Isogloss
Open Journal of Romance Linguistics

Received: 26-02-2024
Accepted: 02-11-2024
Published: 22-03-2025

Branching onsets and empty nuclei: Lessons from the Croissant

Noam Faust
Université Paris 8, CNRS SFL
Noam.faust@univ-paris8.fr

How to cite: Faust, Noam. 2025. Branching onsets and empty nuclei: lessons from the Croissant. In *Romance minority languages: A challenge for linguistic theory*, eds. Michela Russo & Shanti Ulfsbjorninn. Special issue of *Isogloss. Open Journal of Romance Linguistics* 11(3)/7, 1–20.
DOI: https://doi.org/10.5565/rev/isogloss.405

Abstract

In the Croissant Romance variety of Saint-Pierre-le-Bost (SPLB), branching onsets can be split by epenthetic vowels. In addition, epenthetic vowels should be distinguished from weak, non-epenthetic ones, although both are identical in quality – because only the latter can be stressed. The consequences of these facts for autosegmental representations are explored in this paper. The splitting of branching onsets is claimed to argue in favor of a Strict CV view (Lowenstamm 1996, Scheer 2004). An account is proposed within this approach that involves a distinction between two types of featurally empty V-slots. Finally, a parameter is put forth regarding the realization of such V-slots which distinguishes French-type languages from SPLB-type languages.

Keywords: Croissant, branching onset, empty nuclei, French, Government Phonology

> "How many years can some people exist, before they're allowed to be fuh-ree?"
> Bob Dylan at Budokan, 1978

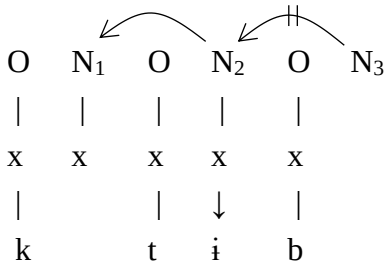
1. Introduction

An axiom of Government Phonology (GP; Kaye et al. 1990) is that segments alternating with zero correspond to lexical, non-alternating skeletal positions. The key concept in this approach is government, a relation between two positions in the representation, whereby the governor is contentful and inhibits the realization of the

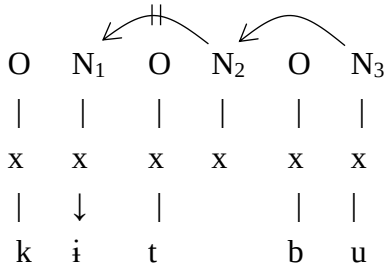
governee. The classic example is Moroccan Arabic [ktɨb] ‘he wrote’ vs. [kɨtbu] ‘they wrote’. In (1), N_3 is empty. Therefore, it does not govern N_2, and the latter is realized and governs N_1. In (1), N_3 is contentful and does govern N_2, which therefore remains unrealized. N_1, in turn, is ungoverned and realized. Both [ɨ]s alternate with zero, but their positions are present in the representation regardless of their realization.[1]

(1) V-Ø alternation in Moroccan Arabic, according to GP

a.



b.



GP and related theories thus distinguish themselves in that positions are not inserted in order for segments to be realized; rather, alternating and epenthetic segments realize existing positions. I will refer to this as the Constant Position Principle (CPP):

(2) The Constant Position Principle
Segments may alternate with zero, but the position they realize is always lexically present.

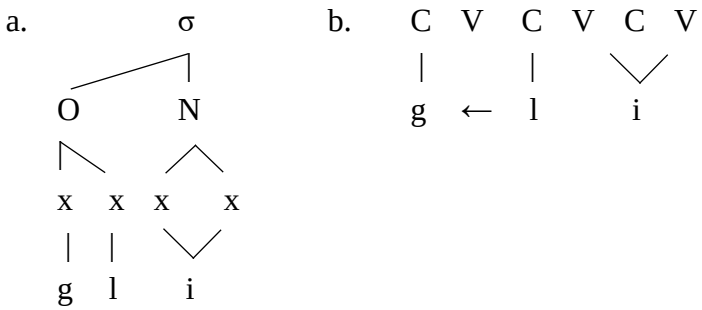
Syllabic theories that are more mainstream than GP do not accept this premise. Instead, they derive epenthesis, syncope, and syllable structure in general from sonority scales, syllabification rules etc., and allow for much more resyllabification than GP and related approaches. For instance, a mainstream account of the facts of Moroccan Arabic will have completely different syllable structures for the stem in [ktɨb] and [kɨtbu]. Possibly as a result of a ban on [ɨ] in open syllables, in [ktɨb], [t] is part of a complex onset, and [b] is a coda; but in [kɨtbu], [t] is a coda and [b] an onset. Such theories will not be considered here.

Two currents can be distinguished within approaches that do adhere to the CPP. In Classic GP, syllable structure is arborescent and hierarchical. Every syllable node branches into an onset and a nucleus, and these constituents, too, may branch. Thus, a representation of English *glee* [gliː] would be as in (3). Strict CV (Lowenstamm 1996, Scheer 2004), an offshoot of GP, attempts to do away with the arborescence. Its basic tenet is that the skeletal tier consists of an iteration of CV units, equivalent to the ON sequences of GP (the difference need not concern us here). In this approach, only segments branch; C and V slots do not. Thus, as shown in (3), branching onsets involve an empty nucleus over which a special relation called “Infrasegmental Government” (signaled by “←”, Scheer 2004) holds, and branching nuclei span over an empty C-slot.[2]

[1] The most common type of government is the inter-nuclear government in (1). For more on the tool of government, see e.g. Kaye et al. 1990 and Scheer 2004.

[2] Branching onsets involve an empty V-slot also in the framework presented in Brandão de Cravalho (2017).

(3) Branching onsets, branching nuclei in GP



Because the two approaches accept the CPP, they make distinct predictions regarding branching onsets. If the Strict CV view is correct, and "branching onsets" involve an empty V-slot, then it is not unimaginable that under certain conditions, such onsets should be severed by an epenthetic vowel. If, in contrast, the Classic GP view is on the right track, then branching onsets should not be severable into CvC; there is simply no position between the Cs to be realized.

In this paper, I bring forth evidence in favor of the Strict CV view from the Croissant variety of Saint-Pierre-le-Bost (SPLB), a village in the north-east of the French department of the Creuse.[3] In addition, I show that the facts from SPLB require a novel representational distinction among two types of empty nuclei: those that involve a root-node and those that do not.

In the next section, I provide the relevant background about the phonology of the variety of SPLB. Section 3 discusses branching onsets and the conditions under which they are severed and their formalization in Strict CV. Section 4 extends the discussion to the representation of weak nuclei. Section 5 concludes, highlighting the difference between French and SPLB. It is claimed that the two in fact represent opposite choices with respect to a parameter regarding the realization of ungoverned empty nuclei.

2. Saint-Pierre-le-Bost

Modern-day geographical France is commonly divided into three main linguistic areas, representing the development of Latin in Gaul (Figure 1): "langue d'Oïl" to the north, "langue d'Oc" or Occitan to the south, and Franco-Provençal in the Alps (Arpitan on the map). Modern French is an Oïl variety. The variety studied in this paper belongs to the contact area of Oïl and Oc, referred to as "the linguistic Croissant" (Ronjat 1913, Guérin 2022), due to its croissant-like shape.

The varieties of the Croissant display characteristics of both language groups. For instance, like Oc languages, many of them distinguish between singular persons in the imperfect and/or present indicative form of the verb, a feature that has been lost to the north of the Croissant. Conversely, Oc languages exhibit final unstressed vowels other than [ə]; but Croissant languages, like Oïl ones, are by and large stress-final (with

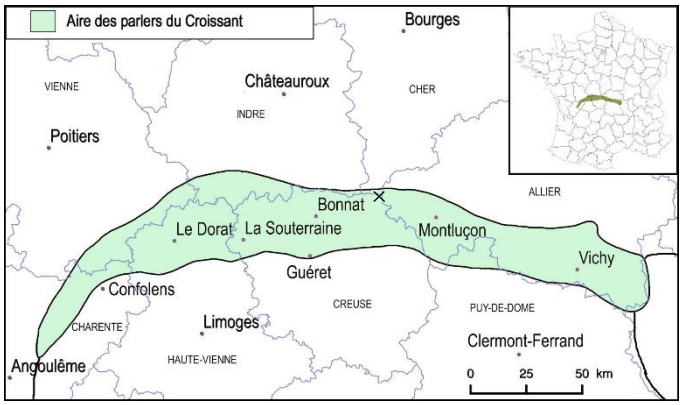
[3] This paper thus joins Scheer (2014), who also argues for the Strict CV view of branching onsets, though not on the basis of epenthesis.

principled exceptions). This feature is the criterion that marks the southern border of the Croissant region.

Figure 1. Linguistic areas of Modern-Day France[4]

The village of Saint-Pierre-le-Bost is situated in the center of the Croissant near the purported border with the langue d'Oïl – the place is marked by X in Figure 2. Accordingly, this variety is more similar to French than to Occitan. Fieldwork has been carried out with one speaker in his late sixties.[5]

Figure 2. Saint-Pierre-le-Bost in the Croissant region[6]



[4] Image taken from https://atlas.limsi.fr/.

[5] A second speaker was also consulted, a woman in her eighties from the close-by village of Nouzerines. Her speech showed the same phenomena.

[6] Image taken from https://parlersducroissant.huma-num.fr/projet.html.

Also helpful was the translation of *Le Petit Prince*, accomplished by the speaker with the help of the linguist Nicolas Quint as part of a large documentation project of the languages of the Croissant (Quint 2021a).[7]

The following paragraphs briefly present the aspects of the phonology of SPLB that are relevant to the present paper. Because of the proximity to French, whose phonological system is well-known, the ways in which SPLB differs from French are highlighted when necessary.

SPLB has the vowel phonemes in (4) below. Some of these are realized differently depending on stress and on the type of syllable they head. The distinction between mid-front /e/ and /ɛ/ is neutralized in unstressed syllables, with only [e] appearing in open ones and only [ɛ] in closed ones. /o/ and /ɔ/ are neutralized in favor of [ɔ] in unstressed closed syllables. /ɔ/ does not undergo an ATR alternation; but in the absolute final position, it is realized [ɒ]. For the sole mid, front rounded vowel /ø/, an ATR alternation holds regardless of stress, with [œ] occurring in closed and [ø] in open syllables. Some, but not all of these distributional properties follow the same generalizations as the *loi de position* of French (e.g. Morin 1986).

As also shown, SPLB distinguishes two low vowels /a/ and /ɑ/. There might be a tendency for the contrast to neutralize into [ɑ] before [ʁ]. Except for that environment, I have not found examples of [ɑ] in closed, unstressed syllables.

(4) SPLB vowel distribution

				Unstressed syllable		Stressed syllable	
				Closed	Open	Closed	Open
Front	High	-Round	/i/	[i]	[i]	[i]	[i]
		+Round	/y/	[y]	[y]	[y]	[y]
	Mid +ATR	-Round	/e/	[ɛ]	[e]	[e]	[e]
		+Round	/ø/	[œ]	[ø]	[œ]	[ø]
	Mid -ATR	+Nasal	/ɛ̃/	[ɛ̃]	[ɛ̃]	[ɛ̃]	[ɛ̃]
		-Nasal	/ɛ/[8]	[ɛ]	[e]	[ɛ]	[ɛ]
Back	High		/u/	[u]	[u]	[u]	[u]
	Mid +Atr	+Nasal	/õ/	[õ]	[õ]	[õ]	[õ]
		-Nasal	/o/	[ɔ]	[o]	[o]	[o]
	Mid -Atr		/ɔ/[9]	[ɔ]	[ɔ]	[ɔ]	[ɒ]
	Low		/ɑ/	[ɑ]?	[ɑ]	[ɑ]	[ɑ]
Central	Mid		/ə/	[œ]	[ø]	[œ]	-
	Low	+Nasal	/ã/	[ã]	[ã]	[ã]	[ã]
		-Nasal	/a/	[a]	[a]	[a]	[a]

[7] https://parlersducroissant.huma-num.fr/index.html. This site features a great wealth of data, including the full conjugation of several verbs in several varieties, among which the variety discussed here.

[8] The contrast between /e/ and /ɛ/ is neutralized in all unstressed syllables before [ʁ]: only [ɛ] is allowed in this environment. It is not neutralized in stressed syllables: [vɛʁ] 'green', [ve-ʁ] 'see-INF'.

[9] The vowel [ɒ] is also found before [ʁ], whether it is an onset or a coda. I assume it is a realization of /ɔ/ there, too.

The distribution of vowel allophones will be important for the discussion of branching onsets.

The vowel designated as /ə/ will occupy a central place in this paper. It is considered to be a "weak" vowel, because it systematically undergoes syncope when it is expected to appear in open syllables. To illustrate, the unsuffixed forms in (5a,c) involve a stressed [œ] in a closed syllable. It is absent before V-initial suffixes, i.e., where it would appear in an open syllable. Note that this vowel cannot be regarded as an epenthetic, non-lexical vowel because, as the unsuffixed forms in (5b,d) illustrate, the phonotactics of the language do not rule out words ending in the relevant clusters. Because of this weakness – and following the tradition in studies of French – this vowel is construed phonemically as /ə/, even though it is never realized as [ə].

(5) Weak stressed vowels in Saint-Pierre-le-Bost

	IND.SG	INF			IND.SG	INF	
a.	kaˈkœt	kakˈt-a	'cackle'	b.	ʁɛsˈpɛkt	ʁɛspɛkˈt-a	'respect'
c.	kaˈʁœl	kaʁˈl-a	'tile'	d.	ˈyʁl	yʁˈl-a	'yell'

In addition, /ə/ has to be distinguished from /ø/, a vowel whose realizations are identical except that it does not undergo syncope in open syllables. For instance, the vowels of [ʒœt] 'throw.IND.SG' and [gœl] 'scream.IND.SG' are identical; but the suffixed forms are [ʒt-a] 'throw-INF' and [gœl-a] 'scream-INF'. The former includes /ə/, the latter /ø/.

To be sure, the weak [œ] occurs in both stressed and unstressed closed syllable, as illustrated by the future forms of the verbs in (5), [ka.k**œ**t.ʁõ] 'they will cackle', [ka.ʁ**œ**l.ʁe] 'I will tile'. /ə/ is realized in open, unstressed syllables only under very special conditions, which I will return to below.

Finally, a comment is due about initial clusters. In SPLB, one finds initial clusters of virtually any sonority slope: [ʒnu] 'knee', [bzwɛ̃] 'need', [ʁʒist] 'log'. This fact is possibly related to the ban on the realization of /ə/ in open syllables; in French, all of these items can be pronounced with [ø] between the two consonants. Such pronunciation are regarded as "foreign" for the speaker I worked with. When words such as French *remorque* [ʁømɔʁk] are pronounced in SPLB, the first vowel is dropped to yield [ʁmɒʁk] 'trailer'.[10]

These basic facts suffice in order to understand the arguments in the next sections.

[10] While French also has a vowel analyzed as /ə/ and realized as [ø/œ], its distribution is different from SPLB /ə/. It never occurs in closed syllables, stressed or not, with the very specific exceptions of (i) the seemingly epenthetic, but stressed vowel [œ] in monosyllabic "verlan" words like [kœf] 'cop' and [ʃœm] 'ugly', from *flic* [flik] and *moche* [mɔʃ] respectively, and (ii) adaptations of words with impermissibly final clusters, e.g., the supermarket chain *LIDL* [liˈdœl]. Still, such vowels do not behave like weak [œ] in that they do not alternate with Ø. In some open unstressed syllables, French weak [ø] can, but does not have to, syncopate – *cheval* 'horse' can be pronounced [ʃøval] or [ʃval].

3. Branching Onsets in SPLB

This section is arranged in the following manner. First, it is shown that SPLB distinguishes TR clusters (where T=obstruent R=liquid) from non-TR clusters, and that the former pattern with singleton onsets. TR clusters must therefore be regarded as branching onsets, regardless of the way these are formalized. Then, evidence is provided to the effect that such branching onsets can be separated by an epenthetic vowel – a fact that is incompatible with the Classic GP view of branching onsets, but *is* compatible with the Strict CV one. A third subsection puts forth a short representational formalization in the latter theory.

3.1. SPLB has branching onsets

TR clusters in SPLB differ from other clusters (TT, RT, RR) in a way that leads to their analysis as branching onsets, i.e. on a par with singleton onsets. At least three facts about SPLB support that claim.

First, the distribution of vowels before TR clusters is the same as before simple onsets. As shown in (4) above, [e] and [o] are not found in unstressed closed syllables, i.e. before RT, RR and TT. But before TR clusters one does find these vowels, e.g. [**o**vʁi-ʁ] ‘open-INF’, [n**o**fʁaʒa] ‘castaway (n.)’, [ʁ**e**fleʃi-ʁ] ‘contemplate-INF’, [**e**klɛʁ-a] ‘illuminate-INF’, [**e**kla] ‘burst (n.)’. In other words, the T of TR clusters is not a coda; the entire cluster belongs to the following syllable.

The second and third pieces of evidence have to do with the behavior of word-initial TR vs RT, RR and TT. As explained above, SPLB allows for all three configurations word-initially: [bʁy] ‘noise’, [pti] ‘small’, [ʁʒist] ‘log’. However, the interaction of such word-initial clusters with their left context is not uniform. A case in point is the MSG definite article. Before a CV-initial word, it is generally realized as [l] regardless of that context (6). The same is true of TR-initial words (6b). In contrast, word-initial TT and RT clusters force a realization [lœ] of the definite article (6c,d). The same distribution holds for other [C(œ)] clitics, such as the possessive [d(œ)].

(6) Interaction of the left edge of a noun/adjective with preceding DEF

	BARE	#DEF+_	
a.	mud	lmud	‘world’
b.	bʁy	lbʁy	‘noise’
c.	pti	lœpti	‘small (one)’
d.	ʁʒist	lœʁʒist	‘log’

Again, TRV sequences pattern with CV sequences, arguing for TRs being branching onsets in this language.

A similar generalization arises from the behavior of verbal stems after the prefix whose realization is [ʁ] before vowels and [aʁ] before consonants. Before CV-initial verbs, [aʁ] is simply added to the verbal form (7). This is also the case before TR-initial verbs (7). But if the verb begins with a non-TR cluster, that cluster is broken

after the prefix [aʁ] (7). This alternation joins the two preceding facts in arguing for a parallel between TR clusters and singleton onsets, as opposed to RT, RR and TT.[11]

(7) Interaction of the left edge of a verb with preceding prefix

	VERB-INF		[aʁ]-VERB-INF	
a.	mõt-a	'go up'	aʁ-mõt-a	'go back up'
	sãbl-a	'seem'	aʁ-sãbl-a	'resemble'
	tuʁn-a	'turn'	aʁ-tuʁn-a	'return'
	kunɛt	'know'	aʁ-kunɛt	'recognize'
b.	pʁesãt-a	'present'	aʁ-pʁesãt-a	'represent'
	tʁuv-a	'find'	aʁ-tʁuv-a	'find again'
	pliʒ-a	'fold'	aʁ-pliʒ-a	'fold back'
	pʁãd	'take'	aʁ-pʁãd	'take up again'
c.	tn-iʁ	'hold'	aʁ-tøn-iʁ	'hold up'
	vn-iʁ	'come'	aʁ-vøn-iʁ	'come back'
	lv-a	'rise, raise'	aʁ-løv-a	'note'
	ʃt-a	'throw'	aʁ-ʒøt-a	'reject'

It must be concluded that TR clusters in SPLB are what is usually referred to as branching onsets. Unlike other CC clusters, both consonants of TR sequences are syllabified with the following vowel, like simplex onsets.

However, under certain conditions, TR clusters, too, can be broken by epenthetic vowels.

3.2. Branching Onsets can be broken

This subsection uses SPLB data to show that, assuming the CPP, TR sequences must involve an empty V-slot between T and R.

The clearest data point concerns a certain sub-group of the so-called "1st group" of verbs – the largest, most productive verbal group, used for instance for most denominal verbs. As shown in (8), the 1st group is characterized by an infinitive ending in [-a]. In denominative verbs, if the base noun ends in a consonant, the IND.SG is identical to it (8). But when the base noun ends in the high vowel [i], as in (8), one finds that the parallel verbal forms all include only the corresponding glide [j]; and the IND.SG exhibits a suffix [e] (in bold). Below it will be shown that the same is true for [u]-final bases.

(8) C-final and V-final verbs in the 1st group and the IND.SG suffix [e]

	a. 'butter'	b. 'spend'	c. 'bet'	d. 'copy'
INF	bøʁ-a	dépãs-a	paʁj-a	kopj-a
IND.SG	bœʁ	dépãs	paʁj-**e**	kopj-**e**
IND-1/3-PL	bøʁ-ã	dépãs-ã	paʁj-ã	kopj-ã
NOUN	bœʁ	dépãs 'expense'	paʁi	kopi

[11] The unsuffixed IND.SG form of the verbs in (7) also contains a realized vowel: [tœn, vœn, lœv, ʒœt]. The first two have an alternate form [tɛ̃,vɛ̃].

Crucially, *there are no vowel-final stems in the 1st group*.[12] The avoidance of such stems might be one reason for the appearance of the suffix [e]: it allows for the gliding of the high vowel of the base, since post-consonantal glides are cross-linguistically marked at the word edge, *[bœʁj]. Further discussion on the nature of this suffix can be found in Faust & Scheer (to appear), as well as in Quint (2021b).

The gliding that occurs in stems with final high vowels interacts with preceding branching onsets. As shown in (9), when the base is of the form TRV (with V being either [i] *or* [u]), the gliding of the high vowel in all verbal forms is accompanied by the breaking of the preceding TR cluster.

(9) TRV-final verbs in the 1st group

	a. 'select'	b. 'shout'	c. 'nail'	d. 'hole'
INF	tœʁj-a	kœʁj-a	kœlw-a	tœʁw-a
IND.SG	tœʁj-**e**	kœʁj-**e**	kœlw**e**	tœʁw-**e**
IND-1/3-PL	tœʁj-ã	kœʁj-ã	kœlw-ã	tœʁw-ã
N	tʁi	kʁi	klu	tʁu

It was shown above that initial TR clusters are branching onsets. In (9), such clusters are broken by epenthetic vowels. Regardless of the motivation for epenthesis in this case – to which I return presently – *any theory that adheres to the CPP must therefore admit that branching onsets involve an internal empty V-slot.* Classic GP is thus wrong on this issue, as it represents branching onsets with no internal V-slot, and therefore predicts that SPLB is not a possible language (again, assuming the CPP).

As for the reason for epenthesis, consider the following. If a base ending in a high vowel like [si] 'saw (n.)' is followed by a vowel-initial suffix, a hiatus is expected /si-a/ 'saw-INF' → [sia]. Such hiatuses are often resolved, in Romance and in other lanaguages, by the gliding of the first vowel: /si-a/ → [sja]. The process is traditionally referred to as "syneresis". Côté (2018) discusses syneresis in French, and describes it as optional: sequences like *sci-a* 'saw-PST.3SG' and *lou-a* 'rent-PST.3SG' can be pronounced either with an intervening glide and no syneresis [sija, luwa] or with syneresis [sja, lwa].[13] In SPLB, however, this process is obligatory: /si-a/ 'saw-INF' can only be pronounced [sja], not *[sija]. To understand the facts in (9), it is useful again to compare them to their French cognates. Syneresis is not possible after branching onsets in French, for reasons that will not concern us here. Possibly because the process is optional in that language, it is simply blocked: *trou-a* 'puncture-PST.3SG' can only be pronounced [tʁuwa], with the vowel remaining intact (and concomitant glide-formation to avoid hiatus). Back to SPLB, one may assume that in this language, too, syneresis cannot apply after branching onsets. But because it is obligatory, it occurs at the expense of the cohesion of that branching onset: /tʁu-a/ is realized [tœʁwa] and /tʁi-a/ is realized [tœʁja].

[12] Stems here are informally defined as the form of the verb without the suffix. To illustrate what the ruled out form would be, consider that French does have such stems, e.g. [saly] 'greet' or [kʁe] 'create', infinitives [saly-e, kʁe-e], or indeed [kopi] 'copy.IND.SG' and [klu] 'nail.IND.SG'.

[13] Tobias Scheer (p.c.) points out that the optionality of syneresis in French is limited to monosyllabic bases. Syneresis is obligatory in polysyllabic bases ending in any consonant except [ʁ]. For instance, /kopi-e/ 'copy-INF' can only be realized as [kopje], not *[kopije].

Branching onsets are also broken *stem-finally*. As shown in (10), some T-final stems in SPLB exhibit an additional liquid and become TR-final when followed by a V-initial suffix (or common collocation, as in (10)). In this case, the TR cluster is demonstrably a branching onset, as shown by the vocalic alternation in (10). It can therefore be concluded that these stems end in a branching onset underlyingly, and that the branching onset is simplified word-finally, e.g. /sabl/ → [sab] 'sand'. When a C-initial suffix is added, the branching onset is not simplified; instead, it is again broken by an epenthetic vowel.

(10) Word final /TR/

	N#		-V		_-C	
a.	kat	'four'	katʁ-uʁ	'four o'clock'	katœʁ-jɛm	'fourth'
b.	sab	'sand'	sabl-a	'sandy (e.g. taste)'	sabœl-jɛʁ	'resemble'
c.	ʁɛg	'rule'	ʁ**e**gl-a	'arranged'	ʁ**e**gœl-mã	'regulation'
d.	mɛg	'thin'	m**e**gʁ-iʁ	'grow.thin-INF'	m**e**gœʁ-jo	'thin person'
e.	ɔf	'offer'	**o**fʁ-iʁ	'propose-INF'	-	

Once again, an epenthetic vowel can break a TR cluster, which is demonstrably realized as a branching onset when possible.

A final observation is that many SPLB words with a sequence [TœR] correspond etymologically to TRV sequences. This is shown in (11) through a comparison to French, which preserved these as [TRø] sequences, the vowel being interpretable as /ə/:

(11) SPLB TœR, French TRø

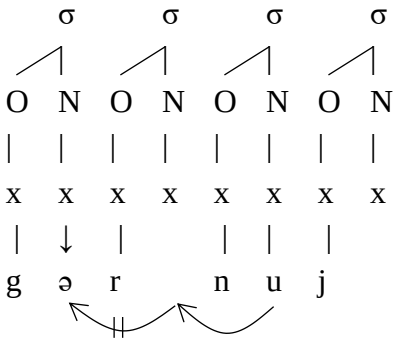
NEC TœR	French TRø	gloss
gœʁnuj	**gʁø**nuj	'frog'
pœʁmje	**pʁø**mje	'first'
vã**dœʁ**di	vã**dʁø**di	'Friday'
ʃø**vœʁ**fœj	ʃɛ**vʁø**fœj	'honeysuckle'
bœʁtœl	**bʁø**tɛl	'suspenders'
ã**gœl**tɛʁ	ã**gløl**tɛʁ	'England'

Non-etymological [TœR] sequences occur here and there with other diachronic sources, e.g. [fœʁnɛt] 'window' or [tœʁʒu] 'always' for French *fenêtre* [fønɛtʁ] and *toujours* [tuʒuʁ]. But in those, the displacement of the rhotic is not systematic. The facts in (11) are systematic: the list is very long, and there are in fact very few [TRœ/ø] sequences in the language. Those that do exist generally involve etymological /ø/, e.g. [kʁøz-a] 'dig-INF'. Since SPLB also exhibits /ø/, such items are easily explicable by assuming underlying /ø/, as opposed to /ə/; the ban on /TRə/ is plausibly synchronic.

Charette (2017) reports on an identical phenomenon in Acadian French (AF). TR clusters in AF never precede /ə/ (she abstracts away from the true surface realization of /ə/, transcribing it as [ə]). Working within classic GP, and assuming that [ə] is the realization of an empty nucleus, Charette proposes that AF does not allow for branching onsets. Instead, TR clusters in AF are like TT clusters: they involve an

intervening V slot which, ungoverned, must be realized. Recall that in classic GP, branching onsets do not involve such a position.

(12) TR clusters in Acadian (Charette 2017)



Since SPLB exhibits the same pattern and *does* allow for branching onsets – as was shown repeatedly – Charette's analysis is probably incorrect. Instead, one may assume that the vowel /ə/ (or an empty nucleus) may not license branching onsets.[14] Underlying /TRə/ (or /TR_/, where _ is an empty nucleus) are realized [TœR].

To summarize, sequences that must be analysed as "branching onsets" can be severed through epenthesis. Among approaches that adhere to the CPP, this fact supports Strict CV, and falsifies the Classic GP approach.

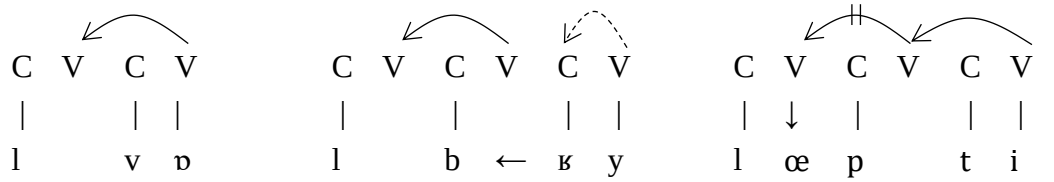
3.3 Representations

To make the proposal explicit, I will now provide representations of several of the data points presented above. Let us begin with a reminder of the standard Strict CV view of the status of the nucleus before the TR cluster: the liquid governs the preceding consonant in a relation called "Infrasegmental Government", which is signaled by the arrow "←". In order for infrasegmental government to hold, the R must be "licensed" by a following contentful nucleus (dashed arrow in 0 below). In addition, Brun-Trigaud & Scheer (2010) argue that the nucleus inside the TR cluster, like a full nucleus, is able to govern. Assume now that the definite article is simply /l/. The empty nucleus after the /l/ is governed by the following contentful nucleus in 0, and by the TR-internal nucleus in 0. In contrast, before the TT cluster in 0, the same nucleus is ungoverned; it therefore must be realized.

[14] That /ə/ is special in its licensing and government properties is well-known. For instance, in Southern French, [ɔ,ɛ] cannot occur in open syllables, except if the next syllable is [Cø] (</Cə/): *autrement* [ɔ.tʁø.mã] 'otherwise' (Scheer 2015). For more on licensing scales see Cyran (2003, 2010), Cavirani & van Oostendorp (2017) and Ulfsbjorninn (2017).

(13) Epenthesis after utterance-initial /l/ before CV, TR, TT

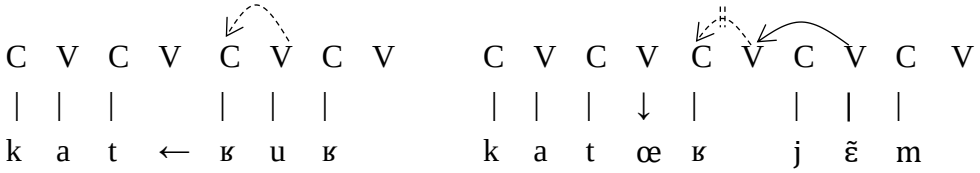
a. [lvɒ] ‘the calf’ a. [lbʁy] ‘the noise’ c. [lœpti] ‘the small one’



Recall that word-finally, TR clusters are not permitted in SPLB. Instead the R is deleted: /katʁ/ => [kat] ‘four’. Faust & Scheer (to appear) discuss the full reasons for this in SPLB, including the impossibility of an epenthetic repair. For the present purpose, it is crucial that the liquid reemerges in derivation. If the next V-slot is occupied (14), the stem-final R is licensed and a branching onset can be established in the stem (as conveyed by the Infrasegmental Government arrow); but if the next nucleus is empty (14), a branching onset cannot be formed, as the R is not licensed. The nucleus inside the TR sequence is ungoverned, and so it must be realized.

(14) Realization and epenthesis in stem-final TR upon suffixation

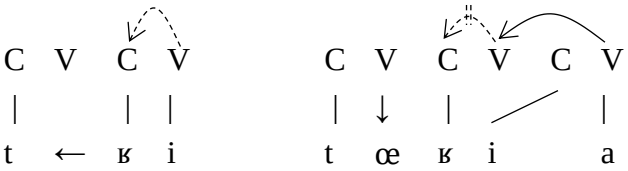
a. [katʁuʁ] ‘four o’clock’ b. [katœʁjɛm] ‘fourth’



Similarly, when a TRV-final noun, such as [tʁi] ‘sorting’ in (15), is verbalized, its final high vowel glides. Because, as explained in the previous subsection, [ijV] is obligatorily avoided through syneresis, this leaves an empty nucleus which requires government. That empty nucleus cannot license the R of the TR sequence, and so the previous V-slot must be realized:

(15) TRV final noun upon verbalization

a. [tʁi] ‘sorting (n.)’ b. [tœʁja] ‘to sort’



(14) and (15) suffice in order to show the conditions under which a V-slot inside a branching onsets is realized. For the case of /TRə/ → [TœR], see Faust & Scheer (to appear).

Before I conclude, in the next section I turn to the formalization of lexical vs. epenthetic weak vowels – another related topic on which the SPLB data sheds new light.

4. On the representation of weak [œ]

In (5) above, I showed that some stressed [œ] vowels are both weak and lexical, i.e., they correspond to non-epenthetic /ə/. The case presented was that of [kakœt] 'cackle.IND.SG', INF [kakt-a]. Since [kt] is a legitimate final cluster – e.g., [ʁɛspɛkt] 'respect.IND.SG' – the stressed [œ] in [kakœt] cannot be epenthetic. Rather, the lexical representation is /kakət/. In contrast, other weak [œ] vowels were designated above as epenthetic, e.g. in /katʁ-jɛm/ => [katœʁjɛm] 'fourth'. This designation is due to the form [kat] that the base takes in isolation: if it were /katər/, we would expect the vowel to be stressed, as in /kakət/=>[kakœt].

This raises the question of how to *represent* the difference between the weak lexical [œ] (underlyingly /ə/) and the epenthetic [œ] (underlyingly an empty nucleus). This question can be broken down into two separate questions. First, why are the two configurations realized identically? And second, what exactly makes /ə/ stronger than an empty nucleus, but weaker than a regular nucleus? The representational difference between these two and stable /ø/ – also realized [œ] in closed syllables – must also be addressed.

I take an Element Theory approach to the representation of vowels (Kaye et al. 1985, Backley 2011). According to this approach, there are three basic elements in vocalic systems – |I|, |U|, |A|. Nasality is marked by an additional element |N|. Vowels other than [i,u,a] are expressed through the combination of two or more elements. Two vocalic expressions can be further distinguished using a head/dependent relation.

With these basic assumptions, the table in (16) presents the vocalic system of SPLB. Heads are underlined. I assume that an important distinction in the system is whether an expression is headed or not. /e,o,ɑ/ are headed, and /ɛ,ɔ,a/ are unheaded. As for /ø/, since its realization as [ø] or [œ] depends on syllable structure, I assume that headedness is attributed to |I| depending on that parameter. Finally, since /ə/ is also realized [ø,œ] depending on the syllable type, I assume it involves the same elements as /ø/. The difference between the two will be explained presently.

(16) SPLB vowel expressed by elements

/i/	I	/ɛ/	IA	/u/	U	/ɑ/	A
/y/	IU	/ɛ̃/	IAN	/o/	UA	/a/	A
/e/	IA	/ø/	IUA	/õ/	UAN	/ã/	AN
		/ə/	IUA	/ɔ/	UA		

At first sight, the analysis of /ə/ in (16) is counter-intuitive, because it depicts this weak vowel as one of the two most complex vowels in the system, i.e. one of the two expressions to employ all three vocalic elements. How could such a complex expression be weak?

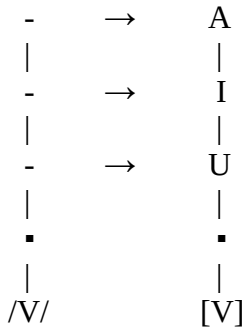
According to Kaye et al. (1985), each element resides on a different tier. In languages without front rounded vowels, the |I|-tier and |U|-tier are fused, such that the two cannot combine. But in languages like SPLB and French, those tiers are not fused. The vowel qualities [œ,ø] are therefore the only ones where all of the three tiers used in the language are occupied.

(17) The representation of [œ,ø]

A
|
I
|
U

In order to explain its weakness, assume now that /ə/ is an empty vowel. It consists only of a root node ▪.[15] All three tiers are lexically empty. Then, a general rule of feature-filling (e.g. Dresher 2009) simultaneously fills all three tiers with the appropriate element, as in (18). In this sense, /ə/ is not complex phonologically; its realization is simply the filling of all three tiers by default.

(18) Feature filling on all three tiers derives [œ,ø]



The same logic can be applied to five-vowel systems with epenthetic/weak [e], such as Modern Hebrew (Faust 2024). In such systems, as mentioned, the I and U tiers are fused, so feature filling inserts only I on the fused tier and A on the other.[16]

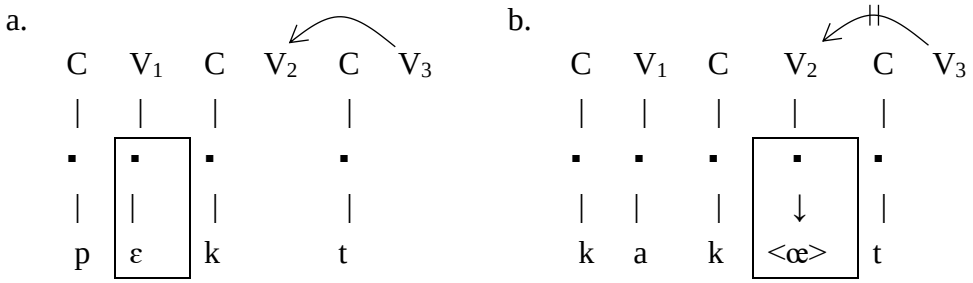
The difference between a lexical /ə/ and an epenthetic [œ] can now be regarded as the presence or absence of the root node ▪: truly empty V-slots carry no root node. To illustrate the difference, the representations of [(ʁɛs)pɛ**kt**] and [ka**kœt**] are provided in (19). Two principles are important for the analysis. First, it is standard in Strict CV to assume that in languages with final clusters, like SPLB, the final empty nucleus can govern. Second, it has been proposed (e.g. Harris 1990) that a governed position may not host an expression that is more complex than the expression in the governing position. Now, the final empty nucleus in both examples in (19) is by assumption entirely empty – it does not carry a root-node. The preceding nucleus is entirely empty in (19), and so the final empty nucleus V_3 may govern V_2 – both are equally "empty". But in (19), V_2 carries a root-node ▪, and is therefore more complex than V_3, thereby blocking the government from V_3. Feature filling, signalled by <œ>, applies to such

[15] Root nodes do not feature prominently in Strict CV, but they are not incompatible with it. They have recently been argued to be necessary Scheer (2022), and have been used in several specific analyses, such as in Ulfsbjorninn (2021) on various phenomena in Ik andCavirani (2022a) on "mobile diphthongs" in Italian.

[16] It seems that |I| almost universally takes precedence over |U| in filling the fused I/U tier: very few systems, if any, employ [o] as weak (though see interesting discussion of a possible counter example in Timugon Murut, in Mackenzie & Ulfsbjorninn, 2023).

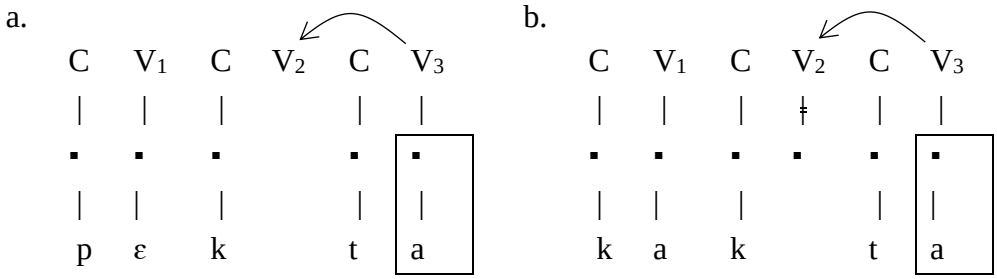
nuclei upon realization – that is, in the translation of phonology to phonetics. Being the last vowel and uninhibited, it is stressed (stressed vowels are framed).[17]

(19) Empty nucleus vs. /ə/ before the final empty nucleus



When, in the suffixed forms, the following nucleus is full, the empty root node is not more complex than its potential governor. Government applies to it and inhibits its realization. Because this inhibition can take place, no feature-filling is required.[18]

(20) Empty nucleus vs. /ə/ before full vowel



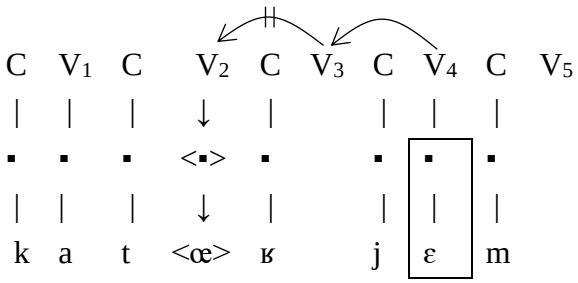
The view of the realization of /ə/ as feature-filling immediately distinguishes it from /ø/. In the latter, the features are present lexically, so the realization of the vowel cannot be blocked. To illustrate, the stem [gœl] 'scream.IND.SG' hosts /ø/, and when the infinitival suffix is added that vowel is not syncopated [gœl-a] 'scream-INF', even though [gla] is not an impossible form in SPLB.

Back to branching onsets, the empty nucleus inside the branching onset of the stem in /katʁ-jɛm/ 'fourth' (V_2 in (21)) is ungoverned, and must therefore be realized. This is achieved through the insertion of a root node (since any segment must have one), as is signalled by the downwards arrow leading from V_2 to <▪>. Once the root node has been inserted, it is unsurprising that feature filling will proceed as for the lexical /ə/.

[17] Cavirani (2022b) uses the concept of "turbidity" in order to define weak, but not empty positions. In this approach, in order for a segment to be realized, two relations must hold between it and its position: association ↑ and pronunciation ↓. A weak vowel is associated to its position through the former but not through the latter. Turbid representations to not obviate root nodes, which are the main analytic tool necessary in the present analysis, so I prefer not to use turbidity. The arrows in the representations in this paper do not imply turbid relations.

[18] I thank Edoardo Cavirani (p.c.) for pointing me to the notion of complexity in government in relation to these data.

(21) Realization of empty nucleus inside a branching onsets



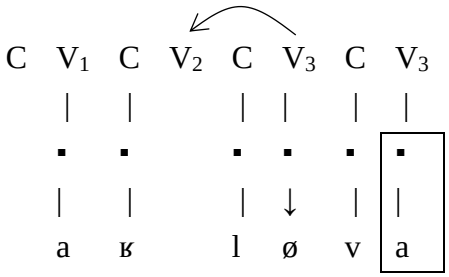
Finally, the distinction between a nucleus associated to an empty root-node and a completely empty nucleus is useful in understanding the effect of the prefix [aʁ]. As repeated in brief in (22), before an initial non-TR cluster, the addition of this prefix causes the breaking of that cluster.

(22) Interaction of the left edge of a verb with preceding prefix (repeated)

	VERB-INF		[aʁ]-VERB-INF	
a.	mõt-a	‘go up’	aʁ-mõt-a	‘go back up’
b.	pʁesãt-a	‘present’	aʁ-pʁesãt-a	‘represent’
c.	tn-iʁ	‘hold’	aʁ-tøn-iʁ	‘hold up’
	vn-iʁ	‘come’	aʁ-vøn-iʁ	‘come back’
	lv-a	‘rise, raise’	aʁ-løv-a	‘note’
	ʃt-a	‘throw’	aʁ-ʒøt-a	‘reject’

As noted in footnote 11 above, the verb stems in (22c) have IND.SG forms with a stressed [œ]: [(aʁ)tœn], [(aʁ)vœn], [(aʁ)lœv], [(aʁ)ʒœt]. This implies that the first nucleus of such verbal stems carries a stressable /ə/, i.e., an empty root-node. As shown in (23), although this nucleus is potentially governed, it is still realized in order to govern the preceding empty, even less complex nucleus. The other possibility, *[aʁœlva], would have violated the complexity condition.

(23) Verb-initial nuclei with an empty root node after [aʁ] prefix



To summarize, the SPLB facts oblige one to distinguish between lexical /ə/ and an empty nucleus. Here, I proposed to distinguish between the two using the notion of a root node, which may be either absent or simply empty.

5. Conclusion, and a comparison to French

In the preceding pages, I have made two claims about autosegmental representations. First, if the CPP is to be adhered to, then branching onsets should be represented with an internal V-slot – as is the standard in Strict CV, but not in Classic Government Phonology. Only in this manner can one account for the severable branching onsets of the Croissant variety of Saint-Pierre-le-Bost. Second, in the representation of weak vowels, a distinction is necessary between fully empty nuclei and empty nuclei that are associated to an empty root node.

I would like to conclude with a short discussion of the differences between SPLB and Standard French. These are summarized in (24). First, French does not sever TR clusters: /TRija/ is realized [TRija] in French, but [TœRja] in SPLB. Second, French does not allow stressed weak [œ] such as SPLB [aʃœt] 'buy.IND.SG'. Third, and relatedly, French never exhibits weak [œ] in closed syllables, whereas that is where one finds this vowel in SPLB. And fourth, SPLB avoids weak [ø] in open syllables (unless the nucleus has a root node and follows an empty V-slot, as discussed in the previous section).

(24) Differences between French and SPLB

	Severable TR	Weak stressed [œ] in closed syllable	Weak unstressed [œ/ø]	
			Closed syllable	(medial) Open syllable
French	–	–	–	+
SPLB	+	+	+	–

All three cases in which SPLB allows weak [œ] have the following characteristic in common: they involve nuclei without associated features in ungoverned positions. This is the case for the severed TR in [TœR_ja]; for the stressed, empty root node in [kakœt] (recall that stressed, empty root nodes are not governed); and naturally for empty nuclei in closed syllables, which always precede other empty nuclei.

Now, it has been proposed that such nuclei are filled through the late, phonetic operation of feature filling. One may therefore set up the following parameter:

(25) The feature-filling parameter of empty nuclei
Can the output of phonology include ungoverned, featurally-empty nuclei?
French: No
SPLB: Yes

The nucleus inside the branching onset /t_ʁia/ is ungoverned if syneresis applies to yield */t_ʁ_ja/. Since the phonology of French cannot leave ungoverned empty nuclei to phonetics, it avoids syneresis in such cases, and the branching onset is not severed. In contrast, since SPLB *may* leave such nuclei to the phonetic interpretation, it does, and the branching nucleus is severed, /t_ʁia/ → /t_ʁ_ja/ → [tœʁja]. Similarly, since SPLB can leave the stressed, empty root-node to the phonetics and the process of feature-filling, it may have a stressed weak [œ]. But French cannot allow for that: a weak stressed vowel in a final closed syllable must be filled in the phonology, i.e., not through feature filling. In other words, it must host a lexical vowel. Presumably, this

is why the form in French is [kakɛt] ‘cackle.IND.SG’, not *[kakœt], but when the same base is suffixed it is still the syncopatable [kak(ø)t-e] ‘cackle-INF’. In this manner, the feature-filling parameter in (25) relates all of the properties in (24) together.

SPLB and French thus illustrate two ways of treating the problematic configuration of ungoverned empty nuclei: either fill them in the phonology or leave them to phonetic feature-filling. But it seems that the proposed parameter setting arguably concerns other languages, too. Biblical Hebrew, for instance, behaves like French, in disallowing [ə] (or the orthographic symbol associated to it in the Hebrew script – the exact pronunciation of Biblical Hebrew is not known for certain) in closed syllables (e.g. Bat-El 2023). Instead, where one expects [ə] in closed syllables, one finds a vowel of a lexical quality: [i] pretonically and [ɛ] post-tonically. Germanic language like Yiddish or English, in turn, have no problem with [ə] in ungoverned nuclei, e.g. Yiddish [χávər-tə] ‘friend-F’. Like the lexical /ə/ of SPLB, these nuclei would be associated with an empty root node in the phonology, and their realization would be a matter of phonetic feature filling.

Acknowledgments

I thank the main speaker I worked with, Mr Gérard Dubac, as well as Madeleine Contarin of the village of Nouzerines. I also thank the following people for having discussed the issues in this paper with me (the order is alphabetical): Joaquim Brandão de Carvalho, Eugeniusz Cyran, Nicolas Quint, Tobias Scheer, Shanti Ulfsbjorninn, Fabian Zuk. I also thank Edoardo Cavirani for reading a draft and commenting on it, as well as two anonymous reviewers.

References

Backley, Phillip 2011. An Introduction to Element Theory. Edinburgh: Edinburgh University Press.

Bat-El Foux, Outi. 2023. Biblical Hebrew phonology from a theoretical perspective. *Radical: A Journal of Phonology* 5: 1-45.

Brun-Trigaud, Guylaine & Tobias Scheer 2010. Lenition in branching onsets in French and in ALF dialects. In P. Karlík (ed), *Development of language through the lens of formal linguistics*, 15-28. Munich: Lincom.

Cavirani, Edoardo. 2022a. Dittongo mobile and other diphthongs: no allomorphy, just phonology. *Rivista di Grammatica Generativa* 44: 1-22.

Cavirani, Edoardo. 2022b. Silent lateral actors: the role of unpronounced nuclei in morpho-phonological analyses. *The Linguistic Review* 39 (4): 615-653.

Cavirani, Edoardo & Marc van Oostendorp. 2017. On silent markedness. In Samuels Bridget. (ed.), *Beyond Markedness in Formal Phonology*. Amsterdam: John Benjamins, pp. 101-120.

Carvalho, Joaquim Brandão de. 2017. Deriving sonority from the structure, not the other way round: A Strict CV approach to consonant clusters. *The Linguistic Review* 34(4): 589-614. https://doi.org/10.1515/tlr-2017-0012

Charette, Monik. 2017. The internal TR clusters of Acadian French: A hint from schwa. In G. Lindsey & A. Nevins (eds.) *Sonic signatures. Studies dedicated to John Harris*, 17-32. Amsterdam: Benjamins. https://doi.org/10.1075/lfab.14.c2

Côté, Marie-Hélène. 2018. Glissantes et diphtongues en français laurentien. *Cahiers de l'ILSL* 56: 41-59. https://doi.org/10.26034/la.cdclsl.2018.211

Cyran, Eugeniusz. 2003. Licensing strength and syllable structure in government phonology. *Poznań Studies in Contemporary Linguistics* 38. 41–73.

Cyran, Eugeniusz. 2010. *Complexity scales and licensing in phonology*. Berlin/New York: Walter de Gruyter.

Dresher, Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Enguehard, Guillaume & Noam Faust 2018. Guttural Ghosts in Modern Hebrew. *Linguistic Inquiry* 49: 685-721. https://doi.org/10.1162/ling_a_00287.

Guérin, Maximilien. 2022. Les parlers du Croissant : des parlers minorisés et marginalisés. In S. Noirard (ed.), *Transmettre les langues minorisées : entre promotion et relégation*, 129-141. Rennes: Presses universitaires de Rennes.

Faust, Noam. 2024. Epenthetic vowel quality cross-linguistically, with focus on Modern Hebrew. *Linguistics Vanguard*. https://doi.org/10.1515/lingvan-2023-0081

Faust, Noam & Tobias Scheer. To appear. Non-final final consonants in the Croissant. *Phonology*.

Harris, John. 1990. Segmental Complexity and Phonological Government. *Phonology* 7(2): 255–300.

Kaye, Jonathan, Lowenstamm, Jean & Jean-Roger Vergnaud 1985. The internal structure of phonological representations: a theory of Charm and Government. *Phonology Yearbook* 2: 305-328.

Kaye, Jonathan, Lowenstamm, Jean & Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology Yearbook* 7: 193-231.

Lowenstamm, Jean 1996. CV as the only syllable type. In J. Durand & B. Laks (eds.), *Current trends in phonology. Models and methods, vol. 2*, 419-441. Salford, Manchester: ESRI. WEB.

Mackenzie, Sara & Shanti Ulfsbjorninn,. 2023. Timugon Murut Vowel Harmony: A foot-based reanalysis. Slides from Linguistics Seminar Series, Memorial University of Newfoundland.

Morin, Yves Charles. 1986. La loi de position ou de l'explication en phonologie historique. *Revue Québecoise de Linguistique* 15: 199-232. https://doi.org/10.7202/602567ar

Quint, Nicolas 2021a. Le regain des études consacrées aux parlers du Croissant. *Langues et cité* 30: 1-15.

Quint, Nicolas 2021b. La question des radicaux asyllabiques et des paradigmes verbaux du 1er groupe dans les parlers du Croissant et dans d'autres variétés occitanes ou romanes. In L. Esher, M. Guérin, N. Quint & M. Russo, *Le Croissant linguistique : entre oc, oïl et franco-provençal - Des mots à la grammaire, des parlers aux aires* by, 229-259. Paris: L'Harmattan.

Ronjat, Jules. 1913. *Essai de Syntaxe des Parlers Provençaux Modernes*. Mâcon: Protat.

Scheer, Tobias 2004. *A lateral theory of phonology. Vol.1: What is CVCV, and why should it be?* Berlin: Mouton de Gruyter.

Scheer, Tobias. 2014. Muta cum Liquida in the Light of Tertenia Sardinian Metathesis and Compensatory Lengthening Lat. tr - Old French Vrr. Variation within and across Romance Languages. In M.-H. Côté & E. Mathieu, *Selected papers from the 41st linguistic symposium on Romance languages, Ottawa*, 77-99. Amsterdam: Benjamins. https://doi.org/10.1075/cilt.333.07sch

Scheer, Tobias. 2015. *Précis de structure syllabique. Accompagné d'un apparat critique.* Lyon: ENS Editions.

Scheer, Tobias. 2022. 3 x Phonology. *Canadian Journal of Linguistics* 67(3): 1–56.

Ulfsbjorninn, Shanti. 2017. Formalising the markedness of syllable structure. In: Samuels, B. *Beyond Markedness in Formal Phonology*. Amsterdam: John Benjamins Publishing Company. pp. 153-190.

Ulfsbjorninn, S. 2021. A Phonological Reanalysis of Morphological Segment Deletion and De-affrication in Ik. *The Linguistic Review* 38(3):483-516.