Geminates and vowel laxing in Quebec French[*]

Anaële Cuerrier & Charles Reiss
Concordia University, Montréal

Submitted to the *Proceedings of LSRL 47*

Abstract

Laxing and harmony in Quebec French (QF) high vowels shows dialectal, register and perhaps even lexical variation. A recent proposal to handle some of the data (Poliquin, 2006) contains a radical innovation to phonological theory concerning long-distance segment interactions. We question the necessity of such an account by pointing out that recognition of geminate sonorants in QF can explain some puzzling forms without recourse to new devices. Our account is supported by phonetic considerations, as well as by recognizing that the alternative both under and overgenerates lax vowels in surface forms.

1   Introduction

Chomsky (2007) poses the question "How little can be attributed to UG while still accounting for the variety of I-languages attained...?" as a way to characterize the 'bottom-up' approach to Universal Grammar. Studying such questions should converge with the 'top-down' approach, the goal of which is "to abstract from the welter of descriptive complexity certain general principles governing computation that would allow the rules of a particular language to be given in very simple forms, with restricted variety" (Chomsky, 2000, p.122).

Whenever a linguistic analysis contains a new formal device or relation, we need to ask ourselves, as bottom-up-ists, whether we really need such an

1

innovation; and as top-down-ists, whether there are any other existing devices that allow us already to handle the phenomena under consideration. In this paper, we argue that the non-local, non-iterative vowel harmony analysis (NLNIVH) offered for Québec French[1] by Poliquin (2006) is not a desirable enhancement of phonological theory from either top-down or bottom-up points of view. A simpler analysis arises from an appreciation of a basic property of QF ignored in many discussions: the language has geminate consonants.

2 Basics of Québec French +High vowels

Québec French has six high vowels on the surface:

(1) Surface high vowels of Québec French[2]

$$
i=\left\{\begin{array}{c} +\text{High} \\ -\text{Back} \\ -\text{Round} \\ +\text{ATR} \end{array}\right\}
\quad
u=\left\{\begin{array}{c} +\text{High} \\ +\text{Back} \\ +\text{Round} \\ +\text{ATR} \end{array}\right\}
\quad
y=\left\{\begin{array}{c} +\text{High} \\ -\text{Back} \\ +\text{Round} \\ +\text{ATR} \end{array}\right\}
$$

$$
\textsc{i}=\left\{\begin{array}{c} +\text{High} \\ -\text{Back} \\ -\text{Round} \\ -\text{ATR} \end{array}\right\}
\quad
\upsilon=\left\{\begin{array}{c} +\text{High} \\ +\text{Back} \\ +\text{Round} \\ -\text{ATR} \end{array}\right\}
\quad
\textsc{y}=\left\{\begin{array}{c} +\text{High} \\ -\text{Back} \\ +\text{Round} \\ -\text{ATR} \end{array}\right\}
$$

As is common in the literature, we treat +Atr as equivalent to Tense and -Atr as equivalent to Lax. A standard assumption is that the lax high vowels are all derived from tense ones. This appears to be true historically. Standard, international French (SF) has no lax high vowels on the surface. As we will see, it is not completely obvious that QF can be analyzed synchronically just one series of high vowels.

As a first approximation, assuming that at least some lax high vowels are derived, there appears to be a rule of allophonic laxing in closed syllables:

---

[1]Poliquin refers to Canadian French, but the subset of Canadian dialects he considers all are represented in the relevant properties in Québec, where all the dialects we consider are spoken. So we refer henceforth to Québec French (QF).

[2]We use the curly brackets of standard set theory to represent segments, for reasons given in Bale et al. (2014); Bale and Reiss (2018).

(2) Allophonic laxing in closed syllables

Open syllable	Closed syllable
petit [ptˢi] 'small-m.'	*petite* [ptˢɪt] 'small-f.'
tout [tu] 'all-m.'	*toute* [tʊt] 'all-f.'
fumer [fyme] 'to smoke'	*(il) fume* [fʏm] 'he smokes'
bru [bry] 'daughter-in-law'	*brute* [brʏt] 'bully'

Morpheme alternants like [fum ∼ fʏm] suggest a synchronic process that laxes high vowels in closed syllables. The pattern is also compatible with the general distribution in unrelated words like [bry, brʏt].

One complication in distribution is due to recent borrowings from English, which in the speech of all our informants, even those who would be considered monolingual francophones, lead to minimal pairs distinguished by tense *vs.* lax high vowels:

(3) Do recent borrowings show lexical lax high vowels?

- *team, pool* [tim, pul] *vs. ultime, poule* [ʏltˢɪm, pʊl] 'last, hen'

The borrowed *team* and *pool* do not have lax vowels, even though the words have high vowels in closed syllables. So, it may be necessary to recognize more than a single series of high vowels underlyingly. The non-alternating ones may be +ATR and the alternating ones could potentially be analyzed as unspecified for ATR. The ultimately correct analysis of these borrowings and the implications for native words is not explored by Poliquin, nor does it bear on our critique of his proposals.

A further complication that can be put aside for our modest purposes is the exact outcome of high vowels in syllables closed by a voiced fricative:

(4) High vowels before coda voiced fricatives:

- QF High vowels are long before voiced fricatives—there seems to be variation and uncertainty with respect to TENSE/LAX in this position

- *rouge, brise, écluse, dire, pur*
  [rʊːʒ, brɪːz, eklyːz, dᶻɪːʁ, pyːʁ] or [ruːʒ, briːz, eklyːz, dᶻiːʁ, pyːʁ]
  'red, breeze, canal lock, for, to say, pure'

Poliquin (p.12) reports that these long vowels are tense on the surface, and he derives them *via* a so-called 'Duke-of-York gambit' (Pullum, 1976) in which

the vowels start out tense, undergo laxing, then become tense again, before being lengthened. In the data we elicited, these lengthened vowels seem to be lax, impressionistically. There may be dialectal variation in the realization of these lengthened vowels, but the issue does not bear on our primary concerns.

The last issue that complicates the distribution of high vowels in QF is vowel harmony. For many, but not all, speakers a high vowel in an open syllable can harmonize with a (derived) lax vowel in a following syllable. Consider some forms of the verb *cuisiner* 'to cook':

(5) Vowel Harmony

  - *cuisiner* [kᵁi.zi.ne] 'to cook' *vs. (il) cuisine* [kᵁɪ.zɪn] 'he cooks'
  - A lax vowels shows up in an **open** syllable, harmonizing with the vowel in the following closed syllable.

In [kᵁi.zi.ne] all the syllables are open, and as expected, the two high vowels surface as tense *i*. In the inflected form [kᵁɪ.zɪn], the word final vowel is in a closed syllable and it surfaces as lax *ɪ*, as expected. However, for some speakers, the first vowel of this form is also lax in harmony with the final vowel. Some speakers do not display harmony, so their pronunciation of *cuisine* is [kᵁi.zɪn].

If we restrict ourselves to two syllable words containing two high vowels and a final closed syllable (closed by anything other than a voiced fricative), there are two possible patterns in QF—either the vowel of the initial syllable is tense as is normal in an open syllable (QF2a), or it harmonizes with that of the final (QF2b) :

(6) Patterns in disyllabics with high vowels and final closed syllable

  - SF with no laxing: *(C)i.CiC.*   [kᵁi.zin]
  - QF2a with laxing, no harmony: *(C)i.CɪC.*   [kᵁi.zɪn]
  - QF2b with laxing and harmony: *Cɪ.CɪC.*   [kᵁɪ.zɪn]

However, once we consider longer words, the question arises of what happens to strings of three or more high vowels. Intuitively, one expects the following patterns for three syllable words with two high vowels in open syllables followed by a high vowel in a final closed syllable:

(7) Patterns in trisyllabics with high vowels and final closed syllable

4

- SF with no laxing: *i-i-i*
- QF3a with laxing, no harmony: *i-i-ɪ*
- QF with laxing and harmony:
  - QF3b with 'iterative' harmony: *ɪ-ɪ-ɪ*
  - QF3c with 'local' harmony: *i-ɪ-ɪ*

In other words, dialects with laxing and harmony will either apply the harmony across the board, yielding a string of three lax vowels (case b), or else, harmony will apply just to the first vowel to the left of the vowel subject to closed syllable laxing (case c). We agree with Poliquin that both dialects are attested.

Our disagreement with Poliquin is about another logical possibility that he identifies which becomes a focal point of his complex analysis. Poliquin claims that there are dialects that have both laxing and harmony, but generate an alternative to QF patterns (b) and (c):

(8)   Poliquin's third option for a dialect with laxing and harmony:
      QF3d with NLNIVH: *ɪ-i-ɪ*

This pattern involves what Poliquin calls 'non-local, non-iterative vowel harmony (NLNIVH)'. Of course, it is hard to find examples of trisyllables or longer words with strings of high vowels that could potentially manifest this pattern, and Poliquin pools the data from the eight potential forms he identifies. The most common form, and the form he uses as a label for this word-shape (trisyllabic with three high vowels, two in open syllables and a final closed syllable) is *illicite* 'illicit'.

Our research agrees with Poliquin's findings of four QF pronunciations of the vowels in this word, in addition to the SF pronunciation, which is also attestable in Québec because of the prestige of that dialect: [3]

(9)   Five pronunciations of the vowels of *illicite*

---

[3]Some of our (non-linguist) informants made comments that displayed a high level of meta-linguistic awareness. For example, one speaker's comments suggested a perceived conflict between the formality of the word *illicite* and the non-prestige of laxing. He had trouble deciding whether the first vowel of *illicite* could be *ɪ*, but he was completely certain that he pronounced the non-formal word *illégal* with initial *ɪ*. As we will see below, this makes no sense under Poliquin's analysis since there is no high lax vowel for the initial vowel of *illégal* to harmonize with.

SF3 *i-i-i* (no laxing, no harmony)

QF3a *i-i-ɪ* (laxing, no harmony)

QF3b *ɪ-ɪ-ɪ* (laxing, across-the-board harmony)

QF3c *i-ɪ-ɪ* (laxing, local harmony)

QF3d *ɪ-i-ɪ* (putative NLNIVH)

So, there is no dispute about the range of variation in the vowels manifested in a word like *illicite*. There are speakers who produce forms with no laxing (and thus no harmony), as in SF. There are speakers with laxing but no harmony, as in QF3a. There are speakers with laxing and across-the-board 'iterative' harmony, as in QF3b. There are speakers with laxing and 'local' harmony, as in QF3c. Finally, there are speakers with the vowels seen in QF3d, and our dispute with Poliquin lies with the explanation for such forms. In brief, Poliquin explains them as manifesting laxing and harmony, but with a non-iterative harmony process that targets the leftmost high vowel in the word, which in the case of *illicite*, is the initial vowel. Our findings suggest that in fact the pronunciation *ɪ-i-ɪ* for the vowels in *illicite* instead reflects the same dialect as manifested in QF3a, showing laxing without harmony. The difference in the two outcomes is a matter of the input representation and resulting syllable structure. The lax initial vowel in QF3d is a result of a closed syllable created by a geminate /l:/. Such forms do not provide evidence of Poliquin's radical claim of NLNIVH.[4]

3 Three strikes against NLNIVH in *illicite*

We must reiterate that the data on vowel distribution in all of Canadian or Québec French is quite complex, however, we are not ready to accept the complex analysis Poliquin offers for the set of forms we call QF3d. We offer a conceptual argument, a logical argument, and an empirical argument, that all combine to cast doubt on Poliquin's proposal.

---

[4]Since geminates do not appear in SF, a form like QF3a could result from a speaker attempting a prestige pronunciation and suppressing non-prestige geminate /l:/, but 'failing' to suppress laxing in the final closed syllable. Our intuition is that this is as likely an explanation of QF3a forms as the idea that in some QF dialects there is no geminate /l:/ (either in this word, or generally).

3.1 'Long-distance' is not 'non-local'

Our conceptual argument is that Poliquin's notion of non-locality is anomalous in the linguistics literature. There are many cases of long-distance phenomena in both the syntactic and phonological literature, but they all refer to dependencies or interactions that *may* happen at a distance, not that *must* happen at a distance. For example, in languages that are reported to have long-distance anaphors, like Icelandic, a long-distance anaphor can cause ambiguity with respect to the antecedent (see Thráinsson, 2007, for discussion and references):

(10) Long-distance anaphors

    a. **English** LOCAL:
       Sigga$_i$ says that Maria$_j$ loves herself$_{*i/j}$

    b. **Icelandic** LONG-DISTANCE (ambiguous):
       Sigga$_i$ segir að **Maria**$_j$ elski sig$_{i/j}$

    c. **Unattested** OBLIGATORILY LONG DISTANCE:
       Sigga$_i$ says that Maria$_j$ loves herself$_{i/*j}$

English anaphors like *herself* in (a) can only have a local antecedent such as the clausemate *Maria$_j$*. The Icelandic anaphor *sig* can be bound by either a local antecedent in the same clause, *Maria$_j$*, or a long-distance antecedent in a higher clause, *Sigga$_i$*. However, case (c) does not exist. There are no obligatory-long-distance binding phenomena where the antecedent *must* be non-local. The same holds for long distance effects like sibilant assimiliation or dissimilation. These relations/processes can apply over a span, from one end of a word to the other, but they also can apply at a closer distance (see Rose and Walker, 2004, for a survey). Poliquin's proposal that a relation holds between the first and last vowels of *illicite*, but not between the medial vowel and the final vowel would be quite extraordinary in the context of other well known long-distance interactions.[5] Long-distance relations seem to *include* local relations. This conceptual argument is not a proof that Poliquin is wrong, but it highlights just how radical his proposal is. Maxime Papillon (p.c.) points out that NLNIVH appears to be an example of what has been

---

[5]Of course, Poliquin's long-distance harmony can be *accidentally* local, for example, in a two-syllable word like *Philippe*. The point is that the domain is defined as *starting* at the beginning of the word, at a distance from the triggering segment.

called *first-last assimilation* in the computational phonology literature (Lai, 2015, for example), a pattern that this literature claims to be unattested for principled reasons.

3.2   NLNIVH overgenerates and undergenerates

As Poliquin points out, and as we illustrated above, disyllabics cannot distinguish between a dialect with long-distance harmony and local harmony; and in fact, disyllabics will also fail to distinguish among those two options and NLNIVH. All three dialects would treat Poliquin's class of words with the shape of *Philippe* the same. Such would surface as [fɪlɪp] in all three dialects.

Crucially, one word of this this shape that Poliquin does not list is morphological related to *illicite*: *licite* 'licit'. Among our informants, there is nobody who produces *illicite* with ɪ-i-ɪ as predicted by NLNIVH, and also produces *licit* with ɪ-ɪ, which would be consistent with NLNIVH as well. Instead, speakers who produce *illicite* with ɪ-i-ɪ consistently produce *licite* with the vowels i-ɪ. So, for the disyllabic, they seem to have laxing, but no harmony. Under Poliquin's analysis, the trisyllabics and disyllabics would have to be subject to different grammars. Poliquin's own logic predicts a different outcome for *licite* in his proposed NLNIVH dialects than what we have seen attested. In brief, the NLNIVH account *overgenerates* lax vowels, since it wrongly predicts a lax vowel in *licite* for the relevant dialects. Poliquin does not seem to have realized that disyllabics were important in this way.

In addition to overgenerating a lax vowel in NLNIVH version of *licite*, Poliquin's account also *undergenerates* lax vowels—it fails to generate lax vowels that actually do surface. Poliquin mentions *illumine* 'illuminates' [ɪ.ly.mɪn], as another example of NLNIVH. However, his model predicts that the infinitive of this verb *illuminer* should be [i.ly.mi.ne], since there are no closed syllables in this form. Poliquin does not cite any pronunciation for *illuminer*, and our informants have a lax initial vowel in both forms of the verb. So, NLNIVH undergenerates the laxing of the initial vowel of the infinitive form. The lax initial vowel of *illégal*, also not discussed by Poliquin, is also undergenerated by the NLNIVH account, since there are no lax high vowels with which the initial vowel could harmonize.

3.3 Geminates to the rescue

A simple, but underappreciated fact about QF phonology explains at least some of the unexpected lax vowels that Poliquin tries to account for with NL-NIVH. This fact also explains the lax vowels that Poliquin fails to address, the ones NLNIVH undergenerates in *illuminer* and *illégal*. It turns out that many, if not all, dialects of QF have some geminate consonants, unlike Standard French (Morin, 2010).[6] In QF geminates appear to be restricted to the sonorants /l,m,n/.

Tautomorphemic (underlying) geminates are present in words like *grammaire*, whereas a word like *amère* has a single [m]. Derived heteromorphemic geminates arise typically in clitic clusters, as in (11):

(11)   Heteromorphemic gemnates in QF

- *elle apprend* [ælaprã] 'she is learning'
- *elle l'apprend* [æl:aprã] 'she is learning it'

Spelling is an unreliable indicator of gemination. The durational difference of single [m] and geminate [m:] in *amère* and *grammaire* is apparent in the spectrogram in (12):

(12)   Spectrogram of *la grammaire, amère*

---
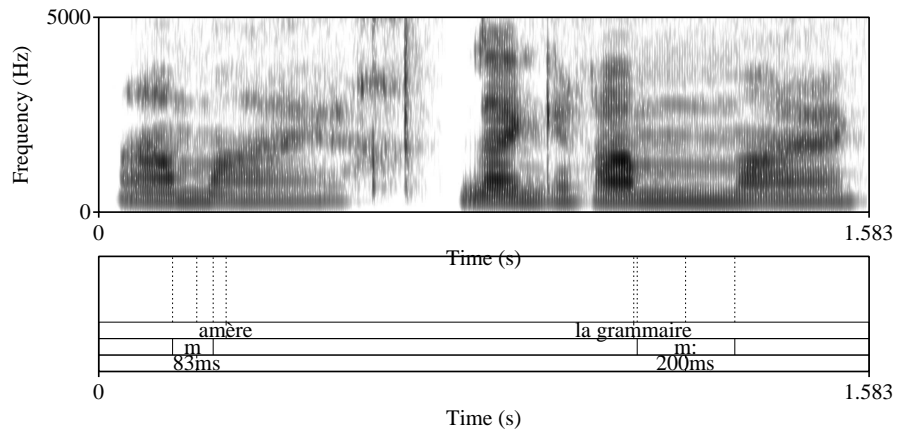
[6]Walker (2001, section 5.1) reports some geminates in SF, but they are irrelevant to our discussion.

9

The single [m] of *amère* is about 83ms., whereas the geminate [m:] of *grammaire* is about 200ms. Such an extreme difference, with the geminate more than twice as long as the single consonant, is not rare for sonorants (Local and Simpson, 1999; Cohn et al., 1999).

Poliquin does not recognize the existence of geminates in Canadian French, however, it is crucial to do so if we are concerned with phenomena that are sensitive to syllable structure. An obvious source of unexpected lax high vowels is the existence of syllables closed by the first half of a geminate.

Consider the spectrograms of three words, *illégal, illicite, il solicite* 'illegal, illicit, he solicits' in (13). The first word has a fairly long [l] of approximately 108 ms. The first vowel is clearly sounds lax, but since there are no other high vowels in the word, harmony cannot be relevant. As noted above, the NLNIVH account of laxing *undergenerates* in this case, because it does not predict the lax initial vowel of *illégal*. In fact, vowel is lax because the long [l] is a geminate, which closes the initial syllable, so the correct analysis is [ɪl.le.gal.] with a geminate /l:/. This is just closed syllable laxing. Poliquin's analysis cannot account for such facts.

The [l] of *illicite* is even longer, about 174 ms. The same explanation is available as for *illégal*. The lax initial vowel is due to laxing in a syllable closed by the first half of the long [l]. In this recording of *illicite*, the medial vowel, in an open syllable, is tense, so this speaker appears to have no harmony from the final vowel. NLNIVH is not needed to account for the initial lax vowel—it is lax because it is in a closed syllable.

In *(il) solicite* the final vowel is lax because the syllable is closed by *t*. The preceding vowel is tense, which is consistent with what we see in *illicite*—this speaker appears to have no harmony, which is why there was a tense medial vowel in *illicite*. The vowel of this syllable is open, since this [l] in *solicite* is clearly a singleton, with a duration of only 38 ms.[7]

(13)   Spectrogram of *illégal, illicite, il solicite*

---

[7]Note that in QF there is no [l] pronounced at all in the subject clitic *il*.
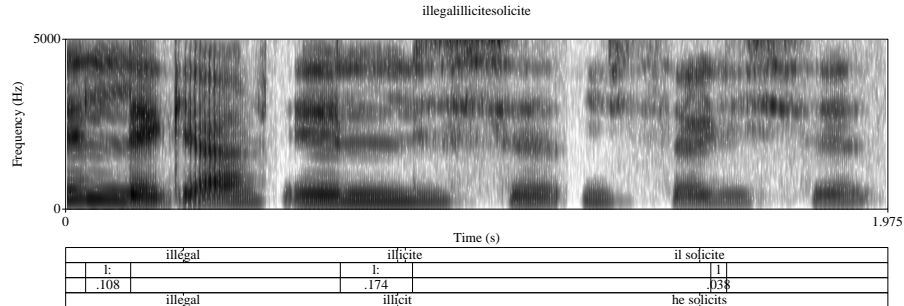
The QF3d pronunciation of *illicite* with the vowels ɪ-i-ɪ reflects the a dialect with no harmony, but a closed initial syllable: ɪl.li.sɪt. There is no reason to posit NLNIVH.

4   Conclusion

This paper is not an account of the messy distribution of lax high vowels in all Canadian French dialects (as well as all forms that reflect attempts to mimic SF and other prestige variants). Rather it is an argument against a particular radical enhancement of phonological theory that would allow for non-local non-iterative vowel harmony. Of course, it is impossible to prove that such phenomena do not exist in any langauge and that phonological theory will never be forced to accept the means to generate them. However, our conceptual argument forces us to accept that NLNIVH is truly different from other long-distance phenomena we are familiar with. Our logical argument points out an error of prediction in that the NLNIVH approach predicts that speakers who have the ɪ-i-ɪ pattern in trisyllabics like *illicite* should have the ɪ-ɪ pattern in disyllabics like *licite*, contrary to our observations—NLNIVH overgenerates. Also, NLNIVH predicts that speakers who have an initial lax vowel in *illumine*, should have a tense initial in *illuminer*, again contrary to our observations. Like our mention of *illégal* above, this shows that NLNIVH undergenerates, since it predicts tense vowels (the absence of laxing), where the data manifests lax vowels. Under our account, the correct transcriptions of *illumine* and *illuminer* are [ɪl.ly.mɪn] and [il.ly.mi.ne], both of which have lax vowels in an initial closed syllable.

11

So, NLNIVH both under- and overgenerates lax vowels under Poliquin's own assumptions. Finally, our empirical instrumental evidence attests to the presence of segmental durations that are compatible with the presence of phonological geminates, which in turn provide a straightforward account of some otherwise unexpected lax high vowels in QF.

Occam's Razor prevents us from accepting Poliquin's enhancement to the *computational* capacities of current models of phonology, an enhancement that would allow NLNIVH. Another recent PhD thesis, Bosworth (2011), also takes on the problem of the distribution of QF lax high vowels. Bosworth proposes an innovation to the *representational* capacities of phonological Universal Grammar, an entity she calls the *hypomora*. Bosworth, like Poliquin, fails to recognize the geminates of QF. Again, we think that radical solutions should not be considered until we have exhausted the resources of current models and paid fuller attention to a salient detail of the language under analysis. Although full diachronic and synchronic accounts of all forms produced by QF speakers would surely be complex and messy, there is no doubt that geminates are implicated in the distribution of high lax vowels in QF.

We close with a somewhat philosophical point. We think that some readers will find our phonetic documentation of geminates and our arguments about over and undergeneration by NLNIVH more convincing than the conceptual argument that NLNIVH is unlike other long-distance phenomena. In terms of this paper, at least, it was our (rationalist) prejudices about what is a possible phonological process that led us to seek out further empirical facts and reject NLNIVH—our biases suggested that Poliquin's 'observations' could not be valid. We believe that this rationalist approach to inquiry is necessary and unavoidable—there are no facts that speak for themselves. This simple example from an accessible, fairly well-studied language suggests that one must be especially careful about accepting 'facts' that lead to radical theoretical innovations.

References

Bale, Alan, Maxime Papillon, and Charles Reiss. 2014. Targeting underspecified segments: A formal analysis of feature changing and feature filling rules. *Lingua* 148:240–253.

Bale, Alan, and Charles Reiss. 2018. *Phonology: A formal introduction*. MIT

Press.

Bosworth, Yulia. 2011. Weight and feet in Québécois. Doctoral Dissertation, University of Texas, Austin.

Chomsky, Noam. 2000. Language as a natural object. In *New horizons in the study of language and mind*, 106–133. Cambridge, MA, US: Cambridge University Press.

Chomsky, Noam. 2007. Approaching UG from below. In *Interfaces + recursion = language?: Chomsky's minimalism and the view from syntax-semantics*, ed. Uli Sauerland and Hans-Martin Gärtner, volume 89, 1–24. Berlin: Mouton de Gruyter.

Cohn, Abigail C, William H Ham, and Robert J Podesva. 1999. The phonetic realization of singleton-geminate contrasts in three languages of indonesia. In *Proceedings of ICPhS*, volume 14, 587–590.

Lai, Regine. 2015. Learnable vs. unlearnable harmony patterns. *Linguistic Inquiry* .

Local, John, and Adrian P Simpson. 1999. Phonetic implemenation of geminates in Malayalam nouns. *ICPhS* 595–9.

Morin, Annick. 2010. Diachrony and synchrony of/l/gemination in Québec french. In *Proceedings of the 2010 annual conference of the Canadian Linguistic Association.*, ed. Melinda Heijl. Online Proceedings: Canadian Linguistic Association.

Poliquin, Gabriel Christophe. 2006. Canadian French vowel harmony. Doctoral Dissertation, Harvard University Cambridge, Massachusetts.

Pullum, Geoffrey K. 1976. The Duke of York Gambit. *Journal of linguistics* 12:83–102.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80:475–531.

Thráinsson, Höskuldur. 2007. *The syntax of icelandic*. Cambridge University Press.

Walker, Douglas C. 2001. *French sound structure*, volume 1. University of Calgary Press.