WORD STRESS IN FINNISH

Arto Anttila
Stanford University

1. Introduction

(1) In Finnish words, primary stress is easy (= always initial), whereas secondary stress is an outstanding puzzle. However, we need to understand the distribution of secondary stress in order to understand the segmental phonology of the language.

(2) Earlier work: Sadeniemi 1949, Carlson 1978, Keyser and Kiparsky 1984, Hanson and Kiparsky 1996, Elenbaas 1999, Elenbaas and Kager 1999, Kiparsky 2003, McCartney 2003, Karvonen 2005, Karttunen 2006.

(3) The stress rule (first approximation):
- (a) Primary stress falls on the initial syllable.
- (b) Secondary stress falls on every other syllable after that
- (c) except that a light syllable is skipped if the syllable after that is heavy
- (d) unless the heavy syllable is final in which case skipping is optional.

(4) The binary pattern

(a)	ká.las.tè.let	'you are fishing'
(b)	ká.las.tè.le.mì.nen	'fishing'
(c)	íl.moit.tàu.tu.mì.nen	'registering'
(d)	ú.jos.tè.le.màt.to.mùu.des.tàn.sa	'from his lack of shyness'

(5) The ternary pattern

(a)	ká.las.te.lèm.me	'we are fishing'
(b)	íl.moit.tàu.tu.mi.sès.ta	'from registering'
(c)	vói.mis.te.lùt.te.le.màs.ta	'from causing to do gymnastics'
(d)	rá.vin.tò.lat ~ rá.vin.to.làt	'restaurants'

(6) (C)V is light; (C)VC, (C)VV, (C)VVC, (C)VCC are heavy.

(7) Problem 1: The skipping rule does not work very well:

(a)	fí.lo.sò.fìs.sa	'philosopher-INE'
(b)	á.te.ri.à.na	'meal-ESS'

(8) Problem 2: Skipping shows variation and gradient judgments:

(a)	ál.ler.gi.às.sa	(?ál.ler.gì.as.sa)	'allergy-INE'
(b)	pró.fes.so.rìs.sa	pró.fes.sò.ris.sa	'professor-INE'
(c)	(?fí.lo.so.fis.sa)	fí.lo.sò.fìs.sa	'philosopher-INE'

(9) Moreover, secondary stress can be hard to hear.

(10) Research strategy: Use segmental alternations as diagnostics for secondary stress. This works because native speakers have better intuitions about segments than about stress. For example, segments are part of the standard orthography, whereas stress is not.

2. The stress diagnostic

(11) Stop Deletion (first approximation): A singleton stop is deleted between short unstressed vowels (Kiparsky 2003:132).

(12) Example: The partitive /-tA/ which freely combines with nominal stems of any length:

(a)	/maa-i-ta/	mái.ta	'country-PL-PAR'
(b)	/talo-i-ta/	tá.lo.ja	'house-PL-PAR'
(c)	/professori-i-ta/	pró.fes.sò.re.ja ~ pró.fes.so.rèi.ta	'professor-PL-PAR'

(13) Recall that skipping is variable in /professori/ 'professor'. This is confirmed by the variation in Stop Deletion, as shown in (12c).

(14) Stop Deletion (second approximation): No extrametrical singleton stops.

(a)	/maa-i-ta/	(mái.ta)	'country-PL-PAR'
(b)	/talo-i-ta/	(tá.lo)ja	'house-PL-PAR'
(c)	/professori-i-ta/	(pró.fes)(sò.re)ja ~ (pró.fes.so)(rèi.ta)	'professor-PL-PAR'

(15) Stop Deletion also applies in verbs, as in the first infinitive:

(a)	/juo-taC/	(júo.da)	'drink-INF'
(b)	/puno-taC/	(pú.no)a	'weave-INF'
(c)	/vakioi-taC/	(vá.ki)(òi.da)	'keep.constant-INF'
(d)	/formalisoi-taC/	(fór.ma.li)(sòi.da)	'formalize-INF'

(16) In the past participle, a singleton stop geminates in the same environment:

(a)	/juo-tu/	(júo.tu)	'drink-PAST.PCP'
(b)	/puno-tu/	(pú.not)tu	'weave-PAST.PCP'
(c)	/vakioi-tu/	(vá.ki)(òi.tu)	'keep.constant-PAST.PCP'
(d)	/formalisoi-tu/	(fór.ma.li)(sòi.tu)	'formalize-PAST.PCP'

(17) How to explain the variation in skipping?

(18) Vowel Sonority: The morphophonemically [+low] vowels /a, ä, o, ö/ are preferably stressed; the morphophonemically [+high] vowels /i, e, u, y/ are preferably unstressed. For cross-linguistic evidence, see e.g. Kenstowicz 1996, de Lacy 2004, Crowhurst and Michael 2005.

(19) Example: /allergia/ 'allergy' vs. /professori/ 'professor'

	BINARY PATTERN	TERNARY PATTERN	3RD	4TH
(a)	?(ál.ler)(gì.o)ja	(ál.ler.gi)(òi.ta)	/i/	/a/
(b)	(pró.fes)(sò.re)ja	(pró.fes.so)(rèi.ta)	/o/	/i/

(20) Prominence Clash. Stress avoids falling next to a heavy syllable (see Inkelas 1999:145 for a similar constraint in Turkish).

(21) Example: /professori/ 'professor' vs. /filosofi/ 'philosopher'

	BINARY PATTERN	TERNARY PATTERN	2ND	3RD	4TH
(a)	(fí.lo)(sò.fe)ja	?(fí.lo.so)(fèi.ta)	light	/o/	/i/
(b)	(pró.fes)(sò.re)ja	(pró.fes.so)(rèi.ta)	heavy	/o/	/i/

(22) Combined effect: /filosofi/ 'philosopher' vs. /allergia/ 'allergy'

	BINARY PATTERN	TERNARY PATTERN	2ND	3RD	4TH
(a)	(fí.lo)(sò.fe)ja	?(fí.lo.so)(fèi.ta)	light	/o/	/i/
(b)	?(ál.ler)(gì.o)ja	(ál.ler.gi)(òi.ta)	heavy	/i/	/a/

(23) Does this work in just these examples or across the entire lexicon?

3. Procedure

(24) The search engine methodology (see e.g. Zuraw 2000, Hayes and Londe 2006):
- (a) Extract all relevant nouns from an unabridged dictionary.
- (b) Generate all hypothetically possible partitive plural forms for each noun.
- (c) Find matching words on Finnish web sites and count the sites.

(25) Step 1: Extract all vowel-final noun stems from *Suomen kielen käänteissanakirja* [Reverse Dictionary of the Finnish Language] (Tuomi 1972) (165,123 lemmas), University of Helsinki Language Corpus Server http://www.ling.helsinki.fi/uhlcs/, about 23,000 stems in all.

(26)

SYLLABLES	TYPE FREQUENCY	SYLLABLES	TYPE FREQUENCY
1	56	6	415
2	5,952	7	80
3	8,801	8	10
4	5,768	9	3
5	1,762	10	1

(27) Step 2: Annotate stems for the number of syllables, syllable weight (H, L) and vowel quality (*a, ä* = A; *o, ö* = O; *i* = I; *u* = U; and *y* = Y), ignoring the first syllable.

(28)

STEM	DECLENSION	SYLLABLES	PHONOLOGY	
tunturisto	S01	4	UHO	'fells'
neulasisto	S01	4	AHO	'needles'
lihaksisto	S01	4	HHO	'muscles'
koordinaatisto	S01	5	IHHO	'coordinate grid'
konsonantisto	S01	5	OHHO	'consonant inventory'

(29) Step 3: Generate all hypothetically possible partitive plural forms by combining all possible variant endings (Karlsson 1982:282) and morphophonological processes.

(30) Candidate partitive plurals for *Esperanto* 'Esperanto' and *salaatti* 'salad':

VARIANT	DECLENSION	SYLLABLES	PHONOLOGY	STOP
esperantoita	S01	4	EHO	T
esperannoita	S01	4	EHO	T
esperantoja	S01	4	EHO	0
salaatteita	S06 *	3	HI	T
salaateita	S06 *	3	HI	T
salaatteja	S06 *	3	HI	0
salaatteitä	S06 *	3	HI	T
salaateitä	S06 *	3	HI	T
salaattejä	S06 *	3	HI	0

(31) Step 4: Exclude (i) stems with unclear syllabification; (ii) uninflectable expressions; (iii) stem with a final heavy where Stop Deletion is never possible; (iv) candidates with potential coding problems, in particular those containing *ä* (= the low front unrounded vowel) and *ö* (= the mid front rounded vowel).

(32) Step 5: Searched Finnish web sites for the remaining 29,526 candidate forms (April 12, 2005) using QueryGoogle. 7,632 forms scored actual hits, about 9 million hits in all.

(33) A persistent problem: Homonymies get counted twice.

(a)	opinnoita	1. /opinta-i-ta/	'learning-PL-PAR'
		2. /opinto-i-ta/	'study-PL-PAR'
(b)	puristeita	1. /puristeC-i-ta/	'compress-PL-PAR'
		2. /puristi-i-ta/	'purist-PL-PAR'

4. Categorical patterns

4.1 The constraints

(34) Feet are minimally disyllabic and maximally trisyllabic.

(35)

(a)	TROCHEE	Feet are left-headed.	(undominated)
(b)	PARSE-STEM	Stems are exhaustively footed.	(undominated)
(c)	*UNARY	No monosyllabic feet.	(undominated)
(d)	MAXφ	No deletion within a foot.	(undominated)
(e)	MAX	No deletion.	
(f)	*T	No stops (**p, *t, *k*).	

(36) The blocking of Stop Deletion: the monosyllabic stem /maa/ 'country'

/maa-i-ta/	TROCHEE	PARSE-STEM	*UNARY	MAXφ	MAX	*T
a. → (mái.ta)						1
b. (máa.ja)				1!	1	
c. (mái)ta			1!			1
d. mai.ta		1!				1
e. (mai.tá)	1!					1

(37) Weight-to-Stress Principle (WSP): Heavy syllables are stressed. (undominated). (Prince 1990, Kiparsky 2003, Karvonen 2005)

(38) Effect 1: Invariant Stop Deletion in disyllabic stems. Example: /talo/ 'house'

/talo-i-ta/	WSP	MAX	*T
a. → (tá.lo)ja		1	
b. (tá.loi)ta	1!		1
c. (tá.loi.ta)	1!		1

(39) Effect 2: Ternary stress in longer stems. Example: /adrenaliini/ 'adrenalin'.

/adrenaliini-i-ta/	WSP	MAX	*T
a. → (ád.re.na)(lìi.ne)ja		1	
b. (ád.re)(nà.lii.ne)ja	1!	1	
c. (ád.re)(nà.lii)(nèi.ta)	1!		1
d. (ád.re.na)(lìi.nei.ta)	1!		1
e. (ád.re.na)(lìi.nei)ta	1!		1
f. (ád.re)(nà.lii.nei)ta	2!		1

(40) A minor complication: CVV attracts stress from CVC (Karvonen 2005:81-94):
(a) (kóor.di)(nàa.tis.to)ja 'coordinate grid-PL-PAR'
(b) (hó.ri.son)(tàa.le)ja 'horizontal-PL-PAR'

(41) WSP/VV: Heavy syllables with a long vowel are stressed. (undominated). (Karvonen 2005:90)

(42) Illustrating WSP/VV

/koordinaatisto-i-ta/	WSP/VV	WSP	MAX	*T
a. → (kóor.di)(nàa.tis)(tòi.ta)		1		1
b. → (kóor.di)(nàa.tis.to)ja		1	1	
c. (kóor.di.naa)(tìs.to)ja	1!	1	1	
/horisontaali-i-ta/	WSP/VV	WSP	MAX	*T
a. (hó.ri)(sòn.taa)(lèi.ta)	1!	1		1
b. (hó.ri)(sòn.taa.le)ja	1!	1	1	
c. → (hó.ri.son)(tàa.le)ja		1	1	

(43) Interim summary:
(a) No variation in Stop Deletion in monosyllabic and disyllabic stems because the foot structure is invariant.
(b) No variation in Stop Deletion in certain longer stems with heavy syllables because the foot structure is invariant.

(44) By hypothesis, all stem syllables except the first are metrically relevant. L = light, H = heavy, I = light [+high], A = light [+low] (ignoring the two types of heavies for now).

(a)	3-syllable stems	X.{L,H}.{I,A}	4 types
(b)	4-syllable stems	X.{L,H}.{I,A,H}.{I,A}	12 types
(c)	5-syllable stems	X.{L,H}.{I,A,H}.{I,A,H}.{I,A}	36 types

4.2 Predictions and observations

(45) We now test the predictions systematically. Only candidates that satisfy the undominated TROCHEE, PARSE-STEM, *UNARY, and MAXφ are included.

(46) Three-syllable stems. All stems are predicted to be variable (4/4).

HI	/poliisi-i-ta/ ‘police’	WSP/VV	WSP
	a. → (pó.lii)(sèi.ta)	1	1
	b. → (pó.lii.se)ja	1	1
	c. (pó.lii.sei)ta	2	2
LI	/kameli-i-ta/ ‘camel’	WSP/VV	WSP
	a. → (ká.me)(lèi.ta)		
	b. → (ká.me.le)ja		
	c. (ká.me.lei)ta	1	1
HA	/korjaamo-i-ta/ ‘repair shop’	WSP/VV	WSP
	a. → (kór.jaa)(mòi.ta)	1	1
	b. → (kór.jaa.mo)ja	1	1
	c. (kór.jaa.moi)ta	2	2
LA	/kamera-i-ta/ ‘camera’	WSP/VV	WSP
	a. → (ká.me)(ròi.ta)		
	b. → (ká.me.ro)ja		
	c. (ká.me.roi)ta	1	1

(47) The dactylic variant, e.g. (*pó.lii.se*)*ja*, conflicts with (3) which only predicts the trochaic variant, e.g. (*pó.lii*)(*sè.ja*). This is because light syllable secondary stresses come in at a later level and thus do not bleed Stop Deletion (Kiparsky 2003:125).

(48) Four syllable stems. No variation in stems with a heavy third syllable (4/12).

HHA	/edustusto-i-ta/ ‘representation’	WSP/VV	WSP
	a. (é.dus)(tùs.toi)ta	1	2
	b. (é.dus.tus)(tòi.ta)		2
	c. (é.dus)(tùs.toi.ta)	1	2
	d. → (é.dus)(tùs.to)ja		1
HHI	/termostaatti-i-ta/ ‘thermostat’	WSP/VV	WSP
	a. (tér.mos)(tàa.tei)ta	1	2
	b. (tér.mos.taa)(tèi.ta)	1	2
	c. (tér.mos)(tàa.tei.ta)	1	2
	d. → (tér.mos)(tàat.te)ja		1
LHA	/affrikaatta-i-ta/ ‘affricate’	WSP/VV	WSP
	a. (áf.fri)(kàa.toi)ta	1	1
	b. (áf.fri.kaa)(tòi.ta)	1	1
	c. (áf.fri)(kàa.toi.ta)	1	1
	d. → (áf.fri)(kàat.to)ja		
LHI	/margariini-i-ta/ ‘margarine’	WSP/VV	WSP
	a. (már.ga)(rìi.nei)ta	1	1
	b. (már.ga.rii)(nèi.ta)	1	1
	c. (már.ga)(rìi.nei.ta)	1	1
	d. → (már.ga)(rìi.ne)ja		

(49) Four-syllable stems: predictions and observations

	TYPE	PREDICTION	DELETION%	GHITS
(a)	HHA	(é.dus)(tùs.to)ja	100.0	92,308
(b)	HHI	(tér.mos)(tàat.te)ja	100.0	80,063
(c)	LHA	(áff.ri)(kàat.to)ja	99.7	13,039
(d)	LHI	(már.ga)(rìi.ne)ja	100.0	392,942

(50) Some counterexamples:

	TYPE	N	
pioneereita	LHI	39	'pioneer-PL-PAR'
politiikoita	LHA	30	'politics-PL-PAR'
seminaareita	LHI	18	'seminar-PL-PAR'
statistiikoita	HHA	17	'statistics-PL-PAR'

(51) Possible explanations:
- (a) Syllabification: *pio.nee.rei.ta* vs. *pi.o.nee.rei.ta* 'pioneer-PL-PAR'
- (b) Production errors: *po.lii.ti.koi.ta* 'politician-PL-PAR'

(52) Five-syllable stems: No variation if the third syllable is lighter than the fourth (12/36).

LIHI	/auktoriteetti/ 'authority'	WSP/VV	WSP
	a. (áuk.to)(rì.tee)(tèi.ta)	1	1
	b. (áuk.to.ri)(tèe.tei.ta)	1	1
	c. (áuk.to.ri)(tèe.tei)ta	1	1
	d. (áuk.to)(rì.tee.tei)ta	2	2
	e. → (áuk.to.ri)(tèet.te)ja		
	f. (áuk.to)(rì.teet.te)ja	1	1
L.VC.VV.I	/horisontaali/ 'horizontal'	WSP/VV	WSP
	a. (hó.ri)(sòn.taa)(lèi.ta)	1	1
	b. (hó.ri.son)(tàa.lei.ta)	1	2
	c. (hó.ri.son)(tàa.lei)ta	1	2
	d. (hó.ri)(sòn.taa.lei)ta	2	2
	e. → (hó.ri.son)(tàa.le)ja		1
	f. (hó.ri)(sòn.taa.le)ja	1	1

(53) Five-syllable stems: predictions and observations

	TYPE	PREDICTION	DEL%	GHITS	
(a)	LIHI	(áuk.to.ri)(tèet.te)ja	100.0	103,312	'authority-PL-PAR'
(b)	LIHA	(hý.po.te)(nùu.so)ja	100.0	65	'hypotenuse-PL-PAR'
(c)	LAHI	(ád.re.na)(lìi.ne)ja	100.0	9,225	'adrenalin-PL-PAR'
(d)	LAHA	(má.te.ma)(tìik.ko)ja	92.2	167	'mathematics-PL-PAR'
(e)	HIHI	(ré.pub.li)(kàa.ne)ja	100.0	1,884	'republican-PL-PAR'
(f)	HIHA	(é.van.ke)(lìs.to)ja	100.0	290	'evangelist-PL-PAR'
(g)	HAHI	(kón.ser.va)(tìi.ve)ja	100.0	969	'conservative-PL-PAR'
(h)	HAHA	(é.dus.ta)(jìs.to)ja	100.0	14	'representation-PL-PAR'
(i)	LHHI	(hó.ri.son)(tàa.le)ja	93.5	1,040	'horizontal-PL-PAR'
(j)	LHHA	(dí.a.lek)(tìik.ko)ja	100.0	2	'dialectic-PL-PAR'
(k)	HHHI	(ós.kil.los)(kòop.pe)ja	100.0	117	'oscilloscope-PL-PAR'
(l)	HHHA	--	--	--	--

(54) Some counterexamples:

	TYPE	GHITS	GLOSS
arkkitehtuureita	LHHI	65	'architecture-PL-PAR'
semifinaaleita	LIHI	14	'semifinal-PL-PAR'
automatiikoita	LAHA	11	'automaticity-PL-PAR'
materiaaleita	LIHI	4	'material-PL-PAR'
automobiileita	LAHI	3	'automobile-PL-PAR'

(55) Of the 328 predicted invariant initial dactyls, 38 seem to allow an initial trochee:

	INITIAL DACTYL		INITIAL TROCHEE	
(a)	(án.ti.bi)(òot.te)ja	'antibiotic'	(án.ti)(sè.miit.te)ja	'antisemite'
(b)	(má.ni.e)(rìs.me)ja	'mannerism'	(má.ni)(kỳ.ris.te)ja	'manicurist'
(c)	(mó.no.te)(ìs.te)ja	'monotheist'	(mó.no)(fỳ.siit.te)ja	'monophysite'
(d)	(má.te.ri)(àa.le)ja	'material'	(sé.mi)(fì.naa.le)ja	'semifinal'

(56) Some of these examples are quasi-compounds, e.g. *arkki=tehtuuri* (cf. Ringen and Heinämäki 1999:313, Kiparsky 1993, 2003, Välimaa-Blum 1999, Karvonen 2005).

(57) Summary:
- (a) The metrical analysis seems essentially correct.
- (b) Many problem stems are plausibly reanalyzable as quasi-compounds.

5. Quantitative patterns

5.1 The extent of variation

(58) The internet data instantiate 6,150 stems 1,475 of which are variable.

STEM	CATEGORICAL		VARIABLE		TOTAL
3 syllables	2,095	(66.4%)	1,058	(33.6%)	3,153
4 syllables	1,808	(83.6%)	355	(16.4%)	2,163
5 syllables	772	(92.6%)	62	(7.4%)	834

5.2 Vowel sonority and prominence clash

(59)
- (a) /a, ä, o, ö/ are preferably stressed, /i, e, u, y/ preferably unstressed.
- (b) Secondary stress avoids falling next to a heavy syllable.

(60) Vowel sonority and prominence clash at work

	TYPE	BINARY	TERNARY	DEL%	GHITS	
(a)	LAI	(fí.lo)(sò.fe)ja	(fí.lo.so)(fèi.ta)	90.7%	23,595	'philosopher'
(b)	HAI	(pró.fes)(sò.re)ja	(pró.fes.so)(rèi.ta)	84.9%	34,612	'professor'
(c)	LIA	(gál.le)(rì.o)ja	(gál.le.ri)(òi.ta)	1.0%	91,598	'gallery'
(d)	HIA	(ál.ler)(gì.o)ja	(ál.ler.gi)(òi.ta)	0.3%	190,416	'allergy'

(61) Vowel sonority and prominence clash constraints

(a)	*x/A	No unstressed light A (= /a, ä, o, ö/)
(b)	*x/AA	No unstressed heavy A (= /a, ä, o, ö/)
(c)	*X/I	No stressed light I (= /e, i, u, y/)
(d)	*X/II	No stressed heavy I (= /e, i, u, y/)
(e)	*H.X	No stress next to a heavy syllable.

(62) Additional constraints (Hanson and Kiparsky 1996, Elenbaas and Kager 1999):

(a)	SWP	Stressed syllables are heavy.
(b)	FTBIN	Feet are disyllabic.
(c)	PARSE-σ	Syllables belong to feet.
(d)	ALL-FEET-LEFT	All feet are at the left edge of the prosodic word.

5.3 Variation in three-syllable stems

(63) The predicted variants:

(a) trochee + trochee, no Stop Deletion: (*pó.lii*)(*sèi.ta*)
(b) dactyl, Stop Deletion: (*pó.lii.se*)*ja*

(64)

	TYPE	PREDICTIONS	DEL%	GHITS	
(a)	HI	(pó.lii)(sèi.ta) ~ (pó.lii.se)ja	98.1%	1,803,692	'police'
(b)	HA	(kór.jaa)(mòi.ta) ~ (kór.jaa.mo)ja	96.1%	2,203,235	'repair shop'
(c)	LI	(ká.me)(lèi.ta) ~ (ká.me.le)ja	36.7%	1,180,225	'camel'
(d)	LA	(ká.me)(ròi.ta) ~ (ká.me.ro)ja	2.2%	2,030,424	'camera'

(65) Possible OT analyses: Multiple Grammars (Kiparsky 1994), Partially Ordered Grammars (Anttila 1997, Anttila and Cho 1998) or Stochastic Optimality Theory (Boersma 1997, Boersma and Hayes 2001).

(66) A more general question: What sorts of quantitative patterns are admitted and excluded, given these constraints? Such questions are routinely asked in generative linguistics in the context of qualitative data.

(67) Constraint violations for three-syllable stems. No rankings are intended.

		FTBIN	*T	SWP	*X.H	*x/A	*x/AA	I/*X	*X/II	ALIGN-L	P-SYLL	MAX
HI	a. (pó.lii)(sèi.ta)		1	1	2	1			1	2		
	b. (pó.lii.se)ja	1		1	1	1					1	1
LI	a. (ká.me)(lèi.ta)		1	1		1			1	2		
	b. (ká.me.le)ja	1		1		1					1	1
HA	a. (kór.jaa)(mòi.ta)		1		2	1	1			2		
	b. (kór.jaa.mo)ja	1			1	2	1				1	1
LA	a. (ká.me)(ròi.ta)		1	1		1				2		
	b. (ká.me.ro)ja	1		1		2					1	1

(68) Factorial typology (OTSOFT, Hayes, Tesar and Zuraw 2003). Stop Deletion shaded.

	Output #1	Output #2	Output #3
HI	(po.lii)(sei.ta)	(po.lii.se)ja	(po.lii.se)ja
LI	(ka.me)(lei.ta)	(ka.me)(lei.ta)	(ka.me)(lei.ta)
HA	(kor.jaa)(moi.ta)	(kor.jaa)(moi.ta)	(kor.jaa.mo)ja
LA	(ka.me)(roi.ta)	(ka.me)(roi.ta)	(ka.me)(roi.ta)

	Output #4	Output #5	Output #6
HI	(po.lii.se)ja	(po.lii.se)ja	(po.lii.se)ja
LI	(ka.me.le)ja	(ka.me.le)ja	(ka.me.le)ja
HA	(kor.jaa)(moi.ta)	(kor.jaa.mo)ja	(kor.jaa.mo)ja
LA	(ka.me)(roi.ta)	(ka.me)(roi.ta)	(ka.me.ro)ja

(69) Typological entailment: For all languages in the predicted typology, if the mapping <$input_1$, $output_1$> belongs to the language, so does the mapping <$input_2$, $output_2$>.

(70) Example: If Stop Deletion applies to LI-stems, it also applies to HI-stems.
`<LI, (ka.me.le)ja> --> <HI, (po.lii.se)ja>`

(71) How does this matter to variation? Assumptions:
- (a) Variation arises from multiple grammars (Kroch 1989, Kiparsky 1994) or the stochastic evaluation of a single grammar (Boersma and Hayes 2001).
- (b) The number of grammars predicting an output is proportional to this output's frequency of occurrence.

(72) Quantitative entailment: Stop Deletion is at least as frequent in HI-stems as in LI-stems.

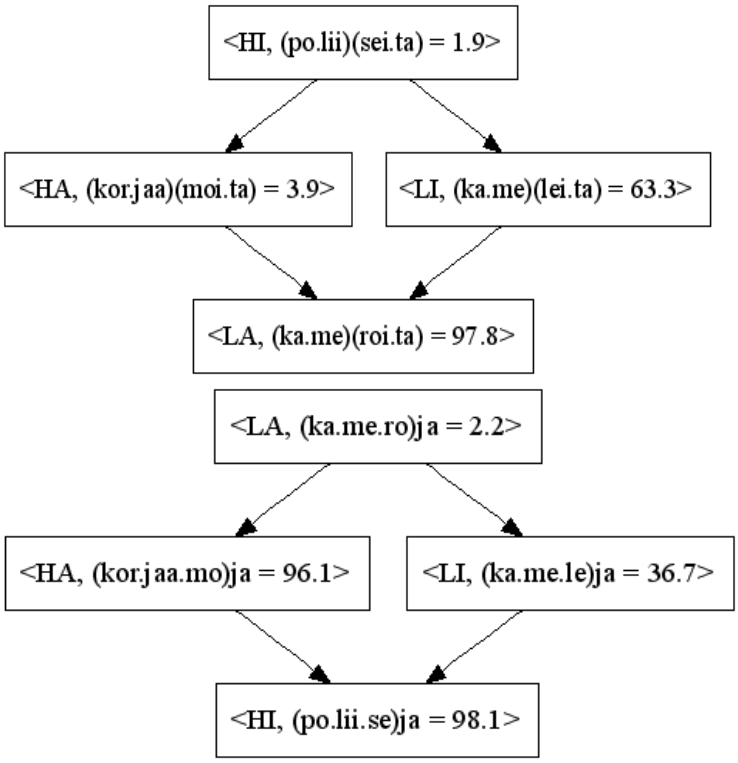
(73) This quantitative prediction is robust: it holds no matter how grammars are selected from the factorial typology during learning or performance.

(74) We call the set of all typological entailments of a grammar a TYPOLOGICAL ORDER (T-ORDER). T-orders can be found using T-Order Generator (Anttila and Andrus 2006).

(75) T-order for three-syllable stems (10 typological entailments)

```
<HI, (po.lii)(sei.ta)>  --> <LA, (ka.me)(roi.ta)>
<LA, (ka.me.ro)ja>      --> <HI, (po.lii.se)ja>
<HI, (po.lii)(sei.ta)>  --> <LI, (ka.me)(lei.ta)>
<LA, (ka.me.ro)ja>      --> <LI, (ka.me.le)ja>
<LI, (ka.me.le)ja>      --> <HI, (po.lii.se)ja>
<LI, (ka.me)(lei.ta)>   --> <LA, (ka.me)(roi.ta)>
<HA, (kor.jaa)(moi.ta)> --> <LA, (ka.me)(roi.ta)>
<HI, (po.lii)(sei.ta)>  --> <HA, (kor.jaa)(moi.ta)>
<LA, (ka.me.ro)ja>      --> <HA, (kor.jaa.mo)ja>
<HA, (kor.jaa.mo)ja>    --> <HI, (po.lii.se)ja>
```

(76) T-order as a directed graph annotated with the relative frequencies of <input, output> mappings. Frequencies are predicted to remain the same or increase as we move along the T-order. This prediction is confirmed.



(77) Two ways of working out the T-order:
- (a) Figure out the entailments from the factorial typology.
- (b) Figure out the Elementary Ranking Conditions (ERCs) for each <input, output> mapping and the entailments among them (Prince 2002a, 2002b, 2006).

(78) How well does the T-order match the corpus data?
- (a) Precision = 1 (100% of the predicted entailments are correct)
- (b) Recall = 0.417 (41.7% of the correct entailments are predicted)

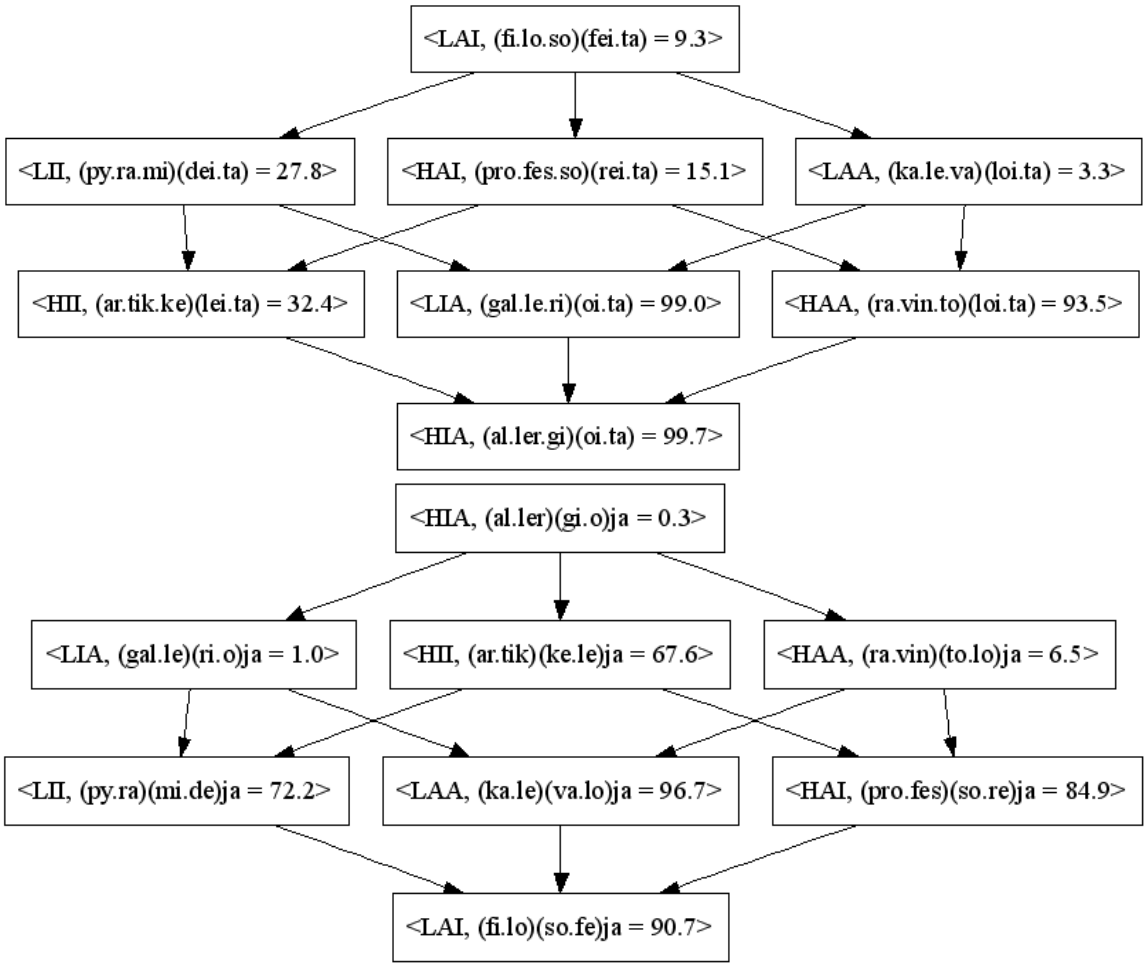
5.4 Variation in four-syllable stems

(79) The predicted variants:
- (a) trochee + trochee, Stop Deletion: (*pró.fes*)(*sò.re*)*ja*
- (b) dactyl + trochee, no Stop Deletion: (*pró.fes.so*)(*rèi.ta*)

(80)

TYPE	PREDICTIONS	DEL%	GHITS	
LAA	(ká.le)(và.lo)ja ~ (ká.le.va)(lòi.ta)	96.7%	30	‘Kalevala’
LAI	(fí.lo)(sò.fe)ja ~ (fi.lo.so)(fèi.ta)	90.7%	23,595	‘philosopher’
HAI	(pró.fes)(sò.re)ja ~ (pró.fes.so)(rèi.ta)	84.9%	34,612	‘professor’
LII	(pý.ra)(mì.de)ja ~ (py.ra.mi)(dèi.ta)	72.2%	4,175	‘pyramid’
HII	(ár.tik)(kè.le)ja ~ (ár.tik.ke)(lèi.ta)	67.6%	630,594	‘article’
HAA	(rá.vin)(tò.lo)ja ~ (rá.vin.to)(lòi.ta)	6.5%	24,479	‘restaurant’
LIA	(gál.le)(rì.o)ja ~ (gál.le.ri)(òi.ta)	1.0%	91,598	‘gallery’
HIA	(ál.ler)(gì.o)ja ~ (ál.ler.gi)(òi.ta)	0.3%	190,416	‘allergy’

(81) T-order for four-syllable stems (only variable cases)



(82) (a) Precision = 0.956 (95.6% of the predicted entailments are correct)
(b) Recall = 0.586 (58.6% of the correct entailments are predicted)

(83) The precision value suffers from one error: LAI (90.7%) is predicted to have a higher Stop Deletion rate than LAA (96.7%),

(84) The LAA-class is small (8 stems, 30 tokens)

	VARIANT	GHITS	GLOSS
(a)	unisonoja	11	‘unison’
	eldoradoja	7	‘eldorado’
	obligatoja	5	‘obbligato’
	autostradoja	2	‘autostrada’
	subimagoja	2	‘subimago’
	telefotoja	1	‘telephoto’
	tuataroja	1	‘tuatara’
(b)	tapioloita	1	‘Tapiola’

(85) Several of these counterexamples seem spurious:
(a) Orthography: long vowel in [eldora:do], [obliga:to], [autostra:da]
(b) Quasi-compounds: *uni=sono*, *tele=foto*

5.5 Variation in five-syllable stems

(86) Only 28/36 types found in the dictionary (temporarily ignoring CIC vs. CAC vs. CVV).

(87) The first predicted type of variation (14 stem types) yields two distinct outputs:
(a) trochee + dactyl, Stop Deletion: (*ák.va*)(*rèl.lis.te*)*ja*
(b) trochee + trochee + trochee, no Stop Deletion: (*ák.va*)(*rèl.lis*)(*tèi.ta*)

(88)

TYPE	PREDICTIONS	DEL%	GHITS	
LHHA	(kóor.di)(nàa.tis.to)ja ~ (kóor.di)(nàa.tis)(tòi.ta)	100	205	'coord. grid'
LHHI	(pró.pa)(gàn.dis.te)ja ~ (pró.pa)(gàn.dis)(tèi.ta)	100	79	'propagandist'
LHHI	(ák.va)(rèl.lis.te)ja ~ (ák.va)(rèl.lis)(tèi.ta)	100	9	'aquarellist'
LAAI	(ánt.ro)(pò.lo.ge)ja ~ (ánt.ro)(pò.lo)(gèi.ta)	99.6	1,990	'anthropologist'
LIII	(sým.po)(sì.u.me)ja ~ (sým.po)(sì.u)(mèi.ta)	98.6	296	'symposium'
LAII	(pó.ly)(à.mi.de)ja ~ (pó.ly)(à.mi)(dèi.ta)	95.7	69	'polyamid'
HHII	(líi.rum)(làa.ru.me)ja ~ (líi.rum)(làa.ru)(mèi.ta)	18.6	43	'nonsense'
LHII	(ín.ku)(nàa.be.le)ja ~ (ín.ku)(nàa.be)(lèi.ta)	9.5	2,139	'incunable'
HHAI	(kóm.men)(tàat.to.re)ja ~ (kóm.men)(tàat.to)(rèi.ta)	1.2	1,525	'commentator'
LHAI	(ó.pe)(ràat.to.re)ja ~ (ó.pe)(ràat.to)(rèi.ta)	0.8	20,316	'operator'
HHIA	(kón.sul)(tàa.ti.o)ja ~ (kón.sul)(tàa.ti)(òi.ta)	0.5	10,126	'consultation'
LIIA	(kóm.mu)(nì.ke.o)ja ~ (kóm.mu)(nì.ke)(òi.ta)	0.3	3,474	'communiqué'
LHIA	(ó.pe)(ràa.ti.o)ja ~ (ó.pe)(ràa.ti)(òi.ta)	0.0	42,237	'operation'
LAIA	(ál.le)(gò.ri.o)ja ~ (ál.le)(gò.ri)(òi.ta)	0.0	63,717	'allegory'

(89) The second predicted type of variation (7 stem types) yields three distinct outputs:
(a) dactyl + trochee, Stop Deletion: (*á.vant.gar*)(*dìs.te*)*ja*,
(b) trochee + dactyl, Stop Deletion, e.g. (*á.vant*)(*gàr.dis.te*)*ja*,
(c) trochee + trochee + trochee, no Stop Deletion, e.g. (*á.vant*)(*gàr.dis*)(*tèi.ta*)

(90)

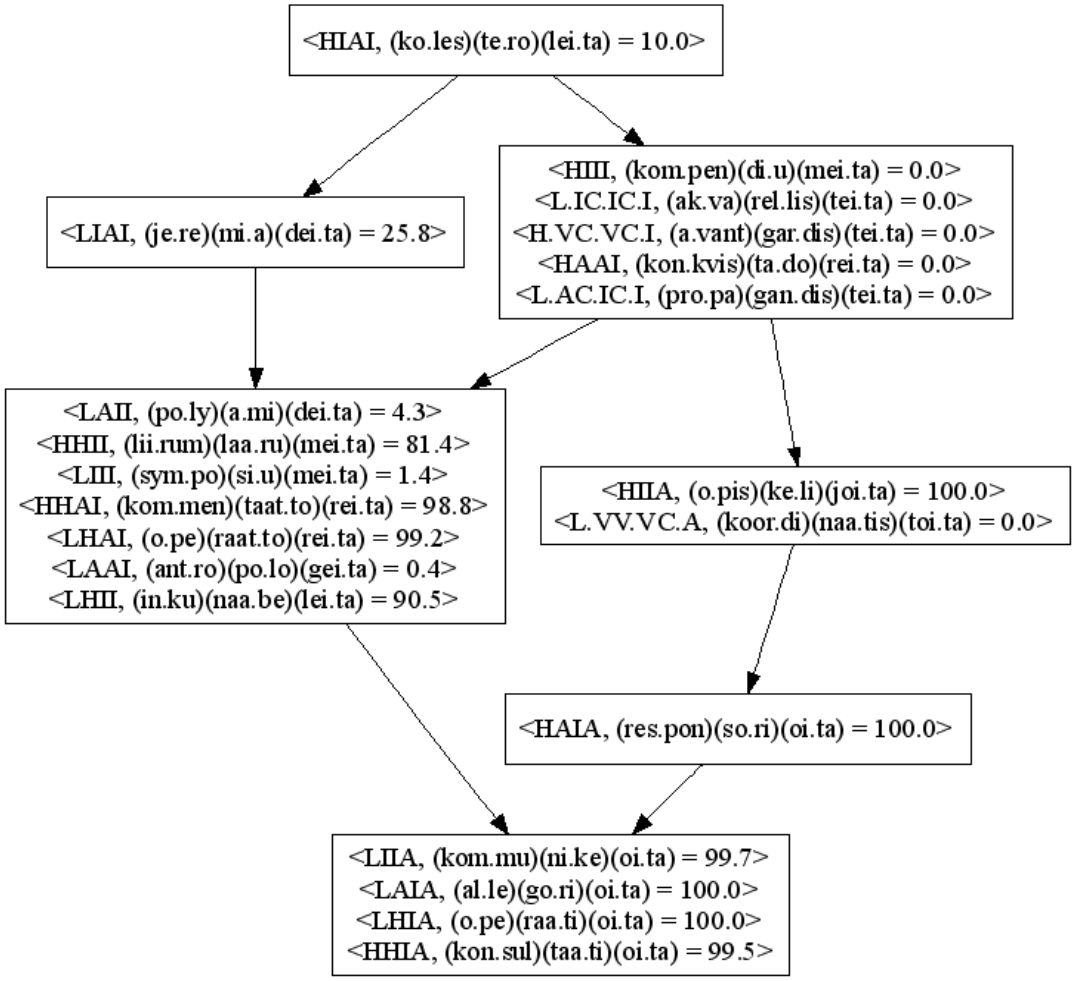
TYPE	PREDICTIONS	DEL%	GHITS	
HAAI	(kón.kvis.ta)(dò.re)ja ~ (kón.kvis)(tà.do.re)ja ~ (kón.kvis)(tà.do)(rèi.ta)	100	23	'conquistador'
HIII	(kóm.pen.di)(ù.me)ja ~ (kóm.pen)(dì.u.me)ja ~ (kóm.pen)(dì.u)(mèi.ta)	100	6	'compendium'
HHHI	(á.vant.gar)(dìs.te)ja ~ (á.vant)(gàr.dis.te)ja ~ (á.vant)(gàr.dis)(tèi.ta)	100	36	'avantgardist'
HIAI	(kó.les.te)(rò.le)ja ~ (kó.les)(tè.ro.le)ja ~ (kó.les)(tè.ro)(lèi.ta)	90.0	50	'cholesterol'
LIAI	(jé.re.mi)(à.de)ja ~ (jé.re)(mì.a.de)ja ~ (jé.re)(mì.a)(dèi.ta)	74.2	66	'jeremiad'
HIIA	(ó.pis.ke)(lì.jo)ja ~ (ó.pis)(kè.li.jo)ja ~ (ó.pis)(kè.li)(jòi.ta)	0.0	221,456	'student'
HAIA	(rés.pon.so)(rì.o)ja ~ (rés.pon)(sò.ri.o)ja ~ (rés.pon)(sò.ri)(òi.ta)	0.0	21	'responsory'

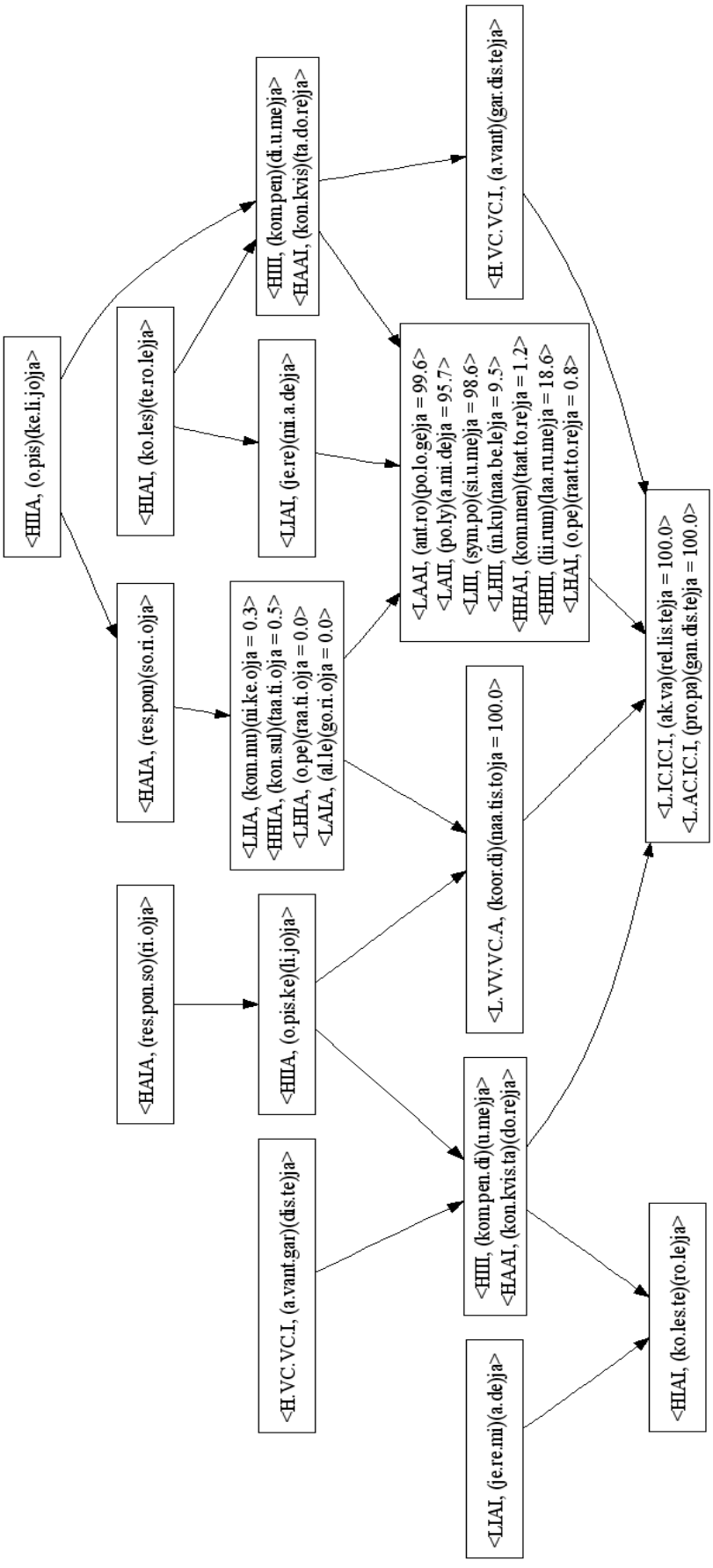
(91) Metrical ambiguity in written data:
- (a) *akvarellisteja* must be (*ák.va*)(*rèl.lis.te*)*ja*
- (b) *avantgardisteja* can be (*á.vant*)(*gàr.dis.te*)*ja* or (*á.vant.gar*)(*dìs.te*)*ja*. Which?

(92) (*á.vant*)(*gàr.dis.te*)*ja* is predicted to have a wider distribution than (*á.vant.gar*)(*dìs.te*)*ja*: they share several typological entailments, but in addition, (*á.vant.gar*)(*dìs.te*)*ja* entails (*kóm.pen.di*)(*ù.me*)*ja* and (*kón.kvis.ta*)(*dò.re*)*ja* and should thus occur under typologically more restricted circumstances.

(93) Conclusion: Exclude metrically ambiguous forms from precision/recall calculations.

(94) T-order for five-syllable stems (only variable cases)




<HIIA, (o.pis)(ke.li.jo)ja>
<HIAI, (ko.les)(te.ro.le)ja>
<HIII, (kom.pen)(di.u.me)ja>
<HAAI, (kon.kvis)(ta.do.re)ja>
<LIAI, (je.re)(mi.a.de)ja>
<H.VC.VC.I, (a.vant)(gar.dis.te)ja>
<HAIA, (res.pon)(so.ri.o)ja>
<LIIA, (kom.mu)(ni.ke.o)ja = 0.3>
<HHIA, (kon.sul)(taa.ti.o)ja = 0.5>
<LHIA, (o.pe)(raa.ti.o)ja = 0.0>
<LAIA, (al.le)(go.ri.o)ja = 0.0>
<LAAI, (ant.ro)(po.lo.ge)ja = 99.6>
<LAII, (po.ly)(a.mi.de)ja = 95.7>
<LIII, (syn.po)(si.u.me)ja = 98.6>
<LHII, (in.ku)(naa.be.le)ja = 9.5>
<HHAI, (kom.men)(taat.to.re)ja = 1.2>
<HHII, (lii.rum)(laa.ru.me)ja = 18.6>
<LHAI, (o.pe)(raat.to.re)ja = 0.8>
<HAIA, (res.pon.so)(ri.o)ja>
<HIIA, (o.pis.ke)(li.jo)ja>
<L.VV.VC.A, (koor.di)(naa.tis.to)ja = 100.0>
<L.IC.IC.I, (ak.va)(rel.lis.te)ja = 100.0>
<L.AC.IC.I, (pro.pa)(gan.dis.te)ja = 100.0>
<H.VC.VC.I, (a.vant.gar)(dis.te)ja>
<HIII, (kom.pen.di)(u.me)ja>
<HAAI, (kon.kvis.ta)(do.re)ja>
<LIAI, (je.re.mi)(a.de)ja>
<HIAI, (ko.les.te)(ro.le)ja>

(95) (a) Precision = 0.860 (86.0% of the predicted entailments are correct)
(b) Recall = 0.607 (60.7% of the correct entailments are predicted)

(96) Mappings enclosed in a box entail one another, i.e. form a cycle. Such mappings are grammatically equivalent. Any differences among them lie beyond the grammar.

(97) The following stems are predicted to behave alike:

TYPE	PREDICTIONS	DEL%	GHITS
LAAI	(ánt.ro)(pò.lo)(gèi.ta) ~ (ánt.ro)(pò.lo.ge)ja	99.6%	1,990
LIII	(sým.po)(sì.u)(mèi.ta) ~ (sým.po)(sì.u.me)ja	98.6%	296
LAII	(pó.ly)(à.mi)(dèi.ta) ~ (pó.ly)(à.mi.de)ja	95.7%	69
HHII	(líi.rum)(làa.ru)(mèi.ta) ~ (líi.rum)(làa.ru.me)ja	18.6%	43
LHII	(ín.ku)(nàa.be)(lèi.ta) ~ (ín.ku)(nàa.be.le)ja	9.5%	2,139
HHAI	(kóm.men)(tàat.to)(rèi.ta) ~ (kóm.men)(tàat.to.re)ja	1.2%	1,525
LHAI	(ó.pe)(ràat.to)(rèi.ta) ~ (ó.pe)(ràat.to.re)ja	0.8%	20,316

(98) In LAAI stems, neither predicted stress pattern is correct: *(*ánt.ro*)(*pò.lo*)(*gèi.ta*) ~ *(*ánt.ro*)(*pò.lo.ge*)*ja*. The correct pattern (*ánt.ro.po*)(*lò.ge*)*ja* is not predicted.

(99) The LAAI class is mostly erudite vocabulary (29 stems):

DELETION	GHITS	NO DELETION	GHITS	GLOSS
aerosoleja	361	aerosoleita	1	‘aerosol’
sosiologeja	306	--		’sociologist’
sosionomeja	295	--		’socionomist’
arkeologeja	284	arkeologeita	1	‘archaeologist’
kosmetologeja	172	--		’cosmetologist’
gynekologeja	150	--		‘gynekologist’
antropologeja	99	--		‘anthropologist’
radiologeja	59	--		’radiologist’
ideologeja	51	--		’ideologist’
antroposofeja	41	--		‘anthroposophist’
ornitologeja	41	--		’ornithologist’
fysiologeja	19	--		‘physiologist’
musikologeja	19	--		’musicologist’
heterotrofeja	13	--		‘heterotroph’
mikrofaradeja	12	mikrofaradeita	1	’microfarad’
magnetofoneja	11	--		’magnetophone’
toreadoreja	9	--		’toreador’
pikofaradeja	8	pikofaradeita	5	’picofarad’
etymologeja	7	--		‘etymologist’
mareografeja	6	--		’marigraph’
genealogeja	4	--		‘genealogist’
papyrologeja	4	--		’papyrologist’
kromatoforeja	3	--		’chromatophore’
bibliografeja	2	--		‘bibliographer’
entomologeja	2	--		‘entomologist’
iktyologeja	1	--		’ichthyologist’
mineralogeja	1	--		’mineralogist’
spermatoforeja	1	--		’spermatophore’
trypanosomeja	1	--		’trypanosoma’

(100) /aerosoli/ ‘aerosol’ may be incorrectly syllabified: *ae.ro.so.li*, not *a.e.ro.so.li*

(101) In *-logi* ‘-logist’ words, the correct pattern, e.g. (*só.si.o*)(*lò.ge*)*ja*, is not predicted and Stop Deletion is (virtually) obligatory, e.g. (*só.si.o*)(*lò.ge*)*ja* / *(*só.si*)(*ò.lo*)(*gèi.ta*). Solution: Fixed stress in /-lógi/ (Hanson and Kiparsky 1996:307) or the pronunciation is really [sosiolo:gi] (= LAHI).

(102) Similarly for *-nomi* ‘-nomist’, *-sofi* ‘-sopher/-sophist’, *-grafi* ‘graph(er)’, *-trofi* ‘-troph’, and *-foni* ‘-phone’.

(103) Other stems where the correct stress pattern is not predicted:

	PREDICTED	CORRECT	
LAII	*(ká.ry)(à.ti)(dèi.ta) ~ *(ká.ry)(à.ti.de)ja	(ká.ry.a)(tì.de)ja	‘karyatid’
LIII	*(ím.pe)(rì.u)(mèi.ta) ~ *(ím.pe)(rì.u.me)ja	(ím.pe.ri)(ù.me)ja	‘empire’
	*(sým.po)(sì.u)(mèi.ta) ~ *(sým.po)(sì.u.me)ja	(sým.po.si)(ù.me)ja	‘symposium’

(104) Overgeneration: All HIAI stems should show three-way variation, but none actually does:

(tés.tos)(tè.ro)(nèi.ta)	(tés.tos)(tè.ro.ne)ja	--	‘testosterone’
--	--	(pró.ges.te)(rò.ne)ja	‘progesterone’
--	--	(kó.les.te)(rò.le)ja	‘cholesterol’
--	--	(ó.lym.pi)(à.de)ja	‘olympiad’

(105) Suggestion: Individual stems can choose a single stress pattern from the variant options predicted by the metrical grammar. The metrical analysis itself seems essentially correct.

6. Does Finnish have lexical accents?

(106) Earlier analyses have posited lexically marked stresses inherited from donor languages:

> [I]n long words of foreign origin, secondary stress is fixed on the syllable where the donor language has primary stress: hence *fílosòfeista*, not *fílosofèista* ‘philosopher-PL-ELA’; *ágronòmeilla*, not *ágronomèilla* ‘agronomist-PL-ADE’; *áleksandrìini*, not *áleksàndriini* ‘alexandrine’. Secondary stress is thus not rhythmical, but fixed. (Sadeniemi 1949:75).

(107) Kiparsky (2003:113): In polysyllabic stems, stress is either FIXED or MOVABLE:

FIXED STRESS		MOVABLE STRESS	
á.la.bà.ma	‘Alabama’	*áp.teek.kà.ri*	‘pharmacist’
á.la.bà.mas.sa	‘Alabama-INE’	*áp.teek.ka.rìs.sa*	‘pharmacist-INE’

(108) A lexical accent in /alabáma/ keeps stress on the third syllable, as in /-lógi/ ‘-logist’. But Kiparsky (2003:114) notes that phonological generalizations may also be involved.

(109) Our proposal: Lexical accents are phonologically predictable (i.e. do not exist).

(110) Kiparsky’s data (2003:113-14). I have omitted stems where the penult is stressed because it is heavy and stems where the penult weight alternates due to Consonant Gradation.

FIXED STRESS		MOVABLE STRESS	
a.la.ba.ma	'Alabama'	ap.teek.ka.ri	'pharmacist'
a.la.do.bi	'aspic from meat'	ar.tik.ke.li	'article'
bar.ri.ka.di	'barricade'	a.ses.so.ri	'assessor'
e.ko.no.mi	'economist'	hant.lan.ka.ri	'helper'
esp.la.na.di	'esplanade'	ka.len.te.ri	'calendar'
e.ta.no.li	'ethanol'	kons.taa.pe.li	'constable'
fi.lo.so.fi	'philosopher'	mak.ros.po.ri	'macrospore'
in.va.li.di	'invalid'	me.an.de.ri	'meandros'
ka.lo.me.li	'mercury chloride'	mo.nok.ke.li	'monocle'
ma.ka.ro.ni	'macaroni'	o.raak.ke.li	'oracle'
me.ga.fo.ni	'megaphone'	par.tik.ke.li	'particle'
or.to.pe.di	'orthopedist'	pro.fes.so.ri	'professor'
pa.ra.gra.fi	'paragraph'	sal.pie.ta.ri	'saltpeter'
por.tu.ga.li	'Portugal'	se.naat.to.ri	'senator'
pri.vi.le.gi	'privilege'	si.lin.te.ri	'top hat'
py.ra.mi.di	'pyramid'	sy.lin.te.ri	'cylinder'
sar.ko.fa.gi	'sarcophagus'	tri.an.ge.li	'triangle'
se.re.na.di	'serenade'	tu.berk.ke.li	'tubercle'
ka.pit.te.li	'chapter'	kor.ri.do.ri	'corridor'
		ka.le.va.la	'Kalevala'

(111) (a) Stems with fixed stress have a light antepenult
(b) Stems with movable stress have a heavy antepenult.

(112) Three apparent exceptions. In the internet corpus, these forms occur as follows:
(a) *kapitteli* once as (*ká.pit.te*)(*lèi.ta*), showing its stress is movable.
(b) *korridori* six times, only as (*kór.ri*)(*dò.re*)*ja*, no evidence for movable stress.
(c) *kalevala* not attested

(113) Sadeniemi's examples also behave as expected:

(a)	fi.lo.sò.feis.ta	LAI	'philosopher-PL-ELA'
	ág.ro.nò.meil.la	LAI	'agronomist-PL-ADE'
(b)	á.lek.sand.rìi.ni	H.CVC. CVV	'alexandrine'

(114) But some irreducible lexical effects remain: /-lógi/ '-logist', (*tés.tos*)(*tè.ro.ne*)*ja* 'testosterone-PL-PAR' vs. (*kó.les.te*)(*rò.le*)*ja* 'cholesterol-PL-PAR'.

(115) Polysyllabic nouns are largely recent borrowings. It is striking how thoroughly they have been nativized, down to the statistical details of word prosody.

7. Conclusions

(116) Finnish word stress is almost entirely rhythmical.

(117) Two phonological tendencies that emerge quantitatively in variable data:
(a) Vowel sonority: low vowels attract stress, high vowels repel stress.
(b) Prominence clash: Secondary stress avoids falling next to a heavy syllable.

(118) Embedded in OT, these principles yield TYPOLOGICAL ENTAILMENTS that constrain the space of possible quantitative patterns, admitting some patterns and excluding others.

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. [ROA-63].

Anttila, Arto and Curtis Andrus. 2006. 'T-order generator' [computer program], Stanford University. http://www.stanford.edu/~anttila/research/software.html. [ROA-873].

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

Boersma, Paul. 1997. 'How we learn variation, optionality, and probability', in *Proceedings of the Institute of Phonetic Sciences* 21, 43-58, University of Amsterdam. [ROA-221].

Boersma, Paul, and Bruce Hayes. 2001. 'Empirical tests of the Gradual Learning Algorithm', *Linguistic Inquiry* 32, 45-86. [ROA-348].

Carlson, Lauri. 1978. 'Word stress in Finnish', Ms., Massachusetts Institute of Technology, Cambridge, Massachusetts.

Crowhurst, Megan J. and Lev Michael. 2005. 'Iterative footing and prominence-driven stress in Nanti (Kampa)', *Language* 81, 47-95.

de Lacy, Paul. 2004. 'Markedness conflation in Optimality Theory', *Phonology* 21, 145-199.

Elenbaas, Nine. 1999. *A Unified Account of Binary and Ternary Stress*, Ph.D. dissertation, University of Utrecht. [ROA-397].

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

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

Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*, Chicago University Press, Chicago.

Hayes, Bruce and Zsuzsa Londe. 2006. 'Stochastic phonological knowledge: the case of Hungarian vowel harmony', *Phonology* 23, 59-104.

Hayes, Bruce, Bruce Tesar, and Kie Zuraw. 2003. 'OTSoft, Version 2.1' [computer program], http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Inkelas, Sharon. 1999. 'Exceptional stress-attracting suffixes in Turkish: representations vs. the grammar', in René Kager, Harry van der Hulst, and Wim Zonneveld (eds.), *The Prosody-Morphology Interface*, Cambridge University Press, Cambridge, pp. 134-87.

Karlsson, Fred. 1982. *Suomen kielen äänne- ja muotorakenne* [*The Phonological and Morphological Structure of Finnish*], Werner Söderström Osakeyhtiö, Helsinki.

Karttunen, Lauri. 2006. 'The Insufficiency of Paper-and-Pencil Linguistics: the Case of Finnish Prosody'. [ROA-818].

Karvonen, Dan. 2005. *Word Prosody in Finnish*, Ph.D. dissertation, University of California, Santa Cruz.

Kenstowicz, Michael. 1996. 'Quality-sensitive stress', *Rivista di linguistica* 9(1), 157-187. [ROA-33].

Keyser, Samuel Jay and Paul Kiparsky. 1984. 'Syllable structure in Finnish Phonology', in Mark Aronoff and Richard Oehrle (eds.), *Language Sound Structure*, MIT Press, Cambridge, Massachusetts, pp. 7-31.

Kiparsky, Paul. 1993. 'Blocking in nonderived environments', in Sharon Hargus and Ellen Kaisse (eds.), *Phonetics and Phonology*, *Vol. 4: Studies in Lexical Phonology*, Academic Press, San Diego, California, pp. 277-313.

Kiparsky, Paul. 1994. 'An OT Perspective on phonological variation', handout from Rutgers Optimality Workshop 1993, also presented at NWAVE 1994, Stanford University. http://www.stanford.edu/~kiparsky/Papers/nwave94.pdf.

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.

McCartney, Steven. 2003. *Ternarity through Binarity*, Ph.D. dissertation, University of Texas at Austin.

Prince, Alan S. 1990. 'Quantitative Consequences of Rhythmic Organization', in Michael Ziolkowski, Manuela Noske, and Karen Deaton (eds.), *Papers from the Chicago Linguistic Society* 26(2), pp. 355-398.

Prince, Alan. 2002a. 'Entailed ranking arguments'. [ROA-500].

Prince, Alan. 2002b. 'Arguing Optimality'. [ROA-562].

Prince, Alan. 2006. 'Implication and impossibility in grammatical systems: What it is and how to find it'. [ROA-880].

Prince, Alan. 2007. 'The pursuit of theory', in Paul de Lacy (ed.), *The Cambridge handbook of phonology*, Cambridge University Press, Cambridge, pp. 33-60.

Prince, Alan and Paul Smolensky 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*, Blackwell Publishing, Malden, Massachusetts. [ROA-537].

Ringen, Catherine, and Orvokki Heinämäki. 1999. 'Variation in Finnish vowel harmony: An OT account', *Natural Language and Linguistic Theory*, 17, 303-337.

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 on the University of Helsinki Language Corpus Server http://www.ling.helsinki.fi/uhlcs/.

Välimaa-Blum, Riitta. 1999. 'A feature geometric description of Finnish vowel harmony covering both loans and native words', *Lingua* 108, 247-268.

Zuraw, Kie. 2000. *Patterned Exceptions in Phonology*, Ph.D. dissertation, UCLA. [ROA-788].

email: anttila@stanford.edu