrical length.

There are two possible versions of the segmental hypothesis. The strong version says that only segments matter and that moras and syllables are irrelevant. The weak version says that segments matter, but so do moras and syllables. The strong version predicts that if we keep the number of segments constant and vary the number of moras, we should see no difference in Assibilation. Relevant examples are given in (42).

(42) Segmentally same, metrically different

			SEGMENTS	MORAS	ASSIBILATION
a.	*vetää*	*yltää*	2	1 vs. 2	*ti* vs. variable
b.	*lentää*	*ääntää*	3	2 vs. 3	both variable

The strong version of the segmental hypothesis is falsified by the contrast between *veti* 'pull-PAST' (no Assibilation) and *ylti ~ ylsi* 'reach-PAST' (variable Assibilation). According to Laalo (1988, p. 103), monomoraic /a, ä/-final verbs never assibilate in any dialect. This is also true in all three newspaper corpora. The bimoraic *yltää* 'reach' is different: among the 11 examples in Laalo's corpus, there are two *si*-forms, both in the first person: *ylsin*(*hän*) 'reach-PAST-1P.SG-CLIT' (Northern Finland). In the newspaper corpora, not only is *ylsi* possible, but it is preferred. Since *vetää* and *yltää* are both bisegmental, the difference must depend on the number of moras. Pairs like *lentää* 'fly' and *ääntää* 'utter' are uninformative because both allow variation.[11]

Laalo (1988, p.c.) subscribes to the weak version of the segmental hypothesis. This hypothesis predicts that if we keep the number of moras constant and vary the

[10] An alternative approach to lexical effects in the midst of global gradience is offered by Zuraw (2000).

[11] As it happens, *yltää* 'reach' is the only bisegmental bimoraic verb of the relevant kind in the language. As for verbs like *lentää* 'fly' and *ääntää* 'utter', it would be interesting to compare their Assibilation rates, but there are too few verbs of the *ääntää* type for any meaningful comparison. Finally, there is no need to examine stems with 4 or more segments because such stems will always have at least 3 moras and all trimoraic and longer stems are predicted to behave alike by our metrical analysis.

number of segments, we should see differences in Assibilation. In particular, syllable onsets should count for length. Relevant examples are given in (43).

(43) Metrically same, segmentally different

			MORAS	SEGMENTS	ASSIBILATION
a.	*itää*	*vetää*	1	1 vs. 2	both *ti*
b.	*yltää*	*lentää*	2	2 vs. 3	both variable
c.	*ääntää*	*kääntää*	3	3 vs. 4	both variable

None of the pairs in (43) support the weak version of the segmental hypothesis: the number of segments never makes a difference. However, (43) still leaves open the possibility that the variable stems exhibit Assibilation at different rates depending on the number of segments. The weak segmental hypothesis predicts that Assibilation should become increasingly common with each additional segment. Examples are shown in (44).

(44) Examples of metrically equivalent stems with different numbers of segments

3 segments: *uurtaa* 'make a groove', *ääntää* 'utter'
4 segments: *kääntää* 'turn', *työntää* 'push', *piirtää* 'draw', *ylentää* 'promote'
5 segments: *rakentaa* 'build', *parantaa* 'improve', *ymmärtää* 'understand'
6 segments: *paimentaa* 'shepherd', *suurentaa* 'enlarge', *korventaa* 'scorch'
7 segments: *viännältää* 'twist onseself'

As shown in (45), there is indeed a slight increase in the rate of Assibilation with each additional segment, ranging from 50.0% in three-segment stems to 56.5% in six-segment stems, but this is not enough to reach statistical significance.

(45) Testing the segmental hypothesis

Number of segments	3	4	5	6	7	Total
ti	5	359	181	67	1	613
si	5	372	214	87	0	678
Assibilation %	50.0	50.9	54.2	56.5	0	1,291

$X^2 = 2.215$, $df = 3$, $p \leq 0.529$ (column 7 omitted)

Our metrical analysis is clearly compatible with the weak segmental hypothesis. It is entirely possible that Assibilation is metrically conditioned, but in addition there is an overlay of segmental effects that arise from ease of word recognition. However, the available evidence does not particularly support such a view. For the moment, it seems safe to conclude that Assibilation is a purely metrically conditioned alternation.

3.6. *Summary*

Based on typological and quantitative evidence, we have argued that Finnish Assibilation is a metrically conditioned alternation. In verbs, Assibilation applies to the onsets of extrametrical syllables. Assibilation is variable or lexically conditioned in verbs that admit multiple metrical parses. In nouns, there is no variation, but Assibilation is either categorical or blocked depending on the noun class. We suggested that this is because different noun classes are associated with different metrical grammars, but did not provide any independent evidence for these metrical differences. We will now turn to such evidence. The place to look is Apocope.

4. POSTLEXICAL LEVEL: APOCOPE

4.1. *Metrical structure and Apocope*

Why does Apocope happen? The standard explanation is that short vowels are deleted in metrically weak positions (see e.g. Kager 1997). If this is so, Apocope provides an independent test for our metrical analysis of Assibilation.

Since Assibilation targets extrametrical onsets, the natural hypothesis is that Apocope targets extrametrical nuclei. This predicts that Assibilation and Apocope should go together: Apocope should be favored in Assibilation contexts because both are conditioned by extrametricality. The predicted pattern is illustrated in (46).

(46) Prediction: Apocope should be favored if Assibilation has applied.

a.	/huuta-i/	→ (húu)si	→ (húu)s	expected
b.	/huuta-i/	→ (húu.ti)	→ (húut)	unexpected

This prediction is confirmed by various kinds of evidence. First, Laalo's corpus reveals a typological gap of the expected kind: there are dialects with Apocope across the board; there are dialects with Apocope only in *si*; there are dialects with no Apocope at all; but there are no dialects with Apocope only in *ti*. In other words, if Apocope applies to metrified syllables, it also applies to extrametrical syllables, but not vice versa. The pattern in regional dialects is shown in (47). Spoken Standard Finnish belongs to group (47b) (Karlsson 1982, p. 147).

(47) Typological confirmation. N = number of attested dialects.[12]

	PATTERN	DESCRIPTION	N	EXAMPLE
a.	*ti si t' s'*	Apocope across the board	6	Southwest-Northern
	-- *si t' s'*	Apocope across the board	2	Vermlanti
	-- *si* -- *s'*	Apocope across the board	1	Ingria
b.	*ti si* -- *s'*	Apocope only in *si*	4	Western Uusimaa
c.	*ti si* -- --	No Apocope	11	Somero and Somerniemi
	ti -- -- --	No Apocope	1	Western Savo

[12] The Ingria pattern where Assibilation is obligatory could be placed under either (47a) "Apocope across the board" or (47b) "Apocope only in *si*". We have arbitrarily placed it under (47a).

Second, the same generalization emerges in Laalo's corpus as a quantitative asymmetry: Apocope is common in *si*, but rare in *ti*. In two dialects (Vermlanti, Southwest-Southern) we find what appears to be the opposite bias, but in each case the difference is too small to be statistically significant.

(48) Quantitative confirmation: individual dialects. N = number of tokens.

REGIONAL DIALECT	PATTERN	APOCOPE *ti* %	APOCOPE *si* %	N
Southeast	*ti si t' s'*	31.7	73.8	273
Central Savo	*ti si t' s'*	9.2	80.5	230
Central Ostrobothnia	*ti si* -- --	--	--	161
Central Häme	*ti si* -- --	--	--	134
South-Eastern Häme	*ti si t' s'*	58.1	89.5	126
Central Finland	*ti si* -- *s'*	--	88.9	85
Northern Carelia	*ti si t' s'*	5.8	25.0	77
Kainuu	*ti si* -- --	--	--	76
Upper Satakunta	*ti si* -- --	--	--	72
Southern Ostrobothnia	*ti si* -- --	--	--	70
Southwest-Northern	*ti si t' s'*	25.0	75.4	69
Northern Finland	*ti si* -- *s'*	--	2.6	62
Western Uusimaa	*ti si* -- *s'*	--	5.0	54
Northern Ostrobothnia	*ti si* -- --	--	--	52
Pori region	*ti si* -- *s'*	--	3.8	37
Western Savo	*ti* -- -- --	--	--	36
Lower Satakunta	*ti si* -- --	--	--	35
Vermlanti	-- *si t' s'*	100.0 [N = 2]	96.8	33
Turun ylämaa	*ti si* -- --	--	--	32
Northwest	*ti si* -- --	--	--	29
Päijät-Häme	*ti si t' s'*	33.3	96.2	29
Southwest-Southern	-- *si t' s'*	100.0 [N= 1]	54.2	25
Ingria	-- *si* -- *s'*	--	70.8	24
Southern Häme	*ti si* -- --	--	--	23
Somero and Somerniemi	*ti si* -- --	--	--	12

The Apocope pattern lends further support to our analysis of Assibilation: the predicted metrical asymmetry emerges in both alternations, typologically as well as quantitatively.

We now turn to Apocope in nouns. Recall that nouns come in two metrical types: regular nouns and monomoraic /e/-final nouns are exhaustively footed, whereas bimoraic /e/-final nouns have an optional extrametrical syllable. The examples in (49) are nominative singular forms where the vowel *i* is not followed by any suffixes.[13]

[13] Nominative singulars of /e/-final nouns are derived by *e*-raising (/vete/ → *veti* 'water') which feeds Assibilation (*veti* → *vesi*) (Kiparsky 1973). This implies that *e*-raising is a stem-level process. However, *e*-raising must be optional since *vete* surfaces with case suffixes other than the nominative, e.g. *vete-nä* 'water-ESS'. Here we assume that *e*-raising is optional and case suffixes select the appropriate stem allomorphs at the word level. A more principled way of ruling out nominative singulars like *vete and

(49) Regular nouns /e/-final nouns

	1	2	3	4	5	6
μ	(la.si)			**(ve.si)**	**(ve.si)**	**(ve.si)**
μμ	(kek.si)			**(kak.si)**	**(kak.si)**	**(kak)si**
σσ	(a.lek.si)			**--**	**--**	**--**

In Stratal Optimality Theory, stem-level metrical structures are visible to subsequent levels, including postlexical phonology. This predicts that Apocope should be favored in bimoraic /e/-final nouns, as shown in (50).

(50) Predictions:

a. (la.si) → (las) Apocope disfavored 'glass'
   (kek.si) → (keks) Apocope disfavored 'cookie'
   (a.lek.si) → (a.leks) Apocope disfavored 'a major shopping street in Helsinki'

b. (ve.si) → (ves) Apocope disfavored 'water'

c. (kak)si → (kak)s Apocope favored 'two'

Is this prediction correct? Since Laalo's corpus only contains verbs, we must seek confirmation elsewhere. Written Standard Finnish will not do because Apocope is a postlexical process and mainly found in spoken registers. What we need is a large corpus of spoken Finnish. The best such corpus known to us is the *Spoken language in the Helsinki area 1972-1974* corpus which consists of interviews with 126 native speakers of Helsinki Finnish, collected and transcribed by Heikki Paunonen and his associates in the early 1970's (Paunonen 1995). The observed pattern is shown in (51).[14]

(51) Apocope in nouns. Data from the *Spoken language in the Helsinki area 1972-1974* corpus (Paunonen 1995)

	Monomoraic /e/-final nouns	Bimoraic /e/-final nouns	Regular nouns
No Apocope	100% [57]	11.0% [158]	100% [101]
Apocope	0% [0]	89.0% [1,282]	0% [0]
Examples	vesi / *ves 'water' käsi / *käs 'hand'	yksi ~ yks 'one' kaksi ~ kaks 'two'	ruotsi/ *ruots 'Sweden' poliisi / *poliis 'police'

essive singulars like **vesi-nä* might be to eliminate **vete* by an independent word-level constraint against final *e* (Keyser and Kiparsky 1984) and **vesi-nä* by homonymy avoidance: *ves-i-nä* 'water-PL-ESS' is the plural form. What we crucially cannot assume is that *e*-raising only applies at the word level (cf. Kiparsky 1993a, p. 283). This would wrongly predict that *e*-raising counterfeeds Assibilation.

[14] The Apocope data were obtained by extracting all *si*-final non-compound noun entries from an unabridged dictionary (Tuomi 1972) and by searching Paunonen's corpus based on their unapocopated and apocopated forms. The following decisions must be noted: (i) *omnibuss* 'bus' and *hospits* 'hospice' were not considered apocopated (cf. *omnibussi*, *hospitsi*), but non-nativized foreign words; (ii) *mies* 'man' was not considered an apocopated form of *miesi* since the latter is obsolete; (iii) the fixed expression *yks kaks* 'suddenly', lit. 'one two' was included although it can hardly occur unapocopated (?*yksi kaksi*).

The main observation is that Apocope is variable in bimoraic /e/-final nouns, where it occurs with high frequency, but categorically blocked elsewhere. This is exactly as expected given our metrical analysis of Finnish noun classes: nouns are exhaustively footed except for bimoraic /e/-final nouns which allow metrical variation, e.g. (*kaksi*) ~ (*kak*)*si* 'two', giving Apocope an opportunity to apply.

The noun data make an important general point. While Apocope is a postlexical process and therefore has no morphological or lexical conditions, it is sensitive to the metrical structures assigned in the lexical phonology. For this reason, Apocope applies differently in /e/-final nouns, e.g. (*kak*)*si* 'two' where the final syllable is extrametrical, and regular nouns, e.g. (*keksi*) 'cookie' which is exhaustively footed. Unlike in verbs where Assibilation served as a convenient surface diagnostic for extrametricality, e.g. (*ylti*) ~ (*yl*)*si* 'ask-PAST', in nouns the metrical structure is covert, but it is just as real, and brought to light by the postlexical process of Apocope.

How about lexical frequency effects? In the case of Assibilation, lexical frequency effects emerged in the metrically variable bimoraic verbs. One might thus expect to see analogous lexical frequency effects in the metrically variable bimoraic /e/-final nouns with respect to Apocope. Assuming that high-frequency words prefer incomplete metrical parsing, we would expect a higher rate of Apocope in high-frequency nouns, e.g. (*yk*)*si* 'one' and a lower rate of Apocope in low-frequency nouns e.g. (*köysi*) 'rope'. The 12 relevant nouns and their Apocope rates are shown in (52). If we compare the log frequencies of the apocopating and non-apocopating nouns, the difference between the means is significant ($t = 4.4891$, $df = 10$, $p = 0.0012$, two tailed).

(52) Apocope in bimoraic /e/-final nouns in Paunonen's corpus. AP% = rate of Apocope, LF = lexical frequency

	AP%	LF			AP%	LF	
yksi	94%	544	'one'	kausi	0%	3	'period'
kaksi	92%	463	'two'	virsi	0%	3	'hymn'
viisi	95%	173	'five'	varsi	50%	2	'shaft'
kuusi	98%	100	'six'	kansi	0%	2	'cover'
vuosi	65%	86	'year'	veitsi	0%	1	'knife'
lapsi	42%	62	'child'	köysi	0%	1	'rope'

4.2. *The analysis*

In Stratal Optimality Theory, the output of lexical phonology is the input to postlexical phonology. This implies that the metrical structures assigned in lexical phonology are visible to postlexical phonology. Given this much, the analysis of Apocope is trivial. The crucial constraints are given in (53). The violation patterns are shown in (54) where the inputs are the possible outputs of stem-level phonology for verbs. No rankings are intended. The resulting factorial typology is shown in (55). The gray cells indicate Apocope.

(53) Constraints:

MAX$_{\varphi}$	Do not delete footed segments.
MAX	Do not delete segments.
*I	No short *i*.

(54) Apocope: The violation patterns

(a) Monomoraic stems

(vé.ti)	MAX$_{\varphi}$	MAX	*I
a. (vé.ti)			*
b. (vét)	*	*	

(b) Bimoraic stems

(húu.ti)	MAX$_{\varphi}$	MAX	*I
a. (húu.ti)			*
b. (húut)	*	*	
(húu)si	MAX$_{\varphi}$	MAX	*I
c. (húu)si			*
d. (húu)s		*	

(c) Disyllabic stems

(pá.ran.ti)	MAX$_{\varphi}$	MAX	*I
a. (pá.ran.ti)			*
b. (pá.rant)	*	*	
(pá.ran)si	MAX$_{\varphi}$	MAX	*I
c. (pá.ran)si			*
d. (pá.ran)s		*	

(55) Factorial typology for Apocope

INPUTS	1	2	3
Exhaustive footing: (ve.ti)	(ve.ti)	(ve.ti)	**(vet)**
Exhaustive footing: (huu.ti)	(huu.ti)	(huu.ti)	**(huut)**
Exhaustive footing: (pa.ran.ti)	(pa.ran.ti)	(pa.ran.ti)	**(pa.rant)**
Extrametricality: (huu)si	(huu)si	**(huu)s**	**(huu)s**
Extrametricality: (pa.ran)si	(pa.ran)si	**(pa.ran)s**	**(pa.ran)s**

The factorial typology contains three distinct dialects: No Apocope, Apocope only in extrametrical syllables, and Apocope across the board. In other words, if Apocope is possible in metrified syllables, it is also possible in extrametrical syllables, but not vice versa. The quantitative prediction is that Apocope should be at least as common in extrametrical syllables as in metrified syllables. Both predictions are confirmed by 23 out of the 25 dialects in Laalo's corpus, as we saw in (48).

There are two dialects that do not fit the predicted pattern: Vermlanti and Southwest-Southern, both of the type *t'*, *si~s'*. However, in each case, the number of tokens is very small: Vermlanti has two tokens of *t'*, Southwest-Southern has only one.

This suggests that we are dealing with sparse data gaps. Indeed, additional data from the *Electronic Morphology Archives for Finnish Dialects* confirms this for the Southwest-Southern dialect where unapocopated *ti* is attested: *pit-i* 'must-PAST' (Halikko); *jätt-i* 'leave-PAST' (Kaarina); *kest-i-vä* 'carry-PAST-3P.PL', *ott-i-va* 'take-PAST-3P.PL', *sièt-i* 'tolerate-PAST' (Sauvo). No additional data from the Vermlanti dialect are currently available.[15]

(56) The predicted dialect typology. N = number of attested dialects.[16]

			SAMPLE DIALECT	N
a.	*ti*	*si*	Central Ostrobothnia	11
b.	*ti*	*s'*	--	
c.	*t'*	*s'*	--	
d.	*ti*	*si~s'*	Western Uusimaa	4
e.	*ti~t'*	*s'*	--	
f.	*ti~t'*	*si~s'*	Southwest-Northern	6

The pattern in (56) shows that the factorial typology is too large in one respect: it predicts dialects with categorical Apocope, but no such dialects are found. In other words, Apocope is always optional. Why should this be? Kawahara (2001) has made the interesting suggestion that opacity may presuppose variation. The gist of Kawahara's proposal is that opacity may arise through faithfulness to coexisting output variants. For example, the overapplication of Assibilation in *huus* would reflect faithfulness to the output variant *huusi* where Assibilation applies transparently. This is very similar to our analysis where *huus* reflects faithfulness to *huusi* which is the output of the lexical phonology. However, in Kawahara's theory, the existence of *huus* implies that *huusi* should also exist as an output in the same dialect, predicting that there should be no dialects with *huus* only. This is indeed what we find in Laalo's corpus.

However, if we look beyond Apocope, Kawahara's theory turns out too restrictive. This is because opacity does not always involve transparent output variants. For example, Degemination counterfeeds Assibilation: *ott-i-n* → *otin* (→ **osin*) 'take-PAST-1P.SG'. Under Stratal Optimality Theory, this is because Assibilation occurs at the stem level and Degemination at the word level. Under Kawahara's theory, *otin* would presumably have to survive on the strength of the geminate *tt* present in some output variant of *otin*. However, **ottin* is not an existing output variant, or even a possible output form, but the output of stem-level phonology. From the point of view of Stratal Optimality Theory, Kawahara's theory is a special case: it only covers those opacities that arise from optional postlexical processes.

[15] The Vermlanti dialect, now extinct, was a transplanted variant of the Savo dialect, spoken by descendants of Finnish settlers who emigrated into the forests of central Scandinavia along the Swedish-Norwegian border in the 17th century (Kettunen 1909).

[16] The numbers add up to 21 instead of 25. This is because two dialects are compatible with more than one predicted pattern: the Western Savo pattern (*ti*) could represent (a), (b), or (d), and the Ingria pattern *si~s'* could represent (d) or (f). Since the data underdetermine the analysis, we have omitted these dialects from the table in (56). We have also omitted the Vermlanti and Southwest-Southern dialects for reasons discussed above.

4.3. *Summary*

Based on typological and quantitative evidence, we have argued that Finnish Apocope is a metrically conditioned postlexical process. Since Apocope is postlexical, it has no morphological or lexical conditions. Apparent lexical effects were shown to arise from metrical structures inherited from the lexical phonology. We can thus maintain the claim that Apocope applies wherever its metrical conditions are met and operates under strictly phonological conditions.

5. IMPLICATIONS FOR PHONOLOGICAL OPACITY

In this section, we follow up on the implications of the Finnish evidence for the analysis of phonological opacity. At the heart of the opacity puzzle lies the following question: why do some phonological processes interact (transparency), but others do not (opacity)? The classical solution is rule ordering (Chomsky and Halle 1968) which predicts two types of interactions: feeding and bleeding, and two types of non-interactions: counterfeeding, and counterbleeding. The picture is very different in Optimality Theory (Prince and Smolensky 1993/2004) which in its simplest form only predicts interactions (feeding, bleeding). This is theoretically more restrictive, hence desirable, but empirically problematic since non-interactions (counterfeeding, counterbleeding) are amply documented in the phonological literature.

Various extensions to standard Optimality Theory have been proposed to accommodate non-interactions. A concise summary can be found in McCarthy 2002, pp. 163-178. The approaches can be divided into three principal groups. Most approaches assume that the problem lies in an inadequate theory of FAITHFULNESS and posit new kinds of faithfulness relations, e.g. Output-Output Correspondence (Benua 1995), Sympathy (McCarthy 1999), and Turbidity (Goldrick 2000). Other approaches assume that the problem lies in an inadequate theory of MARKEDNESS and increase the power of markedness constraints, e.g. Targeted Constraints (Wilson 2001) and Comparative Markedness (McCarthy 2003). Yet other approaches follow the theory where it leads and conclude that phonological constraints indeed always interact transparently, contrary to appearances, and that non-interactions arise from outside phonology proper, in particular morphology. This is the view taken in Stratal Optimality Theory. See e.g. Bermúdez-Otero 1999, Cohn and McCarthy 1994/1998, Hale, Kissock, and Reiss 1998, Itô and Mester 2002, Kenstowicz 1995, Kiparsky 2000, 2003, McCarthy and Prince 1993, Orgun 1996, and Rubach 2000 for proposals of this type.

In this paper, we have shown that the predictions of Stratal Optimality Theory are borne out by the Finnish data: both the transparent and opaque interactions fall out from level ordering. An interesting recent alternative to this approach is the theory of Comparative Markedness (McCarthy 2003) where opacity is derived in a very different way. The basic idea is to divide markedness constraints into two groups: those that register markedness violations created through input-output mapping ($_{new}$M) and those that register markedness violations inherited from the input ($_{old}$M). Old markedness violations are defined as those present in the Fully Faithful Candidate (FFC); all other markedness violations are new. We illustrate this by two simplified examples of Assibilation in (57) and (58).[17]

(57) Vowel Deletion feeds Assibilation (actual situation)

/tietä-i/ 'know-PAST'	*AI	$_{new}$*TI	IDENT	$_{old}$*TI
a. tietäi (= FFC)	*!			*
b. tieti		*!		*
c. → tiesi			*	*
d. sieti		*!	*	
e. siesi			**!	

(58) Vowel Deletion counterfeeds Assibilation (hypothetical situation)

/tietä-i/ 'know-PAST'	*AI	$_{old}$*TI	IDENT	$_{new}$*TI
a. tietäi (= FFC)	*!	*		
b. tieti		*!		*
c. tiesi		*!	*	
d. → sieti			*	*
e. siesi			**!	

The first ranking, $_{new}$*TI >> IDENT(t) >> $_{old}$*TI, predicts that *ti*-sequences created by Vowel Deletion should undergo Assibilation because they are new, but *ti*-sequences inherited from the input should not because they are old. The result is *tiesi*, an attested output. More generally, this ranking schema predicts a pattern that closely resembles derived environment effects (Kiparsky 1973): all and only the derived violations are repaired.[18] The second ranking, $_{old}$*TI >> IDENT(t) >> $_{new}$*TI, predicts the opposite situation: old *ti*-sequences undergo Assibilation, but new *ti*-sequences do not. The result is **sieti*, an unattested output. More generally, this ranking schema predicts a pattern that closely resembles counterfeeding opacity effects: all and only the old violations are repaired. Clearly, the correct ranking for Finnish must be (57). The factorial typology is summarized in (59).

(59) Factorial typology:

a. FAITH >> $_{old}$M, $_{new}$M	No alternation
b. $_{old}$M, $_{new}$M >> FAITH	Alternation everywhere
c. $_{new}$M >> FAITH >> $_{old}$M	Derived environment effects
d. $_{old}$M >> FAITH >> $_{new}$M	Counterfeeding opacity

Comparative Markedness makes the following general prediction: if a process is fed by one process, it is fed by all processes; if a process is counterfed by one process, it is counterfed by all processes. In other words, once fed, always fed; once counterfed, always counterfed (Blumenfeld 2003). For example, the ranking $_{new}$*TI >> IDENT(t)

[17] These examples are simplified in that they abstract away from metrical structure and the dissimilatory constraint on identical onsets.
[18] Finnish Assibilation is a showcase example of nonderived environment blocking (NDEB, Kiparsky 1973, 1993a). The metrical analysis presented here accounts for virtually all the facts usually attributed to NDEB with no reference to derived environments (Anttila 2003).

entails that new *ti*-sequences will always be repaired, no matter what process created them. Similarly, the ranking IDENT(t) >> $_{new}$*TI entails that new *ti*-sequences will never be repaired, no matter what process created them.

This prediction turns out to be a liability when we bring more alternations into the picture. Recall the ordering facts:

(60)	Stem level:	Vowel Deletion	a → ∅ / _i	huuta-i → huuti
		Assibilation (optional)	t → s / _i	huut-i ~ huusi
	Word level:	Degemination	tt → t _VC.	ott-i-n → otin
	Postlexical:	Apocope (optional)	i → ∅	huus-i ~ huus

The ordering in (60) correctly predicts that Vowel Deletion feeds Assibilation and that Degemination counterfeeds Assibilation. This is illustrated in (61).

(61) a. Vowel Deletion feeds Assibilation:
/huuta-i/ → huuti → huusi 'shout-PAST'
b. Degemination counterfeeds Assibilation:
/otta-i-n/ → ottin → otin (→ *osin) 'take-PAST-1P.SG'

In Comparative Markedness, this pattern leads to a contradiction: Vowel Deletion feeds Assibilation, hence the ranking must be $_{new}$*TI >> IDENT(t); Degemination counterfeeds Assibilation, hence the ranking must be IDENT(t) >> $_{new}$*TI. This example demonstrates a general point: whether a markedness violation is repaired or not depends on the process that created the violation, contrary to what Comparative Markedness predicts.

The ordering in (60) also correctly predicts that Vowel Deletion feeds Degemination, but that Apocope counterfeeds Degemination. This is illustrated in (62).

(62) a. Vowel Deletion feeds Degemination:
/otta-i-n/ → ottin → otin 'take-1P.SG'
b. Apocope counterfeeds Degemination:
/hakkat-i/ → hakkasi → hakkas (→ *hakas) 'beat-PAST'

Again, the same contradiction arises: Vowel Deletion feeds Degemination, hence the ranking must be $_{new}$*CCVC >> MAX(μ); Apocope counterfeeds Degemination, hence the ranking must be MAX(μ) >> $_{new}$*CCVC. Again, whether a markedness violation is repaired or not depends on the process that created the violation, contrary to what Comparative Markedness predicts.

The ordering analysis in (60) also generalizes beyond counterfeeding effects: it accounts for counterbleeding effects as well. Thus, it correctly predicts that Apocope counterbleeds Assibilation: *t* → *s* / _ *i* applies even though the triggering *i* is not present on the surface. This is illustrated in (63).

(63) Apocope counterbleeds Assibilation:
/huuta-i/ → huuti → huusi → huus 'shout-PAST'

In contrast, Comparative Markedness makes no predictions about counterbleeding opacity at all. This implies that counterbleeding must fall within the scope of some independent theory, e.g. Sympathy Theory (McCarthy 1999). What makes Stratal Optimality Theory attractive is that it generalizes to both counterfeeding and counterbleeding, with no additional assumptions.

Yet another case of opacity that favors Stratal Optimality Theory over Comparative Markedness is the OCP-effect discussed in Section 2: a sequence of adjacent fricatives is prohibited if created by Assibilation, but not if created by Apocope. The examples are repeated below.

(64) The OCP blocks Assibilation
 a. /hiihtä-i/ → hiihti (*hiihsi) 'ski-PAST'
 b. /varasta-i/ → varasti (*varassi) 'steal-PAST'

(65) The OCP does not block Apocope
 a. /imeltä-i-vät/ → imel<u>sv</u>ät 'sweeten-PAST-3P.PL' (Jämsä)
 b. /piirtä-i-hän/ → piir<u>sh</u>än 'draw-PAST-CLIT' (Siilinjärvi)
 c. /kuumenta-i si-tä/ → kuumen<u>s s</u>itä 'heat-PAST it-PAR' (Ruokolahti)

This is an instance of opacity: the OCP is not surface-true with respect to Apocope. However, this is neither counterfeeding nor counterbleeding since the OCP is not a process, but a constraint. A Stratal Optimality analysis is shown in (66)-(67).

(66) Stem level: OCP blocks Assibilation

/hiihtä-i/ 'ski-PAST'	OCP	IDENT$_{\varphi}$	*TI	PARSE-σ	*TERN
a. → (híi)h.ti			*	*	
b. (híi)h.si	*!			*	

(67) Postlexical level: Apocope overrides OCP

(pii)rsi-hän	MAX$_{\varphi}$	MAX	*I	OCP
a. → (pii)rsihän			*	
b. → (pii)rshän	*	*		*

This kind of opacity falls outside the scope of Comparative Markedness. The theory cannot distinguish between OCP-violations created by Assibilation and OCP-violations created by Apocope because both are new violations: both *hiihti* → **hiihsi* (by Assibilation) and *piirsihän* → *piirshän* (by Apocope) violate the constraint $_{new}$OCP. Some independent theory is thus needed to account for the difference between Assibilation and Apocope.

We conclude that Stratal Optimality Theory correctly predicts all three types of opacity in Finnish (counterfeeding, counterbleeding, OCP), whereas Comparative Markedness predicts none of them. More generally, the Finnish evidence supports the view that phonological constraints always interact transparently and that opacity results from factors outside phonology, in this case morphological level ordering.

6. CONCLUSION

In this paper, we examined two phonological processes that simultaneously exhibit variation and opacity: Assibilation and Apocope in Finnish. Two main conclusions emerged. First, we argued that variation results from the presence of multiple metrical systems within Finnish and that the segmental variation reflects metrical variation. The metrical analysis was shown to generalize to a range of apparently unrelated phenomena, including typological asymmetries across dialects, quantitative asymmetries within dialects, differences between nouns and verbs, differences among noun classes, and the loci of lexical frequency effects. Second, we argued that phonological opacity arises from morphological level ordering. By interleaving transparent phonologies with independently motivated morphosyntactic constituents (stems, words, phrases) we derived the transparent and opaque interactions of four phonological processes, including Assibilation and Apocope. More generally, the present study underlines the importance of quantitative and typological evidence in phonology and the central role that hierarchical morphosyntactic structure plays in shaping phonological generalizations.

ACKNOWLEDGEMENTS

I am deeply indebted to Klaus Laalo who gave me an electronic copy of his thesis, discussed the material with me over email, and commented on an earlier version of this paper. This work has benefited from presentations at UC Berkeley, University of Chicago, Cornell University, McGill University, the Meertens Instituut, MIT, New York University, Stanford University, and Stony Brook University. In particular, I thank the following individuals for their input: Lev Blumenfeld, Jonathan Bobaljik, Ellen Broselow, Wayles Browne, Lisa Davidson, Vivienne Fong, Diamandis Gafos, Andrew Garrett, Jen Hay, Tarja Heinonen, Larry Hyman, Sharon Inkelas, Andrew Joseph, Dan Jurafsky, Kaisu Juusela, Yoonjung Kang, Lauri Karttunen, Dasha Kavitskaya, Michael Kenstowicz, Paul Kiparsky, Jason Merchant, Bruce Morén, K.P. Mohanan, James Myers, Marc van Oostendorp, Joe Pater, Glyne Piggott, Yvan Rose, Michael Wagner, and Alan Yu. Finally, I thank two anonymous reviewers for their insightful comments and helpful advice. I am solely responsible for any errors.

REFERENCES

Anttila, Arto. 1997. 'Deriving variation from grammar', in Frans Hinskens, Roeland van Hout, and Leo Wetzels (eds.), *Variation, Change and Phonological Theory*, John Benjamins, Amsterdam / Philadelphia, pp. 35-68.

Anttila, Arto. 2002a. 'Morphologically Conditioned Phonological Alternations', *Natural Language and Linguistic Theory* 20, 1-42.

Anttila, Arto. 2002b. 'Variation and phonological theory', in Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), *Handbook of Language Variation and Change*, Blackwell, Malden, Massachusetts, and Oxford, U.K., pp. 206-243.

Anttila, Arto. 2003. 'Finnish Assibilation', in Makoto Kadowaki and Shigeto Kawahara (eds.), *Proceedings of NELS 33*, GLSA, Amherst, Massachusetts, pp.13-24.

Anttila, Arto. to appear. ‘Variation and optionality’, in Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, Cambridge University Press, Cambridge.

Anttila, Arto and Young-mee Yu Cho. 1998. ‘Variation and change in Optimality Theory’, *Lingua* 104, 31-56, Special issue on Conflicting Constraints.

Bermúdez-Otero, Ricardo. 1999. *Constraint Interaction in Language Change: Quantity in English and Germanic*, Ph.D. dissertation, University of Manchester.

Benua, Laura. 1995. ‘Identity effects in morphological truncation’, in Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk (eds.), *Papers in Optimality Theory*, University of Massachusetts Occasional Papers 18, GLSA, University of Masssachusetts, Amherst, pp. 77-136.

Blumenfeld, Lev. 2003. ‘Counterfeeding, Derived Environment effects, and Comparative Markedness’, *Theoretical Linguistics* 29, 89-99.

Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*, Ph.D. dissertation, University of Amsterdam, Holland Academic Graphics, The Hague.

Boersma, Paul and Bruce Hayes. 2001. ‘Empirical tests of the Gradual Learning Algorithm’, *Linguistic Inquiry* 32, 45-86.

Bybee, Joan L. 2001. *Phonology and Language Use*, Cambridge University Press, Cambridge, U.K.

Bybee, Joan L. 2002. ‘Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change’, *Language Variation and Change* 14, 261-290.

Bye, Patrik. 1997. ‘Representing “overlength”: Against trimoraic syllables’, in Geert Booij and Jeroen van de Weijer (eds.), *Phonology in Progress – Progress in Phonology*, *HIL Phonology Papers III*, Holland Institute of Generative Linguistics.

Carlson, Lauri. 1978. ‘Word stress in Finnish’, Ms., Massachusetts Institute of Technology.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*, Harper and Row, New York.

Cohn, Abigail and John J. McCarthy. 1994/1998. ‘Alignment and parallelism in Indonesian Phonology’, *Working Papers of the Cornell Phonetics Laboratory* 12, pp. 53-137. Also ROA-25-0894.

Elenbaas, Nine and René Kager. 1999. ‘Ternary Rhythm and the Lapse Constraint’, *Phonology* 16, 273-329.

Fidelholtz, James L. 1975. ‘Word Frequency and Vowel Reduction in English’, in Robin E. Grossman, L. James San, and Timothy J. Vance (eds.), *Papers from the 11th Regional Meeting Chicago Linguistic Society*, pp. 200-213.

Goldrick, Matthew. 2000. ‘Turbid output representations and the unity of opacity’, in M. Hirotani, A. Coetzee, N. Hall, and J.-Y. Kim (eds.), *Proceedings of the 30th annual meeting of the North East Linguistics Society*, *Vol. I*, GLSA, Amherst, Massachusetts, pp. 231-245.

Goldsmith, John. 1976. *Autosegmental phonology*, Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Goldsmith, John. 1993. 'Harmonic Phonology', in John Goldsmith (ed.), *The Last Phonological Rule*, The University of Chicago Press, Chicago and London, pp. 21-60.

Gussenhoven, Carlos and Haike Jacobs. 1998. *Understanding Phonology*, Arnold, London.

Hale, Mark, Madelyn Kissock, and Charles Reiss. 1998. 'Output-output correspondence in OT', in Emily Curtis, James Lyle, and Gabriel Webster (eds.), *The Proceedings of the West Coast Conference on Formal Linguistics 16*, CSLI, Stanford, California, pp. 223-36.

Hall, T.A. 2004. 'Assibilation in modern German', *Lingua* 114, 1035-62.

Hanson, Kristin and Paul Kiparsky. 1996. 'A parametric theory of poetic meter', *Language* 72, 287-335.

Hay, Jennifer. 2003. *Causes and Consequences of Word Structure*, Routledge, New York and London.

Hayes, Bruce, Bruce Tesar, and Kie Zuraw. 2003. OTSoft 2.1, software package. Available at http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Hooper, Joan B. 1976. 'Word frequency in lexical diffusion and the source of morphophonological change', in William McChristie, Jr. (ed.), *Current progress in historical linguistics: Proceedings of the Second International Conference on Historical Linguistics*, North-Holland Publishing Company, Amsterdam, pp. 95-105.

Inkelas, Sharon. 1998. 'The theoretical status of morphologically conditioned phonology: a case study of dominance effects', in Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1997*, Kluwer, Dordrecht, pp. 121-155.

Itkonen, Erkki. 1984. 'Suomalais-ugrilaisen kantakielen äänne- ja muotorakenteesta' [On the phonology and morphology of Proto-Finno-Ugrian], in Heikki Paunonen and Päivi Rintala (eds.), *Nykysuomen rakenne ja kehitys*, Suomalaisen Kirjallisuuden Seura, Helsinki, pp. 55-84.

Itô, Junko and Armin Mester. 1995. 'Japanese Phonology', in John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, Blackwell, Oxford, pp. 817-38.

Itô, Junko and Armin Mester. 1999. 'The phonological lexicon', in Natsuko Tsujimura (ed.), *The Handbook of Japanese Linguistics,* Blackwell, Oxford, pp. 62-100.

Itô, Junko and Armin Mester. 2002. 'On the sources of opacity in OT: Coda processes in German', in Caroline Féry and Ruben van de Vijver (eds.), *The Syllable in Optimality Theory*, Cambridge University Press, Cambridge, pp. 271-303.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. 'Probabilistic Relations between Words: Evidence from Reduction in Lexical Production', in Joan. L. Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, John Benjamins, Amsterdam, pp. 229-254.

Kager, René. 1997. 'Rhythmic vowel deletion in Optimality Theory', in I. Roca (ed.), *Derivations and Constraints in Phonology*, Oxford University Press, Oxford, pp. 463-499.

Kang, Yoonjung. 2003. 'Sound changes affecting noun-final coronal obstruents in Korean', in William McClure (ed.), *Japanese/Korean Linguistics*, Vol 12, pp. 117-127.

Karlsson, Fred. 1982. *Suomen kielen äänne- ja muotorakenne* [*The Phonological and Morphological Structure of Finnish*], Werner Söderström Osakeyhtiö, Helsinki.
Kawahara, Shigeto. 2001. *Similarity among Variants: Output-Variant Correspondence*, Senior Thesis, International Christian University. ROA-497.
Kenstowicz, Michael. 1995. 'Cyclic vs. non-cyclic constraint evaluation', *Phonology* 12, 397-436.
Kettunen, Lauri. 1909. 'Descendenttis-äännehistoriallinen katsaus Keski-Skandinavian metsäsuomalaisten kieleen', *Suomi* IV:8, Helsinki. [non vidi]
Keyser, Samuel Jay and Paul Kiparsky. 1984. 'Syllable structure in Finnish phonology', in Mark Aronoff and Richard T. Oehrle (eds.), *Language Sound Structure. Studies in Phonology Presented to Morris Halle by His Teacher and Students*, Cambridge, Massachusetts, MIT Press, pp. 7-31.
Kim, Hyunsoon. 2001. 'A phonetically based account of phonological stop assibilation', *Phonology* 18, 81-108.
Kiparsky, Paul. 1973. 'Phonological representations', in O. Fujimura (ed.), *Three Dimensions of Linguistic Theory*, TEC, Tokyo.
Kiparsky, Paul. 1982. 'Lexical morphology and phonology', in I.S. Yang (ed.), *Linguistics in the Morning Calm*, Vol. 2, Hanshin, Seoul, pp. 3-91.
Kiparsky, Paul. 1993a. 'Blocking in nonderived environments', in S. Hargus and E. Kaisse (eds.), *Phonetics and Phonology*, *Vol. 4: Studies in Lexical Phonology*, Academic Press, San Diego, California, pp. 277-313.
Kiparsky, Paul. 1993b. 'Variable rules', Handout from Rutgers Optimality Workshop 1.
Kiparsky, Paul. 2000. 'Opacity and cyclicity', *The Linguistic Review*, 17, 351-367.
Kiparsky, Paul. 2003. 'Finnish Noun Inflection', in Satu Manninen and Diane Nelson (eds.), *Generative Approaches to Finnic and Saami Linguistics*, CSLI Publications, Stanford, California. pp. 109-161.
Kroch, Anthony S. 1989. 'Reflexes of grammar in patterns of language change', *Language Variation and Change* 1, 199-244.
Laalo, Klaus. 1988. *Imperfektimuotojen ti~si -vaihtelu suomen kielessä* [*The Past Tense ti~si Alternation in Finnish*], Suomalaisen Kirjallisuuden Seura, Helsinki.
Leben, William R. 1973. *Suprasegmental phonology*, Ph.D. dissertation, Massachusetts Institute of Technology.
McCarthy, John J. 1999. 'Sympathy and Phonological Opacity', *Phonology* 16, 331-399.
McCarthy, John J. 2002. *A Thematic Guide to Optimality Theory*, Cambridge University Press, Cambridge.
McCarthy, John. J. 2003. 'Comparative Markedness', *Theoretical Linguistics* 29, 1-51.
McCarthy, John J. and Alan Prince. 1993. 'Prosodic Morphology I: Constraint Interaction and Satisfaction', Report no. RuCCS-TR-3, Rutgers University Center for Cognitive Science, New Brunswick, NJ.
Mohanan, K.P. 1986. *The Theory of Lexical Phonology*. Reidel, Dordrecht.
Myers, James. 2003. 'Frequency effects in Optimality Theory', Ms., National Chung Cheng University.
Myers, James and Gregory Guy. 1997. 'Frequency effects in Variable Lexical Phonology', *University of Pennsylvania Working Papers in Linguistics* 4(1), 215-227.

Orgun, Cemil Orhan. 1996. *Sign-Based Morphology and Phonology with special attention to Optimality Theory*, Ph.D. dissertation, Department of Linguistics, University of California, Berkeley. ROA-171.

Paunonen, Heikki. 1995. *Suomen kieli Helsingissä* [*The Finnish Language in Helsinki*], Helsingin yliopiston suomen kielen laitos, Helsinki.

Paunonen, Heikki. 1974. 'On free variation', *Suomalais-ugrilaisen seuran aikakauskirja*, 72, 285-300.

Phillips, Betty S. 1984. 'Word frequency and the actuation of sound change', *Language* 60, 320-342.

Phillips, Betty S. 2001. 'Lexical diffusion, lexical frequency, and lexical analysis', in Joan. L. Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, Amsterdam, John Benjamins, pp. 123-136.

Pierrehumbert, Janet. B. 2001. 'Exemplar dynamics: Word frequency, lenition and contrast', in Joan. L. Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, John Benjamins, Amsterdam, pp. 137-157.

Posti, Lauri. 1953. 'From Pre-Finnic to late Proto-Finnic', *Finnisch-Ugrische Forschungen* 31, 1-91.

Prince, Alan and Paul Smolensky 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*, Blackwell Publishing, Malden, Massachusetts.

Raffelsiefen, Renate. 1999. 'Phonological constraints on English word formation' in Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1998*, pp. 225-287.

Rapola, Martti. 1965. *Suomen kirjakielen historia I* [*The History of Literary Finnish, Volume I*], Suomalaisen Kirjallisuuden Seura, Helsinki.

Rapola, Martti. 1966. *Suomen kielen äännehistorian luennot* [*Lectures on the Historical Phonology of Finnish*], Suomalaisen Kirjallisuuden Seura, Helsinki.

Rubach, Jerzy. 2000. 'Backness switch in Russian', *Phonology* 17, 39-64.

Sadeniemi, Matti. 1949. *Metriikkamme perusteet* [*The Fundamentals of Finnish Metrics*]. Suomalaisen Kirjallisuuden Seura, Helsinki.

Tuomi, Tuomo. 1972. *Suomen kielen käänteissanakirja* [*A Reverse Dictionary of Finnish*], Suomalaisen Kirjallisuuden Seuran Toimituksia 274, Suomalaisen Kirjallisuuden Seura, Helsinki. Available at the University of Helsinki Language Corpus Server http://www.ling.helsinki.fi/uhlcs/.

Wilson, Colin. 2001. 'Consonant cluster neutralization and targeted constraints', *Phonology* 18, 147-197.

Zamma, Hideki. 2005. 'Predicting Varieties: Partial Orderings in English Stress Assignment', ROA 712-0205.

Zuraw, Kie. 2000. *Patterned Exceptions in Phonology*, Ph.D. dissertation, Department of Linguistics, University of California, Los Angeles.

Arto Anttila
Department of Linguistics
Stanford University
Stanford, California 94305
email: anttila@stanford.edu