Evidence For Weighted Constraints: Epenthesis In French

Abstract

In French, a process of optional schwa epenthesis is driven by both a prosodic and a segmental constraint. Each of these constraints has an independent effect, but epenthesis is most likely when it satisfies both constraints. The relationship between these constraints is determined by analyzing experimental data collected by Racine (2008), in which speakers rate the likelihood of schwa across different contexts. The constraints' interaction is straightforward captured by a theory of variation with weighted constraints, such as Noisy Harmonic Grammar (Boersma & Pater 2008), in which constraint violations can combine additively to create gang effects.

1. Introduction

This paper presents evidence in favor of an analysis of French epenthesis using weighted constraints. The conclusion is that patterns of variation such as the one in French lend themselves to analysis in weighted-constraint frameworks like Noisy Harmonic Grammar (Noisy HG: Boersma & Pater 2008) or Maximum Entropy Grammars (MaxEnt: Goldwater & Johnson 2003), as opposed to models of variation with ranked constraints. Before explaining the predictions of weighted vs. ranked constraints, I consider the basic pattern of epenthesis which will be discussed throughout the paper.

1.1. The Pattern

In French, a process of variable schwa epenthesis is conditioned by two independent constraints. A segmental constraint requires the insertion site to be after two consonants (VCC_), and a prosodic constraint requires the insertion site to be followed by a single syllable (_σ). Epenthesis is most likely to occur in VCC_σ contexts (1a-b), where it meets both conditions, and less likely to occur in VC_ or _σσ contexts (1c-h), where it meets fewer. The data below are from Charette (1991), who describes an idealized pattern in which epenthesis is completely blocked for the shaded words. The epenthetic vowel will be represented throughout the paper by [ə], although its actual phonetic realization varies between [ø] and [œ] (see Côté 2001).

1. Epenthesis only occurs in VCC_σ[1] (data from Charette 1991, transcriptions my own)

	Word	IPA	Context	...VCC _... ?	... _σ# ?
a.	garde-fou	[gaʁdə + fu]	VCC_σ	√	√
b.	porte-clefs	[poʁtə + kle]	VCC_σ	√	√
c.	port(*e)-manteau	[poʁt + mɑ̃to]	VCC_σσ	√	
d.	gard(*e)-manger	[gaʁd + mɑ̃ʒe]	VCC_σσ	√	
e.	casse(*e)-noix	[kas + nwa]	VC_σ		√
f.	piq(*ue)-nique	[pik + nik]	VC_σ		√
g.	coup(*e)-papier	[kup + papje]	VC_σσ		
h.	pass(*e)-partout	[pas + paʁtu]	VC_σσ		

Although all of the examples in (1) are compounds, the same pattern can be observed between separate words (Noske 1996: film(ə) russe vs. film doublé) and within words at prefix boundaries (as shown in §3: ex(ə)-femme vs. ex-mari).

1.2. Weighted V. Ranked Constraints

A natural way to account for VCC_σ epenthesis is to state the two conditions on epenthesis as weighted constraints which interact cumulatively in Harmonic Grammar (HG: Legendre et al. 1990; Smolensky and Legendre 2006). In Harmonic Grammar, optimality is calculated by summing weighted constraint violations. Because violations are summed, constraints can gang up to force the repair of doubly-marked forms, without forcing the repair of singly-marked forms. The French generalization can be modeled as a gang effect between a segmental constraint, militating against complex clusters, and a prosodic constraint, requiring schwa in the penultimate syllable.

2. The segmental constraint:
   *CLUSTER: Assign one violation mark for every complex coda.

3. The prosodic constraint:
   *PENULT≠ə: Assign one violation mark if the penult vowel is not schwa.

Schwa epenthesis occurs in VCC_σ, where it avoids a violation of both constraints, but not in VC_σ or VCC_σσ, where it only avoids a violation of one constraint. The faithfulness constraint DEP sets a markedness threshold for epenthesis, preventing epenthesis in cases where it only satisfies one of the markedness constraints.

[1] '+' = morpheme boundary; '#' = word boundary; '=' = clitic boundary

4. The faithfulness constraint:
DEP: Assign one violation mark for every segment without an input correspondent.

These three constraints are not sufficient to account for the data in an Optimality Theoretic analysis with ranked constraints (OT: Prince & Smolensky 1993/2004). The problem is OT's requirement for strict constraint domination. In OT, a constraint that favors epenthesis after complex clusters will make no distinction between VCC_σ and VCC_σσ. If *CLUSTER ≫ DEP, epenthesis will occur in both contexts. Likewise, a constraint that favors schwa in the penult will make no distinction between VC_σ and VCC_σ, and epenthesis will occur in both contexts if *PENULT≠ə ≫ DEP. With the three constraints *CLUSTER, *PENULT≠ə, and DEP, there's no ranking in which epenthesis occurs in VCC_σ but not VCC_σσ or VC_σ.

What the OT analysis requires is a constraint that penalizes VCC_σ alone.[2] This constraint must militate against complex clusters in the penultimate syllable. The OT constraint is represented by the constraint below, which is stated as a conjoined constraint (constraint conjunction: Smolensky 1995, Smolensky & Legendre 2006).

5. Example of a constraint needed for an OT analysis
*PENULT≠ə & *CLUSTER: Assign one violation mark if the penult vowel is not schwa and there is a complex coda.

The main reason to doubt this constraint is that it has no external motivation and French, which follows from the fact that it exists only to model a particular pattern of epenthesis.

Putting aside the plausibility of the conjoined constraint, on the surface the weighted constraint and ranked constraint accounts are indistinguishable. They both predict epenthesis in the context VCC_σ, but not elsewhere. When both accounts are possible, the usual way to decide between an HG and OT analysis is typological predictions. The HG account or OT account may be argued to over- or under-generate for a given set of constraints. This paper presents a new sort of argument for using weighted constraints over ranked constraints, using a single phonological phenomenon from French.

The French case is able to bear on the differences between ranked and weighted constraints because it is subject to variation. Contrary to the idealized data in (1), French epenthesis is optional within speakers, and the rate of epenthesis varies across contexts. Variation can be modeled in stochastic variants of HG and OT, — Noisy HG and Stochastic OT (Noisy HG: Boersma & Pater 2008, Pater, Potts, & Bhatt 2007; SOT: Boersma 1997, 1998, Boersma & Hayes 2001). In Noisy HG and SOT, constraint rankings/weightings are evaluated with noise, creating probability distributions over

[2] Alternatives involving other constraints, e.g., licensing constraints, are discussed in §6.

candidates. Unlike the deterministic HG and OT accounts, the stochastic frameworks diverge in their predictions for a case like French schwa.

PREDICTIONS ABOUT INDEPENDENT EFFECTS: In a Noisy HG analysis of French, epenthesis is most likely in VCC_σ contexts due to a gang effect between the segmental and prosodic constraint. Since the process of epenthesis is driven by two independent constraints, we might expect each of the constraints to exhibit an independent effect in the data[3]. In an analysis in which VCC_ σ epenthesis is driven by a single constraint, there is no expectation of an independent effect of *CLUSTER or *PENULT≠ə.

PREDICTIONS ABOUT EFFECT SIZE: The move from standard OT to SOT means ranked constraints are able to capture some probabilistic patterns resembling gang effects, which I'll call *pseudo-gang-effects*. This means that in an SOT grammar, two independent constraints can have an additive effect given the right conditions. The difference between a pseudo-gang-effect in SOT and a gang effect in Noisy HG is that a SOT grammar requires the effect size of the two constraints to be roughly equal for them to gang up. Noisy HG, on the other hand, does not require the effect sizes of the constraints to be equal — one constraint can be high-weighted, another can be low-weighted, and they can still gang up.

1.3. Paper Overview

This paper shows that the data in French are straightforwardly captured in Noisy HG. Both constraints that participate in the gang effect have an independent effect in French, and the effect size of one of these constraints is much greater than the other. As a result, given the same set of constraints, the pattern can be learned by a Noisy HG learner but not a SOT learner.

In Sections 2 and 3, I show that both the segmental constraint and the prosodic constraint exhibit an independent effect on epenthesis. This claim finds support in both the French literature, in which each constraint has been reported to independently condition alternations, and in an analysis of experimental data collected by Racine (2008), in which each constraint has an independent effect on frequency ratings of schwa alternations. The rough probabilities from the experiment are given in (6).

6. p(schwa): VCC_σ > VCC_σσ >> VC_σ > VC_σσ

Schwa epenthesis is more likely in VCC_, regardless of prosodic position. Likewise, schwa epenthesis is more likely when in _σ for both segmental contexts.

[3] Although a Noisy HG grammar can also capture a pattern in which neither constraint shows an independent effect.

In Sections 4 and 5, I show how the findings in (6) are consistent with the weighted constraint account sketched above. The effect of the segmental constraint (comparing VCC_ and VC_) is much larger than the effect of the prosodic constraint (comparing _σ and _σσ). Given the same set of constraints, a Noisy HG learner is able to learn the pattern, while a SOT learner is not.

2. Background On French Schwa

This section presents evidence from the literature supporting the independence of the prosodic and segmental constraints. In Section 5, the independence of these constraints challenges a number of analyses in which VCC_ σ epenthesis is driven by a single constraint.

2.1. The Basic Pattern

The basic pattern of VCC_σ epenthesis was first observed by Léon (1966). Epenthesis requires the insertion site to be preceded by two consonants and followed by one syllable. This observation holds for compound boundaries (7a–7b), independent word boundaries (7c–d), and clitic boundaries (7e–7f). The data in (7) are repeated from Côté (2007).

7. Data from Côté (2007)

	Word	Context	IPA
a.	garde-fou	VCC_σ	[gaʁdə+fu] *[gaʁd+fu]
b.	garde-malade	VCC_σσ	[gaʁd(ə)+malad]
c.	la secte part	VCC_σ	[la=sɛktə#paʁ] ?[la=sɛkt#paʁ]
d.	la secte partait	VCC_σσ	[la=sɛkt(ə)#paʁtɛ]
e.	achète de l'or	VCC_σ	[aʃɛt#də=lɔr] ?[aʃɛt#d=lɔr]
f.	achète de l'ortie	VCC_σ	[aʃɛt#d(ə)=lɔʁti]

VCC_σ epenthesis can also occur within words at prefix boundaries.

8. VCC_σ epenthesis occurs within words (original data)

	Word	Context	IPA
a.	ex-femme	VCC_σ	[ɛks(ə)+fam]
b.	ex-mari	VCC_σσ	[ɛks+maʁi]

The data in (9) are the number of Google search hits for alternative spellings of *ex-femme* and *ex-mari*. Speakers are much more likely to spell *ex-femme* with an extra *e* than *ex-mari*, which may reflect optional epenthesis. In the table below, *une ex(e)-femme* is spelled with an extra *e* nearly eight times more than *un ex(e)-mari*, and *mon ex(e)-femme* is spelled with an extra *e* 15 times more than *mon ex(e)-mari*.

9. VCC_σ epenthesis occurs within words: number of Google hits on 10/18/2011

	Word	Context	mon exe-x	mon ex-x	un(e) exe-x	un(e) ex-x
a.	ex(e̲)-femme	VCC_σ	2540	128M	7	273k
b.	ex(e̲)-mari	VCC_σσ	196	151M	2	620k
	example search phrase		"mon exe femme" site:.fr		"une ex femme" site:.fr	

In all of the cases of VCC_σ epenthesis above, schwa occurs at a morpheme boundary. Schwa epenthesis is completely blocked elsewhere, as shown below.

10. VCC_σ epenthesis only occurs at morpheme boundaries

	Word	Context	IPA
a.	extrême	VCC_CV	[ɛkstʁɛm] *[ɛks(ə)tʁɛm]
b.	ex-femme	VCC_+CV	[ɛks(ə) + fam]
c.	garde-role	VCC_+CV	[gaʁd(ə) + ʁɔl]
d.	gare drôle	VC#C_CV	[gaʁ#dʁɔl] *[gaʁ#d(ə)ʁɔl]

It should also be noted that schwa never occurs at the end of a phrase or adjacent to a vowel (Tranel 1987). While epenthesis is optional in *garde-malade* and *garde-fou*, where schwa is between consonants phrase-medially, epenthesis is forbidden in *garde!* (VCC_#) and *garde au chien* (VCC_V).

2.2. The Segmental Constraint

The distribution of French schwa has been linked to consonant clusters (independent of prosody) since Grammont's (1894) La loi de trois consonnes, which states that schwa is retained when its deletion would result in a sequence of three consonants. A large body of work argues for the existence of cluster-conditioned schwa alternations, including Pulgram (1961), Morin (1974), Cornulier (1975), Bouchard (1981), Anderson (1982), Dell (1973/85), Tranel (1987, 2000), Noske (1988, 1996), and Spa (1988).[4] These analyses are motivated by data like the ones in (11), in which schwa is optionally or obligatorily pronounced in the context VCC_CV, but not in the context VC_CV.

[4] See Côté 2001 for an even longer list.

11. from Côté (2007): no schwa in VC_CV

	Word	Context	IPA
a.	fruiterie	VC_CV	[fʁɥit + ʁi]
b.	gâterais	VC_CV	[gat + ʁɛ]
c.	attaque pénible	VC_CV	[atak#penibl]
d.	lance-fusées	VC_CV	[lɑ̃s + fyze]

12. from Côté (2005): possible/obligatory schwa in VCC_CV

	Word	Context	IPA
a.	justement	VCC_CV	[ʒystə + mɑ̃]
b.	rythmerait	VCC_CV	[ʁitmə + ʁɛ]
c.	Annick le salut	VCC_CV	[anik#lə = saly]
d.	porte-manteau	VCC_CV	[poʁt(ə) + mɑ̃to]

Martinet (1969), Noske (1994), Oostendorp (1998), and Côté (2007) each maintain that at least some of the vowels in (12) must be epenthetic. This conclusion is motivated by forms like those in (13), which show an optional schwa alternation at morpheme boundaries. Although this alternation is optional, it can occur after *any* word ending with a consonant cluster.

13. from Noske (1994) : epenthesis after a complex cluster

	Word	Context	IPA
a.	un contact fugitif	VCC_CV	[œ̃#kɔ̃takt(ə)#fyʒitif]
b.	l'index boursier	VCC_CV	[l = ɑ̃dɛks(ə)#burʁsjɛ]
c.	un film doublé	VCC_CV	[œ̃#film(ə)#duble]

The data in (12c–d) and (13) show that schwa can surface to break up a complex cluster even when the prosodic condition is not met. In all of these cases, schwa occurs in the context VCC_σσ. This is not a controversial claim, and to my knowledge, Charette (1991) is the only source in which cluster-driven schwa epenthesis is ever blocked outside of the penultimate syllable.

2.3. The Prosodic Constraint

Outside of VCC_σ epenthesis, the most evidence for the prosodic constraint comes from schwa deletion. Schwa deletion is less likely when schwa is followed by a single syllable. This is true even in cases where schwa follows a single consonant. The data in (14) and (15) come from Morin (1974) and Côté (2007), although similar observations have been made by others (Léon 1966, Dauses 1973, Lucci 1976, Dell 1973/85, Tranel 1987). Côté (2007) maintains that the essential difference is between one and two syllables, and no real difference between two and more than two syllables exists.

14. Data from Morin (1974), transcriptions my own. schwa more likely in (b) than (c)

	Word	Context	IPA
a.	ce gars	VC_σ	[sə+gaʁ]
b.	ce garçon	VC_σσ	[s(ə)=gaʁsɔ̃]
c.	ce garçon-là	VC_σσσ	[s(ə)=gaʁsɔ̃la]

15. Data from Côté (2007), transcriptions my own. schwa in (a) and (b) are equally likely

	Word	Context	IPA
a.	ce garçon	VC_σσ	s(ə)=gaʁsɔ̃
b.	ce garçonnet	VC_σσσ	s(ə)=gaʁsɔnɛ

The final support comes from Lucci (1976), who looked at the percentage of schwa pronunciation in different speaking styles and found a difference between schwa followed by one syllable and schwa followed by two or three syllables. This difference holds for every speaking style, although the particular percentages vary. All of Lucci's examples contain schwa after a single consonant, suggesting the prosodic effect also occurs in this segmental context.

16. Percentage of schwa pronunciation from Lucci (1976), p. 101

pos.	examples from Lucci (1976)	conference	conference (reading)	newspaper (reading)	interview	conver-sation
_σ	le poids, fréquence nette	68	59	83	75	32
_σσ	le travail, fréquence moyenne	31	45	47	23	12
_σσσ	le phénomène, fréquence aléatoire	23	33	55	37	22

3. Experimental Results

This section presents experimental data supporting the claim that both the segmental and prosodic constraints have an independent effect on schwa, with the segmental constraint having a stronger effect. The data used here are those collected by Racine (2008: 135-223) for a study on regional differences in schwa deletion. Speakers were asked to estimate the frequency of words pronounced with and without schwa. These data are used to approximate the rate of schwa epenthesis in French for the four conditions: VCC_σ, VC_σ, VCC_σσ, VC_σσ. Only data from Nantes speakers were considered.

3.1. Procedure

Twelve native French speakers from Nantes participated in the experiment. Participants were asked (in French) to rate their production frequency for roughly 2,000 nouns pronounced with and without schwa. The experiment used a 7-point scale, where 1 represented not very frequent pronunciations (*Pronunciation très peu fréquente*) and 7 represented very frequent pronunciations (*Pronunciation très fréquente*). A participant was first asked to rate her production frequency for a list of words in which schwa was pronounced. Then, some days later without consulting the first list, the same participant rated her production frequency for the same list of words, this time without pronounced schwas. Pronounced schwas in the first list were indicated with underlined *e* or *ue* (e.g., *le casserole*). Unpronounced schwas in the second list were indicated with an apostrophe (e.g., *le cass'role*). Every noun was presented along with the definite determiner *le*, *la*, or *l'*.

The results were two ratings for each word for each speaker: one rating for the word with schwa pronounced, and one rating for the word without schwa pronounced. These two ratings were used to derive a relative schwa frequency rating for every word, equal to difference between the frequency ratings of the schwaful and schwaless pronunciations. This schwa frequency rating is what I'll consider in the analysis below. The nice thing about this rating is that it captures the relative likelihood of pronunciation of schwa for each word within each speaker. In corpora and forced choice experiments, rates of schwa deletion or epenthesis must often be calculated across speakers instead.

It's important to note that the order of nouns was the same for both lists, and the same for every speaker, although the presentation of lists differed across speakers (half saw the schwa list before the no schwa list, and half saw the no schwa list before the schwa list). The fact that the list had a fixed order across subjects is mitigated by the large number of items (2112 per list), and the fact that the list was randomized initially. There is no strong correlation between item position and ratings, although it seems that speakers did rate schwa as better as the experiment went on. The correlation between item position and rating in the schwaful list is r=0.03 (schwaful pronunciations were rated as more frequent towards the end), and the correlation between item position and rating in the schwaless list is r=-0.03 (schwaless pronunciations were rated as less frequent towards the end).

3.2. Items

The list of items consisted of 2112 nouns with an orthographic *e*. These items were drawn from the BRULEX database (Content, Mousty, Radeau 1990). Every noun with one or more phonetic schwas in its transcription was included. In addition, 143 nouns with unpronounced orthographic *e*s were included (e.g., *rouge-gorge*). These 143 nouns contained a word-final *e* (e.g., *rouge*), which are categorically unpronounced according to traditional analyses of French schwa (Dell 1973/85, Tranel 1987) or represent released stops when they are pronounced (Léon 2005).

EXCLUSIONS: Of the 2112 items, only 1102 were included in the analysis here. Following Racine (2008), 116 items containing more than one schwa were excluded in the analysis for ease of comparison, and 10 items were excluded because they contained an orthographic *e* that was pronounced as a non-schwa vowel.[5] Additionally, 802 items were excluded because they did not contain schwa at a morpheme boundary (see §4.3).

CODING: Each word was supplied with the phonetic transcription and, when possible, written frequency count from the French lexical database, LEXIQUE (New et al 2001). Many of the phonetic transcriptions in LEXIQUE did not contain pronounced schwas, so these were added when necessary. For each word, the transcriptions were used to determine whether the schwa was preceded by a cluster (VC/VCC in the table below), and whether the schwa would appear in penult when pronounced (_σ/ _σσ in the table below). Although *gn* was transcribed as [nj] in LEXIQUE, the sequence was treated as a single sound, consistent with phonological descriptions such as Tranel (1994). The position of schwa was calculated under the assumption that one syllable is equivalent to one vowel in the LEXIQUE transcriptions. The table below contains the number of items for each of the four conditions.

17. Number of items by condition

	VC_	VCC_
_σ	58	50
_σσ	787	207

3.3. Results

The mean schwa frequency ratings for the four conditions are presented in the table below. Again, these numbers are the differences between the schwaful and schwaless ratings for each word within each speaker. They range from –6 (schwa is never pronounced) to +6 (schwa is always pronounced).

18. Mean ratings by condition

	VC_	VCC_
_σσ	–4.96	2.15
_σ	–4.27	4.10

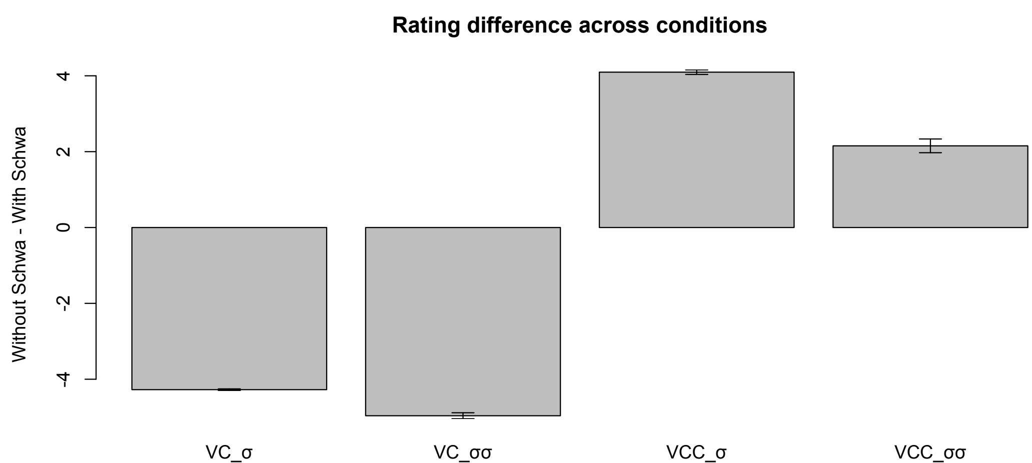
[5] The full list of these nouns: bolchévik, bolchévisme, bolchéviste, cabaretier, cabaretière, édelweiss, éden, répartie, requiem, vilenie

19. Standard errors by condition

	VC_	VCC_
_σσ	0.078	0.181
_σ	0.020	0.061

The graph in (20) illustrates the frequency rating for each of the four conditions. A bar above zero means schwa is preferred in that context, and a bar below zero means schwa is disprefered in that context.

20. Graph of rating difference by condition, 95% CIs, higher scores mean schwa is rated as more likely



These results show an independent effect of both the prosodic and segmental constraints. Schwa is rated as more likely (higher frequency score) in VC_σ words compared to VC_σσ words, when only the prosodic constraint is at stake. Furthermore, schwa is rated as more likely (higher score) in VCC_σσ words compared to VC_ σσ words, showing an independent effect of the segmental constraint.

3.4. Model

The ratings were analyzed using linear mixed effects regression. The goal of this model is to determine whether there is an independent effect of the prosodic and segmental constraints. If there is, both effects should be significant, independently of their interaction. The model contained the following fixed effects

- Cluster: whether there is a cluster before the insertion site
- Penult: whether one or two syllables follows the insertion site

- The interaction of Cluster by Penult
- Frequency: The lexical frequency of the word

Of these fixed effects, Cluster, Penult, and their interaction are of theoretical interest. Frequency was included as a control, as it has been reported to have an effect on the acceptability of schwa (Dell 1985). The coding and distribution of each main effect is presented below.

CLUSTER. Levels are coded using simple-difference coding. VCC_ = 0.5 ; VC_ = –0.5

PENULT.[6] Levels are coded using simple-difference coding. _σ = 0.5 ; _σσ = –0.5

FREQUENCY.[7] Log-transformed frequency counts, taken from LEXIQUE. The counts are taken from film subtitles. Frequency counts in LEXIQUE are words per million, so these are multiplied by 10,000,000 to ensure that the log frequency isn't zero. The mean frequencies for each condition are reported in the table below.

21. Mean log frequency by condition

	VC	VCC
_σσ	7.22	7.24
_σ	7.37	7.40

RANDOM EFFECTS. Random intercepts were included for Speaker and Item, and random slopes for Speaker were included for the predictors of theoretical interest: Cluster, Penult, and Penult x Cluster.[8]

The outcome of the regression is presented in (23). The p-values in the table below are computed from t-values, assuming the t-values represent Z scores (Gelman & Hill 2007).

---

[6] Four other alternative prosodic predictors were tried, but Penult was chosen because it accounted for the most variance. Penult is also the most theoretically motivated of the possible predictors, as the generalization is usually stated as being related to number of following syllables (see Côté 2007 for discussion). These alternative predictors were:

- Number of sylls following insertion cite
- Parity (including determiner): Including la/le/l', whether the word had an even or odd number of syllables
- Parity (without determiner): Same as above, but excluding la/le/l'
- Parity of sylls following the insertion cite: Whether the insertion site was followed by an even or odd number of syllables

[7] Film frequency was chosen over book frequency, since film frequency provides a better fit to the data.

[8] R code: lmer (rating~Cluster*Penult+Freq+(1|Item)+(1+Cluster*Penult|Speaker)

22. Regression results

Fixed effects	Coefficients	S.E.	t-value	p-value
(Intercept)	–0.19	1.20	–0.16	0.87
Cluster	7.69	0.44	17.44	<0.001
Penult	1.44	0.25	5.77	<0.001
Penult x Cluster	1.32	0.54	2.44	<0.05
Frequency	–0.09	0.16	–0.55	0.58

Both the prosodic and segmental constraints have an independent effect on ratings, as shown by the significance of Cluster and Penult, even when their interaction term is included in the model. This finding supports an analysis in which both constraints play a role. The direction of the coefficients shows that both the presence of a cluster and penultimate position increase the schwa frequency score, meaning schwa is more likely in these contexts. The effect of Cluster is about five times greater than the effect of Penult, as shown by the difference in the coefficients for Cluster vs. Penult (cf. 7.69 vs. 1.44). This difference in effect size will be used to support a grammar with weighted constraints over a grammar with ranked constraints, as shown in section 4.

3.5. Epenthesis Or Deletion?

This study only deals with words in which schwa is epenthetic. The approximation used here is that a schwa appearing at a morpheme boundary is epenthetic (see section 2.1). The focus on epenthetic schwas is motivated by the fact that the presence or absence of a schwa in the underlying form may affect performance factors, such as lexical frequency. It has been claimed that frequency of application of a phonological process is conditioned by lexical frequency (Dell 1973/85, Fidelholtz 1975). If both epenthesis and deletion are more likely in frequent words, then schwa will be preferred in a high-frequency item when it is inserted, and disprefered in a high-frequency item when it is deleted. Similarly, speakers may become more or less sensitive to the underlying form over the course of an experiment. The epenthesis/deletion distinction only matters here because the experiment used real words (with lexical frequencies).

4. Theoretical Account

This section shows how the increased probability of epenthsis in the context VCC_σ follows as the result of a gang effect between two constraints: *CLUSTER and *PENULT≠ə. In Noisy HG, the presence of the individual constraints in the grammar allows the analysis to capture the independent prosodic effect (VC_σ vs. VC_σσ) and segmental effect (VC_σσ vs. VCC_σσ). Using the same set of constraints, SOT cannot account for the set of data.

4.1. Constraints

The constraints used in the analysis are straightforward implementations of the two conditions on schwa epenthesis.

23. The segmental constraint:
    *CLUSTER: Assign one violation mark for every complex coda.

24. The prosodic constraint:
    *PENULT≠ə: Assign one violation mark if the penult vowel is not schwa.

In addition, the faithfulness constraint DEP militates against epenthesis.

25. The faithfulness constraint:
    DEP: Assign one violation mark for every segment without an input correspondent.

While *PENULT≠ə appears stipulative, the constraint could be replaced by one with greater theoretical motivation, such as the one below.

26. An alternative prosodic constraint:
    DEPENDENT(FT)=ə: Assign one violation mark if the dependent member of a foot is not schwa.

This constraint has been proposed to account for languages with sonority-driven stress, in which dependents of feet prefer to be low-sonority vowels (de Lacy 2002 and references). It favors schwa in the penultimate syllable under the assumptions that French has a right-aligned iambic foot and schwa is the least sonorous vowel in French. A right-aligned iambic foot has previously been proposed by Charette (1991) on the basis of epenthesis, by Scullen (1993) on the basis of truncation, and on the basis of acquisition by Demuth and Johnson (2003) for Parisian French, and Rose (2000), and Goad and Buckley (2006) for Quebecois French. A right-aligned iambic foot is consistent with the fact that French has phrase-final stress (Dell 1973/85, Tranel 1987). The status of schwa as the least sonorous vowel in French has been argued in Tranel (2000).

One further complication is the possibility that French words contain multiple feet, as argued in Selkirk (1978) and others. There are two ways to limit the effects of the constraint DEPENDENT(FT)=ə to schwa in the penult. The first possibility is that French words only contain a single right-aligned iambic foot, and no other feet. As a result, the penultimate syllable is the only syllable that's in the dependent of a foot. The second possibility is that DEPENDENT(FT)=ə is specific to the head foot of the word. The analysis here makes no claims about foot structure in French beyond the presence of a right-aligned iambic foot.

It should be noted that there are alternative constraints that are consistent with the HG analysis: the constraints *CLASH and PPH-BIN. These are discussed in the appendix.

4.2. Harmonic Grammar Analysis

Before moving on to the experimental data, I'll show how these three constraints, *CLUSTER, *PENULT≠ə, and DEP, can be weighted in an HG analysis to account for the increased rate of epenthesis in VCC_σ contexts. The gang effect presented here will be important in differentiating between Noisy HG and Stochastic OT.

In HG, constraints are weighted instead of ranked. A candidate's Harmony score is equal to the sum of its weighted violations. The optimum is the candidate with the highest Harmony score. Because violations are negative and weights are positive, the optimal candidate is the one with the Harmony score closest to zero.

27. Sample HG tableau

/VC-σ/	DEP w=4	*PENULT≠ə w=3	*CLUSTER w=3	H
V.(Ce.σ)	–1			–4
(VC.σ) ->		–1		–3

In the tableau above, candidate (VC.σ) is optimal, since its harmony (–3) is greater than the harmony of V.(Ce.σ) (–4).

Because constraints are weighted and violations are summed, lower-weighted constraints can combine their violations to overcome higher-weighted constraints. These interactions, called gang effects, provide a way to model differences in the distribution of doubly-marked forms like VCC_σ and singly-marked forms like VCC_σσ versus VC_σ.

The constraints *PENULT≠ə and *CLUSTER gang up against DEP to force epenthesis when both markedness constraints are violated. The epenthesis candidate VC.(Ce.σ) satisfies both lower-weighted markedness constraints at the cost of violating one higher-weighted faithfulness constraint. It satisfies *CLUSTER by removing a cluster, and *PENULT≠ə by pushing the non-schwa vowel out the penult, replacing it with an epenthetic schwa. The weighting conditions required for epenthesis to occur in VCC_σ are given in (28). The gang effect will occur under any set of numerical weights that obeys the weighting conditions.

28. w(DEP) < w(*PENULT≠ə) + w(*CLUSTER)

/VCC-σ/	DEP w=4	*PENULT≠ə w=3	*CLUSTER w=3	H
VC.(Ce.σ) ->	–1			–4
(VCC.σ)		–1	–1	–6

The weighting conditions for (28) are consistent with the weighting required to block epenthesis in words that only violate one (or no) constraint(s): VC_σ, VC_σσ, and VC_σσ. Under the weighting conditions in (29) and (30), epenthesis won't occur when only one markedness constraint is violated, since the weight of DEP is greater than the weight of the individual constraints.

29. w(DEP) > w(*PENULT≠ə)

/VC-σ/	DEP w=4	*PENULT≠ə w=3	*CLUSTER w=3	H
V.(Ce.σ)	–1			–4
(VC.σ) ->		–1		–3

30. w(DEP) > w(*CLUSTER)

/VCC-σσ/	DEP w=4	*PENULT≠ə w=3	*CLUSTER w=3	H
VC.Ce.(σσ)	–1	–1		–7
VCC.(σσ) ->		–1	–1	–6

Finally, there's no reason to epenthesize in VC_σσ contexts, since epenthesis satisfies neither constraint. The epenthetic candidate in this context is harmonically bounded. No matter the weighting of constraints, it will never win.

31. Epenthesis is harmonically bounded

/VC-σσ/	DEP w=4	*PENULT≠ə w=3	*CLUSTER w=3	H
V.Ce.(σσ)	–1	–1		–7
VC.(σσ) ->		–1		–3

In this analysis, schwa epenthesis is the result of a gang effect of two independent markedness constraints. The doubly-marked form VC_σσ is repaired, while the singly-marked forms are not.

4.3. Noisy HG

Noisy HG takes a HG grammar and adds noise to create a probability distribution (Boersma & Pater 2008, Pater, Potts, & Bhatt 2007). At each evaluation, a weight for each constraint is selected from a normal distribution around its mean. Pooling the optima across a number of evaluations gives a probability distribution.

Given the constraint set from the previous section, the experimental data can be captured in Noisy HG. To do this, I assume the transformation the probabilities for French epenthesis given in (34). These probabilities come from the experimental rating data in (32). A rating of –6 is transformed into a probability of 0, and a rating of +6 is transformed into a probability of 1. Since the VC_ σσ context is one in which epenthesis will never occur, under either Noisy HG or SOT, the VC_ σσ context was made 0 by subtracting .09 from each probability. This means that all of the models under consideration have a chance of matching the probabilities from the experiment.

32. Mean ratings

	VC_	VCC_
_σσ	-4.96	2.15
_σ	-4.27	4.09

33. Probabilities from ratings

	VC_	VCC_
_σσ	0.00	0.59
_σ	0.05	0.75

34. p(schwa) = ((rating+6)/12) –.09

The important thing about these probabilities, though, is that the differences between VC_σ and VC_σσ is small, while the difference between VC_σσ and VCC_σσ is great. The difference between Noisy HG and SOT could be shown by any set of input probabilities with this property, in which the prosodic constraint plays a small role, and the segmental constraint plays a large role.

One possible set of weights to account for the French pattern in Noisy HG is presented in (35). These weights were learned in Praat using the HG-GLA and the data in (33) as training data. The weights were then used to generate the probabilities in (36). The learned grammar nearly matches the probabilities from the training data, which are repeated in (37).

35. Weights learned using default settings in Praat
*CLUSTER: 6.884
DEP: 6.319
*PENULT≠ə: 1.686

36. Probabilities generated from weights in (35), 100,000 replications with 2.0 noise

	VC_	VCC_
_σσ	0.00	0.57
_σ	0.05	0.74

37. Probabilities from experimental data, repeated from (33)

	VC_	VCC_
_σσ	0.00	0.59
_σ	0.05	0.75

How do the weights in (35) capture the pattern in (37)? Under the constraint set assumed, there will never be a weighting of constraints in which epenthesis occurs in VC_σσ, so epenthesis never occurs in this context. Epenthesis in VC_σ occurs in 5% of the weightings, because there's a 5% chance that the grammar will sample a weight for *PENULT≠ə that's greater than the weight of DEP (when sampling with a noise of 2.0). Epenthesis in VCC_σσ is more likely because it's more likely for a weight of *CLUSTER that's greater than the weight of DEP to be sampled, since the base weight of *CLUSTER is closer to DEP. The rate of epenthesis in VCC_σ is greatest because epenthesis will occur here when the constraints are weighted to cause epenthesis in VCC_, when the constraints are weighted to cause epenthesis in _σ, and when the constraints are weighted to gang up as in section 4.2.

A MaxEnt grammar is also able to model a gang effect, as shown in the appendix. The use of Noisy HG over MaxEnt is motivated by the ease of comparison between Noisy HG and Stochastic OT, both of which are stochastic models of variation.

4.4. Stochastic OT

Jaeger & Rosenbach (2006) show that psuedo-gang-effects (weak gang effects in Jaeger & Rosenbach) can occur in SOT. However, the pseudo-gang-effects of SOT are limited compared to the gang effects of Noisy HG, and cannot account for the differences in the effect sizes of *PENULT≠ə and *CLUSTER.

Stochastic OT generates a probability distribution over candidates by producing a new ranking of constraints on an utterance-by-utterance basis (SOT: Boersma 1997, 1998, Boersma & Hayes 2001). During each evaluation, a constraint's ranking is determined by taking its rank, a real number on a continuous scale, and adding some random noise according to a normal distribution (with mean 0 and standard deviation 1). The ranking of the constraint after adding noise is the constraint's selection point. For a given evaluation, the constraints are strictly ranked according to their selection points. In a SOT grammar, if two constraints have identical ranks, each will dominate the other 50% of the time. As the distance between two constraints' ranks increases, it's less likely for the lower-ranked constraint to dominate the higher-ranked one.

In a SOT, two constraints can gang up against a third constraint when the rankings of all three are roughly equal. In these pseudo-gang-effects, the doubly-marked form is repaired more often than the singly marked forms, simply because the doubly-marked form is repaired under more rankings. For example, take the three constraints in the French HG analysis, repeated below.

38. The segmental constraint:
*CLUSTER: Assign one violation mark for every complex coda.
39. The prosodic constraint:
*PENULT≠ə: Assign one violation mark if the penult vowel is not schwa.
40. The faithfulness constraint:
DEP: Assign one violation mark for every segment without an input correspondent.

If each of these constraints has an equal rank, all six possible rankings of the constraints will be equally likely. The doubly marked VCC_σ would be repaired under 4/6 of the rankings, when either *CLUSTER or *PENULT≠ə dominate DEP. However, VC_σ would only be repaired in 3/6 rankings, when *PENULT≠ə ≫ DEP. Likewise, VC_σσ would only repaired in 3/6 of the rankings, when *CLUSTER ≫ DEP. The resulting grammar displays a probabilistic gang effect, in which the probability of epenthesis in the doubly-marked context is greater than the probability of epenthesis in either singly-marked context.

41. Probabilities in SOT when constraints have equal ranks

	VC_	VCC_
_σσ	0.00	0.50
_σ	0.50	0.66

While SOT can capture some gang effects, it cannot capture the sort of gang effect demonstrated by the French experimental data. In the French data, the probability of epenthesis in one singly marked context (0.59 in VC_σ) is much greater than the probability of epenthesis in the other singly marked context (0.05 in VCC_σσ). To capture this asymmetry, the rank of *CLUSTER must be greater than the rank of *PENULT≠ə, since *CLUSTER must dominate DEP much more often than *PENULT≠ə dominates DEP. As the ranks of *PENULT≠ə and *CLUSTER become further apart, it becomes less likely for *PENULT≠ə to dominate DEP, and even less likely for *PENULT≠ə to dominate *CLUSTER. As a result, the ranking that separates VCC_σ from VCC_σσ contexts, *PENULT≠ə≫ DEP ≫ *CLUSTER, becomes more and more improbable. The ideal gang effect setup in (41) requires each constraint to have an equal ranking, as the rankings move apart, the size of the gang effect decreases.

This can be shown using an implementation of the Gradual Learning Algorithm in Praat. Using the default settings, the learning data from table (44) and the three constraints in (42), the GLA cannot learn the French pattern. The ranks learned for the input distribution are in (42). The high ranks reflect a failure for the GLA to converge.

42. Ranks learned using default settings in Praat
*CLUSTER: 1220.78
*PENULT≠ə: 1219.66
DEP: 1216.28

43. Probabilities generated from (42), 100,000 replications with 2.0 noise

	VC_	VCC_
_σσ	0.00	0.66
_σ	0.12	0.67

44. Training data, repeated from (18)

	VC_	VCC_
_σσ	0.00	0.59
_σ	0.05	0.75

The ranks learned in (42) practically collapse the difference between epenthesis in the VCC_σσ and VCC_σ contexts. The VCC_σ context gets a small boost over VCC_σσ from the prosodic constraint, since epenthesis in VCC_σ also occurs when *PENULT≠ə dominates DEP. However, this ranking is very unlikely with the ranks in (42). The problem is that *PENULT≠ə must be closer to DEP than *Cluster to capture the weaker effect of the prosodic constraint. As the ranks of *PENULT≠ə and *CLUSTER move apart, the pseudo-gang-effect becomes weaker. This isn't a problem in Noisy HG, in which *PENULT≠ə can gang up with *CLUSTER, even when their weights are far apart.

5. Alternative Constraint Sets

This section outlines other possible OT analyses of the French data. All of these analyses are able to model the increased rate of epenthesis in the VCC_σ context, and each model seems equally plausible given the categorical generalization in which epenthesis is completely blocked in non-VCC_σ contexts.

However, each analysis' constraint set cannot capture the variable data described in Sections 2 and 3. The problem is that the constraints cannot capture the independent effects of both the prosodic and segmental constraints. SOT and Noisy HG create probability distributions by varying the rankings/ weightings of a set of constraints and counting the number of times each candidate is optimal. For a candidate to receive probability mass, it must be optimal under some rankings/weightings. Furthermore, if two candidates are to receive different probabilities, they must have distinct violation profiles.

In each analysis discussed below, there is a ranking/weighting in which VCCəσ is optimal, but there is either no way to distinguish VC_σ from VC_σσ (the independent effect of the prosodic constraint) or no way to distinguish VCC_σσ from VC_σσ (the independent effect of the segmental constraint). Since there is no way to distinguish between these candidates' violation profiles, there is no way to distinguish between their probabilities in Noisy HG or SOT.

As shown in the previous sections, the approach using two independent weighted constraints is able to make all of the necessary distinctions, while maintaining the increased probability of the VCC_σ context.

THE WEIGHT-TO-STRESS PRINCIPLE. The first alternative is a phonologically grounded version of the conjoined constraint account, using a VCC_σ-specific constraint.

45. *PENULT≠ə & *CLUSTER: Assign one violation mark if the penult vowel is not schwa and there is a complex coda.

Instead of using the stipulative conjoined constraint, the Weight-To-Stress analysis accounts for the distribution of epenthetic schwa as the result of a constraint enforcing the Weight-To-Stress Principle.

46. WSP: Assign one violation mark for ever unstressed heavy syllable (If unstressed, then light)

Under the prosodic assumptions outlined below, this constraint is violated by clusters in the penultimate syllable, but not clusters elsewhere. Rose & Dos Santos (2010) argue for a similar constraint, iambic strengthening, to account for vowel lengthening before voiced fricatives in Québecois French.

This analysis makes two assumptions. The first is that French has a right-aligned iambic foot. This is the same assumption made in the HG analysis, and this is well-motivated in the literature. As was the case for the analysis in Section 4, this analysis requires that French words contain only a single foot, or that the WSP constraint only applies in the main foot. The second assumption is that French makes a distinction between the weight of VC (light) and VCC (heavy) syllables.

47. Weight in French
VC, V: < VCC, V:C

This would be a novel claim about weight in French. Scullen (1993) shows that French observes a bimoraic minimality condition in derived words, but in her account, codas count towards fulfilling minimality. Beyond VCC_σ epenthesis, the only evidence for this distinction is phrase-final lengthening, in which VC syllables are lengthened to V:C syllables when stressed, but VCC syllables are not (Tranel 1987, Walker 2001, Féry 2003).

The comparative tableau below shows that this analysis can account for VCC_σ epenthesis. In a comparative tableau, each line presents a pair of candidates, comparing a winning candidate with a losing candidate (Prince 2002). Each *W* and *L* in the tableau indicates whether a constraint favors the desired winner or loser . For the ranking conditions to be consistent, there must be some order of constraints in which every loser-favoring constraint (*L*) is dominated by some winner-favoring constraint (*W*). In other words, an analysis is possible if there's a ranking in which every L is to the right of some W.

48. CT for WSP

winner~loser	WSP	DEP
VC(Cə.σ) ~ (VCC.σ)	W	L
(VCσ) ~ V(Cə.σ)		W
VCC(σσ) ~ VCCə(σσ)		W
VC(σσ) ~ VCə(σσ)		W

The problem with the WSP analysis is the one shared by all of the OT analyses presented here. It cannot account for the independent effects of the segmental and prosodic constraints. In the scale below, the probability of schwa epenthesis is greater in VC_σ than it is in VC_σσ, showing an independent effect of the prosodic constraint.

49. p(schwa): VCC_σ > VCC_σσ > VC_σ > VC_σσ

Under this account, there is no way to differentiate epenthesis in VC_σ from epenthesis in VC_σσ. Neither VC_σ and VC_σσ incur any violations of WSP.

50. No way to differentiate between VC_σ and VC_ σσ

	WSP	DEP
(VC.σ)	Ø	Ø
VC(σ.σ)	Ø	Ø

Prosodic phrasing in Côté 2007. Although Côté's (2007) focus is data showing that French schwa epenthesis is conditioned by the number of preceding syllables, her analysis can be adopted to account for the effect of the number of following syllables.[9] Côté (2007) uses a the model of variation with multiple rankings from Anttila (1997). This model is very similar to SOT. Like SOT, a ranking of constraints is selected from a distribution at each evaluation. The difference is that constraints are either strictly ranked with respect to one another, or unranked within a stratum. At evaluation, the unranked constraints within a stratum are ranked randomly, and this ranking is used to select an optimal.

Under Côte's (2007) analysis, epenthesis in VCC_σ contexts is analyzed as the result of the interaction of prosodic phrasing and perceptually-motivated segmental licensing. Unlike VCC_σσ words ,VCC_σ words must be parsed as one prosodic phrase, and are consequently subject to a segmental constraint applying within phrases.

The analysis rests on the assumption that schwa epenthesis always occurs at phonological word boundaries (prefixes, compounds, words), which can coincide with prosodic phrase (PPh) boundaries. Two phonological words can either be phrased as one PPh or two separate PPhs. The difference between _σ and _σσ contexts is that the two-PPh parse in _σ contexts creates a subminimal PPh, while the two-PPh parse in _σσ contexts does not.

51. Two-PPh parses in Côté (2007)
   a. [VCC]-[σ] = subminimal
   b. [VCC]-[σσ]
   c. [VC]-[σ] = subminimal
   d. [VC]-[σσ]

To avoid the subminimal PPhs, VC-σ and VCC-σ must be parsed as one PPh instead.

[9] Côté (2007) presents some data to support her PPh minimality analysis, arguing that it's the number of syllables—not stress—that determines the acceptability of schwa. Schwa epenthesis is best when it creates a prosodic phrase with four or more syllables.

1. Prosodic phrasing — not foot structure
   a. l'act(e) commence
   b. l'entract(?e) commence
   c. jette d(e) l'ortie
   d. achette d(?e) l'ortie

However, she also notes that stress is an additional factor. In fact, the judgments for words that differ in penult/antepenult are stronger than those that differ in number of syllables.

53. Avoiding subminimal PPhs
a. [VCC-σ]
b. [VCC]-[σσ]
c. [VC-σ]
d. [VC]-[σσ]

The parses above follow from the existence of the constraint *(σ), which militates against monosyllabic PPhs.

54. *(σ): Assign one violation for every PPh with fewer than two syllables.

In Côté's analysis, epenthesis is driven by the constraint is (..C..)↔V, which militates against CCC sequences within a PPh.

55. (..C..)↔V: Assign one violation mark for every phrase-internal consonant that is not adjacent to a vowel.

Phrase-internal CCC sequences only arise under the one-PPh parse. Given that VCC_ σ must be parsed as one PPh, it is more likely to violate (..C..)↔V than VCC_ σσ. The reason epenthesis does not occur in VC_σ or VC_σσ contexts is that these forms do not contain CCC sequences under any prosodic phrasing.

In order to drive epenthesis in VCC_σ contexts, (..C..)↔V must dominate DEP[10]. The comparative tableau contains all of the losing candidates with one PPh.

56. (..C..)↔V ≫ DEP

winner~loser	*(σ)	(..C..)↔V	DEP
(VCCəσ) ~ (VCCσ)		W	L
(VCσ) ~ (VCəσ)			W
(VCC)(σσ) ~ (VCCəσσ)			W
(VC)(σσ) ~ (VCəσσ)			W

In order to make sure (..C..)↔V is satisfied through epenthesis rather than a different prosodic phrasing, *(σ) must dominate DEP. The comparative tableau below contains all of the losing candidates with two PPhs.

[10] which replaces Côté's *SCHWA

57. *(σ) ≫ DEP

winner~loser	*(σ)	(..C..)↔V	DEP
(VCCəσ) ~ (VCC)(σ)	W		L
(VCσ) ~ (VCə)(σ)			W
(VCC)(σσ) ~ (VCCə)(σσ)			W
(VC)(σσ) ~ (VCə)(σσ)			W

In Côté's (2007) analysis, VCC_σ words uniquely condition epenthesis because they cannot be parsed as separate PPhs to avoid violating the segmental constraint.

The problem with the prosodic phrasing analysis is similar to the problem with the WSP analysis. The prosodic phrasing analysis cannot account for the independent effect of the prosodic constraint. In the scale below, the probability of schwa epenthesis is greater in VC_σ than it is in VC_σσ.

58. p(schwa): VCC_σ > VCC_σσ > VC_σ > VC_σσ

Under this account, there is no way to differentiate epenthesis in VC_σ from epenthesis in VC_σσ. Both VC_σ and VC_σσ incur no violations of any constraints when parsed with one PPh.

59. No way to differentiate between VC_σ and VC_ σσ

	*(σ)	(..C..)↔V	DEP
(VCσ)	Ø	Ø	Ø
(VCσσ)	Ø	Ø	Ø

STRESS CLASH: Mazzola (1992) accounts for the increased likelihood of schwa in the context _σ as the result of stress clash avoidance. Every word has final stress. When a word with final stress is followed by a single-syllable word, a stress clash would result, and schwa is pronounced to avoid it.

This analysis finds support in the fact that French also resolves clash through stress shift. According to Mazzola (1994), stress retraction is obligatory before monosyllables, but not before polysyllables. As a result, the word *savant* is pronounced with initial stress in *sávant Súisse*, but not in *savánt Angláis*. One objection, raised by Côté (2007), is that the stress clash analysis cannot make the distinction between VC_σ and VCC_σ, both of which violate stress clash. To account for the distinction in clusters, the analysis must formalize the interaction between the prosodic constraint, stress clash in this case, and the segmental constraint.

Côté's (2007) objection to Mazzola (1992) can now be addressed in a Noisy HG analysis. The constraint *CLASH could replace the constraint *PENULT≠ə. However, there are two reasons to approach such an analysis with hesitation. The first is that the *CLASH analysis is inconsistent with typological claims about stress-epenthesis interactions, which state that epenthesis never occurs to resolve stress clash (Blumenfeld 2006). The second is that the analysis can only account for VCC_σ epenthesis if VCC and σ are separated by a prosodic word boundary, and each carry their own stress. The fact that epenthesis is obligatory within words like *brusquement*, where schwa appears in a VCC_σ context at a suffix boundary, would have to be attributed to an entirely different mechanism.

LICENSING IN CHARETTE 1991: The final alternative is to use a licensing constraint to block the epenthesis of schwa in VCC_σσ, which is driven by the ranking *CLUSTER ≫ DEP. This is similar to Charette's (1991) proposal, although hers is formalized in Government Phonology. The licensing constraint is similar to *PENULT≠ə in that it requires schwa to be associated with a weak prosodic position. The difference is that the licensing constraint militates against any schwa outside of the penult.

60. ə→PENULT: Assign one violation for every schwa that is not in the penultimate syllable.

This constraint can block epenthesis outside of the penultimate syllable under the ranking ə→PENULT≫*CLUSTER≫DEP.

61. ə→PENULT≫*CLUSTER≫DEP.

winner~loser	SCHWA→PENULT	*CLUSTER	DEP
VCCəσ ~ VCCσ		W	L
VCσ ~ VCəσ			W
VCCσσ ~ VCCəσσ	W	L	W
VCσσ ~ VCəσσ	W		W

Like the other OT accounts, this analysis cannot capture the difference between VC_σ and VC_σσ. While schwa epenthesis in VC_σσ is more marked, since it creates a schwa outside the penult, there's still no reason to epenthesize in VC_σ. In fact, epenthesis in VC_σ is harmonically bounded.

62. Violation tableau for VC_σ

	SCHWA→PENULT	*CLUSTER	DEP
VCσ ->			
VCəσ			*

Conclusion

This paper has argued that there are two independent constraints on the epenthesis of schwa in French, and epenthesis is more likely in contexts where both constraints are violated. This pattern is best accounted for in an analysis with weighted constraints, framed here in Noisy HG. Such an account can capture the individual effects of the constraints, as well as the increased probability of epenthesis in contexts in which both constraints are violated.

Other analyses are either unable to capture the independence of both constraints (section 5) or unable to capture differences between their effect sizes (section 4).

Appendix: Alternatives For The Prosodic Constraint

This analysis assumes the prosodic constraint *PENULT≠ə, although it should be noted that there are alternative constraints consistent with the HG analysis: the constraints *CLASH and PPH-BIN. These two constraints are in the spirit of analyses by Mazzola (1992) and Côté (2007) (see §4.5 for a more in-depth discussion of alternatives). Both constraints work because they militate against VC-σ and VCC-σ words, just like *PENULT≠ə. Epenthesis satisfies *CLASH by avoiding a clash, and epenthesis satisfies PPH-BIN by fulfilling minimality.

63. An alternative prosodic constraint:
*CLASH: Assign one violation mark for every sequence of adjacent stressed syllables.
(requires assuming every PWd has final stress)

64. Another alternative prosodic constraint:
PPH-BIN: Assign one violation mark for every the PWd containing one syllable.
(requires assuming the schwa is epenthesized into the monosyllabic PWd)

These two constraints differ from *PENULT≠ə in the predictions they make for schwa at PWd boundaries. Both *CLASH and PPH-BIN are violated in $_{\mathrm{Pwd}}[\mathrm{VCC}]_{\mathrm{Pwd}}[\sigma]_{\mathrm{Pwd}}$, when the insertion site is at a PWd boundary, but not in $_{\mathrm{Pwd}}[\mathrm{VCC\#}\sigma]_{\mathrm{Pwd}}$, when the insertion site is not at a PWd boundary. While epenthesis occurs in the context $_{\mathrm{Pwd}}[\mathrm{VCC\#}\sigma]_{\mathrm{Pwd}}$, supporting *PENULT≠ə, e.g. brusquement [bryskəmɑ] (Dell 1973/85), there is no word of the shape $_{\mathrm{Pwd}}[\mathrm{VCC\#}\sigma\sigma]_{\mathrm{Pwd}}$ to compare it to. This is simply due to an accident of the lexicon. There are CC-final prefixes (entre–, contre–, ex–) that can combine with C-initial disyllabic roots, but there is evidence that prefixes are separate PWds in French, since they do not undergo certain alternations such as glide formation or denasalization (Hannahs 1995). There are CC-final roots in French (brusque, garde), but there are no C-initial disyllabic suffixes to follow them. C-initial suffixes are rare in general, and to my knowledge, there are no cases in which one of the few C-initial suffixes (–ment, –té, –rie) can be followed by additional suffixes.

Appendix: MaxEnt Grammar

One alternative to the Noisy HG grammar is a MaxEnt Grammar (Goldwater & Johnson 2003). Like Noisy HG, this grammar is able to capture the gang effect, along with the difference between the strengths of the two constraints. In a MaxEnt Grammar, the probability assigned to each candidate is proportional to its harmony score. As in HG, constraint violations are summed and weighted to give a harmony score for each candidate. The probability for a candidate is determined by taking the exponent of the candidate's harmony score and dividing it by the sum of the exponents of the harmony scores for the full candidate set.

The solution learned by Praat is presented below.

65. Weights learned using default settings in Praat
*CLUSTER: 4.114
DEP: 3.722
*PENULT≠ə: 0.793

66. Probabilities from weights

	VC_	VCC_
_σσ	0.00	0.60
_σ	0.05	0.77

67. Probabilities from training data

	VC_	VCC_
_σσ	0.00	0.59
_σ	0.05	0.75

References

Anderson, Stephen R. (1982). The analysis of French schwa: or, how to get something from nothing. Language 58:534-573.

Anttila, Arto (1997). Deriving variation from grammar. In Variation, Change, and Phonological Theory, eds. Frans Hinskens, Roeland van Hout and W. Leo Wetzels. Amsterdam/Philadelphia: John Benjamins.

Blumenfeld, Lev (2006). Constraints on phonological interactions. PhD dissertation. Stanford University

Boersma, Paul (1997). Functional Optimality Theory. Proceedings of the Institute of Phonetic Sciences of University of Amsterdam 21, 37–42.

Boersma, Paul (1998). Functional Phonology: Formalizing the Interactions Between Articulatory and Perceptual Drives. PhD dissertation. University of Amsterdam.

Boersma, Paul & Joe Pater (2008). Convergence properties of a gradual learning algorithm for Harmonic Grammar. Ms, University of Amsterdam and University of Massachusetts Amherst. ROA 978

Boersma, Paul, and Hayes, Bruce (2001). Empirical tests of the Gradual Learning Algorithm. Linguistic Inquiry 32:45-86.

Bouchard, Denis (1981). A voice for e muet. Journal of Linguistic Research 1:17-47.

Charette, Monik (1991). Conditions on Phonological Government. Cambridge, UK: Cambridge University Press.

Content, Alain, Philippe Mousty & Monique Radeau (1990). BRULEX: une base de donées lexicales informatisée pour le français écrit et parlé. L'Anée psychologique 4:90: 551-566.

Cornulier, Benoit (1975). Le droit d'e: e et la syllabicite, Cahiers de linguistique, d'orientalisme et de slavistique 5-6: 101-117.

Côté, Marie-Hélène (2001). Consonant Cluster Phonotactics: A Perception-Based Approach. MIT Linguistics Dissertations. Cambridge: MIT Working Papers in Linguistics.

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.

Dauses, August (1973). Etudes sur l'e instable dans le français familier. Tübingen: Niemeyer.

de Lacy, Paul (2002). The formal expression of markedness. PhD Dissertation. University of Massachusetts Amherst.

Demuth, Katherine & Mark Johnson (2003). Truncation to Subminimal Words in Early French. Canadian Journal of Linguistics 48: 211-241.

Fidelholtz, James (1975). Word frequency and vowel reduction in English. In CLS, 11:200-213.

Féry, Caroline (2003). Markedness, Faithfulness, Vowel Quality and Syllable Structure in French. Journal of French Language Studies 13.2.

Gelman, Andrew, & Jennifer Hill (2007). Data Analysis Using Regression and
Multilevel/Hierarchical Models. New York: Cambridge University Press.

Goad, Heather & Meaghan Buckley (2006). Prosodic structure in child French: Evidence for the Foot. Catalan Journal of Linguistics 5: 109–142. (Special issue on the acquisition of Romance languages as first languages.)

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
Jäger, Gerhard & Anette Rosenbach (2006). The winner takes it all - almost: cumulativity in grammatical variation. Linguistics 44(5): 937-971.
Legendre, Geraldine, Yoshiro Miyata & Paul Smolensky (1990). Harmonic Grammar – a formal multi-level connectionist theory of linguistic wellformedness: An application. In Proceedings of the Twelfth Annual Conference of the Cognitive Science Society. Cambridge, Mass.: Lawrence Erlbaum, 884–891.
Léon, Pierre (1966). Apparition, maintien et chute du e caduc. La linguistique 2.111-122.
Léon, Pierre (2005). Phonétisme et prononciations du français. Paris: Armand Colin (4ème édition)
Lucci, Vincent. (1976). Le mécanisme du 'e' muet dans différentes formes de français parlé. La linguistique 12:2, 87-104.
Martinet, André (1969). Le Français sans Fard. Paris: Presses Universitaires de France.
Mazzola, Michael (1992). Stress Clash and Segment Deletion. Theoretical Analyses in Romance Linguistics ed. by Christiane Laeufer & Terrell A. Morgan, 81-97. Amsterdam & Philadelphia: John Benjamins.
Mazzola, Michael (1994). Prosodic Constituency and Intonation. Salford: Linguistics Association of Great Britain, University of Salford.
Morin, Yves-Charles (1974). Règles phonologiques à domaine indéterminé: chute du cheva en français. Cahier de linguistique 4, 69-88.
New, Boris, Christophe Pallier, Ludovic Ferrand, & Rafaël Mattos. 2001. Une base de donées du français contemporain sur internet: LEXIQUE. L'Année psychologique 101: 447-462.
Noske, Roland (1988). La syllabification et les règles de changement de syllabe en français, in Sylvain Paul Verluyten (ed.) La phonologie du schwa français, Linguisticae Investigationes Supplementa. Studies in French and General Linguistics 16, Amsterdam / Philadelphia, PA: John Benjamins: 43-88.
Noske, Roland (1996). Is French optimal? In Jacques Durand and Bernard Laks (eds.), Current Trends in Phonology, eds. : 487-510. Paris: CNRS.
Oostendorp, Marc van (1998). Schwa in Phonological Theory. Glot International, 3(5): 3-8.
Pater, Joe, Rajesh Bhatt & Chris Potts. 2007. Linguistic Optimization. Ms., University of Massachusetts Amherst.
Prince, Alan (2002). Entailed Ranking Arguments. ms. Rutgers University, New Brunswick. ROA-500.
Prince, Alan, and Smolensky, Paul (1993/2004) Optimality Theory: Constraint interaction in generative grammar. Malden, MA: Blackwell.
Pulgram, Ernst (1961). French /ə/: statics and dynamics of linguistic subcodes, Lingua 10: 305-325.
Racine, Isabelle (2008). Les effets de l'effacement du schwa sur la production et la perception de la parole en français. PhD Thesis. Université de Genève, Switzerland.
Rose, Yvan (2000). Headedness and Prosodic Licensing in the L1 Acquisition of Phonology. PhD dissertation. McGill University.
Rose, Yvan and Christophe dos Santos (2010). Stress Domain Effects in French Phonology and Phonological Development. Interactions in Romance: Selected Papers from the 38th Linguistic Symposium on Romance Languages. Karlos Arregi, Zsuzsanna Fagyal, Silvina Montrul & Annie Tremblay (eds.). Amsterdam: Johns Benjamins: 89-104.
Scullen, Mary Ellen (1993). The Prosodic Morphology of French, Ph. D. Diss., Indiana University.

Selkirk, Elisabeth (1978). The French foot: on the status of 'mute' e, Studies in French linguistics 1 (2): 141-150.

Smolensky, Paul (1995). On the structure of the constraint component CON of UG. Talk presented at University of California, Los Angeles. Rutgers Optimality Archive 86

Smolensky, Paul & Geraldine Legendre (2006). The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar. Cambridge, Mass. : MIT Press.

Spa, Jaap (1988). Pourquoi la loi des 3 consonnes? In R. Landheer (ed.) Aspects de linguistique française. Hommage à Q.I.M. Mok, Amsterdam: Rodopi: 161-176.

Tranel, Bernard (1987). The Sounds of French. Cambridge: Cambridge University Press.

Tranel, Bernard (2000). Aspects de la phonologie du français et la théorie de l'optimalité. Langue française 126: 39-72.