Paradigm uniformity effects on French liaison*

Benjamin Storme

February 7, 2022

Abstract

French liaison is a type of external sandhi involving the use of a special allomorph ending in a so-called liaison consonant before vowel-initial words. Liaison consonants are challenging for phonological theory because of evidence that they pattern ambiguously between word-final and word-initial consonants. Specific underlying phonological or lexical representations have been proposed in the literature to account for this ambiguous behavior, including floating consonants, lexical constructions and gradient underlying representations. This paper proposes an alternative analysis of the ambiguous patterning of French liaison as a paradigm uniformity effect, building on Steriade (1999). In a Word1-Word2 sequence, the liaison consonant at the boundary between the two words ends up being ambiguous between a word-final and a word-initial consonant because of a pressure to make contextual variants of Word1 and Word2 similar to their citation forms (i.e. words as pronounced in isolation). The proposal is shown to be able to account for the ambiguity of French liaison both in terms of prosodic attachment and phonetic realization. The paper provides evidence for two key predictions of this analysis, using judgment data on the prosodic attachment of liaison consonants in Swiss French and phonetic data on the interaction between liaison and affrication in Quebec French. The ambiguity of French liaison is modeled using a probabilistic grammar including paradigm-uniformity constraints, without any need for special underlying phonological or lexical representations.

Keywords: French liaison; paradigm uniformity; phonetic detail; constraint-based grammar; laboratory phonology

1 Introduction

French liaison is a type of external sandhi that involves the use of a special variant for some words when they are followed by a vowel-initial word. This variant, known as the

---

*Parts of this work were presented at the *2019 Annual Meeting on Phonology*, at *Going Romance XXIV* in 2020, and at the *Séminaires de Recherches en Phonétique et Phonologie* (SRPP) in 2021. I am particularly grateful to Marie-Hélène Côté, Rachid Ridouane, Cécile Fougeron and Birgit Alber for very useful feedback and suggestions on this project. Special thanks to Aris Xanthos for writing the Python script that I used in Study 1 to obtain lexical frequencies from the OpenSubtitles corpus and to Marie-Hélène Côté for giving me access to the PFC corpus used in Study 2. All supplementary files accompanying the paper may be found in Storme (2022).

word's liaison variant, features a final consonant, called a liaison consonant. For instance, the adjective *grand* 'great' is generally realized as [gʁɑ̃], as shown in (1a) and (1b), but may appear under its liaison variant [gʁɑ̃t] (with a liaison consonant [t]) before vowel-initial words, as shown in (1c).

(1) French liaison

		Word1	Word2		Context
a.	grand	[gʁɑ̃]	(none)	'great'	(citation form)
b.	grand monsieur	[gʁɑ̃]	[məsjø]	'great man'	(before C-initial word)
c.	grand ami	[gʁɑ̃t̬]	[ami]	'great friend'	(before V-initial word)

Liaison is a complex phenomenon that is influenced by a range of linguistic and sociolinguistic variables beyond the basic phonological conditioning described in (1). Due to this complexity, liaison has featured prominently in many theoretical debates over the last decades, including debates on the syntax-phonology interface, the nature of phonological and lexical representations, and the role of frequency (see Côté 2011 for an overview).
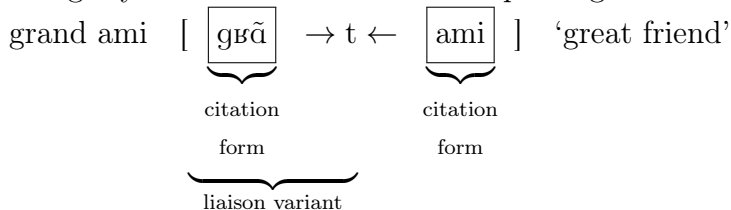
In recent years, there has been renewed interest in a particular challenge that French liaison raises for phonological theory: liaison consonants pattern ambiguously between word-final and word-initial consonants, both prosodically and phonetically. For instance, in the presence of a prosodic break between the liaison variant and the following word, liaison consonants can be attached both at the end of the liaison variant, like word-final consonants (=*liaison non-enchaînée*), and at the beginning of the following word, like word-initial consonants (=*liaison enchaînée*; Encrevé 1988; Durand & Lyche 2008). This ambiguous behavior has led some researchers to propose specific underlying phonological representations for liaison consonants, including floating segments (Encrevé 1988; Tranel 1990) and gradient underlying representations (Smolensky & Goldrick 2016). It has also motivated some researchers to challenge the traditional view according to which liaison consonants are lexically affiliated to Word1, either by positing that they belong to a lexical construction involving both Word1 and Word2 (Bybee 2001) or that they are independently affiliated to both Word1 and Word2 (Smolensky & Goldrick 2016).

This paper proposes an alternative account where the ambiguity of liaison consonants is not captured through inherently ambiguous phonological representations or enriched lexical entries but emerges from the structure of the lexicon as a paradigm uniformity effect, in a similar way to what has been proposed to account for incomplete devoicing in German (Roettger et al. 2014). More specifically, the present proposal builds on a hypothesis put forth by Steriade (1999) and according to which the liaison variant of a word (e.g. [gʁɑ̃t] in (1c)) is attracted to the pronunciation of the corresponding citation form, i.e. the word as pronounced in isolation (e.g. [gʁɑ̃] in (1a)). Crucially, the liaison consonant is typically absent from the citation form. In a Word1-Word2 sequence, uniformity with the citation form of Word1 will push the liaison consonant away from the end of this word, therefore favoring a word-initial behavior.

The present paper extends Steriade's original analysis by hypothesizing that the realization of Word2 is also subject to paradigm uniformity effects. Uniformity with the citation form of Word2 will push the liaison consonant away from the beginning of Word2, therefore favoring a word-final behavior. These two opposite uniformity effects, represented with arrows going in opposite directions in (2), are proposed to underlie the ambiguous realization

of liaison consonants. The forms enclosed in boxes in (2) correspond to the citation forms of the two words involved in the sequence *grand ami*. The hypothesis that the citation form of Word2 also plays a role will be crucial to explain why liaison consonants do not behave just like word-initial consonants but also share properties with word-final consonants, both prosodically and phonetically.

(2)   Ambiguity of liaison consonants as a paradigm uniformity effect

    grand ami   [gʁɑ̃ → t ← ami] 'great friend'

                            citation   citation

                              form   form

                            liaison variant

The paper is organized as follows. Section 2 provides some background on French liaison consonants, with a special focus on their ambiguous prosodic and phonetic realization. Section 3 implements the paradigm-uniformity analysis schematically represented in (2) in a probabilistic constraint-based grammar and shows how this analysis can derive both prosodic ambiguity and phonetic ambiguity. To account for the ambiguous patterning of French liaison at the phonetic level, the paper assumes with Steriade (2000) that phonetic detail may matter in paradigm uniformity effects, in particular that paradigm-uniformity constraints are sensitive to coarticulatory effects between consonants (C) and vowels (V) in CV sequences.

The remainder of the paper provides evidence for two key hypotheses of the analysis. Section 4 presents the results of an experimental study on Swiss French that both confirms the prosodic ambiguity of French liaison and provides evidence for the role of citation forms in this ambiguity. The evidence comes from a comparison of two types of liaison variants differing in their similarity with the corresponding citation forms (epenthetic liaison and suppletive liaison). Section 5 reports on a phonetic study looking at the interaction of affrication and liaison in Quebec French, using data from the *Phonologie du Français Contemporain* (PFC) project. The results of this study provide evidence for the coarticulatory effect that is proposed to underlie the phonetic ambiguity of liaison /t/ before /i/ in Quebec French.

2   Background on French liaison

Two main research questions have been addressed about French liaison: (i) which conditions favor the use of the liaison variant? and (ii) when the liaison variant is used, how is the liaison consonant realized? Section 2.1 briefly presents the first research question. Section 2.2 moves on to the second question, which will be the main focus of this paper. Section 2.3 presents and critically evaluates earlier analyses that have been proposed to account for the data described in section 2.2.

2.1   Question 1: when is the liaison variant used?

The liaison variant of a word is used when the following word starts with a vowel, as illustrated in (1a). However it is not always used when this condition is satisfied and the variant without liaison is still available. Most of the work on French liaison has been dedicated to

understanding which factors beyond the basic phonological conditioning described in (1a) explain the variation between allomorphs with and without liaison. Although this question is not directly relevant to the main issue this paper addresses (namely how is the liaison consonant realized when present?), it is important to briefly review research on this topic in order to understand French liaison in a broader context.

According to previous research, the distribution of the two variants depends on syntactic, phonological and lexical properties of individual words involved in the Word1-Word2 sequence as well as properties of the sequence itself. Sociolinguistic factors such as speech style, language variety and speaker identity have also been found to play a role, as shown in Table 1.

Variables		Source
Morphosyntactic variables	PoS of Word1 and Word2	Fougeron et al. (2001)
Phonological variables	Length of Word1	Fougeron et al. (2001); Kilbourn-Ceron (2017)
Lexical variables	Freq of Word1, Word2 and Word1-Word2	Fougeron et al. (2001); Kilbourn-Ceron (2017)
	Identity of Word1	Côté (2011)
	Identity of Word2	Zuraw & Hayes (2017: section 3)
Sociolinguistic variables	Speech style	Fougeron et al. (2001)
	French variety	Côté (2017)
	Speaker identity	Encrevé (1988)

Table 1: A non-exhaustive list of variables reported to condition the use of the liaison variant in French along with a non-exhaustive list of sources (PoS = part of speech, Freq = frequency).

For instance, the liaison variant is more likely to be used if Word1 is a function word, if it is monosyllabic, and if it has high lexical frequency. The specific identity of Word1 also seems to matter as the aforementioned grammatical and lexical factors do not account for all the variation observed across different types of liaison words (Côté 2011: section 2.2.3). Properties of Word2 also play a role. For instance, Zuraw & Hayes (2017: section 3) found that the likelihood of the liaison variant depends on the lexical identity of the following word (=Word2). More specifically, words with higher lexical frequency have been reported to favor the liaison variant for the preceding word (see Fougeron et al. 2001; but Kilbourn-Ceron 2017: 146). Finally, properties of the Word1-Word2 sequence itself are also relevant. For instance, Fougeron et al. (2001) found that the liaison variant becomes more likely as the frequency of the Word1-Word2 sequence increases, even when controlling for the individual frequencies of Word1 and Word2. Also, Kilbourn-Ceron (2017: 146) found that the frequency of the liaison variant is positively correlated with the conditional probability of Word2 given Word1, i.e. the liaison variant is more likely when the following word is contextually more predictable.

The most recent account of phonological, lexical and syntactic effects on external sandhi is Kilbourn-Ceron's (2017) dissertation. In Chapter 4, she proposed to interpret some of the effects reported above for French liaison in light of the Production Planning Hypothesis (PPH; Wagner 2012; Tanner et al. 2017). According to this hypothesis, an external sandhi process can be blocked if the target of this process (here the liaison word) and the triggering context (here the following vowel-initial word) are not within the same planning window. Under this view, the liaison variant of Word1 is more likely when Word1 and Word2 are more likely to be

encoded together in the same planning window. Whether two words are encoded together phonologically is proposed to depend on a number of language-specific factors (syntactic complexity, lexical frequency, contextual lexical predictability) and cognitive factors (working memory load, cognitive load). The effects of word length, word frequency, and sequence frequency on French liaison are interpreted as compatible with the PPH, on the assumption that short, high-frequency, and contextually predictable words facilitate word form retrieval. Some of the syntactic effects reported by Fougeron et al. (2001) might also be accounted for under this hypothesis. The PPH indeed predicts that words separated by a weaker syntactic boundary are more likely to be encoded together phonologically (Kilbourn-Ceron 2017: chapter 3). Function words in French often entertain a close syntactic relationship with the following word. For instance, the determiner and the following noun form a noun phrase in French (*les enfants* 'the children'). The fact that function words are particularly prone to appear under their liaison variants can then be understood as a phonological reflex of this syntactic proximity.

Beyond the linguistic properties just discussed, the role of sociolinguistic variables such as speech register has long been recognized. For instance, liaison variants are more likely to be used in read speech than in conversational speech (Fougeron et al. 2001). More recently, the role of dialectal variation has also been pointed out by Côté (2017). For instance, liaison variants are overall less likely in French varieties spoken in Louisiana and in Africa than in Canada and Europe. There also seems to be genuine speaker variation, with different individuals having different rates of liaison realization within the same speech register and the same French variety. For instance, Encrevé (1988: 56) reports different rates of liaison in political speeches by different French politicians.

2.2 Question 2: how is the liaison consonant realized when present?

Explaining the distribution of word variants with and without liaison is not enough to have a comprehensive understanding of French liaison. Indeed, there is a further complexity concerning the way liaison consonants are realized when they are present: liaison consonants pattern ambiguously between stable word-final and word-initial consonants. The expression 'stable consonants' refers here to consonants that differ from liaison consonants in being present regardless of the surrounding phonological context (e.g. word-initial [t] and word-final [t] in *trente* [tʁɑ̃t] 'thirty' are stable in this sense). Sections 2.2.1 and 2.2.2 present the evidence that French liaison behaves ambiguously at the prosodic level and at the phonetic level, respectively.

2.2.1 Prosodic ambiguity

In connected speech, stable consonants and liaison consonants generally have the same prosodic behavior when followed by a vowel: they tend to be syllabified with that vowel (Spinelli et al. 2002; Gaskell et al. 2002; Durand & Lyche 2008: section 3.3). This happens whether the vowel is in the same word or in another word. The vowel is in the same word when the consonant is a stable word-initial consonant (e.g. *grand tamis* [gʁɑ̃.#ta.mi] 'big sieve'). The vowel is in a different word when the consonant is a stable word-final consonant (e.g. *trente amis* [tʁɑ̃.t#a.mi] 'thirty friends') or a liaison consonant (e.g. *grand ami*

5

[gʁɑ̃.t#a.mi] 'great friend'). For word-final consonants and liaison consonants, syllabification with the following vowel is described as resyllabification or as *enchaînement*.

In a corpus study, Fougeron & Delais-Roussarie (2004) found that resyllabification of stable word-final consonants actually happens in 60% of cases. In the 40% remaining cases, there is a prosodic break (a pause, a glottal stop, or a hesitation) that intervenes between the consonant and the following vowel. It's in this context that the behavior of liaison consonants and stable consonants actually comes apart. Although lexically dependent on the identity of Word1, the liaison consonant will typically behave like a word-initial consonant prosodically and attach to Word2.

This can be illustrated with the behavior of Word1-Word2 sequences in right dislocations (Tranel 1990; Côté 2005). Right dislocated elements belong to a distinct prosodic unit from their nucleus sentence: their prosody copies the prosody of the nucleus but is characterized by decreased intensity, lower pitch and a flat contour intensity (De Cat 2007: 34-43). When Word2 is right dislocated, the liaison consonant has a very puzzling behavior: it is separated prosodically from the word it is lexically affiliated to (Word1) and is attached to the following word prosodically (Word2). In this case, the liaison consonant is described as resyllabified or *enchaînée*. For instance, in (3a), liaison [t] is separated from its lexical host [gʁɑ̃] by a prosodic boundary[1] and attaches to [elefɑ̃]. The prosodic boundary is materialized by a space in the phonetic transcription. By contrast, stable consonants remain prosodically attached to their lexical hosts, as shown in (3b) for word-final consonants (they attach to Word1) and in (3c) for word-initial consonants (they attach to Word2).

(3)   Liaison vs. stable consonants in right dislocations
      a.   Liaison consonant   J'en ai un grand, éléphant.   [gʁɑ̃ **t**elefɑ̃]
                                   'I have a big one, elephant.'
      b.   Word-final consonant   J'en ai trente, éléphants.   [tʁɑ̃**t** elefɑ̃]
                                   'I have thirty of them, elephants.'
      c.   Word-initial consonant J'en ai un beau, tabouret.   [bo **t**abuʁɛ]
                                   'I have a beautiful one, stool.'

Although resyllabification of liaison consonants is reported as the preferred option in the presence of a prosodic break, liaison consonants may still be attached prosodically at the end of the first word. In this case, the liaison consonant is said to be *non-enchaînée*. The availability of *liaison non-enchaînée* has been famously described by Encrevé (1988), using a corpus of political speeches by prominent French politicians. Encrevé's data are summarized in Table 2 for all politicians included in the study. The data show that the liaison variant is used in about half of the potential sites for liaison (49%). When present, it is mostly syllabified as an onset at the beginning of Word2 (= *liaison enchaînée*) but it can also be syllabified as a coda at the end of Word1 (= *liaison non-enchaînée*).

An example of *liaison non-enchaînée* by the late French president Jacques Chirac is provided in (4) (cited from Durand & Lyche 2008: 51).

---

[1]The liaison consonant is lexically dependent on Word1 because the choice of a specific liaison consonant among the set of potential liaison consonants (i.e. /t z n ʁ/) covaries only with Word1, at least in adult speech (see Chevrot et al. 2009 on child speech). The liaison consonant is /t/ in the sequence *grand ami* 'great friend' but it becomes /n/ if *grand* is replaced by indefinite determiner *un* in the sequence *un ami* 'a friend'.

		Count	Frequency
Liaison present	enchaînée	2488	43%
	non-enchaînée	321	6%
Liaison absent		2978	51%

Table 2: Whether and how French liaison is realized in a corpus of political speeches (based on Encrevé 1988: 56)

(4)   Liaison enchaînée in a speech by Jacques Chirac (November 14, 2005)
       font honneur [fɔ̃t̪ ɔnœʁ] 'do.PRES.3PL honor'

Although *liaison non-enchaînée* has sometimes been described as uniquely confined to high register and planned speech, Durand & Lyche (2008: 50-51) found examples occurring in natural daily interactions, in particular in the contexts of prosodic breaks involving hesitations, as shown in (5a), or repetitions, as shown in (5b). Durand & Lyche (2008) write: 'These examples seem to us extremely interesting: despite the clear predominance of *liaison enchaînée* in our corpus, they provide possible evidence against an analysis which simply treats a liaison consonant as an onset of W[ord] 2.' Indeed, this behavior is not reported for stable word-initial consonants at the beginning of Word2: these are not allowed to be attached at the end of Word1 across a prosodic break occurring in the middle of a Word1-Word2 sequence (see Section 4 for experimental evidence).

(5)   Examples of *liaison non-enchaînée* in conversational speech (Durand & Lyche 2008: 50-51)
     a.   vingt euh   [vɛ̃t̪ œ]   (hesitation)
     b.   un, un Aveyronnais   [œ̃n̪ œ̃naverone]   (repetition)

2.2.2   Phonetic ambiguity

Even in case of resyllabification/ *enchaînement*, the distinction between liaison consonants, stable word-final consonants and stable word-initial consonants is not completely neutralized. There remain phonetic cues that distinguish the three types of consonants, as will be reviewed in this section. In particular, liaison consonants have a phonetic patterning that is intermediary between that of stable word-final and word-initial consonants.

Fougeron (2007) showed that word-final consonants before a vowel (VC#V) do not have the same acoustic realization as word-initial consonants (V#CV), even in contexts that are traditionally treated as involving resyllabification of the final consonant. In particular, she found that word-final consonants tend to be shorter than word-initial consonants (Fougeron 2007: 13).

Liaison consonants are also reported to behave distinctly from both stable word-final and word-initial consonants phonetically. For instance, an early study by Durand (1936: 238) found that stable word-final consonants (e.g. final [t] of *petite* [pətit] 'small.FEM' in *une petite orange* 'a small orange') differ from liaison consonants (e.g. liaison [t] at the end of *petit* [pətit] 'small.MASC' in *un petit orage* 'a small storm') in retaining some cues of their implosive/coda nature. Liaison consonants have also been found to differ from word-initial

consonants and in particular to be characterized by a shorter duration on average (Gaskell et al. 2002; Spinelli et al. 2002, 2003).

More targeted studies found that the effect of the consonant's lexical type (word-final, liaison, word-initial) on phonetic realization might differ by consonant, as summarized in Table 3. For /t/, the duration is longer for word-initial consonants than for both word-final and liaison consonants, but without clear durational difference between the latter two types. For /z/, a study by Nguyen et al. (2007) found a shorter duration for word-final consonants. But this result was not replicated by Bagou et al. (2009). Neither study found a significant durational difference between word-initial and liaison /z/. For /n/, phonetic realization does not seem to be affected by the lexical status of the consonant, according to available studies: the duration of the consonant does not significantly differ whether the consonant is word-final, liaison or word-initial. Overall, taken together, the liaison consonant appears to pattern ambiguously between word-final and word-initial consonants in phonetic realization, with some differences depending on the specific liaison consonant.

	word-final		liaison		word-initial
/t/	dur(C)	$=^a$	dur(C)	$<^{a,d}$	dur(C)
/z/	dur(C)	$<^b/=^a$	dur(C)	$=^{a,b}$	dur(C)
/n/	dur(C)	$=^b$	dur(C)	$=^{b,c}$	dur(C)

[a]Bagou et al. 2009, [b]Nguyen et al. 2007: 12
[c]Wauquier-Gravelines 1996, [d]Dejean de la Bâtie & Bradley 1995: footnote 2

Table 3: Phonetic realization (duration) of consonants /t z n/ as a function of lexical type (stable word-final consonant, liaison consonant, stable word-initial consonant)

The clearest evidence for a phonetic ambiguity of liaison consonants comes from data on affrication in Quebec French (Côté 2014). Quebec French has a process of affrication that turns /t d/ into [ts dz] before /i y j ɥ/. But this process affects differently liaison consonants, stable word-final consonants, and stable word-initial consonants, as shown in Table 4. More specifically, liaison /t/ has a rate of affrication that is intermediary between stable word-final /t/ and word-initial /t/: liaison /t/ is more prone to affrication than stable word-final /t/ but less so than stable word-initial /t/. Rates of affrication were obtained by Côté (2014) on a perceptive basis, using data from the PFC project (Côté 2016).

	Consonant is absent	Consonant is present	
Consonant type		Affrication	No affrication
Liaison consonant	21 (17.0%)	68 (55.0%)	35 (28.0%)
Stable word-final consonant		31 (36.5%)	54 (63.5%)
Word-initial consonant		715 (99.2%)	6 (0.8%)

Table 4: Affrication before /i y j ɥ/ in the PFC Trois-Rivières survey: count and frequency data (Côté 2014: 38)

It is important to note that only contexts involving resyllabification/*enchaînement* were included by Côté in the data reported in Table 4. For instance, among the 131 occur-

rences involving a stable word-final consonant before /i/ in the corpus, only 85 were included in the analysis because the remaining 36 occurrences did not involve resyllabification/*enchaînement*.[2] This means that the differences in affrication observed in Table 4 cannot be explained away as a by-product of prosodic ambiguity: all contexts involve onset consonants. This is important from a theoretical perspective: it means that analyses that are tailored to account for prosodic ambiguity do not automatically account for phonetic ambiguity, as will be further discussed in Section 2.3.

2.3   Previous analyses

Several analyses have been proposed to account for the ambiguity of French liaison, focusing mainly on prosodic ambiguity (see Côté 2011 for an overview). These analyses can be classified according to whether they require to enrich the French phoneme inventory or the French lexicon, as shown in Table 5. This table is only meant as an indicative, non-exhaustive list of the theoretical options and their sources. Discussion of the paradigm-uniformity analysis is delayed to Section 3.

	Enriched phoneme inventory	Enriched lexicon	
Floating consonants	yes	no	Encrevé 1988; Tranel 1990, 2000
Lexical constructions	no	yes	Bybee 2001; Chevrot et al. 2009
Gradient representations	yes	yes	Smolensky & Goldrick 2016
Paradigm uniformity	no	no	Steriade 1999

Table 5: Analyses of French liaison: a typology

2.3.1   The approach using floating consonants

Liaison consonants have been analyzed as floating segments by several researchers, including by Encrevé (1988: 169-173) and Tranel (1990: 183-184) in the framework of autosegmental phonology and by Tranel (2000: 49-52) in the framework of Optimality Theory. In these approaches, liaison consonants are lexically affiliated to the first word but differ from stable word-final consonants in not being attached to the word's early prosodic structure. This property allows them to be associated at a later stage of prosodic-structure building to either Word1 or Word2 in a Word1-Word2 sequence (Encrevé 1988: 182).

The specific proposal advanced by Tranel (1990) is represented in (6). In this analysis, the liaison consonant /z/ does not project a skeletal slot and is therefore 'floating' at the end of Word1 in the early prosodic structure. When the two words are combined together, the liaison consonant has to be attached somewhere prosodically. Rightward syllabification attaches it at the beginning of Word2, making it a *liaison enchaînée*. Leftward syllabification attaches it at the end of Word1, making it a *liaison non-enchaînée*.

---

[2]The rate of *non-enchaînement* (30%) is similar to that reported by Fougeron & Delais-Roussarie (2004) in another corpus (40%).

(6)   Liaison consonants as floating segments (based on Tranel 1990: 184)

Word1   Word2

σ ‑ ‑ ‑ ‑ ‑ ‑ ‑ ‑   σ   σ

x   x   x]   [x   x   x

g   ʁ   o   z   a   n   o
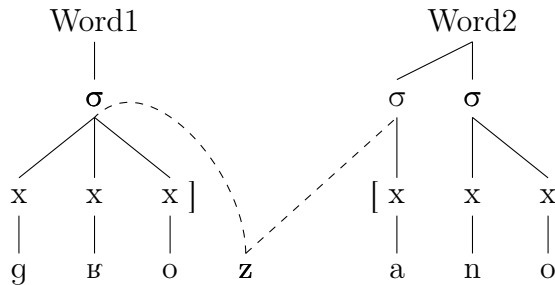
Analyses using floating segments require to enrich the phoneme inventory of French with a new set of phonemes. For instance, Tranel (2000: 51-52) introduces a phonological feature to distinguish stable consonants (noted as C) from liaison consonants (noted as L). As will be further discussed in Section 2.3.4, this raises some issues. But there is a more direct empirical problem that remains unaddressed by this approach: it accounts for the prosodic ambiguity of French liaison but not for its phonetic ambiguity. Indeed, if liaison consonants become identical to onset consonants after rightward resyllabification, then it is unclear why they should pattern differently from stable word-initial consonants phonetically in this context, as documented in Section 2.2.2. An additional mechanism is therefore needed to account for phonetic ambiguity.

2.3.2   The approach using lexical constructions

In the approach using lexical constructions, the liaison consonant belongs neither to Word1 nor to Word2 but to a construction involving the two words (Bybee 2001). For instance, there is a lexical construction of the form /gʁɑ̃$_{Adj}$ t X$_{V\text{-initial N}}$/ 'great N', where X$_{V\text{-initial}}$ is a vowel-initial noun and /t/ a consonant occurring between the Adj and the N. Nouns that are more frequently associated with the adjective *grand* 'great' are more likely to be stored under this frame, explaining for instance why the likelihood of the liaison consonant increases with the frequency of the Word1-Word2 sequence (see Section 2.1).

Although lexical constructions are primarily motivated by the type of frequency effects reported in Section 2.1, Bybee mentioned in passing that they can also account for the prosodic ambiguity of French liaison. She argues that a prosodic break may intervene in the middle of a lexical construction in the same way as it may intervene in the middle of a word. For instance, it it possible to say *un élé phant* [ɛ̃n#ele fɑ̃] 'an ele (prosodic break) fant' with a prosodic break in the middle of the word *éléphant*. *Liaison non-enchaînée* and *liaison enchaînée* would then correspond to situations where the prosodic break within a lexical construction intervenes after and before the liaison consonant, respectively.

This approach suffers from the same limit as the approach using floating consonants in that it does not account for phonetic ambiguity. Moreover, it potentially presents another problem. Stable word-final consonants and word-initial consonants do not seem to be separable from their lexical host prosodically, even in high-frequency two-word sequences. For instance, a prosodic break seems much more natural after the stable final consonant of Word1 than before it in the compound *porte-avion* 'aircraft carrier' (*porte euh avion* [pɔʁt œ avjɔ̃]/*por euh tavion* [pɔk œ tavjɔ̃]). Bybee (2001) sketches an explanation for why it does

10

not happen: 'Since the words of a construction are usually associated with other instances of the same word, their identity as words is known, and the point between two words is a possible place to pause.' In other words, a pause is more likely to occur between words than within words inside a multiple-word construction because the word forms inside this construction stand in correspondence with their base forms (which are independently stored outside of any construction). In other words, stable word-final [t] in *porte-avion* cannot be resyllabified across a prosodic break because there is a pressure from the base form *porte* to maintain the [t] at the end of Word1. Liaison consonants are not subject to the same pressure because they are absent from the base forms of Word1 and Word2. When fleshed out, this explanation actually clearly refers to principles of paradigm uniformity among morphologically related forms. As will be shown in Section 3, paradigm uniformity is actually sufficient to derive the ambiguity of French liaison consonants and lexical constructions are not necessary for this purpose.

2.3.3   The approach using gradient representations

In the approach using gradient representations, liaison consonants are characterized by an activity degree that is smaller than that of stable consonants (Smolensky & Goldrick 2016). This lower activity degree is the mechanism that explains why liaison consonants do not always surface. Moreover, liaison consonants are assumed to be stored both at the end of liaison words and at the beginning of all vowel-initial words. This explains why they might be realized both at the end of Word1 (*liaison non-enchaînée*) and at the beginning of Word2 (*liaison enchaînée*).

   This analysis is represented in (7), where the degree of activity of liaison consonants is indicated as a subscript (for stable consonants, the degree of activity is always equal to 1). A word like *ami* [ami] 'friend' is stored with all the possible liaison consonants that can be attached to it as first segment. When the words *petit* and *ami* are combined, the activity level of liaison /t/ increases, allowing it to surface. Because the /t/ is underlyingly present in both words, it can surface either at the end of Word1 or at the beginning of Word2.

(7)   Liaison consonants as gradient phonemes affiliated to both Word1 and Word2 (Smolensky & Goldrick 2016)
   /pəti$\underline{t_{0.48}}$/ + /{$\underline{t_{0.09}}$, $\underline{z_{0.09}}$, $\underline{n_{0.09}}$}ami/

   Like the two approaches previously reviewed, the approach using gradient underlying representations does not account for the phonetic ambiguity of French liaison. Indeed, the activity degree of phonemes determines how likely they are to surface but does not determine their segmental or phonetic realization. Another mechanism is required in addition to derive the phonetic ambiguity of French liaison.

2.3.4   General discussion

The three approaches just reviewed can derive the prosodic ambiguity of French liaison but not its phonetic ambiguity. Because phonetic ambiguity is not just a by-product of prosodic ambiguity, additional mechanisms are required. Beyond this problem, these approaches also require non-trivial theoretical innovations.

The approaches using floating consonants and gradient representations require to enrich the phoneme inventory of French with a new type of phonemes. At the same time, these new phonemes have to be very constrained in their extension and in their distribution, without any clear principled motivation for why these restrictions would hold. Floating consonants and gradient representations are limited to the specific set of consonants that behave as liaison consonants in French (/t/, /z/, /n/, /ʁ/). Moreover, only word-final segments are floating in the floating-segment approach and only word-final and word-initial segments are gradient in the approach based on gradient underlying representations. Usually such restrictions on feature combinations and their contextual distribution have external motivations. For instance, vowel backness and rounding often correlate in vowel inventories because these two features contribute to differences along the second vowel formant (Flemming 2004). Place features are more likely to be neutralized word-finally because there are less cues available to signal place contrasts in this position (Jun 2004). In the present case, it is not clear why only these particular phonemes would be floating or gradient, and only in these specific positions. In the paradigm-uniformity approach that will be proposed in Section 3, the puzzling behavior of liaison consonants will not follow from the structure of the French phoneme inventory but will ultimately result from the existence of competing allomorphs for some words. This will make it possible to attribute the idiosyncrasies of French liaison to the lexicon rather to the phonology. This solution seems more satisfactory as allomorphy is pervasive in the lexicons of the world's languages.

The approaches using lexical constructions and gradient underlying representations require to enrich the French lexicon by positing a very large amount of allomorphy. In the approach using gradient representations, vowel-initial words must be stored with as many consonant-initial allomorphs as there are liaison consonants in French. In the approach using lexical constructions, many multiple-word sequences must be stored in the lexicon. Frequency effects documented in Section 2.1 provide evidence for these lexical constructions. However, as will be shown further in Section 3, the ambiguity of French liaison can be derived without multiplying allomorphs or lexical constructions under the paradigm-uniformity analysis. It will be enough to say that only liaison words have listed allomorphs.

3 The ambiguity of French liaison as a paradigm uniformity effect

This section proposes an analysis of the ambiguous patterning of French liaison as resulting from a pressure to make contextual variants of words similar to their citation forms (i.e. words as pronounced in isolation). It operates on traditional phonological representations for phonemes and assumes that liaison words come with two listed allomorphs. The analysis extends Steriade's (1999) analysis in assuming that not only the citation form of Word1 is relevant but also the citation form of Word2. This extension will make it possible to explain why the liaison consonant may still be attached at the end of Word1. The analysis also extends Steriade's analysis in that it can derive not only prosodic ambiguity but also phonetic ambiguity.

The analysis is implemented in a constraint-based grammar including input-output (IO) and output-output (OO) faithfulness constraints evaluated in parallel, according to the gen-

eral schema in Figure 1. Input-output faithfulness constraints evaluate the similarity between a surface form and the corresponding underlying representation. Output-output faithfulness constraints evaluate the similarity between the contextual variants of a word and the corresponding citation form. The model assumes base priority (Benua 1997: 240): the phonology of the citation form is computed first and then the resulting output form is used in the evaluation of contextual variants.
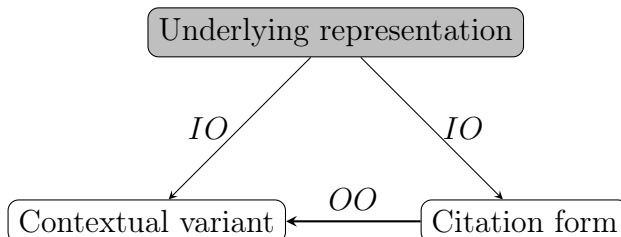
```
            Underlying representation
              /                    \
           IO                        IO
            /                          \
   Contextual variant  ←——OO——  Citation form
```

Figure 1: Correspondence relations assumed in the analysis (IO = input-output correspondence, OO = output-output correspondence)

The key insight on which the analysis builds is that the liaison consonant is present in the underlying representation of a word with a liaison variant but (generally) absent from the corresponding citation form. By contrast, stable word-final and word-initial consonants are systematically present in both representations. Section 3.1 shows how this insight can help derive the prosodic ambiguity of liaison consonants across a prosodic break. Section 3.2 shows how the analysis can extend to the phonetic ambiguity of liaison consonants, using Quebec French as a case study.

The analysis focuses on deriving the main pronunciations of liaison consonants (presence/absence, *liaison enchaînée/non-enchaînée*, affrication/no affrication) but does not derive all the effects reported to condition the presence/absence of French liaison in Section 2.1 (e.g. frequency effects, dialectal effects, etc.). The main reason is that these effects seem to be largely orthogonal to the question of the prosodic/phonetic ambiguity of French liaison. They could probably be derived in a more complex analysis that would for instance combine the constraint-based grammar proposed here with the Production Planning Hypothesis discussed in Section 2.1.

3.1   A model of the prosodic ambiguity of French liaison

Section 3.1.1 shows how the analysis derives citation forms for words with liaison variants and for words that lack such variants. Section 3.1.2 shows how the prosodic ambiguity of liaison consonants can be derived as a result of paradigm uniformity with citation forms. Section 3.1.2 implements the analysis in a probabilistic grammar and shows that it provides a very good fit to the distribution of *liaison enchaînée* and *liaison non-enchaînée* in Encrevé (1988).

3.1.1 Citation forms

Liaison words are assumed to come with two underlyingly listed allomorphs (see also Gaatone 1978; Steriade 1999): the allomorph without liaison (e.g. *grand* /gʁɑ̃/) and the allomorph with liaison (e.g. *grand* /gʁɑ̃t/). By contrast, words with stable final/initial consonants come with a single allomorph.

Table 6 shows how to derive the absence of liaison consonants in citation forms for liaison words (Table 6a) and the presence of stable consonants in citation forms, whether they are word-final (Table 6b) or word-initial (Table 6c). The analysis uses three well motivated constraints: (i) two input-output faithfulness constraints ($\text{Dep}_{\text{IO}}(C)$, $\text{Max}_{\text{IO}}(C)$), penalizing consonant epenthesis and deletion, respectively, and (ii) a markedness constraint penalizing utterance-final consonants (*C#).

UR:	/gʁɑ̃, gʁɑ̃t/	$\text{Dep}_{\text{IO}}(C)$	$\text{Max}_{\text{IO}}(C)$	*C#
a. ☞	/gʁɑ̃/ → [gʁɑ̃]			
b.	/gʁɑ̃t/ → [gʁɑ̃t]			1!
c.	/gʁɑ̃/ → [gʁɑ̃t]	!1		1
d.	/gʁɑ̃t/ → [gʁɑ̃]		!1	

(a) Liaison words (*grand* 'great')

UR:	/tʁɑ̃t/	$\text{Dep}_{\text{IO}}(C)$	$\text{Max}_{\text{IO}}(C)$	*C#
a. ☞	[tʁɑ̃t]			1
b.	[tʁɑ̃]		1!	

(b) Words with stable final consonants (*trente* 'thirty')

UR:	/tablo/	$\text{Dep}_{\text{IO}}(C)$	$\text{Max}_{\text{IO}}(C)$	*C#
a. ☞	[tablo]			
b.	[ablo]		1!	

(c) Words with stable initial consonants (*tableau* 'table')

Table 6: Constraint ranking for citation forms

A technical note about listed allomorphs is in order before motivating the analysis in more details. When the phonology of liaison words (e.g. *grand*) is evaluated, as in Table 6a, the two listed allomorphs can serve as inputs. If one considers [gʁɑ̃] and [gʁɑ̃t] as potential surface candidates, as in Table 6a, this means that a total of four mappings must be evaluated: the two faithful mappings (a and b) and the two unfaithful mappings (c and d). However, for a given output, a mapping from a listed allomorph that involves a smaller number of faithfulness violations (e.g. candidate b: /gʁɑ̃t/ → [gʁɑ̃t]) will always beat a mapping from a listed allomorph that involves a larger number of faithfulness violations (e.g. candidate c: /gʁɑ̃/ → [gʁɑ̃t]) because it implies a stricter subset of constraint violations (it violates the same number of markedness constraints but fewer faithfulness constraints). This means that it is not necessary to evaluate all possible mappings from listed allomorphs to outputs, as done in Table 6a. Instead, for a given output, only the input-output mapping that involves the smaller input-output distance needs to be evaluated. For instance, only mappings a and

b need to be evaluated in Table 6a since candidates c and d are harmonically bounded by those candidates. In what follows, the analysis of liaison words will omit mappings that are harmonically bounded.

In citation forms, the liaison consonant does not appear. The preference for the vowel-final variant can be attributed to a markedness constraint that bans utterance-final consonants (*C#), as shown in Table 6a. In order to block consonant deletion for words with final consonants but no vowel-final listed allomorph, a constraint that bans consonant deletion ($Max_{IO}(C)$) must outrank *C#, as shown in Table 6b. This ranking ensures that a vowel-final allomorph is preferred in citation forms only in case it is listed underlyingly. In other words, the preference for the vowel-final allomorph can be analyzed as a case of emergence of the unmarked. Tables 6c shows how the same grammar also blocks consonant deletion at the beginning of words that begin with a consonant underlyingly.

3.1.2 Contextual variants

In connected speech, markedness constraints that are not relevant utterance-finally will play a role and drive alternations. The crucial markedness constraint that will motivate the external sandhi in French liaison is the anti-hiatus constraint *VV (see Steriade 1999; Tranel 2000). In addition to input-output faithfulness, paradigm uniformity with citation forms derived in section 3.1.1 will also play a role. The analysis here focuses specifically on the context where liaison consonants and stable word-final/word-initial consonants have different behaviors prosodically, namely in Word1-Word2 sequences with a prosodic break between the two words.

Table 7 shows how to derive the prosodic ambiguity of liaison consonants in Word1-Word2 sequences (see Table 7a) and the absence of prosodic ambiguity for stable word-final and word-initial consonants (see Tables 7b and 7c, respectively). % is used to indicate free variation among several candidates. ☞ indicates the categorical winner candidate according to the analysis. The analysis uses four constraints that are well motivated: (a) the anti-hiatus constraint *VV, (b) an input-output faithfulness constraint that penalizes consonant epenthesis in the input-output dimension ($Dep_{IO}(C)$) and (c) two output-output faithfulness constraints (Left-Anchor$_{OO}$, Right-Anchor$_{OO}$) that will be further motivated below.

URs: Citation forms:	/gʁɑ̃, gʁɑ̃t/ + /ami/ [gʁɑ̃] + [ami]	Dep$_{IO}$(C)	*VV	L-Anchor$_{OO}$	R-Anchor$_{OO}$
a. %	gʁɑ̃ ami		1		
b. %	gʁɑ̃t ami				1
c. %	gʁɑ̃ tami			1	
d.	gʁɑ̃l ami	1!			1
e.	gʁɑ̃ lami	1!		1	

(a) Liaison words before vowel-initial words (*grand ami* 'great friend')

URs: Citation forms:	/tʁɑ̃t/ + /ami/ [tʁɑ̃t] + [ami]	Dep$_{IO}$(C)	*VV	L-Anchor$_{OO}$	R-Anchor$_{OO}$
a. ☞	tʁɑ̃t ami				
b.	tʁɑ̃ tami			1	1

(b) Words with stable final consonants before vowel-initial words (*trente amis* 'thirty friends')

URs: Citation forms:	/vʁɛ/ + /tablo/ [vʁɛ] + [tablo]	Dep$_{IO}$(C)	*VV	L-Anchor$_{OO}$	R-Anchor$_{OO}$
a.	vʁɛt ablo			1	1
b. ☞	vʁɛ tablo				

(c) Words with stable initial consonants after vowel-final words (*vrai tableau* 'true table')

Table 7: Constraint ranking for contextual variants

The analysis in Table 7a assumes that the liaison consonant is used to avoid a sequence of two vowels (*VV), even across a prosodic boundary. This configuration cannot be avoided through the epenthesis of any kind of consonant, as this would violate the higher-ranked Dep$_{IO}$(C) constraint. For instance, [l]-epenthesis (present in candidates d and e) is penalized because [l] belongs neither to *grand* nor to *ami* underlyingly.

Crucially, liaison [t] is present in the underlying representation of *grand* and therefore the candidates with liaison [t] (candidates b and c in Table 7a) do not violate Dep$_{IO}$(C). This makes them available to break the hiatus at the boundary between the two words. However liaison variants still involve a departure from the corresponding citation forms in the output-output dimension. Indeed, the candidates with liaison [t] (candidates b and c) feature a consonant that is absent in the citation forms of *grand* and *ami* ([gʁɑ̃] and [ami], respectively). To penalize epenthesis at word edges, Right-Anchor and Left-Anchor constraints are used (Kager 1999: 251). R(ight)-Anchor$_{OO}$ bans epenthesis at the right edge of the word and therefore penalizes *liaison non-enchaînée* (candidate b). L(eft)-Anchor$_{OO}$ bans epenthesis at the left edge of the word and therefore penalizes *liaison enchaînée* (candidate c).

The free variation observed by Encrevé (1988) between absence of liaison (candidate a), *liaison enchaînée* (candidate c) and *liaison non-enchaînée* (candidate b) can be derived if *VV and the two paradigm-uniformity constraints L-Anchor$_{OO}$ and R-Anchor$_{OO}$ are freely ranked, as shown in Table 7a.

This constraint ranking predicts prosodic ambiguity for liaison consonants but not for stable word-final and word-initial consonants. This is shown in Tables 7b and 7c, respectively. In Table 7b, candidate b deletes word-final [t] in the citation form of *trente* and epenthesizes

a [t] at the beginning of the citation form of *ami*, hence violating the two output-output faithfulness constraints L(eft)-Anchor$_{OO}$ and R(ight)-Anchor$_{OO}$. By contrast, candidate a does not violate any constraint and therefore harmonically bounds candidate b. In other words, the analysis predicts that stable word-final consonants should be categorically attached to Word1 prosodically.

In Table 7c, candidate a epenthesizes a [t] at the end of the citation form of *vrai* and deletes word-initial [t] in the citation form of *tableau*, hence violating the two output-output faithfulness constraints L(eft)-Anchor$_{OO}$ and R(ight)-Anchor$_{OO}$. By contrast, candidate b does not violate any constraint and therefore harmonically bounds candidate a. In other words, the analysis predicts that stable word-initial consonants should be categorically attached to Word2 prosodically.

3.1.3   Modeling study: Encrevé (1988)

The analysis captures the ambiguous prosodic behavior of liaison consonants at a conceptual level. To test whether it can also match the specific rates of *liaison enchaînée* and *liaison non-enchaînée* attested in French, the grammatical model described in section 3.1.2 was fit to Encrevé's count data shown in Table 2, using Maxent (Hayes & Wilson 2008) as framework for probabilistic grammars and the software OT-Soft (Hayes et al. 2013) to infer constraint weights. In OT-Soft, the number of iterations was set to 100, the minimum weight to 0, and the maximum weight to 50. The file that was used for the analysis can be found in Storme (2022) under the name `prosodic-ambiguity.txt`.

Three assumptions were made in this analysis: (i) citation forms were assumed to be categorically realized as in section 3.1.1, in particular the liaison consonant was assumed to be categorically absent from the citation form of liaison words, (ii) stable word-final and word-initial consonants were assumed to have a categorical prosodic realization, attaching categorically to Word1 and Word2, respectively (this assumption will receive empirical support in Section 4), and (iii) candidates with [l]-epenthesis in liaison words (candidates d and e in Table 7a) were assumed to be unattested.

Table 8a shows the weights inferred for each of the four constraints used in the analysis. Table 8b shows that the frequencies predicted for each surface form perfectly match the frequencies attested in the corpus, therefore providing a quantitative argument for the paradigm-uniformity analysis. The fact that *liaison enchaînée* is more common than *liaison non-enchaînée* is captured in the difference between the weights of the two paradigm-uniformity constraints: epenthesis at the right edge is more penalized than epenthesis at the left edge.

Constraint	Weight
Dep$_{IO}$(C)	50.00
*VV	1.82
L-Anchor$_{OO}$	2.00
R-Anchor$_{OO}$	4.04

(a) Constraint weights

	Candidates	Attested frequency	Predicted frequency
grand ami	[gʁɑ̃ ami]	0.51	0.51
	[gʁɑ̃t ami]	0.06	0.06
	[gʁɑ̃ tami]	0.43	0.43
	[gʁɑ̃l ami]	0.00	0.00
	[gʁɑ̃ lami]	0.00	0.00
trente amis	[tʁɑ̃t ami]	1.00	1.00
	[tʁɑ̃ tami]	0.00	0.00
vrai tableau	[vʁɛt ablo]	0.00	0.00
	[vʁɛ tablo]	1.00	1.00

(b) Attested and predicted frequencies

Table 8: Results of the modeling study

3.2 A model of the phonetic ambiguity of French liaison

This section shows how the model proposed for the prosodic ambiguity of French in section 3.1 liaison can be extended to deal with phonetic ambiguity if some aspects of phonetic detail are integrated in the analysis (Steriade 2000). This section focuses on the interaction of liaison and affrication in Quebec French as this case study provides the clearest case of ambiguous realization for French liaison.[3]

The key ingredient in the analysis will be the observation that coarticulation is bidirectional, namely it affects both the realization of C and V in a CV sequence. Bidirectionality of coarticulation predicts that a change on C (e.g. affrication in the case of Quebec French) correlates with a change on the following V (e.g. vowel deletion/reduction in Quebec French). In combination with paradigm uniformity, this correlation will be crucial to explain why liaison consonants might pattern ambiguously between word-final and word-initial consonants phonetically. In a nutshell, CV-coarticulation at word boundaries will potentially result in violations of paradigm uniformity for both words in a Word1-Word2 sequence. However there will be less violations for a coarticulated liaison consonant than for a coarticulated word-final consonant due to the liaison consonant being absent from the corresponding citation form.

Section 3.2.1 provides some background on the bidirectionality of coarticulation, and shows how it applies in the case of Quebec French affrication. Building on these results, section 3.2.2 shows how the analysis derives citation forms for words with liaison consonants and stable word-final and word-initial consonants. Section 3.2.3 shows how the phonetic ambiguity of liaison consonants can be derived as a paradigm uniformity effect, assuming bidirectionality of coarticulation. Section 3.2.4 implements the analysis in a probabilistic grammar and shows that it provides a very good fit to the distribution of affrication in Côté (2014).

3.2.1 Bidirectionality of coarticulation in CV sequences

Probably the best studied case of coarticulation is the assimilation in second formant (F2) frequency between consonants and vowels (see Flemming 2001: 16-23). A large number of

---

[3]The analysis presented in this section supersedes the analysis presented in Storme (2020).

studies have shown that C assimilates to V in CV, in particular F2 at consonant release can be described as an increasing linear function of F2 in the middle of the vowel: as the F2 in the middle of the vowel increases, the F2 at consonant release also increases (Lindblom 1963; Sussman et al. 1991). In turn, V has also been found to assimilate to C in CV, with F2 in the middle in the vowel being higher when F2 at consonant release is higher (Lindblom 1963; Broad & Clermont 1987). These results suggest that coarticulation is bidirectional in CV sequences: both C and V are affected when the two sounds are combined in a CV sequence, and any change affecting one of the two sounds should also affect the other one.

Bidirectionality of coarticulation extends beyond this well studied case and applies in particular to Quebec French affrication. In Quebec French, affrication of /t d/ before high front vowels and glides applies almost categorically morpheme-internally (Côté 2014). Phonetically, affrication involves a change in consonant manner: the stop burst is followed by a frication noise (Stevens 1998: 412). But affrication before high vowels does not only affect the realization of the consonant. It also correlates with changes in the following vowel. In particular, Cedergren & Simoneau (1985: 72-80) report that high vowels tend to be reduced/deleted in the vicinity of fricatives, including before affricates. This effect is stronger with voiceless fricatives/affricates. Because /t/ maps to a voiceless affricate [ts] after affrication, high-vowel reduction is expected to be particularly common after this sound. In the remainder of this paper, [i̥] will be used to note this reduced/deleted high vowel. In other words, an underlying sequence /ti/ tends to be realized as [tsi̥] on the surface in Quebec French, with both affrication and high-vowel reduction.

This coarticulatory pattern involving fricatives/affricates and high vowels is found in other languages such as Japanese (Beckman & Shoji 1984; Whang 2018). For instance, Whang (2018: 1166) found a positive correlation between lengthening of [tʃ] (from an underlying /t/) and high-vowel devoicing in [tʃi] sequences in Japanese. This result suggests that the affricate gets more affricated (the frication noise gets longer) as the following high vowel devoices/reduces in Japanese, in line with what has been found in Quebec French. Based on such parallels, Cedergren & Simoneau (1985: 189) propose that this interaction stems from a universal phonetic constraint, but without providing more details. One possible mechanism relating the two changes is compensatory lengthening/shortening: there is a trading relationship between the duration of C and V such that if C lengthens then V shortens and conversely (see Whang 2018: 1160 and literature therein). The present paper will remain agnostic as to what the precise coarticulatory mechanism underlying this pattern is. In what follows, the constraint that drives the interaction between affrication and high-vowel reduction will be noted descriptively as *tsi. Assuming that affrication of /t/ to [ts] is independently motivated before [i] in a language (by a markedness constraint *ti), the constraint *tsi will favor a candidate [tsi̥] involving a concomitant change in vowel quality over a candidate [tsi] involving affrication but no high-vowel reduction.

3.2.2 Citation forms

To derive categorical affrication before high front vocoids morpheme-internally in Quebec French, the markedness constraint penalizing the non-affricated candidate (noted *ti) must outrank the faithfulness constraint penalizing a change in consonant continuancy (noted Ident$_{IO}$(cont)). Assuming that affrication is always accompanied by high-vowel reduction,

*tsi must outrank the faithfulness constraint penalizing vowel reduction (noted $\text{Ident}_{\text{IO}}(\text{voi})$). Table 9a illustrates how these ranking conditions predict that an underlying /ti/ sequence in *timide* /timid/ 'shy' should be realized with both affrication and high-vowel reduction.

UR:	/timid/	*ti	*tsi	$\text{Ident}_{\text{IO}}(\text{cont})$	$\text{Ident}_{\text{IO}}(\text{voi})$
a.	[timid]	!1			
b.	[tsimid]		!1	1	
c. ☞	[tsi̥mid]			1	1

(a) Words with initial /ti/ (*timide* 'shy')

UR:	/tʁɑ̃t/	*ti	*tsi	$\text{Ident}_{\text{IO}}(\text{cont})$	$\text{Ident}_{\text{IO}}(\text{voi})$
a. ☞	[tʁɑ̃t]				
b.	[tʁɑ̃ts]			!1	

(b) Words with stable final consonants (*trente* 'thirty')

UR:	/inosɑ̃/	*ti	*tsi	$\text{Ident}_{\text{IO}}(\text{cont})$	$\text{Ident}_{\text{IO}}(\text{voi})$
a. ☞	[inosɑ̃]				
b.	[i̥nosɑ̃]				!1

(c) Words with initial /i/ (*innocents* 'innocent')

Table 9: Constraint ranking for citation forms

Tables 9b and 9c show that the same ranking conditions predict that affrication and high-vowel reduction should not apply in the citation forms of words with stable word-final /t/ and words with word-initial /i/, respectively. Indeed, in these cases, the relevant markedness constraints are not violated and therefore nothing motivates any change in consonant continuancy or vowel quality on the surface.

Moreover, if the same constraint ranking $\text{Max}_{\text{IO}}(\text{C}) \gg *\text{C}\#$ is assumed as in section 3.1.1, liaison consonants are predicted to be categorically absent from citation forms whereas stable word-final/word-initial consonants are predicted to be categorically present in citation forms.

3.2.3 Contextual variants

At the boundary between two words, affrication will potentially result in changes in both Word1 and Word2, due to the bidirectional nature of coarticulation. But this will have different implications in terms of paradigm uniformity with the corresponding citation forms depending on the type of consonant, as shown in Table 10. The forms enclosed in boxes in Table 10 correspond to the citation forms of the two words. The liaison consonant differs from stable word-final/word-initial consonants in being absent from the corresponding citation form.

20

		Affrication	Citation form	Feature changes	Rate of affrication (Côté 2014)
(a)	Word-final C	tʁãts | i̥nosã	tʁãt | inosã	2	36.5%
(b)	Liaison C	gʁã | ts | i̥nosã	gʁã | inosã	1	66.0%
(c)	Word-initial C	gʁã | tsi̥mid	gʁã | tsi̥mid	0	99.2%

Table 10: How affrication affects the similarity with citation forms depending on the type of consonant

For liaison consonants, affrication implies a number of feature changes that is intermediary between the number of feature changes for stable word-final and word-initial consonants. Indeed, only the feature change affecting vowel quality at the beginning of Word2 ([i] - [i̥]) is penalized by paradigm uniformity. The change in consonant continuancy at the end of Word1 ([t] - [ts]) is not penalized by paradigm-uniformity constraints because the liaison consonant is missing from the corresponding citation form. For stable word-final consonants, affrication implies two feature changes relative to the corresponding citation form (one on the consonant at the end of Word1 and another one on the vowel at the beginning of Word2). For stable word-initial consonants, affrication does not imply any feature change relative to the corresponding citation form, since affrication already applies categorically in this form.

As shown in the last two columns of Table 10, the rate of affrication is inversely correlated with the number of feature changes implied by affrication across the three types of consonants. This can be understood as a paradigm uniformity effect: the grammar militates for uniformity between contextual and citation forms, resulting in less affrication for forms that imply more changes.

Table 11 shows how to derive the phonetic ambiguity of liaison consonants in Word1-Word2 sequences. Two new faithfulness constraints play a role in the analysis: $\text{Ident}_{OO}(\text{cont})$ and $\text{Ident}_{OO}(\text{voi})$. They correspond to $\text{Ident}_{IO}(\text{cont})$ and $\text{Ident}_{IO}(\text{voi})$ used in the analysis of morpheme-internal affrication, but in the output-output dimension: they penalize dissimilarities between contextual variants and the corresponding citation forms in terms of consonant continuancy and vowel quality, respectively.

URs:	/gʁɑ̃, gʁɑ̃t/ + /inosɑ̃/	Dep$_{IO}$(C)	*VV	*ti	*tsi	Ident$_{OO}$(cont)	Ident$_{OO}$(voi)
Citation forms:	[gʁɑ̃] + [inosɑ̃]						
a. %	gʁɑ̃inosɑ̃		1				
b. %	gʁɑ̃tinosɑ̃			1			
c. %	gʁɑ̃tsinosɑ̃				1		
d. %	gʁɑ̃tsinosɑ̃̊						1
e.	gʁɑ̃linosɑ̃	1!					

(a) Liaison words before /i/-initial words (*grand innocent* 'great innocent')

URs:	/tʁɑ̃t/ + /inosɑ̃/	Dep$_{IO}$(C)	*VV	*ti	*tsi	Ident$_{OO}$(cont)	Ident$_{OO}$(voi)
Citation forms:	[tʁɑ̃t] + [inosɑ̃]						
a. %	tʁɑ̃tinosɑ̃			1			
b. %	tʁɑ̃tsinosɑ̃				1	1	
c. %	tʁɑ̃tsinosɑ̃̊					1	1

(b) Words with stable final consonant before /i/-initial words (*trente innocents* 'thirty innocent (persons)')

URs:	/vʁɛ/ + /timid/	Dep$_{IO}$(C)	*VV	*ti	*tsi	Ident$_{OO}$(cont)	Ident$_{OO}$(voi)
Citation forms:	[vʁɛ] + [tsi̥mid]						
a.	vʁɛtimid			1		1	1
b.	vʁɛtsimid				1		1
c. ☞	vʁɛtsi̥mid						

(c) Words with initial /ti/ (*vrai timide* 'truly shy (person)')

Table 11: Constraint ranking for contextual variants

The analysis derives all three realizations attested in Côté (2014) for liaison words, as shown in Table 11a: absence of liaison consonant (candidate a), liaison without affrication (candidate b), and liaison with affrication (candidates c and d). Côté (2014) does not include a phonetic analysis, therefore it is not possible to determine whether vowel reduction always accompanies affrication (candidate d) or not (candidate c). A strict correlation between affrication and reduction could be derived by assuming that *tsi is high ranked.

The analysis also derives the two realizations attested in Côté (2014) for stable word-final consonants, as shown in Table 11b: blocking of affrication (candidate a) and regular application of affrication (candidates b and c). Moreover candidates with affrication are more penalized by faithfulness in the case of word-final consonants (candidates b and c in Table 11b) than in the case of liaison (candidates c and d in Table 11a), as discussed above: this means that affrication should be less likely for stable word-final consonants than for liaison consonants.

Finally, the analysis derives categorical affrication for stable word-initial consonants before /i/ in connected speech, as shown in Table 11c: the candidate with affrication and vowel reduction (candidate c) is better than the other candidates both in terms of markedness and uniformity with the citation form (where affrication and vowel reduction have already applied).

3.2.4 Modeling study: Côté (2014)

The analysis captures the ambiguous phonetic behavior of liaison consonants at a conceptual level. To test whether it can also match the specific rates of affrication attested in Quebec French, the grammatical model described in section 3.2.3 was fit to Côté's count data shown in Table 4, using the same methods as in the modeling of prosodic ambiguity. The file that was used for the analysis can be found in Storme (2022) under the name `phonetic-ambiguity.txt`.

Three assumptions were made in this analysis: (i) citation forms were assumed to be categorically realized as in section 3.2.2, (ii) affrication was assumed to be categorically associated with high-vowel reduction (i.e. candidates with [tsi] were assigned a null probability),[4] and (iii) the candidate with [l]-epenthesis in liaison words (candidate e in Table 11a) was assumed to be unattested.

Table 12a shows the weights inferred for each constraint. Table 12b shows that the frequencies predicted for each surface form match the frequencies attested in the corpus very well, therefore providing a quantitative argument for the paradigm-uniformity analysis.

Constraint	Weight
$Dep_{IO}(C)$	50.00
*VV	2.32
*ti	1.74
*tsi	50.00
$Ident_{OO}(cont)$	1.33
$Ident_{OO}(voi)$	1.02

(a) Constraint weights

	Candidates	Attested frequency	Predicted frequency
grand innocent	[gʁɑ̃inosɑ̃]	0.15	0.15
	[gʁɑ̃tinosɑ̃]	0.29	0.28
	[gʁɑ̃tsinosɑ̃]	0.00	0.00
	[gʁɑ̃tsi̥nosɑ̃]	0.56	0.57
	[gʁɑ̃linosɑ̃]	0.00	0.00
trente innocents	[tʁɑ̃tinosɑ̃]	0.64	0.65
	[tʁɑ̃tsinosɑ̃]	0.00	0.00
	[tʁɑ̃tsi̥nosɑ̃]	0.36	0.35
vrai timide	[vʁɛtimid]	0.01	0.02
	[vʁɛtsimid]	0.00	0.00
	[vʁɛtsi̥mid]	0.99	0.98

(b) Attested and predicted frequencies

Table 12: Results of the modeling study

4 Study 1: epenthetic and suppletive liaison

In the paradigm-uniformity analysis of French liaison, the crucial ingredient that explains the ambiguous patterning of liaison consonants is the fact that they are absent from the corresponding citation form. As pointed out by Steriade (1999), the paradigm-uniformity analysis predicts that liaison consonants that are present in citation forms should pattern unambiguously like word-final consonants. The goal of this section is to test this prediction experimentally.

Section 4.1 introduces the distinction between epenthetic liaison and suppletive liaison, explaining how this distinction provides a testing ground for the role of citation forms in

---

[4]This assumption is not crucial for the analysis. It makes it possible to simplify the analysis by having just one candidate corresponding to affrication ([tsi̥]) instead of two ([tsi̥] and [tsi]).

the realization of liaison consonants. Section 4.2 presents the methods used to test this prediction. Section 4.3 presents the results. Section 4.4 concludes with a brief discussion. The data and code for Study 1 are available in Storme (2022) under the names `study1-data.csv` and `study1-code.R`, respectively.

4.1 Epenthetic and suppletive liaison

Epenthetic liaison describes cases where the liaison variant contains the morphologically corresponding citation form as a substring, with the liaison consonant being epenthesized after this substring (e.g. $[\text{gʁɑ̃t}]_{\text{liaison variant}} = [\text{gʁɑ̃}]_{\text{citation form}} + [\text{t}]$). Suppletive liaison describes cases where the liaison variant does not contain the morphologically corresponding citation form as a substring but is based on a morphologically distinct form in the paradigm. For instance, the adjective *beau* [bo] 'beautiful.MASC' uses the form [bɛl] as a liaison variant (e.g. *bel ami* [bɛl#ami] 'beautiful friend.MASC'). This form cannot be analyzed as the masculine citation form plus an epenthetic consonant. Rather it corresponds to the feminine form of the adjective (*belle* [bɛl] 'beautiful.FEM').

The distinction between the two types of liaison is well-known and has been discussed by Delattre (1947: 150) and Tranel (1990, 2000) among others. Its relevance for the hypothesis of paradigm-uniformity effects has been first discussed by Steriade (1999). The paradigm-uniformity analysis predicts that only epenthetic liaison should pattern ambiguously between stable word-final and word-initial consonants. Indeed, for suppletive liaison, the liaison consonant is present at the end of the corresponding citation form (e.g. the [l] in *bel* [bɛl] is present at the end of the feminine citation form *belle* [bɛl]) and this word-final attachment should be enforced in contextual realizations by paradigm uniformity. Table 13 shows that the same constraint ranking that derived prosodic ambiguity for epenthetic liaison in Table 7a indeed predicts an unambiguously word-final behavior for suppletive liaison: candidate e (with liaison [l] attaching to Word2 across a prosodic break) is harmonically bounded by candidate d (with liaison [l] attaching to Word1). Epenthetic liaison (candidates b and c) is also ruled out because /t/ is not present in any listed allomorph for the adjective *beau*.[5]

URs:	/bo, bɛl$_{\text{FEM}}$/ + /ami/	Dep$_{\text{IO}}$(C)	*VV	L-Anchor$_{\text{OO}}$	R-Anchor$_{\text{OO}}$
Citation forms:	[bo, bɛl$_{\text{FEM}}$] + [ami]				
a.	bo ami		1		
b.	bot ami	1!			1
c.	bo tami	1!		1	
d. ☞	bɛl ami				
e.	bɛ lami			1!	1

Table 13: Suppletive liaison consonants are predicted to behave unambiguously like word-final consonants and attach to Word1 (*bel ami* 'beautiful friend').

---

[5]In Table 13, no constraint favors the candidate without liaison (candidate a) over the candidate with suppletive liaison (candidate d). This contrasts with the situation in Table 7a, where epenthetic liaison was penalized by paradigm-uniformity constraints. However the candidate without liaison (candidate a in Table 13) could be favored over the candidate with suppletive liaison (candidate d) by a constraint penalizing suppletion, e.g. a constraint requiring that a masculine allomorph be used in a grammatically masculine context (see Steriade 1999).

There is preliminary evidence for the prediction that epenthetic liaison and suppletive liaison differ in this way, as pointed out by Steriade (1999). In right-dislocation contexts, Tranel (1990) reports that epenthetic liaison consonants attach to Word2 whereas suppletive liaison consonants attach to Word1, as illustrated in (8).

(8)   Epenthetic liaison vs. suppletive liaison in right dislocations
     a. Epenthetic liaison   J'en ai un grand, éléphant.   [gʁɑ̃ <u>t</u>elefɑ̃]
                                      'I have a big one, elephant.'
     b. Suppletive liaison   J'en ai un bel, éléphant.   [bɛ<u>l</u> elefɑ̃]
                                        'I have a beautiful one, elephant.'

However, right-dislocation contexts are probably not the most appropriate context to make a case for the ambiguous patterning of epenthetic liaison, as they seem to very strongly favor a prosodic attachment to Word2 (=*liaison enchaînée*). As noted by Durand & Lyche (2008: 50), the contexts where epenthetic liaison consonants are more readily found to attach to Word1 (= *liaison non-enchaînée*) involve a hesitation between Word1 and Word2 (see Section 2.2.1).[6] The goal of Study 1 is therefore to compare the behavior of liaison consonants (both epenthetic and suppletive) and stable consonants (both word-final and word-initial) across a prosodic break involving a hesitation.

4.2   Methods

Adjective-noun (Adj-N) sequences were chosen as Word1-Word2 sequences. This choice was motivated by the fact that both epenthetic liaison (e.g. *grand*) and suppletive liaison (e.g. *beau/bel*) can be found among adjectives. Each of the four experimental conditions (epenthetic liaison, suppletive liaison, stable word-final consonants, stable word-initial consonants) was represented by 12 Adj-N sequences, for a total of 48 Adj-N sequences. Six adjectives were used by condition, as shown in Table 14, and each adjective appeared in two Adj-N sequences varying by the strength of their collocation. For instance, *petit* appeared both in *petit ami* 'boyfriend' (more frequent) and in *petit anneau* 'small ring' (less frequent).[7] This manipulation was meant to control for potential effects of the following noun on the behavior of liaison consonants, as this variable has been shown to influence some aspects of French liaison in previous research (see Section 2.1).

---

[6]In the paradigm-uniformity analysis, the difference between right-dislocation contexts and hesitations could be captured by indexing paradigm-uniformity constraints (e.g. R(ight)-Anchor$_{OO}$) to prosodic domains. In right dislocations such as (8), the adjective (Word1) is under focus. When there is a hesitation between Word1 and Word2, Word1 is not necessarily under focus. If the requirement to be similar to the citation form is more strongly enforced under focus, then this might explain why epenthetic liaison consonants are more likely to be pushed onto Word2 in right dislocations.

[7]The strength of the collocation was measured as the conditional probability of N given Adj in the corpus of movie and TV subtitles OpenSubtitles (Lison & Tiedemann 2016).

Condition	Adjectives
Word-final C	énorme, jeune, large, magnifique, meilleur, superbe
Epenthetic liaison	faux, grand, gros, mauvais, parfait, petit
Suppletive liaison	ancien, bel, bon, prochain, vieil, nouvel
Word-initial C	affreux, charmant, gentil, joli, long, vrai

Table 14: Adjectives used in Study 1

A French native speaker (the author) read each of the 48 Adj-N sequences twice, with a hesitation (*euh* [œ]) between the two words. The two pronunciations varied in the prosodic attachment of the consonant between the two words. In one pronunciation, the consonant was pronounced at the end of Word1 before the hesitation. This corresponds to a case of *liaison non-enchaînée* for liaison conditions. In the other pronunciation, the consonant was pronounced at the beginning of Word2 after the hesitation. This corresponds to a case of *liaison enchaînée*. Examples are shown in Table 15 for each of the four experimental conditions.

| | Pronunciation | | |
	Word1 attachment	Word2 attachment	
Stable word-final C	magnifi[k] euh hôtel	magnifi euh [k]hôtel	'magnificent hotel'
Epenthetic liaison	grand[t] euh hommage	grand euh [t]hommage	'great tribute'
Suppletive liaison	be[l] euh appartement	be euh [l]appartement	'beautiful apartment'
Stable word-initial C	joli[s] euh ourire	joli euh [s]ourire	'nice smile'

Table 15: Experimental items

Twenty-three Swiss French speakers were recruited among university students to participate in an online study. The 48 Adj-N sequences were presented to participants in random order. For each Adj-N sequence, the two pronunciations were presented one after the other, with the pronunciation involving prosodic attachment to Word1 always preceding the pronunciation involving prosodic attachment to Word2. Participants were asked to indicate which of the two pronunciations sounded more natural to them. The target Adj-N sequence was not presented graphically to participant in order to avoid any explicit orthographic bias. Liaison consonants (both epenthetic and suppletive) appear at the end of Word1 in the spelling and this could directly bias participants towards a word-final attachment. Participants were invited to wear headphones while taking the study. The LimeSurvey platform (LimeSurvey 2012) was used to carry out the online study. The participants provided their informed consent to participate in the research and agreed to make their data available online. No sensitive information about participants was collected.

A Bayesian hierarchical logistic regression was fit to the participants' responses as a function of the dummy-coded factor Consonant (reference level 'stable word-final consonant'). Consonant has four levels, corresponding to the four types of consonants (stable word-final consonant, suppletive liaison, epenthetic liaison, stable word-initial consonant). The random effect structure included a random intercept for each participant, a by-participant random slope for the effect of Consonant, and a random intercept for each Adj-N sequence. The logistic regression was fit using the `brms` package (Bürkner 2017) in R (R Core Team 2020). For
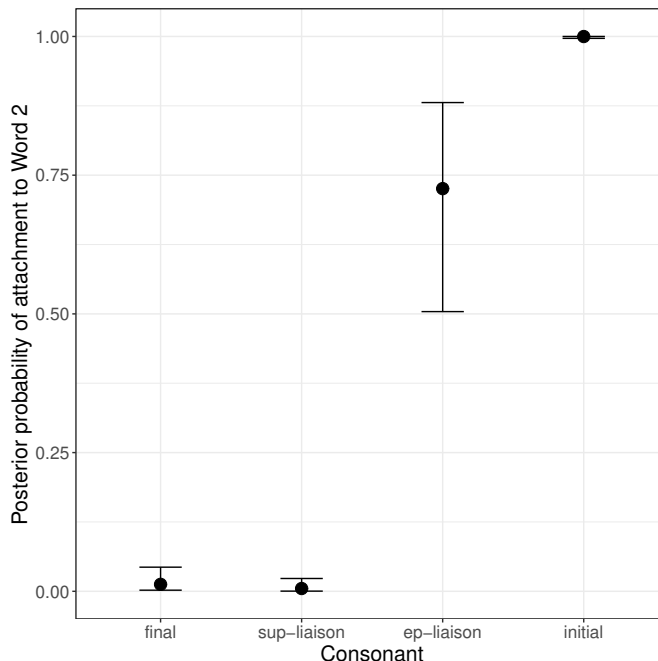
Figure 2: Posterior probability of attachment to Word2 as a function of consonant type (mean and 95% CI)

hypothesis testing, the difference Δ in the posterior log-odds ratios of attachment to Word2 was computed for the two relevant conditions (e.g. epenthetic liaison vs. stable word-final consonants). Compelling evidence for a difference between two conditions was considered to be provided only in case zero was outside of the posterior 95% Credible Interval (CI) for Δ. Credible Intervals were obtained using the ETI (Equal-tailed Interval) method and the package `bayestestR` (Makowski et al. 2019).

4.3 Results

As predicted by the paradigm-uniformity analysis, suppletive liaison and epenthetic liaison pattern differently (see Figure 2). Suppletive liaison behaves like stable word-final consonants ($\Delta_{\text{sup liaison - final}} = -1.01$, $CI = [-3.62, 0.73]$), favoring an attachment to Word1 almost categorically. Epenthetic liaison has an intermediary rate of attachment to Word2 between stable word-final consonants ($\Delta_{\text{ep liaison - final}} = 5.41$, $CI = [3.81, 7.36]$) and word-initial consonants ($\Delta_{\text{ep liaison - initial}} = -8.92$, $CI = [-18.61, -4.59]$).

To make sure that ambiguity of epenthetic liaison does not just result from averaging across participants and Adj-N sequences, exploratory analyses of individual variation (Figure 3a) and lexical variation (Figure 3b) were conducted. Inspection of Figures 3a and 3b reveals that there is variation in the treatment of epenthetic liaison across participants and Adj-N sequences. However epenthetic liaison is generally treated ambiguously by individual participants (except maybe participants 10 and 11 who are close to categorical). Also, epenthetic liaison is treated ambiguously in each Adj-N sequence taken individually: for each sequence, the mean is distinct from 0 and 1 and the 95% CI includes neither 0 nor 1.

In other words, epenthetic liaison is genuinely treated as ambiguous by most participants and across all Adj-N sequences.

4.4 Discussion

An important prediction of the paradigm-uniformity was corroborated by the results of Study 1: liaison behaves ambiguously if the liaison consonant can be analyzed as epenthesized at the end of the citation form but not if the liaison variant is suppletive. This study is to the author's knowledge the first controlled study that establishes this difference between epenthetic and suppletive liaison.

The results of Study 1 also show that, although epenthetic liaison is ambiguous between stable word-final and word-initial consonants, it is more often treated like a word-initial consonant (see Figure 2). This is in line with the results of previous research according to which *liaison enchaînée* is more likely than *liaison non-enchaînée* (see Section 2.2.1). However the results focusing on individual variation also show that *liaison non-enchaînée* is preferred by some speakers (Participants 9, 8 and 12) and is more likely overall for one Adj-N sequence (*parfait inconnu*), at least in the specific context elicited in this study. This variability could be captured by the grammatical model proposed in Section 3.1 if speaker-specific and sequence-specific weights were set for the two paradigm-uniformity constraints.

5 Study 2 : liaison and affrication in Quebec French

The key ingredient in the analysis of the phonetic ambiguity of French liaison proposed in section 3.2 was the bidirectionality of coarticulation. For Quebec French specifically, bidirectionality of coarticulation means that affrication of /t/ correlates with a reduction of /i/ in /ti/ sequences. The goal of this section is to test whether affrication does indeed correlate with vowel reduction across a word boundary, for both stable word-final consonants and liaison consonants. Section 5.1 presents the methods used to test the hypothesis. Section 5.2 presents the results. Section 5.3 concludes with a brief discussion. The data and code for Study 2 are available in Storme (2022) under the names `study2-data.csv` and `study2-code.R`, respectively.

5.1 Methods

Data from the Quebec PFC project (Côté 2016) were used to investigate this question. The analysis focuses on two quasi-minimal pairs from the PFC word lists that feature an underlying sequence /ti/ at the boundary between two words: *grand innocent* 'great innocent' (with liaison /t/) and *trente innocents* 'thirty innocent (people)' (with stable word-final /t/). These data are particularly interesting because they make it possible to test both whether affrication and high-vowel reduction are correlated and how this correlation might differ for liaison consonants and stable word-final consonants. The analysis does not focus on word-initial /ti/ sequences as the PFC word lists do not include minimal pairs allowing for a controlled comparison with liaison and word-final consonants.
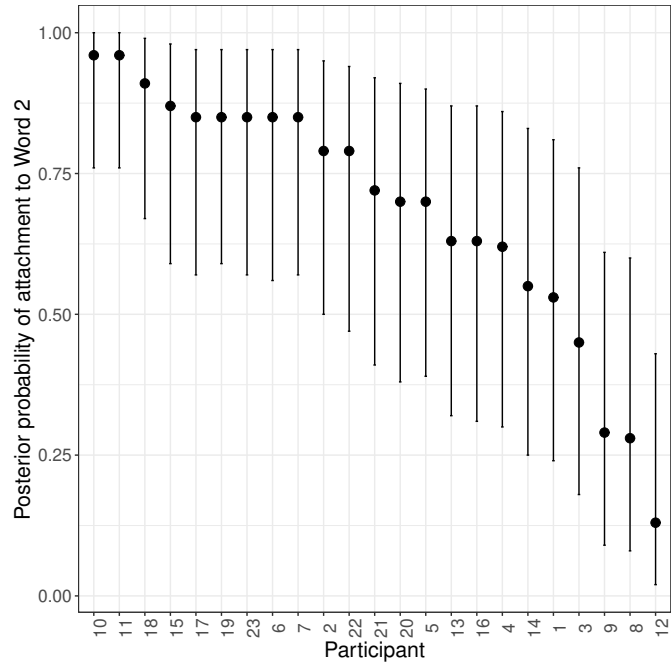
(a) Individual variation
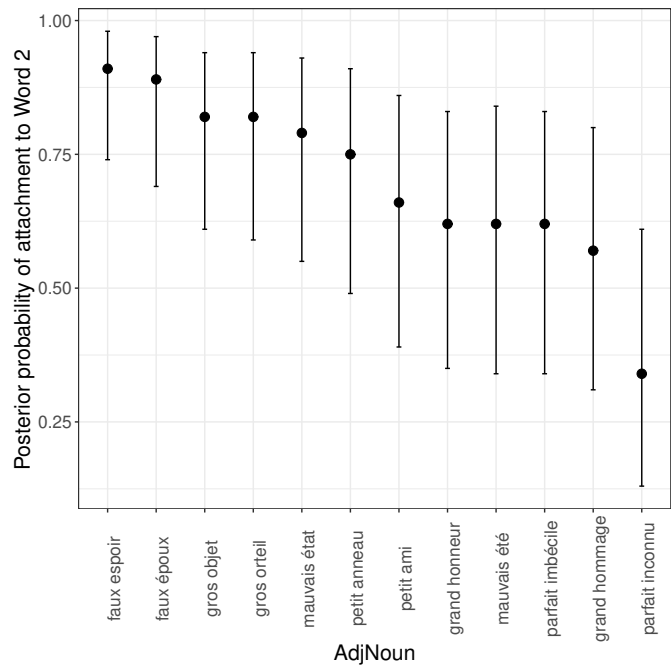
(b) Lexical variation

Figure 3: Individual variation and lexical variation in the treatment of epenthetic liaison

The data from all locations available in the corpus in 2021 were selected, corresponding to a total of 394 participants.[8] Annotations were done manually in Praat (Boersma & Weenink 2021). /t/ duration was used as acoustic correlate for affrication (an underlying /t/ that is affricated on the surface should be longer than an underlying /t/ that is not affricated). The duration of /t/ included the burst and/or frication noise, following Whang (2018: 1163). Vowel reduction was also annotated, using the presence of formant structure as a criterion. In the absence of clear formant structure, no vowel /i/ was included on the corresponding tier. This does not mean that the vowel is completely absent phonetically as phonetic reflexes of /i/ could be present in the burst or frication noise of /t/. Pauses and schwas that sometimes occurred between /t/ and /i/ were also annotated, as well as cases of non-conventional consonant realizations (for instance, some participants pronounced a [z] between *trente* and *innocents*) and cases where the liaison consonant was absent (in these cases, no consonant was annotated on the corresponding tier). Segment durations were extracted automatically using a Praat script.

Only sequences that involve a [t] on the surface (affricated or not) and no pause between the consonant and the vowel were included in the final analyses, corresponding to a total of 322 participants and 494 occurrences of consonants (243 liaison consonants and 251 stable word-final consonants). Two statistical analyses were conducted. A Bayesian logistic regression was fit to the data using `brms` in R, with Vowel (present, absent) as as dependent variable and Consonant (liaison, final), Consonant duration and their interaction as independent variables. The goal of this first analysis was to test whether vowel deletion correlates with lengthening of /t/, as expected under the hypothesis that affrication results in high-vowel deletion/reduction. A Bayesian linear regression was also fit to the data, with Consonant duration as dependent variable and Consonant (liaison, final) as independent variable. The goal of this second analysis was to test whether liaison /t/ is phonetically longer than stable word-final /t/. A greater duration for liaison /t/ is expected if liaison /t/ is more affricated than stable word-final /t/, as reported by Côté (2014) on a perceptive basis, and if liaison consonants are generally longer than stable word-final consonants (see section 2.2.2). The analyses did not include random effects because there was at most one occurrence of each type of consonant (liaison, final) per speaker.

5.2 Results

The results of the logistic regression confirm the hypothesis that /t/-lengthening correlates with a higher likelihood of /i/-deletion, as shown in Figure 4. An increase of 1 ms in /t/ duration corresponds to a decrease of 0.08 unit ($CI = [0.06, 0.11]$) in the posterior log-odds ratio of /i/-presence. This result was found to hold for both liaison and stable word-final consonants, as the interaction term between duration and consonant type was not significantly different from zero ($\beta = 0.02, CI = [-0.01, 0.05]$). Moreover, liaison /t/ was found to favor /i/-deletion more than than word-final /t/ ($\beta = -2.25, CI = [-4.31, -0.27]$), independently from the effect of duration.

The results of the linear regression show that liaison /t/ is longer on average than word-final /t/ ($\beta = 11.79, CI = [7.66, 16.01]$), as shown in Figure 5. This lengthening corresponds

---

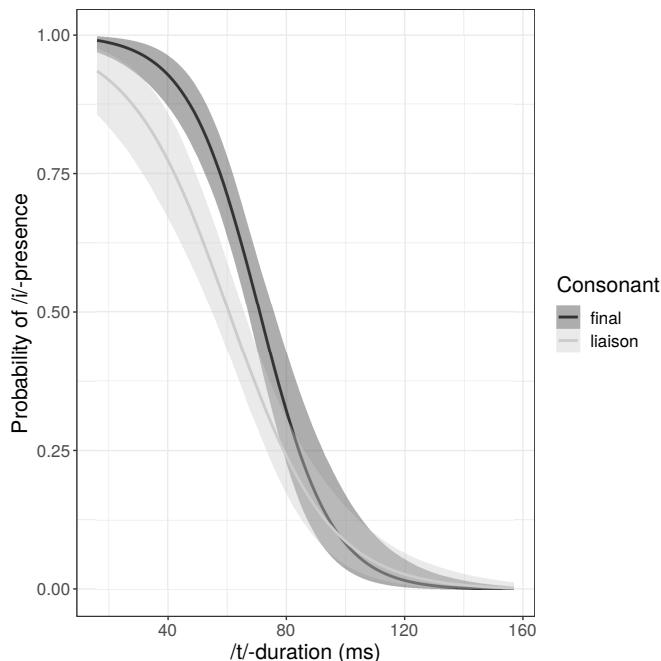[8]I am grateful to Marie-Hélène for making the data available to me.

Figure 4: Posterior probability of /i/-deletion as a function of /t/-duration and consonant type (liaison, final)

to an increase of 19% in duration. This is compatible with the observation in Côté (2014) that liaison /t/ is more affricated than stable word-final /t/ on average. This is also compatible with earlier observations about the relative duration of liaison consonants and stable word-final consonants more generally.

5.3 Discussion

The results of Study 2 support a key hypothesis of the paradigm-uniformity account of the phonetic ambiguity of French liaison: a change in C correlates with a change in V in CV sequences. More specifically, the results provide evidence for the hypothesis that affrication at a word boundary results in reduction/deletion of the following vowel. This hypothesis was crucial to explain why the rate of affrication of liaison /t/ is intermediary between word-final /t/ and word-initial /t/. Moreover, the results also support the hypothesis that liaison /t/ is more prone to affricate than stable word-final /t/. In the paradigm-uniformity analysis, this follows from the effect of the corresponding citation form. Stable word-final /t/ is influenced by the corresponding unaffricated [t] in the citation form. Liaison /t/ does not correspond to any [t] in the citation form and therefore is less likely to resist affrication.

6 Conclusion

Liaison consonants have been shown in previous research to pattern ambiguously between stable word-final and word-initial consonants. The present paper has shown that it is not necessary to attribute this behavior to differences in the phonological underlying status
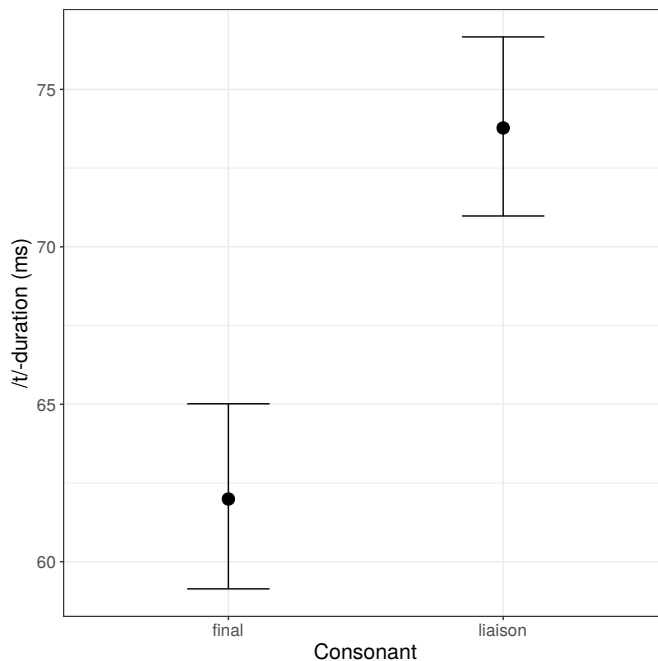
Figure 5: Posterior distribution of consonant duration (ms) for stable word-final /t/ and liaison /t/ (mean and 95% CI)

of liaison consonants. Rather it can be derived 'for free' from the observation that liaison words come under two variants (with and without liaison) and from independently motivated principles of uniformity among paradigmatically related forms (contextual variants of a word and the corresponding citation form). Also, the ambiguous behavior of liaison consonants can be derived without positing lexical constructions or massive allomorphy in the lexicon. It is sufficient to assume that only liaison words have two listed allomorphs. An explicit implementation of the analysis in a probabilistic constraint-based grammar was proposed and shown to be able to derive both prosodic ambiguity and phonetic ambiguity of French liaison. Crucially, the analysis assumed standard lexical and phonological representations as inputs.

Quantitative evidence was provided for two important hypotheses of the paradigm uniformity analysis. Study 1 showed that liaison consonants are not ambiguous in themselves but only if they are absent from the corresponding citation form, thus making a clear argument for the role of paradigm uniformity with citation forms. Study 2 provided evidence for the phonetic mechanism that underlies the paradigm-uniformity analysis of the phonetic ambiguity of liaison in Quebec French. Affrication of /t/ was found to correlate with a higher likelihood of high-vowel reduction in Quebec French, for both liaison and stable word-final consonants. This result is in line with the hypothesis that affrication at word boundaries has consequences for uniformity with the citation forms of both Word1 (through a change affecting word-final /t/) and Word2 (through a change affecting word-initial /i/). This hypothesis was key to explain why liaison /t/ is more likely to affricate than final /t/ before /i/.

The analysis presented in this paper focused on deriving the ambiguous realization of

French liaison. But a comprehensive model of French liaison should include additional elements in order to account for the role of other linguistic and sociolinguistic variables (e.g. syntactic and lexical properties of the words involved in two-word sequences, speech register, speaker identity, etc.). Combining the kind of paradigm-uniformity model presented in this paper with a model of phonological encoding based on the Production Planning Hypothesis (Kilbourn-Ceron 2017) looks like a promising avenue for future research on these topics.

References

Bagou, Odile, Violaine Michel & Marina Laganaro. 2009. On the production of sandhi phenomena in French: Psycholinguistic and acoustic data. In *Tenth Annual Conference of the International Speech Communication Association*, 452–455.

Dejean de la Bâtie, Bernadette & Dianne C Bradley. 1995. Resolving word boundaries in spoken French: Native and non-native strategies. *Applied Psycholinguistics* 16:59–81.

Beckman, Mary & Atsuko Shoji. 1984. Spectral and perceptual evidence for CV coarticulation in devoiced /si/ and /syu/ in Japanese. *Phonetica* 41:61–71.

Benua, Laure. 1997. Transderivational identity: phonological relations between words. Doctoral Dissertation, University of Massachusetts, Amherst.

Boersma, Paul & David Weenink. 2021. Praat: doing phonetics by computer [computer program]. version 6.1.41. URL http://www.praat.org/, retrieved 25 March 2021.

Broad, David J & Frantz Clermont. 1987. A methodology for modeling vowel formant contours in CVC context. *Journal of the Acoustical Society of America* 81:155–165.

Bürkner, Paul-Christian. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80:1–28.

Bybee, Joan L. 2001. Frequency effects on French liaison. In *Frequency and the emergence of linguistic structure*, ed. Joan L. Bybee & Paul J. Hopper, 337–359. Amsterdam: John Benjamins Publishing Company.

Cedergren, Henrietta J. & Louise Simoneau. 1985. La chute des voyelles hautes en français de Montréal: "As-tu entendu la belle syncope?". In *Les tendances dynamiques du français parlé à Montréal*, ed. Monique Lemieux & Henrietta J. Cedergren, volume 1, 57–145. Quebec: Office de la langue française.

Chevrot, Jean-Pierre, Céline Dugua & Michel Fayol. 2009. Liaison acquisition, word segmentation and construction in French: A usage-based account. *Journal of Child Language* 36:557–596.

Côté, Marie-Hélène. 2005. Le statut lexical des consonnes de liaison. *Langages* 158:66–78.

Côté, Marie-Hélène. 2011. French liaison. In *The Blackwell Companion to Phonology*, ed. Marc van Oostendorp, Colin J. Ewen, Elizabeth V. Hume, & Keren Rice, 1–26. Chicester, UK: John Wiley and Sons Ltd.

Côté, Marie-Hélène. 2014. Liaison et assibilation en français laurentien. In *La liaison: approches contemporaines*, ed. Christiane Soum-Favaro, Annelise Coquillon, & Jean-Pierre Chevrot, 9–32. Berne: Peter Lang.

Côté, Marie-Hélène. 2016. Variation in Canada: Trois-Rivières in Quebec. In *Varieties of spoken French*, ed. Sylvain Detey, Jacques Durand, Bernard Laks, & Chantal Lyche, 449–462. Oxford University Press.

Côté, Marie-Hélène. 2017. La liaison en diatopie: esquisse d'une typologie. *Journal of French Language Studies* 27:13–25.

De Cat, Cécile. 2007. *French dislocation. interpretation, syntax, acquisition*. Oxford: Oxford University Press.

Delattre, Pierre C. 1947. La liaison en français, tendances et classification. *The French Review* 21:148–157.

Durand, Jacques & Chantal Lyche. 2008. French liaison in the light of corpus data. *Journal of French Language Studies* 18:33–66.

Durand, Marguerite. 1936. *Le genre grammatical en français parlé Paris et en région parisienne*. Paris: Bibliothèque du Français Moderne.

Encrevé, Pierre. 1988. *La liaison avec et sans enchaînement: phonologie tridimensionnelle et usages du français*. Paris: Seuil.

Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18:7–44.

Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In *Phonetically based phonology*, ed. Bruce Hayes, Robert Kirchner, & Donca Steriade, 232–276. Cambridge: Cambridge University Press.

Fougeron, Cécile. 2007. Word boundaries and contrast neutralization in the case of enchaînement in French. In *Papers in Laboratory Phonology 9: Change in Phonology*, ed. Jennifer Cole & José Ignacio Hualde, 609–642. Berlin: Mouton de Gruyter.

Fougeron, Cécile & Elisabeth Delais-Roussarie. 2004. Liaisons et enchaînements «fais_en à fez_en parlant». *Actes des Journées d'Etudes sur la Parole* 221–224.

Fougeron, Cécile, Jean-Philippe Goldman, Alicia Dart, Laurence Guélat & Clémentine Jeager. 2001. Influence de facteurs stylistiques, syntaxiques et lexicaux sur la réalisation de la liaison en français. In *Actes de la 8ème conférence sur le Traitement Automatique des Langues Naturelles. Articles longs*, 172–181. Tours, France: ATALA.

Gaatone, David. 1978. Forme sous-jacente unique ou liste d'allomorphes? (à propos des consonnes de liaison en français). *Linguistics* 214:33–54.

Gaskell, M. Gareth, Elsa Spinelli & Fanny Meunier. 2002. Perception of resyllabification in french. *Memory & Cognition* 30:798–810.

Hayes, Bruce, Bruce Tesar & Kie Zuraw. 2013. OTSoft 2.5. URL http://www.linguistics.ucla.edu/people/hayes/otsoft/.

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:379–440.

Jun, Jongho. 2004. Place assimilation. In *Phonetically based phonology*, ed. Bruce Hayes, Robert Kirchner, & Donca Steriade, 58–86. Cambridge: Cambridge University Press.

Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.

Kilbourn-Ceron, Oriana. 2017. Speech production planning affects variation in external sandhi. Doctoral Dissertation, McGill University.

LimeSurvey. 2012. LimeSurvey: An open source survey tool. URL http://www.limesurvey.org.

Lindblom, Björn. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35:1773–1781.

Lison, Pierre & Jörg Tiedemann. 2016. OpenSubtitles2016: Extracting large parallel corpora from movie and TV subtitles. In *Proceedings of the 10th International Conference on Language Resources and Evaluation*. European Language Resources Association.

Makowski, Dominique, Mattan S. Ben-Shachar & Daniel Lüdecke. 2019. bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software* 4:1541.

Nguyen, Noël, Sophie Wauquier-Gravelines, Leonardo Lancia & Betty Tuller. 2007. Detection of liaison consonants in speech processing in French: Experimental data and theoretical implications. In *Segmental and prosodic issues in Romance phonology*, ed. Pilar Prieto, Joan Mascaró, & Marie-Josep Solé, 3–23. Amsterdam/Philadelphia: John Benjamins Publishing Company.

R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Roettger, Timo B., Bodo Winter, Sven Grawunder, James Kirby & Martine Grice. 2014. Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics* 43:11–25.

Smolensky, Paul & Matthew Goldrick. 2016. Gradient symbolic representations in grammar: The case of French liaison. *Ms. Johns Hopkins University and Northwestern University. Available as ROA* 1286.

Spinelli, Elsa, Anne Cutler & James M. McQueen. 2002. Resolution of liaison for lexical access in French. *Revue française de linguistique appliquée* 7:83–96.

Spinelli, Elsa, James M. McQueen & Anne Cutler. 2003. Processing resyllabified words in French. *Journal of Memory and Language* 48:233–254.

Steriade, Donca. 1999. Lexical conservatism in French adjectival liaison. In *Formal perspectives in Romance linguistics. Selected papers from the 28th linguistic symposium on Romance languages*, ed. Jean-Marc Authier, Barbara E. Bullock, & Lisa A. Reed, 243–270. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, ed. Michael B. Broe & Janet B. Pierrehumbert. Cambridge: Cambridge University Press.

Stevens, Kenneth N. 1998. *Acoustic phonetics*. Cambridge, MA/London: The MIT Press.

Storme, Benjamin. 2020. Gradient behavior without gradient underlying representations: the case of French liaison. In *Proceedings of the 2019 Annual Meeting on Phonology*, ed. Hyunah Baek, Chikako Takahashi, & Alex Hong-Lun Yeung. Washington, DC: Linguistic Society of America.

Storme, Benjamin. 2022. Paradigm uniformity effects on French liaison. [Data set]. URL https://doi.org/10.5281/zenodo.5987806.

Sussman, Harvey M., Helen A. McCaffrey & Sandra A. Matthews. 1991. An investigation of locus equations as a source of relational invariance for stop place categorization. *Journal of the Acoustical Society of America* 90:1309–1325.

Tanner, James, Morgan Sonderegger & Michael Wagner. 2017. Production planning and coronal stop deletion in spontaneous speech. *Laboratory Phonology* 8.

Tranel, Bernard. 1990. On suppletion and French liaison. *Probus* 2:87–104.

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

Wagner, Michael. 2012. Locality in phonology and production planning. *McGill Working Papers in Linguistics* 22:1–18.

Wauquier-Gravelines, Sophie. 1996. Organisation phonologique et traitement de la parole continue. Doctoral Dissertation, Université Paris 7.

Whang, James. 2018. Recoverability-driven coarticulation: Acoustic evidence from Japanese high vowel devoicing. *Journal of the Acoustical Society of America* 143:1159–1172.

Zuraw, Kie & Bruce Hayes. 2017. Intersecting constraint families: an argument for Harmonic Grammar. *Language* 93:497–548.