Intrusive vowels preceding /R/ in Quebec French

Mélanie Lancien & Marie-Hélène Côté

Section SLI, Université de Lausanne
melanie.lancien@unil.ch , marie-helene.cote@unil.ch

ABSTRACT

It has been observed (e.g. [2]) that an intrusive vowel (*Vi*) is frequently produced in obstruent+/R/ clusters, which is interpreted as a cluster simplification process.

Here we look at *Vi* next to /R/ in a broader variety of contexts, in a corpus of 12,981 words read by 103 speakers of Quebec French [3, 4].

LMM shows that *Vi* is more likely to appear word-initially than in clusters, which questions the interpretation of *Vi* in terms of simplification. *Vi* is also more likely to appear with apical than dorsal /R/ (both attested in Quebec French) and, in clusters, after voiced obstruents. *Vi* is significantly more frequent in monosyllabic than in polysyllabic words. Regarding the quality of *Vi*, t-tests between *Vi* and the following vowel's mean F1 and F2 (Hz) and duration show that there is little to no difference between the quality of the two vowels, though *Vi* is significantly shorter.

Keywords: Vowel intrusion, Quebec French, Rhotics, Consonant clusters.

1. INTRODUCTION

French phonotactics allows a variety of consonant clusters: obstruent-obstruent (1a-c), obstruent-sonorant (1d), and sonorant-obstruent (1e), which can appear in branching onsets (1a), coda-onset sequences (1b), or branching codas (1c).

(1)	a.	*sport*	[spɔχ]	'sport'
	b.	*fructose*	[fχyk.toz]	'fructose'
	c.	*pacte*	[pakt]	'pact'
	d.	*fleur*	[flœχ]	'flower'
	e.	*sultan*	[syl.tɑ̃]	'sultan'

Clusters are often reduced by consonant deletion or vowel epenthesis, and simplification tends to be related to the degree of similarity between the consonants in the cluster [1]. Furthermore, [2] show for Quebec French (QF) that vocalic elements may appear in onset obstruent+/l, R/ clusters, where epenthesis had not been observed, as in *poudrer* 'to powder' pronounced [pudeʁe]. They find that these intrusive vowels (*Vi*) appear almost exclusively with /R/, not /l/. They looked at the frequency of vowel intrusion and the duration and quality of the vowel, with respect to a variety of factors (the obstruent's manner and place of articulation and voicing; position of the cluster in the word; stress; quality of the preceding and following vowels). They observed, inter alia, that *Vi* durations ranged from 27 to 50ms and *Vi*'s shared features with adjacent vowels. Vowel intrusion was also more frequent when the obstruent was voiced. However, [2]'s results are not globally congruent with the idea that intrusion is favored in clusters composed of more similar consonants.

Here, we further explore vowel intrusion in the context of /R/ in QF, building on the results of [2], but with more speakers, more phonotactic contexts, and more types of /R/. We observe intrusion before word-initial /R/, a fact also reminiscent of initial vowels in words like *regarde* 'look' [əʁgaʁd] and *demande* 'ask' [ədmɑ̃d] [3]. Intrusion not being necessarily dependent on clusters, we look at it both in clusters and word-initially. Knowing that QF exhibits variation between apical and dorsal realizations of /R/ [4], we also investigate the effect of the place of articulation of /R/ on the frequency of intrusion, in addition to the manner of articulation and voicing of the preceding obstruent, as well as position in the word. Finally we report on the duration and spectral features of *Vi* and compare them to those of the following vowel (*Vw* hereafter).

2. METHOD

2.1. Corpus

The data selected for this study is a subpart of the Canadian portion [4] of the Phonologie du Français Contemporain (PFC) project [5]. We considered 103 native speakers of QF (μ=51.6 years, σ=22.4) from ten different locations in Quebec and one in Ontario.

Among other tasks, each speaker read two lists of isolated words, and a total number of 13,724 tokens of /R/ were produced in 163 different word types. These tokens appeared in six broad contexts:

- #_V: *roc* [ʁɔk] 'rock'
- OBS_V: *gris* [gʁi] 'grey', *mettre* [mɛtχə] 'put'
- V_C: *turban* [tyʁbɑ̃] 'turban'
- V_V: *pourri* [puʁi] 'rotten'
- V_#: *port* [pɔχ] 'port'
- C_#: *mettre* [mɛtχ] 'put'

After a semi-automatic alignment [6] in Praat [7], each token was manually annotated for the type of /R/

(manner and place of articulation, and voicing, based on one annotator's perception and inspection of the spectrogram). When an intrusive vowel was present, its quality, based on F1 and F2 values on the spectrogram, was also indicated. A total of 2437 *Vi*'s were identified (μ=23 per speaker, σ=11.3). They all occurred in #_V, OBS_V and, marginally, V_C contexts. In #_V and OBS_V, the *Vi* always preceded /R/, in V_C it followed it. From now on, we will only focus on these three contexts (N=9304).

2.2. Data processing

For each annotated intrusive vowel Praat scripts extracted the duration (s) and mean F1 and F2 (Hz), as well as the word to which the vowel belonged, its phonemic context, and the duration (s), and mean F1 and F2 (Hz) of the following vowel. The preceding vowel was not considered, as it is irrelevant in the case of word-initial /R/. The extracted values were manually checked and corrected.

For instance, for every exemplar (one per speaker) of the word *creuse* [kχøz] 'hollow.FEM', our script created a column indicating the presence or absence of an intrusive vowel between [k] and [χ] and, if there was one, the script extracted its duration, mean F1, and mean F2, the preceding context [k], the following context [χ], as well as the duration, mean F1, and mean F2 of the vowel [ø]. The preceding and following phones were grouped into broad categories distinguishing sonorants (SON), obstruents (OBS), and word boundaries (#). The preceding context was also coded for voicing: [+/-voice] or irrelevant (for word boundaries). Finally, the number of syllables and position in the word were indicated (monosyllabic, 1st syllable of a dissyllable, 2nd syllable of a dissyllable, etc.).

We used a linear mixed model completed by Tuckey posthoc tests to analyse vowel epenthesis. The presence of a *Vi* was set as the dependant variable. Segmental context, syllabic position, voicing, and place of articulation of /R/ were set as fixed effects, and word and speaker were set as random effects.

3. RESULTS

3.1. Frequency and context of appearance of *Vi*

3.1.1. The effect of place and manner of articulation

Out of the 9304 positions for a potential *Vi*, 2437 *Vi*'s were actually realized. *Vi* occurred in three contexts: #_/R/, OBS_/R/, and marginally, /R/_OBS. *Vi*'s were significantly more frequent word-initially (32%) than in clusters (23%) ($p < 0.001$). This observation isn't congruent with the assumption that *Vi*'s are meant to simplify complex consonant structures [1, 2]; this will be discussed in section 4.

As our corpus exhibited variation between apical and dorsal types of /R/, we looked at the effect of place of articulation on *Vi*'s and noticed that intrusion is significantly more frequent with a following apical /R/ ($p<0.001$). As can be seen in Table 1, 36% of apical realizations of /R/ were preceded by a *Vi*, as opposed to 23% for dorsal ones.

Table 1 : Number of *Vi*'s as a function of the place of articulation of /R/ (NR = unrealized /R/, e.g. *mettre* [mɛt] 'put'; "unrealized *Vi*" = *Vi* that did not occur)

Place of articulation	Nb of realized *Vi*	Nb of unrealized *Vi*	Total
Apical	926	1629	2555
Dorsal	1511	5096	6607
NR	_	142	142
Total	**2437**	**6867**	**9304**

Apical realizations of /R/ are always SON, while dorsal ones alternate between SON and OBS. The fact that apical /R/'s attract more *Vi*'s could indicate that SON /R/'s are more likely to trigger intrusion than OBS /R/'s. We therefore looked at the behavior of dorsal /R/ alone (see Table 2) and observed that significantly more *Vi*'s were produced in #_OBS than in #_SON ($p<0.05$). We also found significantly more intrusions in OBS_SON than in OBS_OBS ($p<0.001$).

Thus, *Vi*'s seem to behave differently in clusters than in word-initial position when /R/ is dorsal. They appear more often with initial obstruent /R/'s, but preferably before a SON in clusters. However, DorsSON realizations of /R/ in clusters seem to appear mainly after a voiced cluster-initial consonant (C1), a context that is known to favor intrusion [2]. We therefore further look at the effect of C1 voicing.

Table 2 : Number of *Vi*'s per place and manner of articulation of /R/.

Place/ manner of articulation	Dorsal /R/	Apical /R/	Total *Vi*	Total nb of tokens
#_OBS	450		450	1297
#_SON	134	173	307	1003
OBS_OBS	502		5052	3343
OBS_SON	419	753	1172	3396
SON_OBS	3		3	265
Grand Total	**873**	**919**	**2437**	**9304**

3.1.2. The effect of voicing

As shown in Table 3, we had a similar number of voiceless and voiced C1, but we observed twice as many *Vi*'s in voiced (32%) than in voiceless (16%) contexts ($p<0.001$). This replicates [2]'s results.

Table 3 : Number of *Vi*'s preceded by a [+/-voice] consonant. 'Irrelevant' applies to word-initial /R/'s (#_R).

C1 voicing	Nb of *Vi*	Nb of unrealized *Vi*	Total
Voiceless	555	2886	3441
Voiced	1125	2355	3480
Irrelevant	757	1626	2383
Total	**2437**	**6867**	**9304**

We also looked at the interaction between C1 voicing and manner of articulation for dorsal /R/ (as mentioned in 3.1.1.). Among the 6607 occurrences involving a dorsal /R/, 4789 were located in clusters (see Table 4). In the OBS[-voice]_SON context, 10% of *Vi* were realized, vs. 37% for OBS[+voice]_SON, ($p<0.001$). Likewise, 7% of *Vi*'s were realized in OBS[-voice]_OBS, vs. 30% for OBS[+voice]_OBS ($p< 0.001$). In OBS[+voice] contexts only, we found 30% of intrusion before DorsOBS /R/ vs. 37% before sonorants ($p<0.001$). This indicates a clear impact of C1 voicing on the realization of a *Vi*, as well as an impact of manner of articulation (at least in voiced contexts), but the former appears more important. Thus, the larger number of *Vi*'s in OBS_DorsSON (Table 4) could be nothing but a collateral effect of C1 voicing in those clusters.

Table 4 : Interaction between voicing and manner of articulation for *Vi* with dorsal /R/.

	-Voi		**-Voi Tot.**	**+Voi**		**+Voi Tot.**	**Tot.**
	No *Vi*	***Vi***		**No *Vi***	***Vi***		
OBS OBS	1769	124	1893	899	381	1280	3173
OBS SON	400	45	445	633	374	1007	1452
SON OBS				161	3	164	164
Tot.	**2169**	**169**	**2338**	**1693**	**758**	**2451**	**4789**

In this corpus, the large span and variability of consonantal contexts does not allow us to further explore the interaction between C1 voicing and the phonetic properties of the following /R/. Thus a more controlled set of data is needed.

3.1.3. Mono vs. polysyllabic words

As can be seen in Table 5, there were 3865 tokens of monosyllabic words involving /R/ word-initially or in a cluster. Among those, 34% were realized with a *Vi*, as opposed to only 20% in polysyllabic words ($p<0.001$).

We further explored the influence of syllable position on the realization of *Vi*. We looked at dissyllabic words only ('2σ' hereafter), as longer words were highly underrepresented. Overall, 23% of 2σ were realized with a *Vi*. As the #_R context is only possible in initial syllable, we only compared OBS_R contexts in σ1 and σ2 of dissyllabic words. Results showed that 26% of potential *Vi*'s were realized in σ1, against only 14% in σ2. Thus, with regard to clusters, *Vi* appeared significantly more often in 1st syllables of 2σ than in 2nd syllables ($p<0.001$), leading to the hypothesis that word-initial syllables could favor intrusion more than final ones. However this assumption needs to be tested on a balanced experimental design including words with more syllables.

Table 5 : Number of *Vi* per position in the word. E.g.: 1_2 : *Vi* was in σ1 of a dissyllabic word.

Syl_Position	Nb of *Vi*	Nb of unrealized *Vi*	Total
1_1	1331	2534	3865
1_2	745	2097	2842
2_2	125	789	914
Total	**2201**	**5420**	**7621**

3.2. Spectral and temporal differences between *Vi* and *Vw*

3.2.1. Spectral measures

One goal of this experiment was to investigate the link between the quality and length of *Vi*'s and those of *Vw*'s. Overall, we had between 38 and 417 tokens of each category among 13 oral vowel types ($\mu=187.46$, $\sigma=119.01$). For the paired comparisons of *Vi* and *Vw* properties, we used a subsample of this corpus that included 560 tokens of *Vi*-*Vw* pairs belonging to /a, e, u, ɔ, ø, o, ə, i, ɑ/.

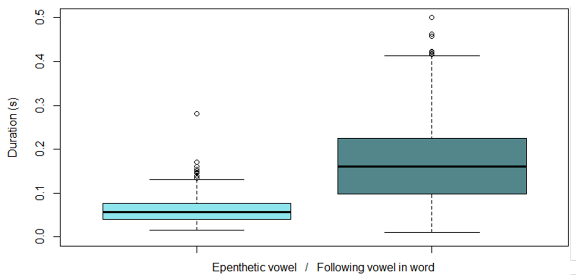
All vowel categories pooled, we observed no significant difference between the mean F1 and mean F2 (Hz) of *Vi*'s and *Vw*'s ($t(558)=-1.14$, $p=0.2$; $t(558)=-2.03$, $p=0.07$ respectively). However, T-tests do not allow us to conclude that there is no difference between the *Vi* and *Vw* groups. An equivalence test (Toster package of R [8]) confirmed, all vowels pooled, the equivalence between the mean F1 and mean F2 values of *Vi*'s and *Vw*'s ($p<0.001$).

In addition, we performed a Linear Discriminant Analysis (LDA). We trained the model on *Vw*'s and then observed how *Vi*'s were classified. When there were misclassifications, /o/ was classified as /ɔ/; /i/ as /e,ə,ø/; /ɑ/ as /a,o,ɔ,ə/; and /ə/ as almost any other category. These confusions can be explained by the qualitative examination of the position of *Vi*'s and *Vw*'s in the F1/F2 space, that shows a slightly more centralized space for *Vi*'s (which is likely due to variation in F2 if we take into account the t-test results for F2). More specifically, misclassified /o/'s might also be related to their diphthongized quality moving from [ɔ] to a mid-high vowel [9]. Misclassified /ɑ/'s always appeared word-finally, where the quality of the vowel is known to be variable [9]; the same explanation may apply to misclassified /ə/'s.

3.2.2. Durations

We first compared *Vi* and *Vw* durations. As expected *Vi*'s are far shorter (μ=0.061s, σ=0.02) than *Vw*'s (μ=0.171s, σ=0.098), as shown in Figure 2 ($t(571)$=-25.814, p<0.001).

Figure 2: Distribution of *Vi*'s (left) and *Vw*'s (right) duration (s).

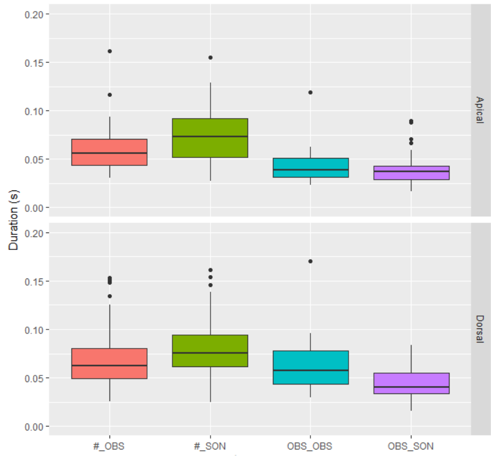


We also explored *Vi* durations across contexts. We computed an analysis of variance (ANOVA) with three factors: the manner of articulation of the preceding and following segments (see Table 2), the voicing of C1, and the place of articulation of /R/ (dorsal vs. apical). We also computed the interactions between these three factors.

Manner of articulation had a significant impact on *Vi* duration ($F(3,564)$=50.38, p<0.001), a posthoc test (Tuckey HSD) then confirmed that this difference was significant between all context categories, except for OBS_OBS and #_OBS. As expected, #_SON induced longer *Vi*'s than contexts where C2 was an obstruent. Surprisingly OBS_SON clusters showed the shortest *Vi*'s, while we could have expected OBS_OBS clusters to induce shorter *Vi*'s.

The interaction between manner and place of articulation of /R/ was also significant ($F(1,550)$= 3.825, p<0.05). As seen in Figure 3, generally speaking, *Vi*'s followed by a dorsal /R/ were a bit longer than those followed by an apical /R/ (all contexts pooled), and word-initial contexts favored longer *Vi*'s. For both apical and dorsal /R/'s, OBS_SON clusters showed the shortest *Vi*'s.

Figure 3: Distribution of *Vi* duration (s) in each context and followed by apical (top) versus dorsal (down) /R/.



4. DISCUSSION

We pursued the work of [2] by adding more contexts involving /R/. In addition to clusters in different syllabic position (see 3.1.3), we looked at word-initial /R/. We considered the place and manner of articulation of /R/ (apical vs. dorsal, SON vs. OBS), as well as C1 voicing and position in the word. Finally we investigated interactions between the multiple features of *Vi*'s contexts and their acoustical cues.

Overall our results showed fewer *Vi*'s than [2]. For instance, [2] observed 89% of *Vi* after OBS[+voice], against only 32% in our study. The source of this gap is unclear. Intrusion was found to be more frequent with apical realizations of /R/, which is consistent with [2]'s comparison between French (whose /R/ is dorsal) and Spanish (whose /R/ is apical). Intrusion is also favored in monosyllables. This is compatible with a word-minimality effect, whereby monosyllabic words tend to be augmented by epenthesis; see [9] for examples in French.

From the acoustical point of view, we observed longer *Vi*'s than [2]. Furthermore, we established that *Vi*'s were longer in #_SON and when /R/ was dorsal. The quality of *Vi* and *Vw* was also compared, showing similar F1 and F2 values, though *Vi*'s were slightly reduced. This similarity suggests a vowel harmony process between the two vowels.

Our results suggest that vowel intrusion does not only serve to simplify clusters, as we observed more *Vi*'s word-initially. It seems that the source of *Vi*'s is to be found in articulatory or perceptual constraints specific to /R/, but it remains to be determined whether intrusion in clusters and word-initially share the same motivation.

5. REFERENCES

[1] Côté, M. H. 2000. *Consonant Cluster Phonotactics: A Perceptual Approach*. PhD thesis, MIT.

[2] Colantoni, L., Steele, J. 2005. Phonetically-driven epenthesis asymmetries in French and Spanish obstruent-liquid clusters. In: Gess, R., Rubin, E. J. (eds), *Theoretical and Experimental Approaches to Romance Linguistics: Selected papers from the 34th Linguistic Symposium on Romance Languages (LSRL), Salt Lake City, March 2004*. Amsterdam: John Benjamins, 77–96.

[3] Picard, M.-P. 2017. *Schwa initial en français laurentien: distribution et nature*. MA thesis, Université Laval.

[4] Côté, M.-H. 2014. Le projet PFC et la géophonologie du français laurentien. In: Durand, J., Kristoffersen, G., Laks, B. (eds), *La phonologie du français: normes, périphéries, modélisation*. Nanterre: Presses Universitaires de Paris Ouest, 175–198.

[5] Durand, J., Laks, B., Lyche, C. 2002. La phonologie du français contemporain: usages, variétés et structure. In: Pusch, C., Raible, W. (eds), *Romanistische Korpuslinguistik – Korpora und gesprochene Sprache / Romance Corpus Linguistics – Corpora and Spoken Language*. Tübingen: Gunter Narr Verlag, 93–106.

[6] Milne, P. 2014. *The Variable Pronunciations of Word-Final Consonant Clusters in a Force Aligned Corpus of Spoken French*. PhD thesis, University of Ottawa.

[7] Boersma, P. 2002. Praat, a system for doing phonetics by computer. *Glot international* 5.

[8] Lakens, D. 2017. Equivalence tests: a practical primer for t Tests, correlations, and meta-analyses. *Social Psychological and Personality Science* 8, 355–362.

[9] Côté, M.-H. 2007. Rhythmic constraints on the distribution of schwa in French. In: Camacho, J., Flores-Ferrán, N., Sánchez, L., Déprez, V., Cabrera M. J. (eds), *Romance Linguistics 2006*. Amsterdam: John Benjamins, 79–92.