French schwa in Harmonic Grammar

Brian Smith
UCLA

Joe Pater
UMass Amherst

4/1/2016
Linguistic Symposium on Romance Languages
Stonybrook (SUNY)

- Ongoing collaboration with Joe Pater

Today

- Experimental data on the phonological conditioning of two optional processes in French
  - Schwa epenthesis
  - Schwa deletion
- Both processes can be accounted for with weighted constraints, and both demonstrate ganging effects
- Takeaway: weighted constraints provide a natural fit for the data, capture both *cumulative* and *independent* effects of constraints

Plan

- The idealized (categorical) epenthesis pattern
  - Analysis in HG
  - Alternatives
- Experiment: variation in epenthesis and deletion
- A MaxEnt-HG model of variable epenthesis and deletion

Schwa background

- Definition: front rounded mid vowel that alternates with zero
- Phonetically [ø], [œ], or somewhere between

Background

- Dell (1973/1985) describes three levels of optionality
  - Forbidden schwa — Jacques l*ø* achète
  - Optional schwa — Marie l(e) vend
  - Obligatory schwa — Jacques l<u>e</u> vend
- I'll follow his notation, and mostly use his examples (all IPA transcriptions are by me)

Schwa epenthesis

- Described as obligatory (Léon 1966); occurs iff:
  - The epenthesis site is after a complex cluster
  - The site is followed by exactly one syllable
- Data for Verb+Noun compounds

	Word	IPA	Context of ə	Meaning
a.	gard<u>ə̲</u>-fou	**gaʁdə̲+fu**	VCC_σ	railing
b.	port<u>ə̲</u>-clefs	**poʁtə̲+kle**	VCC_σ	keychain

Schwa epenthesis

- No schwa if epenthesis site is followed by 2+ syllables

	Word	IPA	Context of ə	Meaning
c.	portɐ-manteau	**poʁt+mãto**	VCC_σσ	coat rack
d.	gardɐ-manger	**gaʁd+mãʒe**	VCC_σσ	cold kitchen

Schwa epenthesis

- No schwa if epenthesis site isn't preceded by a cluster

	Word	IPA	Context of ə	Meaning
e.	cassə-noix	kas+nwa	VC_σ	nutcracker
f.	piquə-nique	pik+nik	VC_σ	picnic
g.	coupə-papier	kup+papje	VC_σσ	paper cutter
h.	passə-partout	pas+paʁtu	VC_σσ	master key

Why analyze as epenthesis?

- Not just in compounds. Occurs at every morpheme boundary (if CC_σ), even if there's no orthographic 'e'

(i)	une veste rouge	[yn vɛstə̲ ʁuʒ]	(Dell: 224)
	une veste̸ rouge et blanc	[yn vɛst ʁuʒ ɛ blã]	
(j)	exacte+ment	[ɛgzaktə̲+mã]	(Dell: 228)
	massive̸+ment	[masiv+mã]	
(k)	un short vert	[ɛ̃ ʃɔʁtə̲ vɛʁ]	(Dell:237)

- Completely predictable → epenthesis (not deletion)

One more requirement

- Epenthesis is forbidden before a vowel

(I) notrø âmes /notʁ am/
our souls [notʁ am] *[notʁə am]

- Cannot create a schwa-V sequence (which is avoided throughout French, excepting h-aspiré)

An HG analysis of schwa epenthesis

- Lends itself to analysis with weighted constraints in Maximum Entropy Harmonic Grammar (MaxEnt: Goldwater & Johnson 2003)

- HG is like OT — with constraints and candidate sets — but constraints are weighted instead of ranked

- Results in **ganging:** one strong constraint can be overtaken by two weaker constraints together

- MaxEnt is a probabilistic variant of HG: outputs a probability distribution

Constraints

- Two independent requirements:
  - Requirement 1: Schwa must be in the penult
  - Requirement 2: Schwa must be after a cluster
- Translated into two constraints:
  - **PENULT = SCHWA**
  - ***CLUSTER**

Constraints

- ***CLUSTER**: Assign one violation for every coda cluster.
  - Well-documented effects across French, most famously Grammont's (1894) *La Loi de Trois Consonnes* (An early constraint: *CCC)
- Abstracting away from the effects of sonority, which have been noticed as early as Grammont
  - (The cluster in *livre* **[livʁ]** is more marked than *peste* **[pɛst]**)

Constraints

- **PENULT = SCHWA**: Assign one violation if the penultimate syllable of the Phonological Phrase is a non-schwa vowel

- Restated: pre-tonic syllable should contain the least sonorous vowel

  - Common across stress systems (de Lacy 2006)

  - Stress is phrase-final in French, and schwa ([ø]~[œ]) is the closest thing in French to a mid central vowel

Calculating Harmony

/gaʁd+fu/	*CLUSTER w=20	PENULT=ə w=10	DEP w=25	Harmony
gaʁ.d**ə̲**.fu	0	0	–1	**–25** (0*20)+(0*10)+(–1*25)
gaʁd.fu	–1	–1	0	**–30** (–1*20)+(–1*10)+(0*25)

Exponentiating

/gaɹd+fu/	*CLUSTER w=20	PENULT=ə w=10	DEP w=25	Harmony	$e^{Harmony}$
gaɹ.d**ə̲**.fu	0	0	–1	–25	**1.39×10^{-11}**
gaɹd.fu	–1	–1	0	–30	**9.35×10^{-14}**

Natural exponential function

Probabilities

/gaʁd+fu/	*CLUSTER w=20	PENULT=ə w=10	DEP w=25	Harmony	$e^{Harmony}$	Probability
gaʁ.d**ə̲**.fu	0	0	–1	–25	$1.39\ e^{-11}$	**0.99**
gaʁd.fu	–1	–1	0	–30	$9.35\ e^{-14}$	**<0.01**

Normalize: divide each candidate's e^{H} by sum of all candidates' in set

Both constraints violated: epenthesis

/gaʁd+fu/	*CLUSTER w=20	PENULT=ə w=10	DEP w=25	Harmony	$e^{Harmony}$	Probability
gaʁ.d**ə̲**.fu	0	0	–1	–25	$1.39\ e^{-11}$	**0.99**
gaʁd.fu	**–1**	**–1**	0	–30	$9.35\ e^{-14}$	**<0.01**

Ganging: two weaker constraints (*CLUSTER and PENULT=Ə) overcoming stronger constraint (DEP)

Just PENULT=ə: no epenthesis

/kas+nwa/	*CLUSTER w=20	PENULT=ə w=10	DEP w=25	Harmony	Probability
ka.s**ə̲**.nwa	0	0	–1	–25	**<0.01**
kas.nwa	0	**–1**	0	–10	**0.99**

Just *CLUSTER: no epenthesis

/gaɹd+malad/	*CLUSTER w=20	PENULT=ə w=10	DEP w=25	Harmony	Probability
gaɹ.d**ə̲**.ma.lad	0	–1	–1	–35	**<0.01**
gaɹd.ma.lad	**–1**	–1	0	–30	**0.99**

Summary of the MaxEnt analysis

- Two independent markedness constraints
  - *CLUSTER
  - PENULT = ə
- Weighted to produce a ganging effect: epenthesis only applies if it avoids violations of **both** constraints

Alternative accounts

- To compare: other accounts of French epenthesis capture the pattern without ganging or cumulativity
  - Charette 1991: (in GP) Epenthesis occurs after clusters, and epenthetic schwa is only licensed in the penultimate syllable
  - Coté 2007: (in OT) Cluster-driven epenthesis only occurs within PWds. *Garde-fou* is parsed as one PWd, *garde-malade* is parsed as two.

Why MaxEnt?

- All three accounts can handle the basic pattern
  - MaxEnt with two constraints and a ganging effect
  - Licensing in Charette (1991)
  - Prosodic analysis in Côté (2007)

Why MaxEnt?

- The MaxEnt analysis captures the pattern through the cumulative interaction of two independent constraints
  - If we find independent evidence for *CLUSTER and PENULT=SCHWA in French, the MaxEnt account is on the right track
  - If we need *CLUSTER and PENULT=SCHWA independently, why not take advantage of their cumulative interaction?

Why MaxEnt?
Empirical arguments

- **The rest of today**: both *CLUSTER and PENULT=SCHWA play a role in variable epenthesis and deletion
  - Independently of each other
  - Outside of the context VCC_σ
- We need **both** of the constraints to capture the full set of data

Variable epenthesis and deletion

Variation and epenthesis

- Côté (2007) describes epenthesis as variable

(m) la secte part	[la sɛktə̲ paʁ]	CCe̲ σ
la sect(e) partait	[la sɛkt(ə) paʁtɛ]	CC(e) σσ
l'Aztèqu**ɸ** part	[l aztɛk paʁ]	C**ɸ** σ

- Epenthesis most likely after clusters and followed by one syllable
- Generally occurs after clusters, regardless of position
- No independent effect of position

Clusters and deletion

- Schwa deletion is optional after a single consonant, but only if it doesn't create a coda cluster

  - tu le retrouves (Dell: 248)
    /ty lə ʁətʁuv/ [ty l_ ʁətʁuv] [ty lə ʁ_tʁuv]
    [ty lə ʁətʁuv] *[ty l ʁtʁuv]

- Doesn't matter where the resulting cluster is

  - la queue de ce renard (Dell: 248)
    /la kø də sə ʁənaʁ/ *[la kø d_ s_ ʁənaʁ]

Cluster and position

- Some coda clusters are possible outcomes of deletion, and these clusters show an effect of prosodic position (Dell: 231, citing Morin 1974)
  - la terre s̲e̲ vend /la tɛʁ sə̲ vã/
  - la terre s(e) vend bien /la tɛʁ s(ə) vã bjɛ̃/

Position alone

- There's an effect of position outside of coda clusters, although this effect is much more subtle

venez in Dell: 227:

/vəne isi/	/vəne/
*v**e**nez ici*	*v**e**nez*
<– Less schwa	More schwa –>

Summary of previous literature

- In both epenthesis and deletion:
  - Schwa is most likely to be pronounced in CC_σ
  - Schwa is generally more likely after clusters
- In deletion: schwa is more likely in C_σ than C_σσ (but the effect is weak)
- In epenthesis: C_σ and C_σσ are equal

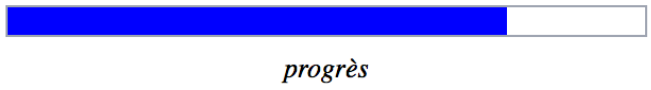
Next

- Although there are hints of independent effects of both constraints in descriptions, we don't know the actual probabilities of schwa
- An experiment to estimate the rates of deletion and epenthesis

Experiment

Experiment

- Web-based, through IbexFarm
- Two alternative forced choice, with confidence rating



Imaginez que vous parlez avec un ami. Est-ce que vous prononceriez le 'e'?

C'est un texte court

C'est un text' court

certainement **un texte court**	*probablement* **un texte court**	*probablement* **un text' court**	*certainement* **un text' court**

Design

- 2 x 2 x 2 factorial design
  - Cluster (C_ or CC_)
  - Position (_σ or _σσ)
  - Epenthesis / Deletion

Design: epenthesis

- **Noun** + **Adjective**
  - **Noun**: C-final or CC-final, all final Cs obstruents
  - **Adjective**: σ or σσ, all obstruent-initial

	C_	CC_
_σ	une bott**(e)** jaune [yn bɔt _ ʒon]	une vest**(e)** jaune [yn vɛst _ ʒon]
_σσ	une bott**(e)** chinoise [yn bɔt _ ʃinwaz]	une vest**(e)** chinoise [yn vɛst _ ʃinwaz]

Design: deletion

- **Name** + te + **Verb** (e.g. Maurice te cite)
  - **Name**: C-final or V-final, all final Cs obstruents
  - **Verb**: σ (present) or σσ (imperfect), all obstruent-initial

	C_	CC_
_σ	Eva t**(e)** choque [evat _ ʃok]	Maurice t**(e)** cite [moʁist _ sit]
_σσ	Eva t**(e)** choquait [evat _ ʃokɛ]	Maurice t**(e)** citait [moʁist _ sitɛ]

Design

- 78 judgments per participant
  - 24 deletion (6 per condition, no name or verb repeated)
  - 24 epenthesis (6 per condition, no adj. or noun repeated)
  - 30 fillers
    - Different tenses (future, past) and contexts (V_, _V, _σσσ)
    - 20 fillers for deletion (e.g. Anna s(e) est levée)
    - 10 fillers for epenthesis (e.g. un iguan(e) solitaire)

Predictions

- **Cumulativity**: schwa is most likely when it avoids violations of both constraints
- **Independence**: schwa is more likely when it avoids a violation of a single constraint
- Predicted probabilities of schwa:

C_σσ	<	CC_σσ, C_σ	<	CC_σ
Neither	<	*CLUSTER, PENULT=ə	<	Both

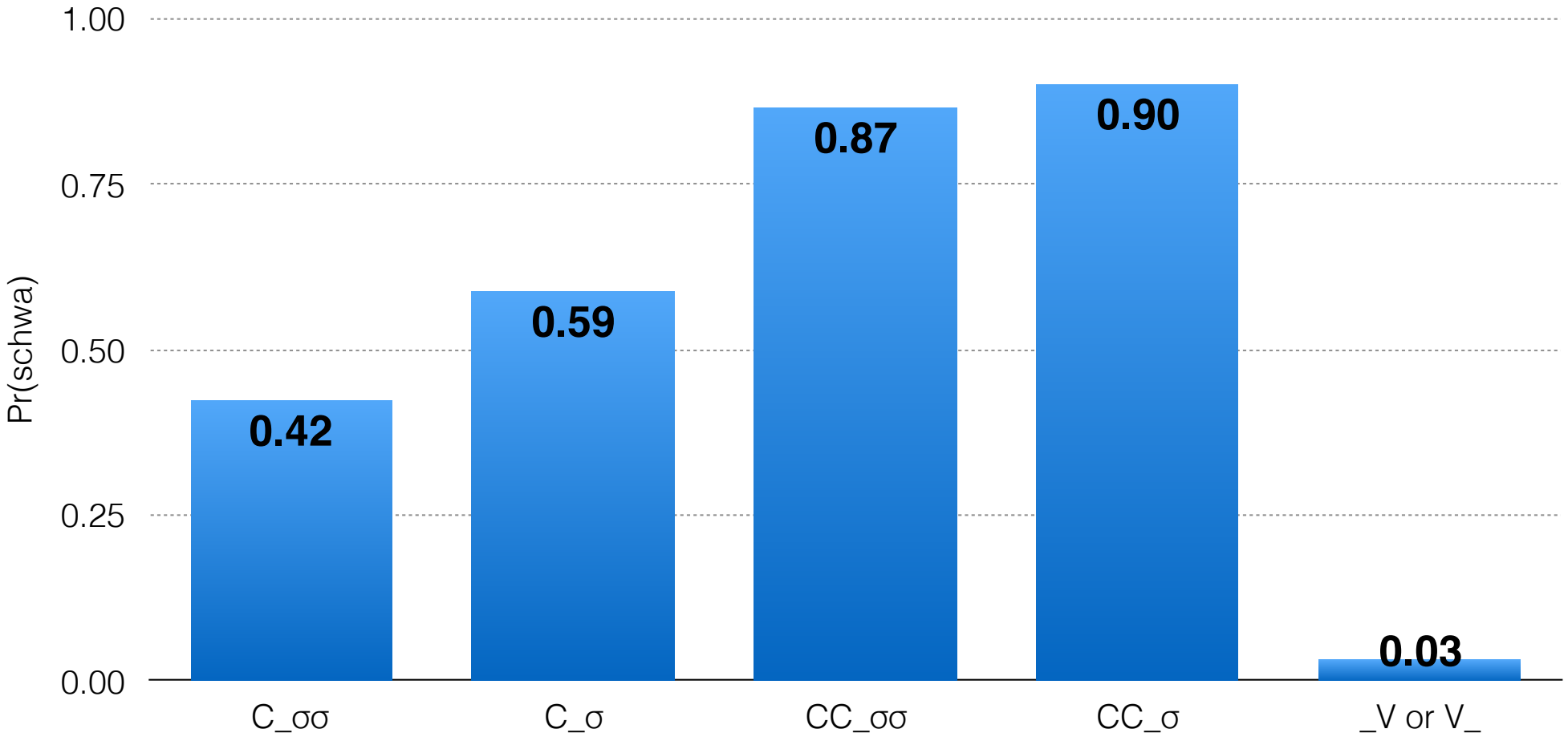
Participants

- Recruited online through word of mouth
- 51 respondents (ongoing)
- Preliminary results for 33 native French speakers who aren't from Canada

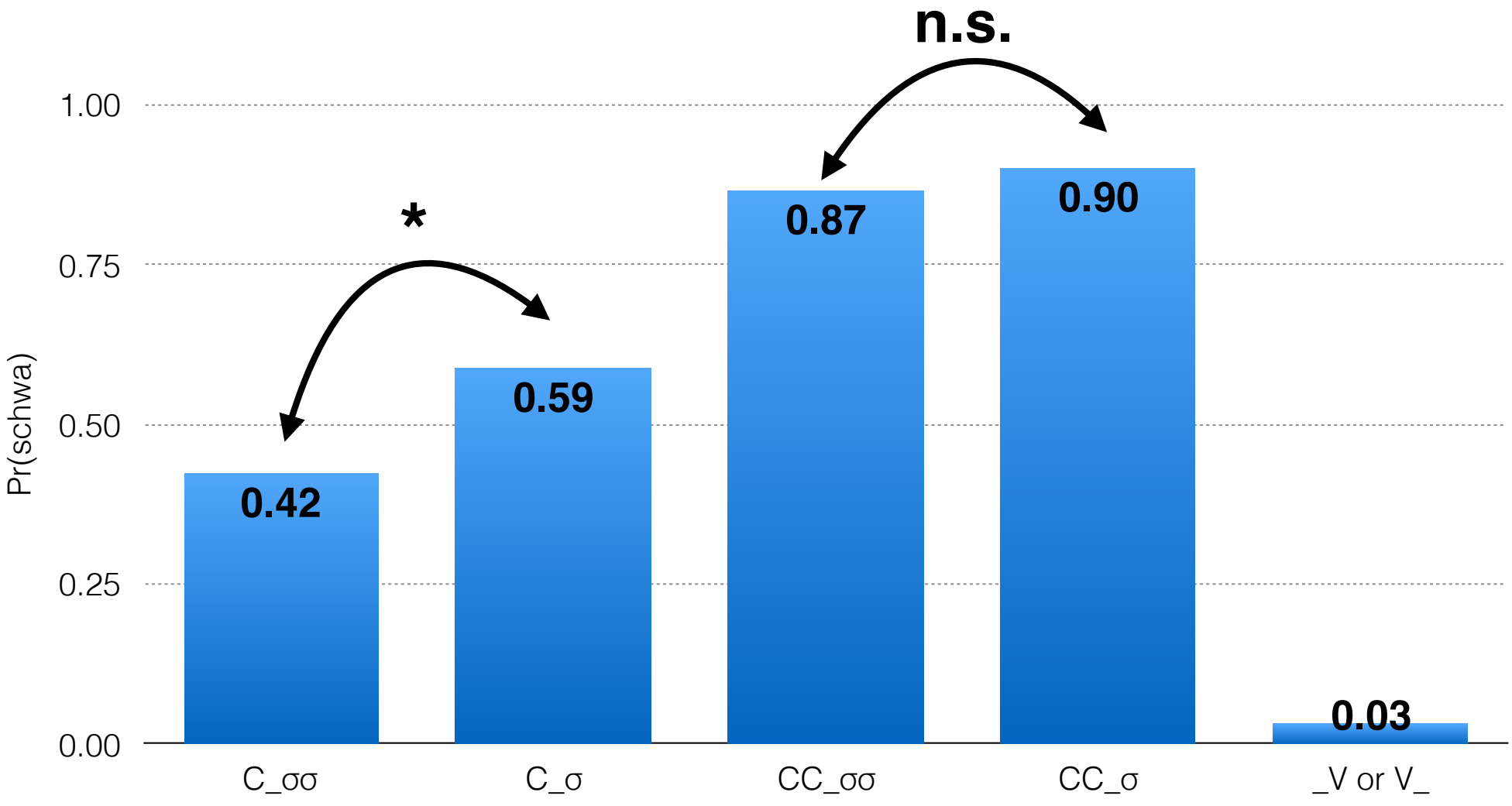
Results

- A lot of experimental noise — schwa is conditioned by geography, social factors, register
- Reaction time cutoffs
  - min RT = 100 ms, max RT = 9.3 hrs
  - Only considered responses between 3000s-7500s

Rate of schwa from experiment: deletion

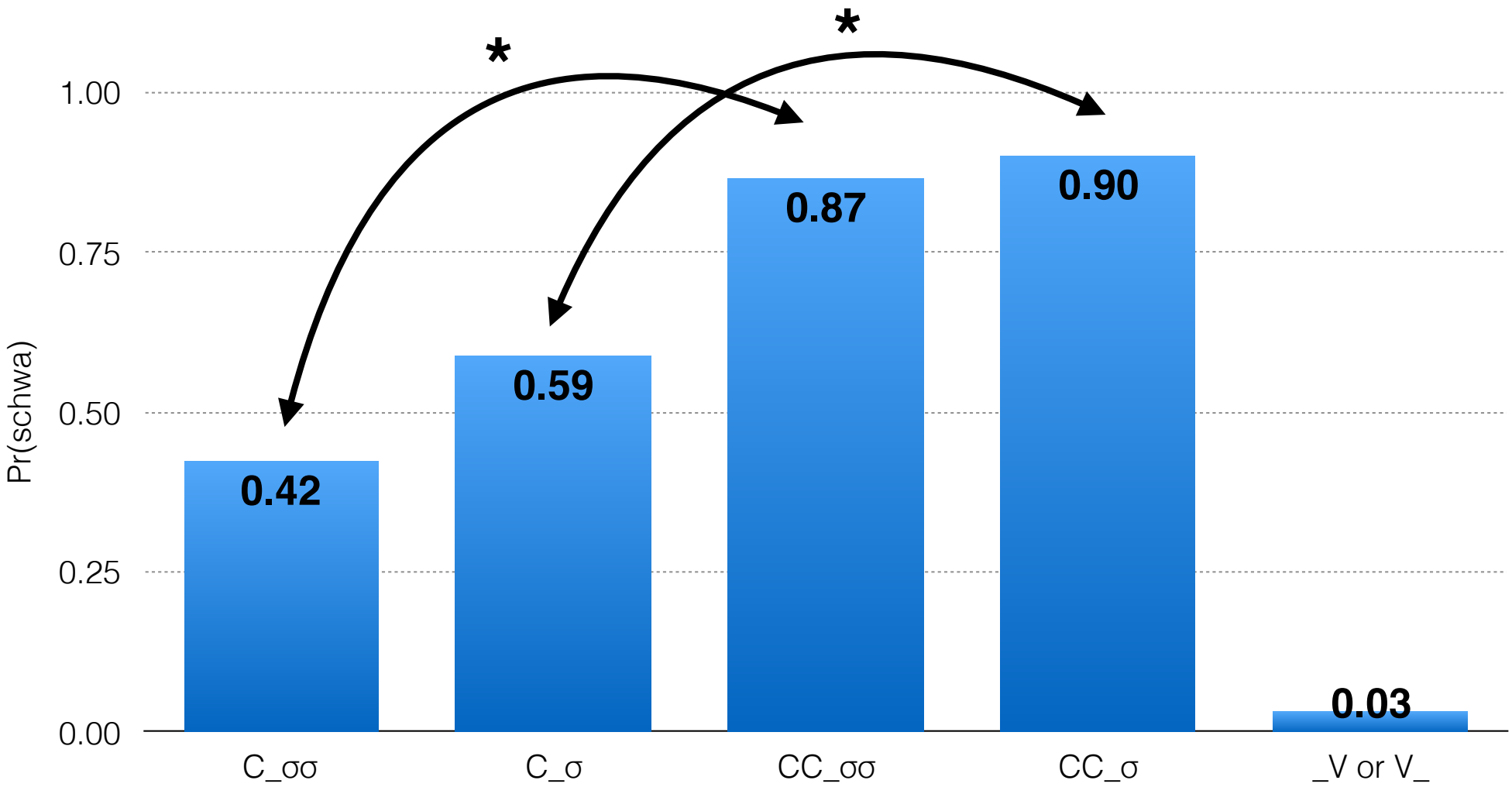


Effect of position in deletion



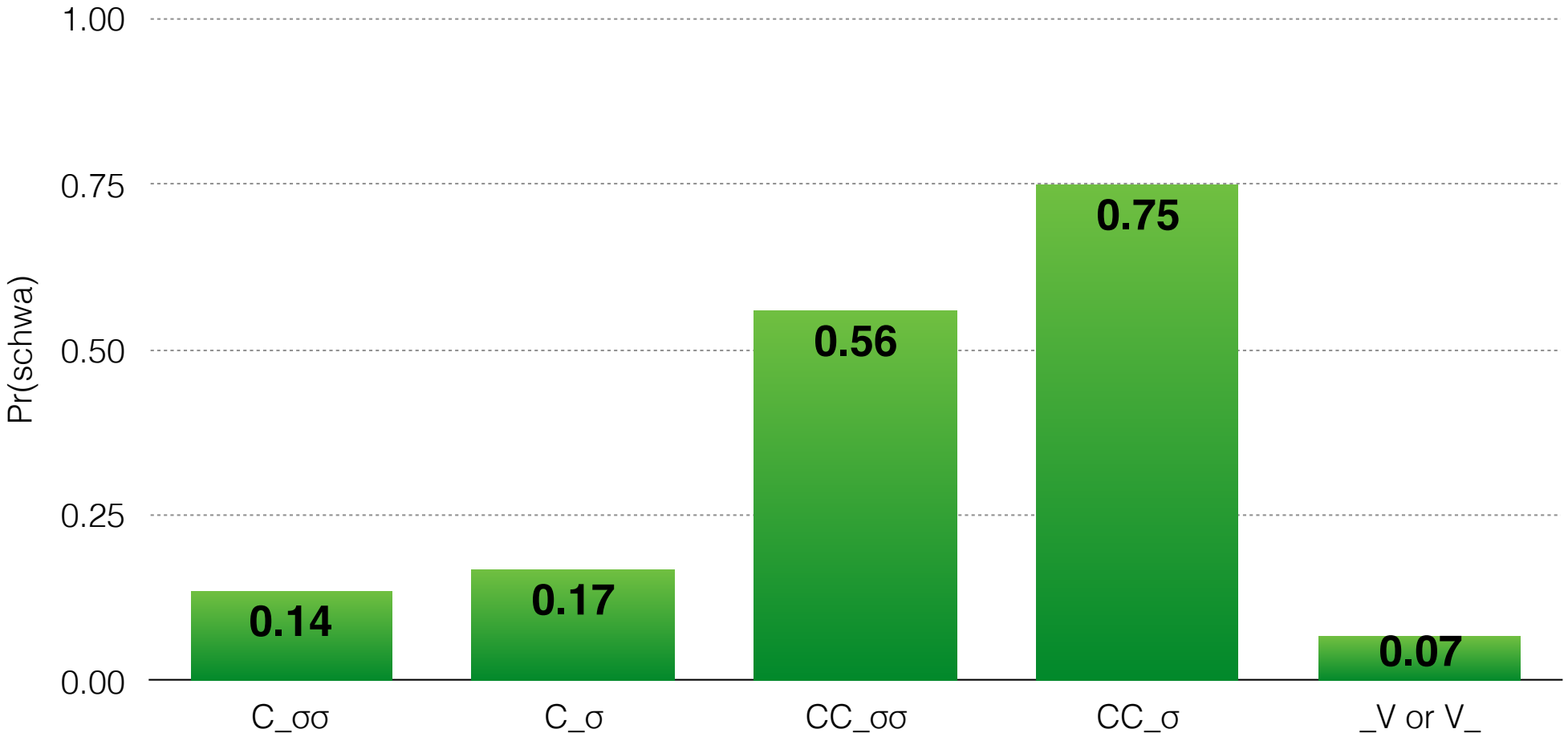
* = p<0.05 in chi-square test

Effect of cluster in deletion

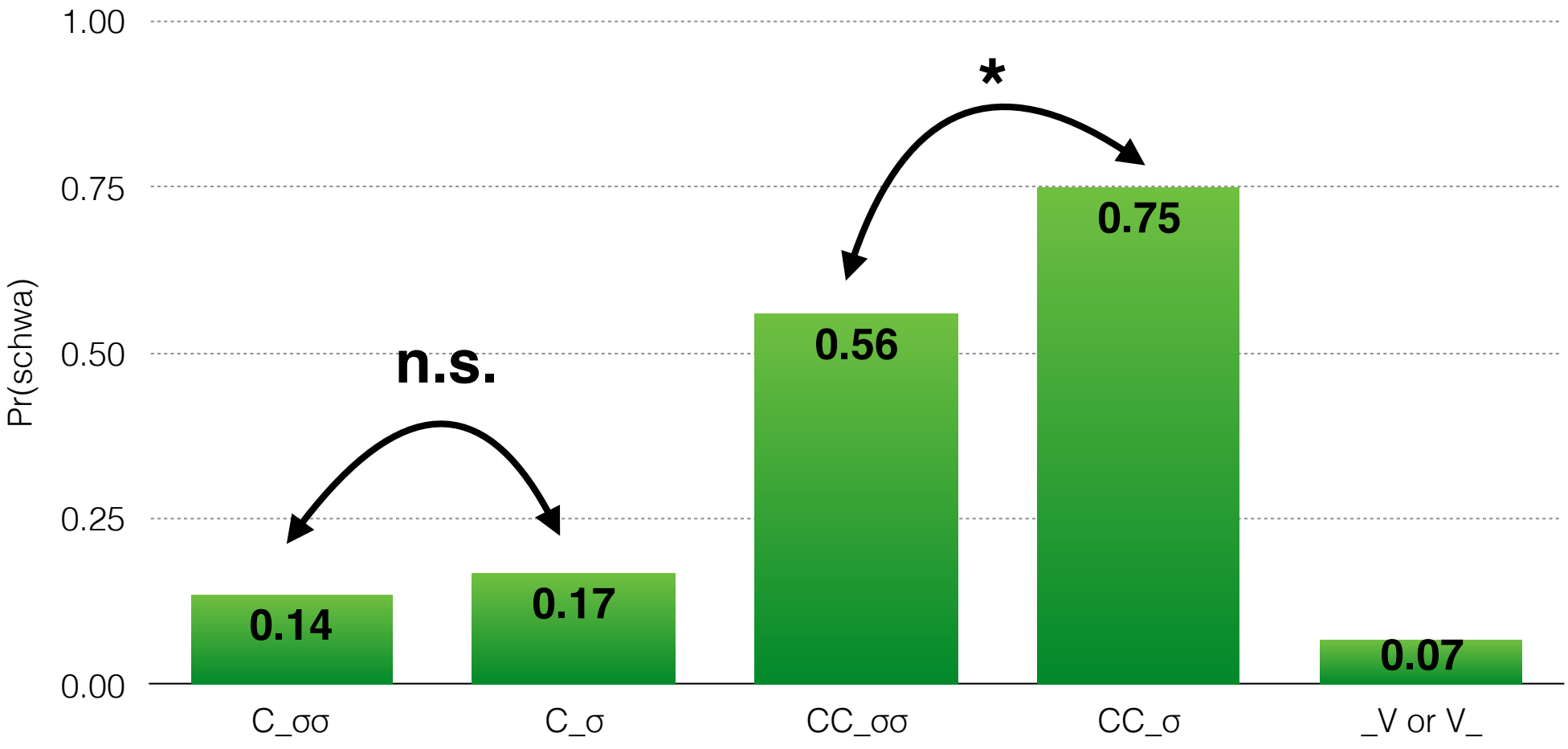


* = p<0.05 in chi-square test

Rate of schwa from experiment: epenthesis

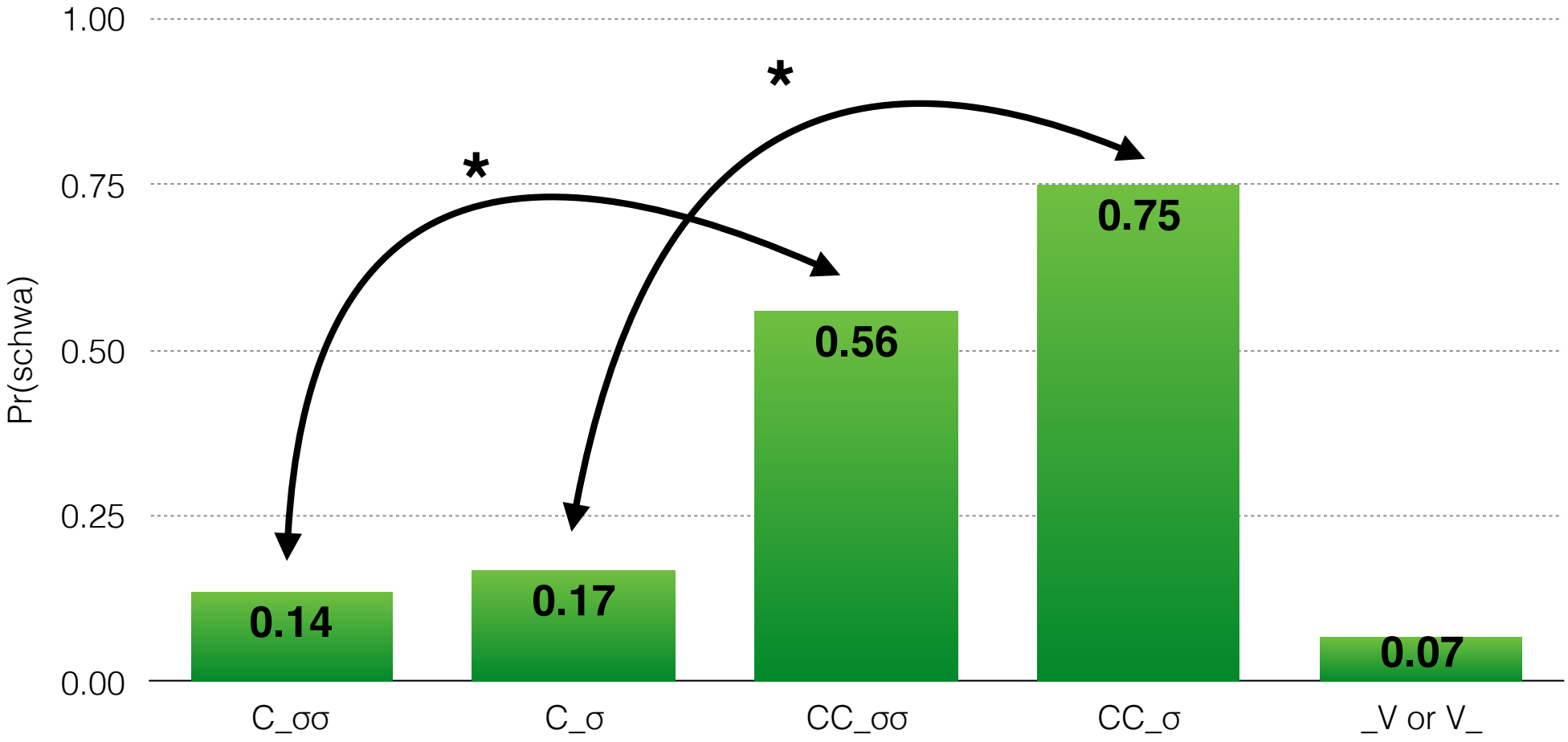


Effect of position



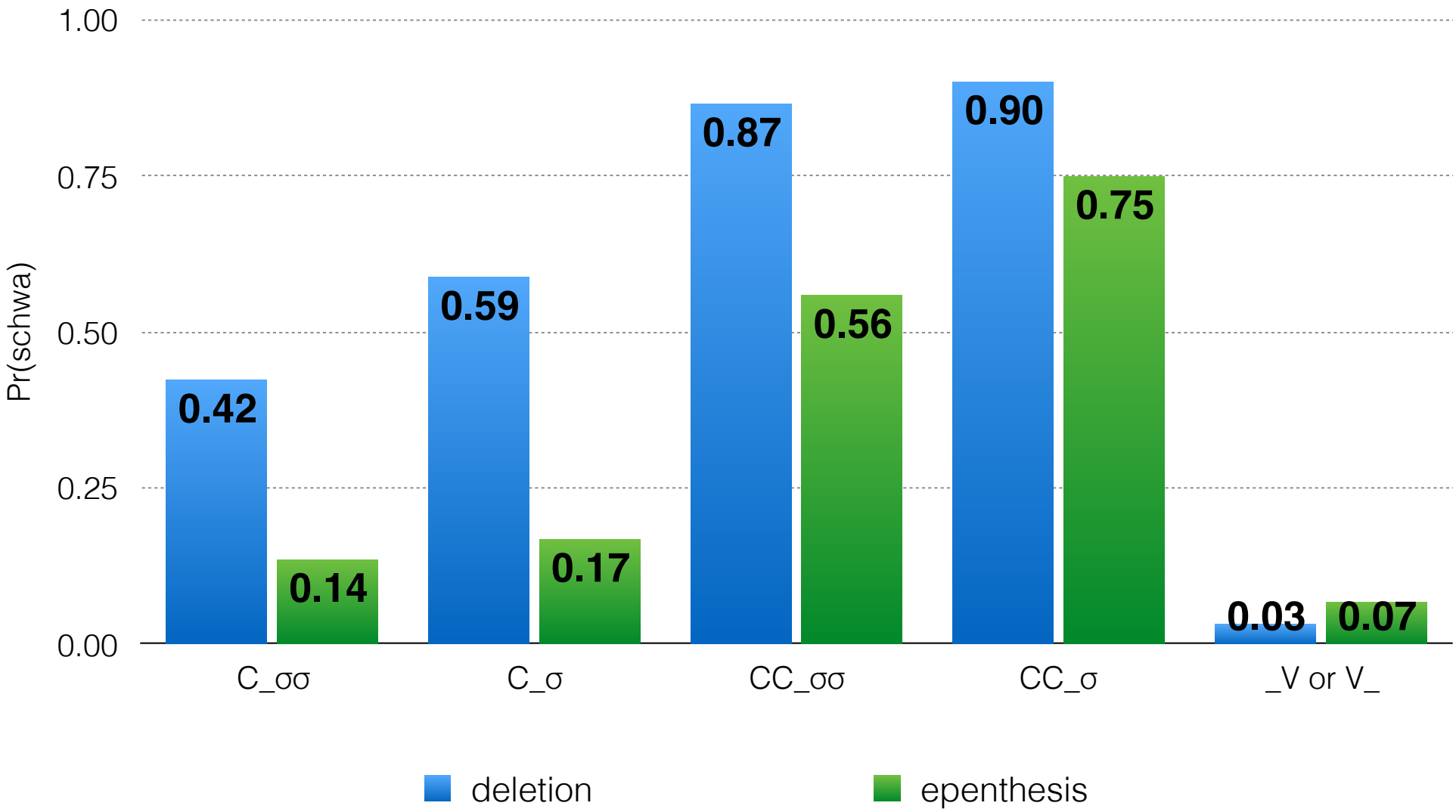
* = p<0.05 in chi-square test

Effect of cluster



* = p<0.05 in chi-square test

Epenthesis and deletion



Mixed effects logistic regression

- Fixed effects:
  - Epenthesis/deletion
  - Cluster
  - Position
  - Cluster x Position

- Random effects: intercepts for Subject & Item, random slopes for Subject for all fixed effects

Findings

- sig. effect for Cluster
  Pr(schwa): CC_ > C_
- sig effect of Position
  Pr(schwa): _σ > _σσ
- sig effect of Deletion/Epenthesis:
  Pr(schwa): deletion > epenthesis
- Effect of Cluster is greater than the effect of Position
- Interaction of Cluster x Position is not significant

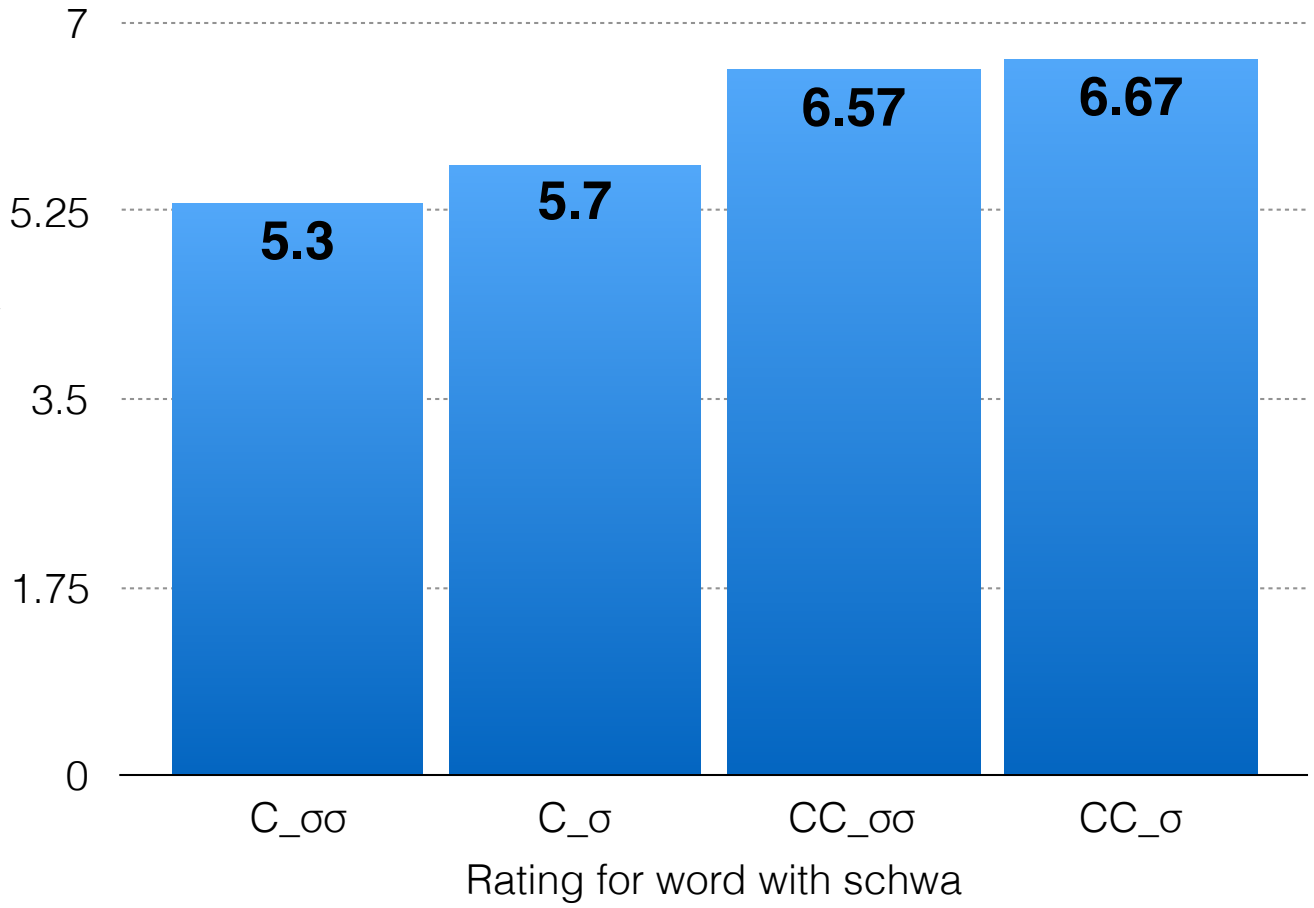
Data from Racine (2008)

- 12 speakers from Nantes
- Frequency judgments for single words with schwa
    - 1 = infrequent schwa, 7 = very frequent schwa
- Judgments for nearly 2,000 words with orthographic ‘e’

Ratings for deletion

Words in which schwa isn't at a morpheme boundary

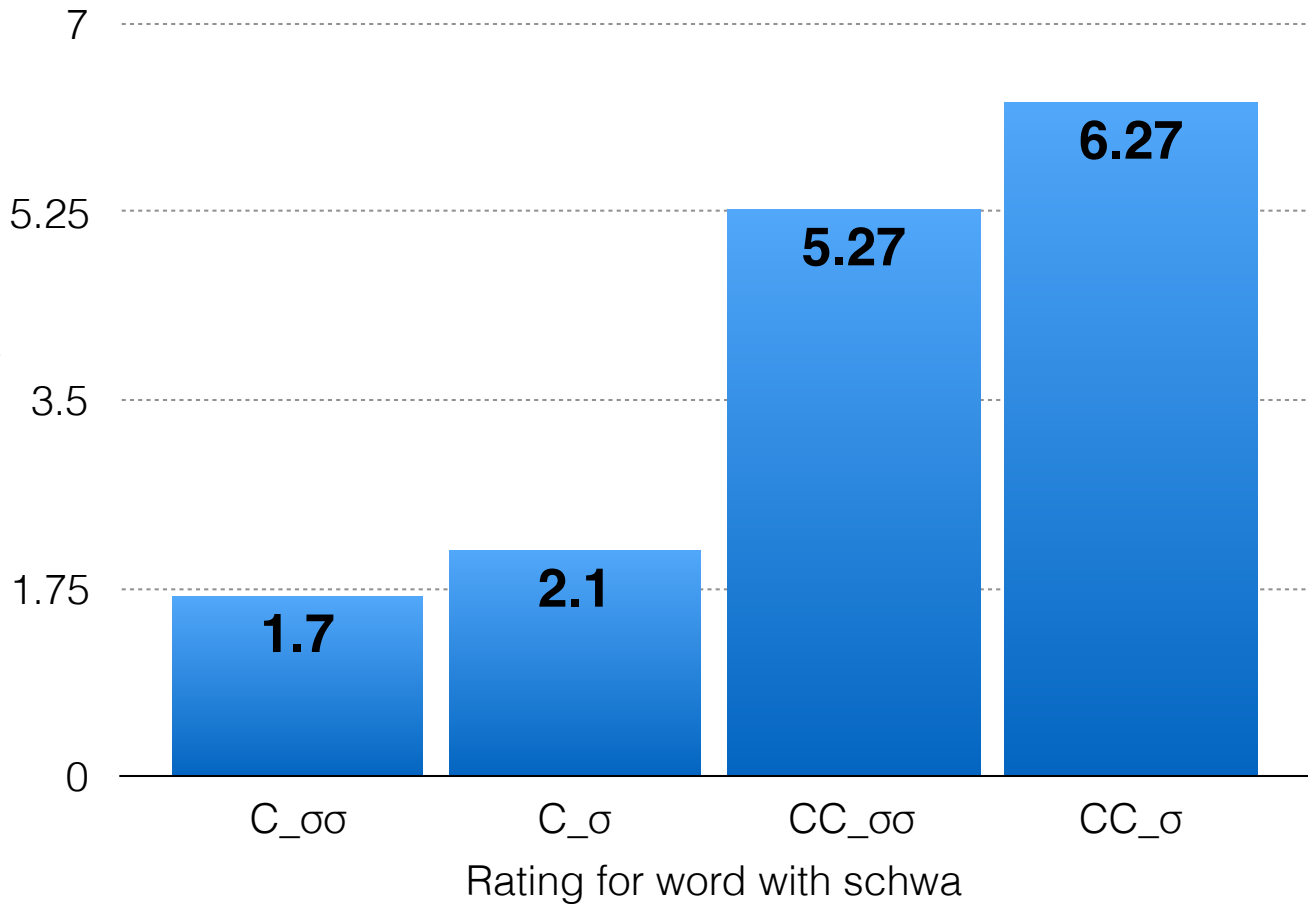
"Internal" schwas, e.g. d**e**v**e**nir



Ratings for epenthesis

Words in which schwa is at a morpheme boundary

e.g. brusq**ue**ment



A model of variable epenthesis and deletion

Epenthesis and deletion

- Epenthesis and deletion obey similar tendencies with respect to schwa
    - Not often modeled together
- A single weighted constraint grammar can model both epenthesis and deletion
    - The model captures all target generalizations
    - and generates probabilities close to experiment probabilities (within 5 percentage points)

Qualitative goals of the model

- Pr(schwa)
  C_σσ < C_σ < CC_σσ < CC_σ

  - **Cumulativity and independence:**
    Schwa is most likely when two conditions are met
    Schwa is least likely when zero conditions are met
    Both requirements have an effect (across all contexts)

- Cluster plays a bigger role than position

- Pr(schwa)
  Epenthesis < Deletion

Fitting the actual values

- *CLUSTER
- PENULT = ə
- *SCHWA Need some constraint to drive deletion
- MAX Constraint against deletion
- DEP Constraint against epenthesis

Target probabilities

Taken directly from experiment

Deletion: p(schwa)

	_σσ	_σ
CC_	**0.86**	**0.90**
C_	**0.42**	**0.59**

Epenthesis: p(schwa)

	_σσ	_σ
CC_	**0.56**	**0.75**
C_	**0.13**	**0.17**

Target probabilities

Taken directly from experiment

Deletion: p(schwa)

	_σσ	_σ
CC_	**0.86**	**0.90**
C_	**0.42**	**0.59**

Epenthesis: p(schwa)

	_σσ	_σ
CC_	**0.56**	**0.75**
C_	**0.13**	**0.17**

Learning

- Constraint weights found using MaxEnt grammar tool
  - http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/
- Learner is supplied target probabilities, inputs, outputs, and constraint violations
- Objective: maximize likelihood: find a set of weights that matches the target probabilities as closely as possible

Model probabilities (Target probabilities)

Constraint	Weight
*CLUSTER	2.16
PENULT=Ə	0.69
*SCHWA	0.27
MAX	0.00
DEP	1.66

Deletion: Pr(schwa)

	_σσ	_σ
CC_	**0.87** (0.86)	**0.93** (0.90)
C_	**0.43** (0.42)	**0.60** (0.59)

Epenthesis: Pr(schwa)

	_σσ	_σ
CC_	**0.55** (0.56)	**0.71** (0.75)
C_	**0.13** (0.13)	**0.22** (0.17)

Constraints have weaker effects at the margins

- In the data and model, PENULT=Ə has a weaker effect when the rate of schwa is closer to 0% and 100%, and a stronger effect when closer to 50%

- This falls out of the math of the MaxEnt model, without interaction terms or special constraints (see McPherson & Hayes 2015 for an application of this)

Deletion (model)

	_σσ	_σ	Diff
CC_	0.87	0.93	**0.06**
C_	0.43	0.60	**0.27**

Epenthesis (model)

	_σσ	_σ	Diff
CC_	0.55	0.71	**0.16**
C_	0.13	0.22	**0.09**

Modeling summary

- Model captures cumulativity in both epenthesis and deletion
  - Both processes are conditioned by *CLUSTER and PENULT=Ə, most likely when both constraints are applicable
  - Both constraints have independent effects
- The fact that the constraints have weaker effects at margins falls out of MaxEnt

Conclusion

- Two types of analysis for epenthesis in VCC_σ
  - MaxEnt: Two constraints and ganging effects
  - Others: Constraints specific to VCC_σ
- For the categorical data, these approaches are equal
- For the variable data, MaxEnt captures the fact that both constraints have independent effects
  - Condition both epenthesis and deletion, outside of context VCC_σ

General conclusion

- Weighted constraints allow us to capture patterns with fewer constraints
  - Thanks to ganging effects
  - In this case, matching probabilities for 8 inputs with a simple 4 constraint grammar
- Weighted constraints provide a straightforward model of variation, with machine-learnable parameters

Thank you

http://www.linguistics.ucla.edu/people/bsmith/

Acknowledgments

- Thanks to many French speakers and learners for help with stimuli and recruitment, especially:
  - Magda Oiry
  - Isabelle Lin
  - Yu Tanaka
  - Kie Zuraw
- Thanks to Isabelle Racine for sharing French data

- Charette, Monik (1991). Conditions on Phonological Government. Cambridge, UK: Cambridge University Press.
- Côté, Marie-Hélène (2007). Rhythmic constraints on the distribution of schwa in French. In Romance Linguistics 2006, José Camacho, Nydia Flores-Ferrán, Liliana Sánchez, Viviane Déprez & María José Cabrera, eds. Amsterdam: John Benjamins, 79-92.
- de Lacy, Paul (2002). The formal expression of markedness. PhD Dissertation. University of Massachusetts Amherst.
- Dell, François (1985). Les règles et les sons, 2nd ed. Paris: Hermann.
- Goldwater, Sharon, & Mark Johnson (2003). Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson, and Osten Dahl (eds.), Proceedings of the Stockholm Workshop on Variation within Optimality Theory, 111–120.
- Grammont, Maurice (1894). La loi des trois consonnes. Mémoires de la société de linguistique de Paris 8, 53-90.
- Léon, Pierre (1966). Apparition, maintien et chute du e caduc. La linguistique 2, 111-122.

Thank you

http://www.linguistics.ucla.edu/people/bsmith/