Finnish Consonant Gradation is a Stochastic Phonotactic Constraint

Arto Anttila

Stanford University

Fonologi i Norden (FiNo), Helsinki/Zoom

February 26, 2021

Finnish Consonant Gradation

Lenition

/p, t, k/ weaken in the onsets of closed syllables:

/ma**t**o-n/	ˈma.**d**on	‘worm-GEN’
/auriŋ**k**o-n/	ˈau.riŋ.**ŋ**on	‘sun-GEN’

Finnish Consonant Gradation

Lenition

/p, t, k/ weaken in the onsets of closed syllables

/ma**t**o-n/	ˈma.**d**on	‘worm-GEN’
/auriŋ**k**o-n/	ˈau.riŋ.**ŋ**on	‘sun-GEN’

Degemination

/pp, tt, kk/ shorten in the onsets of closed syllables

/ma**tt**o-n/	ˈma.**t**on	‘mat-GEN’
/ula**pp**a-n/	ˈu.la.**p**an	‘open.sea-GEN’

Analogical extensions (Hakulinen 1961:60, Kiparsky 2003)

Both alternations have spread outside their historical environment

- from **onsets of closed syllables (.CVC)**
- to **onsets of stressed heavy syllables (.ˈCVV).**

Analogical extensions (Hakulinen 1961:60, Kiparsky 2003)

Both alternations have spread outside their historical environment

- from **onsets of closed syllables (.CVC)**
- to **onsets of stressed heavy syllables (.ˈCVV)**

where

- they often remain variable (strong ~ weak)
- morphology seems to matter (e.g., case differences)
- lenition and degemination don't work alike

Puzzle 1: Apparent morphological conditions on lenition

Lenition is possible in the **partitive** and **genitive** plural…

(a) /auriŋko-i-ta/ ‘sun-PL-PAR’
ʔˈau.riŋ.ˌŋoi.ta ~ ˈau.riŋ.ko.ja (Lauri Viita)
/auriŋko-i-ten/ ‘sun-PL-GEN’
ʔˈau.riŋ.ˌŋoi.den ~ ˈau.riŋ.ko.jen

… but not in the **essive** and **illative** plural:

(b) /auriŋko-i-na/ ‘sun-PL-ESS’
*ˈau.riŋ.ˌŋoi.na / ˈau.riŋ.koi.na
/auriŋko-i-hVn/ ‘sun-PL-ILL’
*ˈau.riŋ.ˌŋoi.hin / ˈau.riŋ.koi.hin

Puzzle 2: No morphological conditions on degemination

Degemination is optionally possible everywhere:

(a) /logiikka-i-ta/ ‘logic-PL-PAR’
ˈlo.gii.koi.ta ~ ˈlo.giik.ko.ja
/logiikka-i-ten/ ‘logic-PL-GEN’
ˈlo.gii.koi.den ~ ˈlo.giik.ko.jen

(b) /logiikka-i-na/ ‘logic-PL-ESS’
ˈlo.gii.koi.na ~ ˈlo.giik.koi.na
/logiikka-i-hVn/ ‘logic-PL-ILL’
ˈlo.gii.koi.hin ~ ˈlo.giik.koi.hin

Summary

- Puzzle 1:

  Why does morphological case matter (e.g., PAR vs. ESS)?

  Answer: In fact, it doesn't. It is all phonology.

- Puzzle 2:

  Why do these effects only emerge in lenition?

  Answer: This is morphology (stems vs. words).

- Puzzle 3:

  Where does the variation come from?

  Answer: Probabilistic ranking/weighting.

The constraints

Σ = prominent syllable, i.e., either stressed or heavy

H = heavy syllable

The lenition constraint:

*Σ[−voice]Σ ‘Assign a violation to a short voiceless consonant between adjacent prominent syllables.’

The degemination constraint:

*HCCH ‘Assign a violation to a long consonant or a consonant cluster between adjacent heavies.’

Consonant gradation is a phonotactic constraint

ˈX.H.L.X		ˈX.H.ˌH.X	
ˈma.nee.re.ja	~	ˈma.neeˌrei.ta	‘mannerism-PL-PAR’
ˈma.nee.se.ja	~	??ˈma.neeˌsei.ta	‘manège-PL-PAR’

*Σ[−voice]Σ

Consonant gradation is a phonotactic constraint

ˈX.H.L.X		ˈX.H.ˌH.X	
ˈma.nee.re.ja	~	ˈma.neeˌrei.ta	‘mannerism-PL-PAR’
ˈma.nee.se.ja	~	??ˈma.neeˌsei.ta	‘manège-PL-PAR’

*Σ[−voice]Σ

ˈX.H.L.X		ˈX.H.ˌH.X	
?ˈsai.raa.lo.ja	~	ˈsai.raaˌloi.ta	‘hospital-PL-PAR’
ˈa.lus.to.ja	~	?ˈa.lus.ˌtoi.ta	‘base-PAR-PL’

*HCCH

What are your intuitions?

(a)	huokailuita	?huokaisuita	'sighing'
	kirmailuita	?kirmaisuita	'sprinting'
(b)	tehtailuita		'manufacturing'
	mahtailuita		'ostentation'
	puuhailuita		'tinkering'
(c)	?kirkaisuita		'screaming'
	?potkaisuita		'kicking'
	?karjaisuita		'roaring'

Preliminary statistical evidence

Preliminary support from an internet-based corpus of partitive plurals collected on April 12, 2005:

- 6,148 noun stems
- 9,280,395 partitive plural tokens, e.g., *maneereja* ~ *maneereita*

cluster = **TRUE** if the stem penult-ultima interlude has a consonant cluster in the *t*-retention form, else **cluster** = **FALSE**.

voice = **TRUE** if the stem penult-ultima interlude has no voiceless consonants in the *t*-retention form, else **voice** = **FALSE**.

Regression modeling

A logistic regression model (R Core Team 2019):

- Response:

  Partitive plural variant: *t*-deletion vs. *t*-retention

- Predictors:

  voice (= Σ[±voice]Σ)

  cluster (= H[CC/no CC]H)

- Controls:

  rv.del: syllable weight and vowel height

  (Anttila, Borgeson, and Magri 2019)

  total.freq.log: lexical frequency (logged)

Model summary

```
finn3.glm <- glm(cbind(finndata$del.freq, finndata$ret.freq) ~
   rv.del + total.freq.log + cluster + voice,
   data = finndata, family = binomial(link="logit"))
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-6.140e+00	1.362e-02	-450.74	<2e-16 ***
rv.del	1.413e-02	1.944e-05	726.92	<2e-16 ***
total.freq.log	4.163e-02	4.710e-04	88.37	<2e-16 ***
clusterTRUE	4.544e+00	5.866e-03	774.67	<2e-16 ***
voiceTRUE	-2.384e+00	4.366e-03	-546.07	<2e-16 ***

Interpretation

(a) In general, *t*-retention is favored (= faithfulness), but
[–voice] disfavors it, preferring *t*-deletion.

[+voice]: ˈ*ma.nee.*ˌ***r****ei.ta* favored, ˈ*ma.nee.****r****e.ja* disfavored

[–voice]: ˈ*ma.nee.****s****e.ja* favored, ˈ*ma.nee.*ˌ***s****ei.ta* disfavored

violates *Σ[–voice]Σ

Interpretation

(a) In general, *t*-retention is favored (= faithfulness), but
[–voice] disfavors it, preferring *t*-deletion.

[+voice]: ˈ*ma.nee.*ˌ***r****ei.ta* favored, ˈ*ma.nee.****r****e.ja* disfavored

[–voice]: ˈ*ma.nee.****s****e.ja* favored, ˈ*ma.nee.*ˌ***s****ei.ta* disfavored

violates *Σ[–voice]Σ

(b) In general, *t*-retention is favored (= faithfulness), but
CC-cluster disfavors it, preferring *t*-deletion.

No cluster: ˈ*sai.raa.*ˌ***l****oi.ta* favored, ˈ*sai.raa.****l****o.ja* disfavored

CC-cluster: ˈ*a.lu****s.t****o.ja* favored, ˈ*a.lu****s.***ˌ***t****oi.ta* disfavored

violates *HCCH

Puzzle 1: Why are partitive and essive different?

This is a side effect of SUFFIXAL CONSONANT GRADATION: a short voiceless stop /p, t, k/ is deleted outside a metrical foot (Keyser and Kiparsky 1984, Kiparsky 2003, Anttila 2012).

/maa-i-tA/	→	(mái.ta)	‘land-PL-PAR’
/talo-i-tA/	→	(tá.lo)ja	‘house-PL-PAR’
/korjaamo-i-tA/			‘repair shop-PL-PAR’
	→	(kór.jaa)(mòi.ta) ~ (kór.jaa.mo)ja	

The deletion applies to /t/ (partitive), but not to /n/ (essive).

The essive vs. partitive difference

Why optional lenition in the partitive, but not in the essive?

?(au.rin)(ŋoi.ta) ~ (au.rin.ko)ja	‘sun-PL-PAR’
*(au.rin)(ŋoi.na) ~ (au.rin.koi)na	‘sun-PL-ESS’

The essive vs. partitive difference

Why optional lenition in the partitive, but not in the essive?

?(au.rin)(ŋoi.ta) ~ (au.rin.ko)ja	'sun-PL-PAR'
*(au.rin)(ŋoi.na) ~ (au.rin.koi)na	'sun-PL-ESS'

Answer:

- We can leave /n/ outside a foot, but we can't leave /t/.
- Instead, we must scramble for less perfect options:
  - (a) Keep the suffixal/t/ and do gradation: ?(*au.rin*)(*ŋoi.ta*)
  - (b) Delete the suffixal /t/: (*au.rin.ko*)*ja*

The essive vs. partitive at the stem level: /aurinko/ ‘sun’

1. /aurinko-i-na/	*)t	*Σ[−voice]Σ	MAX(C)	ALIGN-L	ID([voice])
a. (ˈau.rin)(ˌkoi.na)		1!		2	
☞ b. (ˈau.rin.koi)na					
c. (ˈau.rin)(ˌgoi.na)				2	1
d. (ˈau.rin.goi)na					1
2. /aurinko-i-ta/	***)t**	***Σ[−voice]Σ**	**MAX(C)**	**ALIGN-L**	**ID([voice])**
a. (ˈau.rin)(ˌkoi.ta)		1!		2	
b. (ˈau.rin.koi)ta	1!				
☞ c. (ˈau.rin)(ˌgoi.ta)				2	1
d. (ˈau.rin.goi)ta	1!				1
☞ e. (ˈau.rin.ko)ja			1		

/n/ survives outside of a foot, /t/ deletes.

Outcome: Two partitives in free variation, deletion preferred.

Assumption: Stressed /oi/ is heavy, unstressed /oi/ is light (Keyser and Kiparsky 1984).

Puzzle 2: Why does this only happen in lenition?

This time the difference is truly morphological:

- *Σ[−voice]Σ (= lenition) is located at the stem level.
- *HCCH (= degemination) is located at the word level.

Puzzle 2: Why are lenition and degemination different?

This time the difference is truly morphological:

- *Σ[−voice]Σ (= lenition) is located at the stem level.
- *HCCH (= degemination) is located at the word level.

Prima facie evidence: Degemination counterfeeds lenition.

	/mato-n/ ‘worm-GEN’	/matto-n/ ‘mat-GEN’	
Stem level:	ˈma.don	--	
Word level:	--	ˈma.ton	
	ˈma.don	*ˈma.ton*	(→ **ˈma.don*)

No degemination at the stem level: /mellakka/ ‘riot’

1. /mellakka-i-na/	*)t	*Σ[−voice]Σ	MAX(C)	ALIGN-L	ID(clos)
a. (mel.lak)(koi.na)				2	
☞ b. (mel.lak.koi)na					
2. /mellakka-i-ta/	*)t	*Σ[−voice]Σ	MAX(C)	ALIGN-L	ID(clos)
☞ a. (mel.lak)(koi.ta)				2	
b. (mel.lak.koi)ta	1!				
☞ c. (mel.lak.ko)ja			1		

Here stem-level grammar simply takes care of footing.
No lenition, no degemination.

No degemination at the stem level: /mellakka/ ‘riot’

1. /mellakka-i-na/	*)t	*Σ[−voice]Σ	MAX(C)	ALIGN-L	ID(clos)
a. (mel.lak)(koi.na)				2	
☞ b. (mel.lak.koi)na					
2. /mellakka-i-ta/	***)t**	***Σ[−voice]Σ**	**MAX(C)**	**ALIGN-L**	**ID(clos)**
☞ a. (mel.lak)(koi.ta)				2	
b. (mel.lak.koi)ta	1!				
☞ c. (mel.lak.ko)ja			1		

- Two of these forms further degeminate at the word level:
  1b. *mel.lak.koi.na* ~ *mel.la.koi.na* (ESS) optionally
  2a. **mel.lak.koi.ta* / *mel.la.koi.ta* (PAR) obligatorily
- At the word level we have two new sources of variation: faithfulness to stem-level feet (FAITH-FT) and foot binarity (FTBIN) (Kiparsky 2003).

Optional degemination at the word level

(mel.lak)(koi.ta)	FTBIN	FAITH-FT	*HCCH	MAX(C)	ALIGN-L	ID(clos)
(mel.lak)(koi.ta)			2!		2	
☞(mel.la(k)oi.ta)			1		2	1
(mel.lak.koi)ta	1!	1!	1			
(mel.la(k)oi)ta	1	1	1			1
(mel.lak.ko)ja	FTBIN	FAITH-FT	*HCCH	MAX(C)	ALIGN-L	ID(clos)
☞ (mel.lak.ko)ja	1		1			
(mel.la(k)o)ja	1		1			1
☞ (mel.lak)(ko.ja)		1	1		2	
(mel.la(k)o.ja)		1	1		2	1
(mel.lak.koi)na	FTBIN	FAITH-FT	*HCCH	MAX(C)	ALIGN-L	ID(clos)
☞ (mel.lak.koi)na	1		1			
(mel.la(k)oi)na	1		1			1
(mel.lak)(koi.na)		1	2!		2	
☞ (mel.la(k)oi.na)		1	1		2	1

A somewhat unsettling discovery

This implies that degemination does not change syllable weight: it satisfies *HCCH, but keeps the preceding syllable heavy. In other words, the shortened geminate (k) is ambisyllabic.

☞(mel.la(k)oi.ta) = ˈH.H.ˌH.L (not ˈH.L.ˌH.L)
☞(mel.la(k)oi.na) = ˈH.H.ˌH.L (not ˈH.L.ˌH.L)

We learn that Finnish has “short geminates” that are moraic and ambisyllabic. This is good as it explains the puzzling opaque degemination in ˈ*ha.tu.ton* ‘hatless’ (cf. Kiparsky 1993).

☞(ha(t)u(t)on) = ˈH.H.H (not ˈL.L.H)

Conclusion: What is consonant gradation (CG)?

- In modern Finnish, CG has become a PHONOTACTIC CONSTRAINT that applies as a filter, beyond segmental alternations.

Conclusion: What is consonant gradation (CG)?

- In modern Finnish, CG has become a PHONOTACTIC CONSTRAINT that applies as a filter, beyond segmental alternations.
- CG targets the consonantal interlude between two PROMINENT SYLLABLES: Σ.Σ at the stem level, H.H at the word level.

Conclusion: What is consonant gradation (CG)?

- In modern Finnish, CG has become a PHONOTACTIC CONSTRAINT that applies as a filter, beyond segmental alternations.
- CG targets the consonantal interlude between two PROMINENT SYLLABLES: Σ.Σ at the stem level, H.H at the word level.
- Lenition happens at the STEM LEVEL, degemination happens at the WORD LEVEL.

Conclusion: What is consonant gradation (CG)?

- In modern Finnish, CG has become a PHONOTACTIC CONSTRAINT that applies as a filter, beyond segmental alternations.
- CG targets the consonantal interlude between two PROMINENT SYLLABLES: Σ.Σ at the stem level, H.H at the word level.
- Lenition happens at the STEM LEVEL, degemination happens at the WORD LEVEL.
- The constraint ranking/weighting at both levels is STOCHASTIC, which results in free variation.

References

Anttila, Arto. 2012. Modeling phonological variation. In Abigail C. Cohn, Cécile Fougeron, and Marie Huffman (eds), *The Oxford Handbook of Laboratory Phonology*, Oxford University Press, Oxford. pp. 76-91.

Anttila, Arto, Scott Borgeson, and Giorgio Magri. 2019. Equiprobable mappings in weighted constraint grammars. In *SIGMORPHON 2019: Proceedings of the 16th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*. https://arxiv.org/abs/1907.05839

Hakulinen, Lauri. 1961. *Suomen kielen rakenne ja kehitys* [*The Structure and Development of the Finnish Language*], 2nd revised edition, Otava, Helsinki.

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 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. 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.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.