Laurentian French Affrication in External Sandhi: The Facts, and a CVCV Analysis

Marie-Michèle Beausoleil¹, Heather Newell¹

¹Université du Québec à Montréal

1 Introduction

This article describes and analyzes the behaviour of Laurentian French Affrication (LFA) in external sandhi. This phenomenon is recognized as one of the most important markers of Laurentian French (Dolbec & Paradis 2008). It is quite well-documented (ex. Gendron 1970, Santerre 1976, Dumas 1987, Friesner 2010, Brassard 2019, Dow 2019, Pöchtrager 2021), but there is little consensus on many of its properties, notably its phonological conditioning, its diachronic, synchronic, and sociolinguistic variation, and its phonetic description. The literature agrees that affrication is obligatory inside words, as in the /ti/ sequence inside petit (1a), and that it is variable in external sandhi; the /t#i/ sequence in (1a, Figure 1) does not trigger affrication, while the /t#i/ sequence in (1b, Figure 2) does. Nonetheless, there is very little data on the conditioning of the variable nature of LFA in external sandhi. This article aims to begin to fill this gap.

(1) a. / la petit iɲɔʁ/ [lapatsitɲɔʁ]
   DEFDET.FEM little.FEM not.know
   ‘the little one doesn’t know...’

   b. / sɛt imaʒ/ [satsimaʒ]
   DEM.FEM image
   ‘this image’

Here we aim at an initial description and analysis of the variable nature of LFA in external sandhi. We will demonstrate that this variation has syntactic conditioning similar to that which is found with liaison (for an overview see Côté 2011). LFA, we propose, may offer a cleaner view of the syntax-phonology interface in French than liaison does, as, unlike with liaison in formal speech, there is no sociolinguistic motivation to over-apply LFA. In the sections below we show that the LFA’s variation in external sandhi is not strictly optional. In certain environments LFA is obligatory in external sandhi, in others it is banned, and in still others it does display optionality.

Section 2 reviews the main descriptions of LFA in the literature. Then, in §3 we describe the methodology used to collect and analyze the discourse of 35 native speakers of Laurentian French as well as...
the results obtained. Section 4 then presents a brief analysis of the syntactic and phonological conditions that underlie the patterns in the data in §3. Section 5 concludes.

Figure 1: Absence of affrication at the word boundary of petite and ignore

Figure 2: Affrication at the word boundary of cette and image

2 Affrication in Laurentian French

LFA targets coronal stops (/t, d/) when followed by a high, front vowel or glide (/i, y, j,ɥ/), as in (1,2).

\[(2) \quad \begin{align*}
    &a. /dy/ \; [\tilde{d}zy] \; \text{‘due’} \\
    &b. /diz/ \; [\tilde{d}zi:z] \; \text{‘say’} \\
    &c. /tjil/ \; [\tilde{t}ʃil] \; \text{‘tile’}
\end{align*}\]
Affrication is also described as palatalization or assimilation in the literature. Walker (1980, 1984) opts for the latter description, and Brasseur (2019) notes that Laurentian French assimilation is only partial. The link between assimilation and palatalization is due to the triggering environment: high front vowels and glides, typical palatalizing segments, cause LFA. But, as Laurentian French also allows for a distinct operation of palatalization, we follow Dolbec & Paradis (2008) and Pöchtrager (2021) in denying the palatal nature of LFA. We opt for the term affrication throughout, as this captures the partial nature of the assimilation process.

There are several potential symbolic representations for LFA. We follow Brasseur (2019) in using the symbols [ts] and [dz], as it appears that LFA is a double articulation (an articulatory target) rather than being the result of coarticulation (which we would transcribe as [tʰ] and [dʃ]).

3 Data collection and analysis

3.1 Participants In order to compensate for the lack of empirical data currently available on LFA in external sandhi, 35 native speakers of Laurentian French (20 women and 15 men, between the ages of 26 and 75) were asked to record themselves reading sets of provided sentences. The participants came from various regions of Québec (the Eastern Townships, Lanaudière, the Laurentians, and the greater Quebec City and Montreal areas) and from a variety of socio-cultural and professional backgrounds (ex. students, artists, self-employed, employees in various fields, and retirees).

3.2 Stimuli Two lists of sentences were created to facilitate potential LFA in various distinct syntactic environments where a word ending in a coronal stop was followed by a word beginning with a high front vowel or glide. These same sentences also included obligatory word internal LFA environments to control for the base rate of affrication of each participant. Reading each set of sentences took roughly 2 minutes.

The advantage of this methodology is that it produces identical data for all participants, allowing for clear comparison and statistical analysis. A disadvantage is that reading tasks may trigger the use of a formal register, leading to distinctions between the data obtained and spontaneous speech. Fortunately, the rate of affrication generally decreases in formal Laurentian French, therefore the presence of affrication in the reading task predicts its presence in the same environment in spontaneous speech. The data here therefore likely underpredict the rate of LFA in Laurentian French.

3.3 Procedure Participants were recruited via social networks. Volunteers were asked to record themselves reading a list of sentences as naturally as possible. In case of a self-perceived error, participants were permitted to pause or reread the problematic sentence from the beginning but had to continue without stopping the recording. The recordings were taken on cell phones and sent via email.

3.4 Data analysis The stimuli were analysed by the first author (a native speaker of Laurentian French), listening to recordings with a Beyerdynamic DT 770 Pro headset and using Praat (Boersma & Weenink 2022). Affrication was coded as present or absent. As the level of affrication across instances is variable and the recordings obtained were not of equivalent quality, any ambiguous data was excluded. It is of note that LFA is not an invariant phenomenon (Dow 2019), even within words. Several factors, such as the acoustic quality of adjacent segments and the position of segments within a syllable, influence the surface qualities of affrication. Moreover, in cases of flagrant error (wrong word, marked pause, language tic, etc.), the stimuli were considered invalid and were excluded from analysis.

3.5 Results

3.5.1 Systematic affrication Affrication was, as expected, systematic, at 100%, within words. It was also systematic in the following four environments: between a demonstrative determiner and a noun (3a), between an adjective and a noun (3b), between a verb and the determiner of its object (3c), and between an epenthetic [t] and a following pronoun (3d).

1 There was one instance of word-internal non-affrication; in an expression associated with Hexagonal French, a dialect that does not include affrication. There was also one instance of non-affrication between a determiner and a noun; there was an unnaturally long pause between the two words. These data were excluded.
Laurentian French Affrication in External Sandhi

Beausoleil and Newell

3.5.2 Absent affrication  Affrication was blocked, occurring at 0%, between a subject nominal, either common (4a) or proper (4b), and a following verb.

3.5.3 Optional affrication  When a potential affrication target failed to emerge with affrication more than once, but did emerge as affricated more than once, affrication for that target was coded as optional. Environments with optional affrication were between a verb and a proper noun (5a), between a noun and an adverb (5b), between a numeral and a noun (5c), between a demonstrative determiner and an h-aspiré (5d), and between the indefinite determiner d’ and a noun beginning with [ɥ] (5e).

(3)  a.  [sɛt̪simaz]  
Il est évident que cette image vaut mille mots.  
‘It is clear that this picture is worth a thousand words’

b.  [pɔtsɪˈvʁɔm]  
Ce n’est vraiment pas correct de traiter quelqu’un de petit ivrogne.  
‘It’s really not nice to call someone a little drunk.’

c.  [sɛtsyn]  
C’est une carotte identique.  
‘It’s an identical carrot.’

d.  [tsil]  
Marche-t-il toujours avec son énorme chien?  
‘Does he always walk with his enormous dog?’

(4)  a.  [lapɔstɪˈtɪmɔʁ]  
Mario a dit que la petite ignore ce qui l’attend.  
Mario be.3SG say that DEFDET.FEM girl not.know that which her’await.3SG  
‘Mario said that the girl doesn’t know what to expect.’

b.  [ivɛtɪˈmazn]  
Yvette imagine toujours le pire.  
‘Yvette always imagines the worst.’

(5)  a.  [pʁɛzɑtɪˈtɪmɔʁ] ~ [pʁɛzɑtɪˈsɪvɔʁ]  
Présente Yvon aux enfants s’il est là demain.  
introduce.3SG Yvon to.the children if’he be.3SG.PRES there tomorrow  
‘Introduce Yvon to the children if he is there tomorrow.’

b.  [pɔstiˈʁ] ~ [pɔstɪˈʃɔʁ]  
La lettre, je l’ai envoyée à la poste hier.  
DEFDET.FEM letter 1SG.PRON it’have.1SG sent to DEFDEM.FEM post office yesterday  
‘I sent the letter to the post office yesterday.’

c.  [sɛtɪma3], [yitɪma3] ~ [sɛtɪma3], [yitɪma3]  
Il y avait sept images, pas huit images.  
It there have.3SG.PROG seven images not eight images  
‘There were seven images, not eight images.’
Laurentian French Affrication in External Sandhi

In sum, the rate of affrication across all data can be found in Figure 3.

![Figure 3: Rates of LFA in different syntactic environments](image)

### 3.5.4 Potential Confounding Factors

Among the collected results, some data stood out as containing potential confounding factors for the rate of affrication: focus, the glide /q/, sept-cette homophony, and h-aspiré. We discuss each briefly in the sections that follow, noting that their effects need to be examined further in future work.

#### 3.5.4.1 Focus

In two cases we saw an inhibiting effect of focus stress on the rate of affrication. A sequence of two words that created a potential environment for affrication triggered LFA at a high rate in the absence of focus (6) but evidenced LFA at a much lower rate in the presence of focus stress (7).

(6) a. [tɛ̃ʃmä] (85%)
Il avait une tête immense ou une immense tête?
‘He had a head that was huge or a huge head?’

b. [sɛ̃ʃmä] (95%)
J’ai bien compté, il y avait huit images, pas sept images.
‘I counted properly, there were eight pictures not seven pictures.’

(7) a. [tɛ̃ʃmä] (0%)
J’te jure, elle avait une tête IMMENSE.
‘I assure you she had a HUGE head.’
b. [sctimaʒ] (50%)  
J’ai bien compté, il y avait sept images, pas huit images.  
‘I counted properly, there were seven images not eight images.’

This distinction is not surprising, as stress is often associated with a strengthening of prosodic boundaries which in turn may interfere with phonological processes.

3.5.4.2 The glide /ɥ/ Affrication is found variably (65%) between the preposition d’ ‘of’ and the glide /ɥ/ (8), but systematically (100%) between this preposition and a noun beginning with a vocalic trigger (9).

(8)  a. [d̛i̯l] ~ [dʒi̯l]  
Ne mettez aucun gras sur la poêle; pas d’huile, pas de beurre!  
‘Do not put any fat on the pan: no oil, no butter!’

(9)  a. [d̛ylɛʁ]  
Mon docteur me l’a souvent dit : pas d’ulcère, pas de problème.  
‘My doctor often told me: no ulcer, no problem.’

French allows for the syllabification of glides as consonants or vowels, although the former is generally restricted to loan words. If the glide in the above (8) may be syllabified as a consonant, and if consonants do not trigger affrication, we might explain the variability seen. The above pattern warrants further investigation.

3.5.4.3 Numerals The numbers sept ‘seven’ and huit ‘eight’ figured among the initial words in potentially affricating word-pairs. Notably, pairs that included sept followed by a high front vowel or glide showed 95%, affrication (modulo the effect of focus in §3.5.4.1). The final consonant of huit, however, only affricated at a rate of 62.5%. We propose that this confound is due to the homophony between the numeral sept and the demonstrative cette. The latter evidenced systematic (100%) affrication (as function words do in the data, when no confounding factors are present). This hypothesis requires further investigation.

3.5.4.4 h-aspiré h-aspiré is known to block liaison (ex., Robert and Rey 2017). However, this ‘unpronounced’ consonant is not present in the lexicon of all speakers of French in a stable manner; some words that begin with an h-aspiré for some speakers are truly vowel-initial for others. For example, although the majority of speakers pronounced the sentence (10a) with an h-aspiré (and no liaison) (10b), a few, unsurprisingly, treated the same word as having no initial onset and applied liaison between the indefinite determiner un and the following noun (10c).

(10)  a. il y avait un hippie révolutionnaire au conseil des ministres.  
‘There was a revolutionary hippie in the cabinet.’

(11)  a. Je l’aime bien, cette hippie révolutionnaire.  
‘I quite like her, that revolutionary hippie.’
b. [sɔ̃si̯pi] (55%)
c. [s trìpi] (45%)

3.5.5 Conclusions of the data-analysis
In sum, if we abstract away from the confounding factors in §3.5.4, LFA is systematic in the environments in (12), absent in the environments in (13), and variable in the environments in (14).

(12) Systematic affrication
a. word-internally
b. epenthetic [t] - pronoun
c. demonstrative determiner – noun
d. adjective - noun

e. verb – determiner of object
f. d’ - noun

(13) No affrication
a. lexical subject – verb
b. noun–focused adj.

(14) Optional affrication
a. huit – noun
b. verb – proper noun
c. noun – adverb
d. noun–non-focused adj.

It is clear that LFA in external sandhi is not optional generally. There are patterns in the data to be explained. In the following section we offer an analysis of (12) and (13) that appeals to the theory of syntactic phases (Chomsky 2001) and the representational framework of CVCV phonology (Lowenstamm 1996, Scheer 2004, 2012). We briefly discuss paths to follow in the analysis of (14) in §4.1.3.

4 LFA and the Syntax-Phonology Interface

4.1 Phases and Phonological Realization
In the literature on morpho-syntactic cyclicity it is proposed that only certain chunks of morpho-syntactic structure undergo phonological processing as units. Whether two words or morphemes are processed together in the phonology then has an impact on their phonological derivation. The current cyclic hypothesis in the generative literature is that syntactic cycles, phases, are triggered at vP and CP, and also by DP, PP and categorizing heads (see ex. Citko 2014 for an overview of the phase literature).

Once a phase has been sent to the PF branch (we ignore LF here), it will undergo Vocabulary Insertion (in realizational theories like Distributed Morphology; Halle & Marantz 1993, 1994), and then the linearized string of morphemes will be sent to the phonology-proper. There, the underlying phonological forms of these morphemes are subjected to phonological processing. Theories of the interface between the syntax and the phonology propose that each string that arrives at the phonology-proper be marked as belonging to a phonological domain. Prosodic Phonology (Selkirk 1986, Nespor & Vogel 1986 and innumerable subsequent works) proposes that a string be marked as belonging to a prosodic constituent (ex. Prosodic Word (PWd), Prosodic Phrase (PPh)), as in (15), where the subject and the verb are sent to PF separately according to phase-theory. Note that this discussion of Prosodic Phonology is very simplified. It is assumed that the reader is familiar with the Prosodic Hierarchy; this brief exposition serves as a comparison with the representational theory we will adopt in (16).

(15) [MariePWd] [runsPWd]

An alternate proposal of phonological domain formation can be found in CVCV phonology (Lowenstamm 1996, Scheer 2004, and much subsequent literature). This theory proposes that prosodic domains are not constituents; there is no hierarchical structure in the phonology. Instead, each linear string that is sent to the phonology from the syntax at the spellout of a phase is endowed with an empty CV (akin to an empty syllable) at its left edge, as in (14).

(16) CVMarie CVruns

The following sections will demonstrate how phases, combined with a CVCV phonology approach to the
interface, can offer an analysis of the patterns in §3.

4.1.1 Nested Phases and LFA  Let us consider determiner-noun sequences. We saw above that LFA is systematic in these environments. Let us add to the discussion a consideration of the syntactic structure of these constructions.

(17)   a.  

cette image  ‘this image’  [sʁtʃimaʒ]

   b.  

   \[
   \text{DP} \\
   \text{D} \quad \text{nP} \\
   | \quad \text{cette} \\
   | \quad \text{image}
   \]

   Here, the first phase is the (categorizing) nP. When the noun undergoes spell-out, it will be sent to the phonology, and a left-edge CV will be inserted (18a). However, in CVCV phonology, it is proposed that empty VC sequences be deleted (18b, indicated by diagonal hatching; following Gussmann & Kaye 1993). The output of the nP phase is therefore as in (18c).

(18)  a.  

   \[
   \begin{array}{ccccccc}
   \text{C} & \text{V} & \text{C} & \text{V} & \text{C} & \text{V} \\
   | & | & | & | & | \\
   s & ε & t & i & m & a & ʒ
   \end{array}
   \]

   b.  

   \[
   \begin{array}{ccccccc}
   \text{C} & \text{V} & \text{C} & \text{V} & \text{C} & \text{V} \\
   | & | & | & | & | \\
   i & m & a & ʒ
   \end{array}
   \]

   c.  

   \[
   \begin{array}{ccccccc}
   \text{C} & \text{V} & \text{C} & \text{V} \\
   | & | & | \\
   s & ε & t & i & m & a & ʒ
   \end{array}
   \]

   The subsequent phase is the DP. When the DP undergoes spell-out, the entire DP is visible to the phonology: the new material (in D) and the phonological representation that was the output of the previous phase (18c). The input to the phonology in this cycle is therefore as in (19a). It undergoes VC deletion in (19b) and has the output in (19c).

(19)  a.  

   \[
   \begin{array}{ccccccc}
   \text{C} & \text{V} & \text{C} & \text{V} & \text{C} & \text{V} & \text{C} & \text{V} \\
   | & | & | & | & | & | & | \\
   s & ε & t & i & m & a & ʒ
   \end{array}
   \]

   b.  

   \[
   \begin{array}{ccccccc}
   \text{C} & \text{V} & \text{C} & \text{V} & \text{C} & \text{V} & \text{C} & \text{V} \\
   | & | & | & | & | & | & | \\
   s & ε & t & i & m & a & ʒ
   \end{array}
   \]

   c.  

   \[
   \begin{array}{ccccccc}
   \text{C} & \text{V} & \text{C} & \text{V} & \text{C} & \text{V} & \text{C} & \text{V} \\
   | & | & | & | & | & | & | \\
   s & ε & t & i & m & a & ʒ
   \end{array}
   \]

   Notably, in (19c) the sequence \textit{cette image} has the phonological structure of a monomorphic string; no syllabic structure intervenes between the final consonant of the first word, and the initial vowel of the second (in grey). It is in this configuration that LFA is systematically triggered.

4.1.2 Non-nested Phases and LFA  Let us now consider lexical (non-pronominal) subject-verb sequences. In these cases, LFA does not occur. The reason for this stems again from an interaction between the syntactic structure and the phonological representation. Consider the relevant syntactic structure of \textit{Yvette imagine}. 
Laurentian French Affrication in External Sandhi

Beausoleil and Newell

(20) a. Yvette imagine ‘Yvette thinks’ [ivetimaʒin]

b. CP
   ┌─────┐
   │ C │ TP
   └─────┘
   ┌─────┐
   │ DP │ TP
   └─────┘
   ┌─────┐
   │ Yvette │ imagine
   └─────┘

In the derivation of (20b) there are again two relevant phases: CP and DP. As DP is a syntactic left-branch (a specifier) it is spelled out separately from the tree it merges with (Uriagereka 1999). Its derivation at PF is as in (21).

(21) a. C V C V C V C V
    ┌─────┐
    │ i v e t
    └─────┘

b. C V C V C V C V
    ┌─────┐
    │ i v e t
    └─────┘

c. C V C V C V C V
    ┌─────┐
    │ i v e t
    └─────┘

Now, at the CP phase it appears that the verb might not be endowed with a left-edge empty CV, as it is not leftmost in the CP. However, it has been shown that the spell-out of CP may ignore the content in specifiers (Compton & Pittman 2010); the spellout of phases must take into account nested information as in (19) but may ignore left branches. Therefore, the spell-out of CP targets only imagine, as in (22). The two independent phases are then linearized (23).

(22) a. C V C V C V C V C V
    ┌─────┐
    │ i m a ʒin
    └─────┘

b. C V C V C V C V C V
    ┌─────┐
    │ i m a ʒin
    └─────┘

c. C V C V C V C V
    ┌─────┐
    │ i m a ʒin
    └─────┘

(23) a. C V C V C V C V C V
    ┌─────┐
    │ i v e t
    └─────┘

In (23) the VC-sequence in grey is not deleted. VC sequences only undergo deletion if they undergo phonological processing within a cycle. Here, the linearized string is determined post-cyclically, and therefore the structure that emerges from each cycle (as indicated by the boxes, which have no representational value in the phonology) is maintained. We can see that the final consonant of Yvette and the initial vowel of imagine are therefore separated by syllabic space; they are not appropriately local. LFA is triggered only under adjacency on the melodic (segmental) tier, and therefore is not triggered in (23).

4.1.2 Optional LFA Recall that optional affrication occurred in our data in sentences like the following (repeated from 5, where the glosses can be found);

(24) a. Présente Yvon aux enfants s’il est là demain.

verb – proper noun (71.4%)
b. Il y avait sept images, pas **huit images**.
   huit – noun  (62.5%)

c. La lettre, je l’ai envoyé à la **poste hier**.
   noun – adverb  (47.4%)

d. Il avait une **tête immense** ou une immense tête?
   noun – non focused adjective  (85%)

We do not have a worked-out analysis of why these constructions should evidence variable LFA. We can note, however, that verb-object structures are considered to be nested, and therefore the analysis presented in §4.1.1 predicts LFA between the verb and Yvon in (24a); it is therefore of interest that this variable LFA construction has the highest rate of optional LFA in (24). In (22b) we must contend with the fact that the syntactic literature allows for multiple analyses of numeral constructions (ex. Miechowicz-Mathiasen (2011) on the head vs. specifier classes of numerals in Polish); if some speakers analyse a numeral as a pre-nominal head, LFA should occur, while if it is analysed as a specifier of a numP, LFA should not occur. Clearly further investigation is required.

In the cases of (24c,d) our algorithm seems to predict that LFA should not occur. Uriagereka’s (1999) Multiple Spell-Out algorithm predicts both specifiers and adjuncts to spell out before merger to the main spine of the tree. As **hier** and **immense** are adverbal and adjectival adjuncts, they should be exempt from LFA. However, with regard to (24c), French adverbs can be divided into (at least) three groups (Abeillé & Godar 2003); non-light, light, and weak. Notably these adverbs have distinct syntactic properties, and lighter adverbs have been argued to be simplex heads. As **hier** is light, it might not be the kind of XP-adjunct discussed in Uriagereka (1999). Again, we have work to do at future stages of this project.

Another consideration when looking at optionality is production planning (ex. Wagner 2012). Wagner notes that Morin & Kaye (1982) report a difference in frequency of liaison between the verb and an adverbial modifier depending on the complexity of the latter.

   ‘Usually, they eat after they have finished their chess game.’

   b. Ils travaillent d’abord et mangent après
   ‘They work first and eat later.’

Morin & Kaye (1982, 296)

Notably, there is more likely to be liaison between **mangent** and **après** in (25b) than in (25a). As the adverbial modifier in (25a) is itself more complex than the one in (25b), it is proposed that the former takes longer to plan, predicting a stronger left-edge prosodic break. Looking at the rates of LFA in the environment of adverbials of different sizes will be part of our future work.

5 Conclusions

The data presented in this paper offer the first systematic look at the behaviour of Laurentian French Affrication in external sandhi in the literature. Although we present a pilot study here, the data already lead to the conclusion that LFA is not ‘optional’ in general, but rather is obligatory, optional, and banned in particular environments. This is reminiscent of studies on French liaison (see Côté 2011 for an overview), and it appears to pattern similarly (but note the mismatch observed in §3.5.4.4). However, the literature takes care to note that the synchronic status of liaison might be (at least sometimes) fossilized, and it is sometimes subject to sociolinguistic factors that lead to its over-application (and therefore optionality) in certain environments in formal registers. LFA, however, is not a prestige dialect feature (in the past, Québécois radio
announcers were trained to remove LFA from their speech, for example); its presence is never “affected”. It is also a stable feature of Laurentian French and does not appear to be fossilized. It therefore offers a clear look at the phonology-syntax interface in French.

References