Evidence for suffix cohesion in French

Benjamin Storme (Université de Lausanne)

Abstract

In a language, suffix cohesion refers to the fact that suffixed words behave phonologically as simple or complex units depending on the suffix they are built with. This paper uncovers a previously undescribed pattern of suffix cohesion in French, where words suffixed with vowel- and glide-initial suffixes behave phonologically like simple units (e.g. *fêtiez* [fet-je] 'you partied') and words built with other suffixes behave phonologically like complex units (e.g. *fêterez* [fɛt-ʁe] 'you will party'). The evidence comes from a reassessment of well-known data on [ə]-[ɛ] stem alternations and from an acoustic study of [e]-[ɛ] and [o]-[ɔ] stem alternations in suffixed words as pronounced by 10 speakers living in the Paris area. The suffix's phonological shape is found to provide the best account of the data among a set of factors that have been argued to be relevant to suffix cohesion in previous works on other languages (resyllabification, suffix's morphological function, and relative frequency of the base with respect to the suffixed word).

1. Introduction

Some languages distinguish two types of suffixes depending on how the suffixed word behaves phonologically: cohering and non-cohering suffixes (Dixon, 1977; van Oostendorp, 2004; Haspelmath and Sims 2010; Raffelsiefen, 2015; Plag 2018). With cohering suffixes, the shape of the suffixed word is determined by the regular phonotactics of the language: the suffixed word behaves phonologically like a monomorphemic word. English suffix *-al* provides an example of cohering suffix: words suffixed with *-al* behave like monomorphemic words with respect to the position of stress in the word. For instance, stress falls on the heavy penult in *parént-al*, as expected phonotactically (compare to monomorphemic *uténsil).* With non-cohering suffixes, the shape of the suffixed word is determined by paradigm uniformity: the stem in the suffixed word behaves phonologically like the base, i.e. the

word-form containing the stem without any affix. English suffix *-hood* provides an example of non-cohering suffix. For instance, in *párent-hood,* stress falls on the stem-initial syllable, as expected under paradigm uniformity with the base *párent* and in violation of regular stress phonotactics, which predicts **parént-hood*.

The distinction between cohering and non-cohering suffixes often correlates with phonological properties of the relevant suffixes. For instance, in English and Dutch, cohering suffixes tend to be vowel-initial and non-cohering suffixes consonant-initial (e.g. English *-al* vs. *-hood*; see Raffelsiefen, 2015:901 on English; van Oostendorp, 2004:1376 on Dutch). In Yidiny derivational and inflectional paradigms, cohering suffixes are monosyllabic and non-cohering suffixes are disyllabic (Dixon, 1977:27).

This paper contributes to enriching this typology by uncovering a previously undescribed pattern of suffix cohesion in French. More specifically, the paper provides evidence that, in French, vowel- and glide-initial suffixes are cohering, as illustrated in (1a),[1] while other suffixes are non-cohering, as illustrated in (1b). The evidence comes from a reassessment of well-known data on [ə]-[ɛ] stem alternations and from an acoustic study of [e]-[ɛ] and [o]-[ɔ] stem alternations in suffixed words as pronounced by 10 speakers living in the Paris area. The suffix's phonological shape is found to provide the best account of the data among a set of factors that have been argued to play a role in suffix cohesion in previous works on other languages (resyllabification, suffix's morphological function, and relative frequency of the base with respect to the suffixed word).

(1) Suffix cohesion in French

a. *fêtiez* 'party-IMP.2PL' [fet-je] (form predicted by phonotactics)

*[fɛt-je](form predicted by uniformity with *fête* [fɛt] 'party.PRES.3SG')

1 IMP: imperfect; FUT: future; PRES: present; 2PL/SG: second person plural/singular; 3SG: third person singular.

b. *fêterez* ‘party-FUT.2PL’ *[fet-ʁe] (form predicted by phonotactics)

[fɛt-ʁe] (form predicted by uniformity with *fête* [fɛt] ‘party.PRES.3SG’

Section 2 provides preliminary evidence for the hypothesis that French distinguishes cohering and non-cohering suffixes and that this distinction can be explained in terms of the suffix-initial segment’s phonological category. Section 3 presents the study that was conducted to further test this hypothesis. Section 4 describes the methods. Section 5 reports the study’s results. Section 6 proposes an interpretation of suffix cohesion in French, using the theory and results presented in Storme (2017c).

2. Background

2.1. Cohering vs. non-cohering suffixes in French: preliminary data and hypothesis

Data on [ə]-[ɛ] stem alternations in suffixed words (e.g. Dell, 1985:198-214) point to a pattern of suffix cohesion in French, where the phonological shape of the suffix determines whether the suffixed word is phonotactically regular or paradigmatically uniform. The relevant data are shown in (2), using forms of the inflectional paradigm of *jeter* ‘throw’ to illustrate. In the paradigm of *jeter*, [ə] (a close-mid vowel) alternates with [ɛ] (an open-mid vowel) in stems (see (2b,c) vs. (2a,d,e)). [ə] is treated as a close-mid vowel because it patterns as such phonotactically and phonetically: it is banned in closed syllables, like the close-mid vowel [e] (Dell, 1985; Walker, 1993), and it is similar to the close-mid vowel [ø] acoustically, at least in the variety of European French that the present study focuses on, i.e. the standard variety spoken in Paris (Malécot and Chollet, 1977; Fougeron et al 2007). Note that, in (2e), the word-medial schwa is not part of the suffix or the stem underlyingly: it is epenthetic and motivated by the need to avoid the triconsonantal cluster [tʁj] (see Dell, 1978 on schwa epenthesis at morpheme boundaries).

(2) [ə]-[ɛ] stem-alternations in French

a. *il jette*	[ʒɛt]	_C#	‘he throws’
b. *vous jetez*	[ʒət-e]	_CV	‘you.PL throw’
c. *vous jetiez*	[ʒət-je]	_CGV	‘you.PL threw’
d. *vous jetterez*	[ʒɛt-ʁe]/*[ʒət-ʁe]	_CLV	‘you.PL will throw’
e. *vous jetteriez*	[ʒɛt-ə-ʁje]/*[ʒət-ə-ʁje]	_CV	‘you.PL would throw’

Forms (2a-c) are phonotactically regular: they can be explained by reference to the *loi de position*, which regulates the distribution of mid vowels (as close-mid or open-mid) as a function of the following segmental context in morphologically simple words (Tranel, 1987). Schwa does not occur before word-final consonants (_C#), explaining why [ɛ] is preferred to [ə] in (2a). Close-mid vowels are preferred to open-mid vowels before consonant-vowel and consonant-glide-vowel sequences (_CV, _CGV), explaining why [ə] is preferred to [ɛ] in (2b-c).

However, forms (2d,e) are not easily accounted for by the regular phonotactics of French. Close-mid vowels are preferred to open-mid vowels before consonant-liquid-vowel sequences (_CLV; the French liquids are [l ʁ]). Hence, on purely phonotactic grounds, [ʒət-ʁe] should be preferred to the attested [ʒɛt-ʁe] in (2d). Close-mid vowels are preferred to open-mid vowels before consonant-vowel sequences, including before consonant-schwa sequences in Northern/Standard French varieties.[2] Hence, on purely phonotactic grounds, [ʒət-ə-ʁje] should be preferred to the attested [ʒɛt-ə-ʁje] in (2e).

One way to explain these alternations, and in particular the phonotactically problematic forms in (2d-e), is hypothesis (H1).

2 This differs in Southern French, where open-mid vowels are preferred before Cə morpheme-internally; see Eychenne (2014) and Storme (2017b). For instance, *hôtel-ier* ‘relative to hotels’ is expected to be realized as [otəl-je] in Northern French but as [ɔtəl-je] in Southern French (Storme, 2017b).

(H1) Phonologically-conditioned paradigm uniformity in French

In suffixed words, the phonological shape of the stem is:

- phonotactically regular when the suffix is underlyingly vowel- or glide-initial (e.g. in (2b,c))
- paradigmatically uniform with the base with other consonant-initial suffixes (e.g. in (2d,e)).

The base is the word-form containing the stem without any affix (e.g. (2a)).

In other words, this hypothesis states that underlyingly vowel- and glide-initial suffixes are cohering and other consonant-initial suffixes are non-cohering. According to this hypothesis, suffixed forms (2b,c) are phonotactically regular for mid-vowel quality because they are built with underlyingly vowel-initial and glide-initial suffixes (*-ez* [e] and *-iez* [je], respectively). Suffixed forms (2d,e) are paradigmatically uniform with the base *jette* [ʒɛt] in (2a) along mid-vowel quality because they are built with suffixes which start with a non-glide consonant underlyingly (*-rez* [ʁe] and *-riez* [ʁje], respectively). As noted above, the word-medial schwa in (2e) is epenthetic and not underlying: therefore the suffix in (2e) does not qualify as a cohering suffix according to (H1).

To the author's knowledge, the data in (2) have not been analyzed in terms of suffix cohesion before. For instance, Dell (1985:198-214) accounted for forms (2d,e) without reference to paradigm uniformity. Instead, he proposed two rules of schwa-lowering. A rule lowering /ə/ to [ɛ] before two consonants (including before consonant-liquid sequences) accounts for (2d), i.e. /ʒətʁe/ → [ʒɛtʁe]. A rule lowering /ə/ to [ɛ] before a syllable containing a schwa accounts for (2e), i.e. /ʒətəʁje/ → [ʒɛtəʁje]. However, these rules are not phonotactically motivated because (i) schwa can occur before consonant-liquid clusters morpheme-internally (e.g. *sevr-ez* [səvʁ-e] 'wean-PRES.2PL') and (ii) schwas can occur in two adjacent syllables morpheme-internally (e.g. *Genev-ois* [ʒənəv-wa] 'Genevan'). Therefore, to avoid overgenerating lowering morpheme-internally (e.g. *sevr-ez* [səvʁ-e]/*[sɛvʁ-e], *Genev-ois* [ʒənəv-wa]/*[ʒɛnəv-wa]), Dell's two lowering rules need to be restricted to apply across morpheme boundaries.[3] This account is very different from the one proposed in (H1) because it does

3 Alternatively, for (2e), the lowering rule could be restated as applying only in case the following schwa is epenthetic. Indeed, the schwa is epenthetic in [ʒɛt-ə-ʁje] but not in [ʒənəv-wa].

not make any reference to paradigm uniformity: the open-mid quality in suffixed forms (2d,e) is not explained in terms of copying the quality of the corresponding vowel in the base in (2a), but in terms of *ad hoc* rules of lowering.

If hypothesis (H1) was confirmed, then French would have a distinction between cohering and non-cohering suffixes very similar to the one found in Germanic languages such as English and Dutch, where vowel-initial suffixes tend to be cohering and consonant-initial suffixes non-cohering. One remarkable difference concerns the behavior of glide-initial suffixes. According to (H1), glide-initial suffixes pattern as cohering in French: like vowel-initial suffixes, they trigger regular phonotactics in the stem, and this despite the fact that glides are consonants phonemically.[4] But in English, glide-initial suffixes pattern with other consonant-initial suffixes as non-cohering (see Raffelsiefen, 2015:901 on *-ward* [wəɹd] and *-ior/-iure* [jəɹ]).

Another interesting consequence of (H1) is that resyllabification cannot straightforwardly be the mechanism underlying the distinction between cohering and non-cohering suffixes in French. Theories of suffix cohesion proposed for English and Dutch generally consider that a suffix is cohering if and only if the final consonant in the stem is resyllabified as a consequence of suffixation with this suffix (e.g. van Oostendorp, 2004). Under this view, English *-al* is cohering because it imposes different syllabifications for the stem-final consonant and the corresponding base-final consonant (e.g. [t] is in onset position in *paren.t-al* but in coda position in the base *parent.*). English *-hood* is non-cohering because it imposes identical syllabifications for the two consonants (e.g. [t] is in coda position in both *parent.-hood* and *parent.*).

However, this analysis cannot straightforwardly account for suffix cohesion in French as defined in (H1). Indeed, there are two immediate problems for an account where surface syllabification matters

4 Glides are distinct from vowels underlyingly, as shown by minimal pairs like *oui* [wi] ‘yes’ - *houille* [uj] ‘coal’. Suppose by contradiction that all surface glides are vowels underlyingly. Then, there is a single underlying representation for *oui* and *houille*: /ui/. But then the two words are predicted to have the same surface realization, contrary to fact. This means that not all surface glides are vowels underlyingly in French.

for suffix cohesion. First, it fails to derive the different behaviors of glide- and liquid-initial suffixes. CG and CL are syllabified as onset clusters in French (Dell 1985, Goslin and Frauenfelder 2000). Therefore, stem-final consonants are expected to be resyllabified before both glide- and liquid-initial suffixes. For instance, [t] is expected to form an onset cluster with the following consonant in both *jetiez* [ʒətje] in (2c) and *jetterez* [ʒɛtʁe] in (2d). Yet only glide-initial suffixes are cohering according to (H1).[5]

Second, it fails to derive the different behaviors of suffixes that are underlyingly vowel-initial (e.g. *jetiez* [ʒət-e] in (2b)) and suffixes that are vowel-initial in the surface as a result of schwa epenthesis (*jetteriez* [ʒɛt-ə-ʁje] in (2e)). All prevocalic consonants are syllabified as onsets in French. Therefore, stem-final consonants are expected to be resyllabified before both vowel-initial suffixes and epenthetic schwa (e.g. [t] is expected to be syllabified as an onset in both *jetez* [ʒə.te] and *jetteriez* [ʒɛ.tə.ʁje]). Yet only suffixes that are underlyingly vowel-initial are cohering according to (H1).

This discussion immediately raises the following questions. If resyllabification is not the mechanism underlying suffix cohesion in French, what is it then? Why should glides pattern with vowels rather with other consonants? Why should the difference between underlying and epenthetic vowels matter to suffix cohesion? Answering these questions in detail is beyond the scope of this paper, whose primary goal is to provide evidence for (H1). However, section 6 will summarize the proposal that I made in Storme (2017c) and where I argue that consonants' phonetic realization rather than syllabification matters to determine suffix cohesion in French.

2.2. **[e]-[ɛ] and [o]-[ɔ] alternations: predictions**

The goal of this paper is to further test hypothesis (H1), using [e]-[ɛ] and [o]-[ɔ] alternations. These alternations are particularly relevant because they are more natural phonologically and phonetically

5 One could argue that derived C-L clusters are not syllabified identically as underlying CL sequences in French (whereas derived and underlying C(-)V and C(-)G sequences would be syllabified identically). However, there is no conclusive evidence that underlying clusters and derived clusters differ in the way they are syllabified (see Storme, 2017c:204-205) or even in their phonetic realization (see Bürki et al., 2009 contra Fougeron and Steriade, 1997).

than [ə]-[ɛ] alternations. Phonologically, [e]-[ɛ] and [o]-[ɔ] alternations involve changes along a single dimension (height) whereas [ə]-[ɛ] alternations involve changes along two dimensions: [ə]-[ɛ] differ in height but also in rounding (French [ə] is rounded). Phonetically, [e]-[ɛ] and [o]-[ɔ] are also closer than [ə]-[ɛ]. The acoustic distances among the members of each of these three pairs are as follows: d([e],[ɛ])=179 Hz, d([o],[ɔ])=206 Hz, and d([ə],[ɛ])=310 Hz. The distance between [ə] and [ɛ] is about 100 Hz larger than the distance between [e] and [ɛ] and the distance between [o] and [ɔ]. These distances were calculated using Gendrot and Adda-Decker's (2005) first three formants' data on [e ɛ o ɔ ø], assuming that [ə] is phonetically identical to [ø] (see Malécot and Cholet, 1977 and Fougeron et al., 2007) and averaging across male and female speakers. Alternations that are more natural phonologically and phonetically are also more likely to be more productive. Therefore, if [e]-[ɛ] and [o]-[ɔ] alternations also patterned according to (H1), this would provide stronger evidence for (H1) as a productive generalization in French. This section describes how these alternations are expected to behave if (H1) also applies to them.

Data on [o]-[ɔ] alternations are particularly relevant because [o]-[ɔ] contrasts are available in base-final syllables before consonants (e.g. *côte* [kot] 'coast' vs. *cote* [kɔt] 'rating') but not word-medially in monomorphemic words (Tranel, 1987). If these contrasts were maintained in suffixed words with consonant-initial suffixes but were not maintained in suffixed words with vowel- and glide-initial suffixes, this would provide strong evidence in favor of (H1). This kind of strong evidence is not available for [ə]-[ɛ] alternations because the two vowels do no contrast in word-final syllables, where only [ɛ] is available before consonants (Tranel, 1987). Note that the maintenance of [o]-[ɔ] contrasts before consonant-initial suffixes is not predicted by Dell's (1985) analysis of [ə]-[ɛ] alternations presented in section 2.1. Indeed, Dell's analysis does not apply to back rounded vowels [o]-[ɔ]. If Dell's two rules of lowering discussed in section 2.1 were to be extended to all mid vowels including back rounded vowels, then this would predict a single open-mid quality [ɔ] in stems containing back rounded vowels and suffixed with consonant-initial suffixes (e.g. *côte*- and *cote*- would neutralize to [kɔt-] before consonant-initial suffixes).

Like [ə]-[ɛ], [e] and [ɛ] do not contrast in base-final syllables before consonants, where only [ɛ] is available (Tranel, 1987). Therefore, paradigm uniformity cannot be diagnosed by contrast maintenance for this mid-vowel pair. However, it can be diagnosed by other means. As in the case of [ə]-[ɛ] alternations, paradigm uniformity and regular phonotactics predict different qualities in some contexts. Paradigm uniformity before consonant-initial suffixes predicts a preference for [ɛ] before morphologically-derived C-LV sequences and before epenthetic schwa in C-ə-CV sequences (in violation of the language's phonotactics, which requires [e] in these contexts). Regular phonotactics before vowel- and glide-initial suffixes predicts a preference for [e] before morphologically-derived C-GV and C-V sequences (in violation of paradigm uniformity, which requires [ɛ]). In other words, one expects [e]-[ɛ] alternations to pattern identically to [ə]-[ɛ] alternations.

The predictions of hypothesis (H1) for [e]-[ɛ] and [o]-[ɔ] alternations are summarized in (3). Here it is assumed that the effect of phonotactics on stem mid-vowels before vowel- and glide-initial suffixes is mid-vowel raising, according to the hypothesis that the *loi de position* is the main factor explaining the distribution of close-mid vs. open-mid allophones in non-final positions (see Nguyen and Fagyal, 2008:23). The predictions of (H1) for [ə]-[ɛ] alternations (see section 2.1) are summarized in the last column in (3).

(3) Underlyingly vowel-/glide-initial suffixes as cohering vs. other consonant-initial suffixes as non-cohering (=H1)

		e/ɛ	o/ɔ	ə/ɛ	
Base	VC#	ɛ	o-ɔ	ɛ	
Suffixed word	VC-V/GV	e	o	ə	(Phonotactics; *loi de position*)
	VC-(ə)-LV/CV	ɛ	o-ɔ	ɛ	(Paradigm uniformity)

2.3. Previous studies

2.3.1. Previous studies of mid-vowel quality in French

The predictions concerning [e]-[ɛ] and [o]-[ɔ] alternations are not as easy to test as in the case of [ə]-[ɛ] and this is why this paper uses an acoustic study rather pronunciation dictionaries (or speakers' metalinguistic judgments). Indeed, pronunciation dictionaries do not always provide a single variant for the realization of non-final mid vowels in the relevant stems (whereas they do for [ə]-[ɛ]) and it is not very clear what the source of this variation is. For instance, Martinet and Walter (1973) list both [e] and [ɛ] realizations for the stem mid-vowel in the suffixed word *fêt-ard* 'reveller', derived from the base *fête* [fɛt] 'party'. The authors indicate that the choice between the two realizations is speaker-dependent. However, in the absence of systematic acoustic data, it is difficult to conclude whether one is really dealing with speaker variation or other sources of noise, for instance a difficulty with discriminating phonetically close allophones on the part of transcribers. Indeed, [e]-[ɛ] and [o]-[ɔ] are closer in the acoustic space than [ə]-[ɛ] (see section 2.2), and therefore probably harder to distinguish perceptually, as the first three formants are the main dimensions of the similarity space for oral vowels (Shepard, 1972). This difficulty should be increased in the context under consideration, namely in non-final syllables. Indeed, there is substantial target undershoot in non-final syllables as compared to word-final syllables in French, resulting in close-mid and open-mid vowels being closer in the acoustic space non-finally than word-finally (Storme, 2017a). Moreover, close-mid and open-mid vowels do not contrast in non-final syllables in morphologically simple words (Tranel, 1987), making it potentially harder to detect differences between the two vowels in this context. Indeed, there is evidence that pairs of sounds are harder to discriminate if the two sounds are in allophonic distribution in one's language than if they contrast (Boomershine et al., 2008).

Concretely, this means that any study of [e]-[ɛ] and [o]-[ɔ] alternations should include several speakers (to control for speaker variation) and be based on acoustic data (to bypass potential biases introduced by transcriptions). The study that comes closest to satisfy these requirements is Nguyen and Fagyal's (2008) acoustic study on vowel harmony in suffixed words (with six speakers). Nguyen

and Fagyal (2008:23) found that, in stems suffixed with vowel-initial suffixes (e.g. *fêt-ard*), stem mid-vowels tend to be closer to a close-mid realization than to an open-mid vowel realization in the F1/F2 space. This would be in line with the prediction that regular phonotactics (in particular, the *loi de position*) are privileged over paradigm uniformity in stems suffixed with vowel-initial suffixes.

However, this study was not meant to test hypotheses about suffix cohesion and therefore does not include the data relevant for a comprehensive test of these hypotheses. In particular, it does not include data on suffixes other than vowel-initial. In addition, the study does not include paradigm uniformity and phonotactic baselines to compare stem mid-vowels to: mid-vowel realizations in suffixed words are not compared to mid-vowel realizations in the corresponding bases (e.g. *fête* for *fêt-ard*) or to mid-vowel realizations in morphologically simple phonological neighbours (e.g. *feta* 'feta cheese' for *fêt-ard*). Therefore, the study does not provide a direct test of whether phonotactics are favored over paradigm uniformity, even in stems suffixed with vowel-initial suffixes.

2.3.2. Previous studies of suffix cohesion

The current paper proposes that suffix cohesion in French is determined by the phonological nature of the suffix's initial segment. But other factors have been argued to be relevant to suffix cohesion in previous works on other languages. This section reviews some of these other factors and describes how mid-vowel alternations should pattern in French if there was no suffix cohesion or if suffix cohesion was conditioned by these other factors. These different hypotheses will be compared quantitatively against the acoustic data collected in this study.

First, we need to consider the possibility that all stems are treated identically, all according to regular phonotactics or all according to paradigm uniformity. The predictions of these two hypotheses are summarized in (4) and (5), respectively.

(4)

		e/ɛ	o/ɔ	
Base	VC#	ɛ	o-ɔ	
Suffixed word	VC-V/GV/LV	e	o	(Phonotactics; *loi de position*)
	VC-CV	ɛ	ɔ	(Phonotactics; *loi de position*)

(5)

		e/ɛ	o/ɔ	
Base	VC#	ɛ	o-ɔ	
Suffixed word	VC-V/GV/LV/CV	ɛ	o-ɔ	(Paradigm uniformity)

Next, we consider models assuming different conditionings for suffix cohesion. The first obvious hypothesis to consider is the one according to which cohering suffixes coincide with resyllabifying suffixes and non-cohering suffixes with non-resyllabifying suffixes. This hypothesis is indeed central to accounts of suffix cohesion in Germanic languages (see section 2.1). The predictions of the syllable-based conditioning on suffix cohesion are summarised in (6). This hypothesis distinguishes two sets of contexts: before CV, CGV, and CLV, where mid vowels behave regularly due to resyllabification of the stem-final consonant, and before other consonantal clusters (CCV), where mid vowels behave according to paradigm uniformity due to the asbence of resyllabification. In this hypothesis, consonant-initial suffixes preceded by epenthetic schwa pattern with underlyingly vowel-initial suffixes because epenthetic schwa also induces a resyllabification of the stem-final consonant.

(6)

		e/ɛ	o/ɔ	
Base	VC#	ɛ	o-ɔ	
Suffixed word	VC-V/GV/LV	e	o	(Phonotactics)
	VC-CV	ɛ	o-ɔ	(Paradigm uniformity)

Paradigm uniformity may also be conditioned by non-phonological factors. In dual-route models of lexical access, the relative frequency of a suffixed word with respect to its base is expected to affect the derivative's pronunciation (Hay, 2001; Cohen, 2014; Zuraw and Peperkamp, 2015). Suffixed words that are frequent as compared to their base favor the direct route (i.e. they are analyzed as a single morphological unit) and therefore are expected to be more likely to behave according to regular phonotactics. In other words, the base is not sufficiently salient in speakers' memory so as to be activated when the suffixed word is accessed, resulting in the suffixed word taking on a life of its own as a morphologically simple word. Suffixed words that are not frequent compared to their base favor the decomposition route (i.e. they are analyzed as a complex morphological unit) and therefore are more likely to behave according to paradigm uniformity. In other words, the base is sufficiently salient in speakers' memory so as to be activated when the suffixed word is accessed, resulting in the phonology of the suffixed word being influenced by the phonology of its base.

The morphological function of suffixes as inflectional or derivational is also predicted to be relevant to the phonology of suffixed words in dual-route models of lexical access. Because inflectional paradigms are generally more transparent semantically than derivational paradigms (Haspelmath and Sims 2010: chapter 5), dual-route models of lexical access predict that the base will be more salient when accessing an inflected word than when accessing a derived word. In accordance with the hypothesis of a greater cohesiveness of inflectional paradigms as compared to derivational ones, Schriefers et al. (1992) found priming effects among forms belonging to the same inflectional paradigm but not among forms belonging to the same derivational paradigm. From a phonological

perspective, the greater cohesiveness of inflectional paradigms is expected to translate into a greater pressure for paradigm uniformity in words suffixed with inflectional morphemes than in words suffixed with derivational morphemes.

It should be noted that, although models resorting to frequency and morphological function are plausible conceptually, the data on [ə]-[ɛ] stem-alternations described in section 2.1 do not make them particularly plausible for French. All suffixed forms *jetez, jetiez, jetterez,* and *jetteriez in* (2b-d) are less frequent than their base *jette* in (2a) (New et al 2007) and yet they do not pattern identically in terms of suffix cohesion, contrary to the predictions of the frequency-based model of suffix cohesion. Also, all forms *jetez, jetiez, jetterez,* and *jetteriez* are inflected and yet they do not pattern identically in terms of suffix cohesion, contrary to the predictions of the morphological model of suffix cohesion. One also finds cohering and non-cohering suffixes among derivational suffixes. For instance, *lunetier* [lynət-je] ‘eyewear manufacturer’ (built with the glide-initial, derivational suffix -*ier* [je]) behaves regulary phonologically whereas *lunetterie* [lynɛt-ʁi] ‘eyewear industry’ (built with the liquid-initial, derivational suffix -*rie* [ʁi]) behaves according to paradigm uniformity, i.e. it copies the vowel quality of its derivational base *lunettes* [lynɛt] ‘glasses’. These data suggest that suffix cohesion is orthogonal to the suffix’s morphological function in French.

Although these models of suffix cohesion are not particularly plausible to explain [ə]-[ɛ] stem-alternations, they will still be considered as alternatives to explain [e]-[ɛ] and [o]-[ɔ] in this paper. One reason for that is that it is possible that the two types of alternations ([ə]-[ɛ] on one hand and [e]-[ɛ] and [o]-[ɔ] on the other hand) behave differently. Indeed, as explained above, [e]-[ɛ] and [o]-[ɔ] alternations are more natural phonologically and phonetically than [ə]-[ɛ] alternations and this difference could translate into different patterns of alternations. In particular, due to their greater naturalness, [e]-[ɛ] and [o]-[ɔ] alternations could pattern in a way that is more consistent with the predictions of the dual-route model of lexical access.

3. Methods

The current study aims to compare hypotheses about suffix cohesion in French using acoustic data on the realization of mid vowels in suffixed stems, in the corresponding bases, and in morphologically simple words. The use of an acoustic measure for mid-vowel quality makes it possible to avoid transcription biases. The study includes several participants, making it possible to control for speaker-dependent variation. The use of phonotactic and paradigm uniformity baselines makes it possible to directly assess (i) whether the suffixed word is realized according to regular phonotactics or to paradigm uniformity, (ii) whether this effect depends on the suffix, as predicted under suffix cohesion, and (iii) in case it does, what properties of the suffixes are relevant to determine their cohesion (suffix's initial segment, resyllabification, morphological function, or base saliency).

3.1. Participants

Ten native French speakers living in the Paris area at the time of the recordings (6 females, 4 males; aged 21-30, mean = 25.3 years, sd = 3.1) participated in the study on a voluntary basis. All of them were university students at the time of the recordings. To make sure that the participants had the relevant mid-vowel contrasts in bases (in particular, [o]-[ɔ] in word-final syllables before consonants), they were asked (after being recorded) whether they distinguish the pronunciation of *taupe* [top] and *top* [tɔp] in their daily speech. They all answered positively.

3.2. Stimuli and recordings

To test hypotheses about the realization of stem mid vowels as phonotactically regular or paradigmatically uniform, three forms are needed: the suffixed word (e.g. *fêt-ard* 'partier'), the corresponding base (e.g. *fête* 'party'), and the phonotactic baseline (e.g. *feta* 'feta cheese'). In what follows, these forms will be referred to as the derivative, the base, and the neighbour, respectively.

Note that the term ‘derivative’ is used as a cover term for both inflected and derived forms. The stimuli used in the study can be found in the appendix at the end of this paper.

42 derivative/base/neighbour triplets were selected (see the appendix for a list of all triplets). There were 14 bases for each vowel among [ɛ o ɔ] ([e] is not represented because it does not occur in final syllables of consonant-final bases). Derivatives were built with V-initial, G-initial, L-initial, and C-initial suffixes (where C is not a glide or a liquid). For each base mid vowel among [ɛ o ɔ], there were four vowel-initial suffixes, three glide-initial suffixes, three liquid-initial suffixes, three other consonant-initial suffixes, and one suffix requiring an epenthetic schwa between the stem and the suffix (*maigr-e-let* for [ɛ], *pauvr-e-té* for [o], *sobr-e-ment* for [ɔ]). The C-initial suffix *-ment* is represented in several suffixed words (see the appendix) because it was difficult to find actual examples of words with other suffixes. This means that the results for this condition might be heavily influenced by this suffix. See section 5.1 for further discussion.

The broad segmental context following the target mid vowel was the same in the derivative and in the neighbour. If the target mid vowel was in an open syllable in the derivative (i.e. followed by CV, CGV, or CLV), it was also in an open syllable in the neighbour. For instance, the target mid vowel is in an open syllable in both *fêt-ard* and *feta*. If it was in a closed syllable in the derivative (i.e. followed by CCV), it was also in a closed syllable in the neighbour. For instance, the target mid vowel is in a closed syllable in both *vête-ment* ‘clothe’ and *Etna*.

In two cases, the neighbour was actually a morphologically complex word, i.e. in the triplets *coquelet/coq/mot-clé* and *faussiez/fausse/dossier*, where *mot-clé* ‘keyword’ is a compound and *dossier* is a suffixed word (etymologically derived from *dos* [do] ‘back’ with the suffix *-ier*). However, in these words, regular phonotactics and paradigm uniformity predict the same realization. The mid vowel in *mot-clé* should be realized as [o] according to both phonotactics (due to the *loi de position* before CLV) and paradigm uniformity (because the base *mot* is pronounced as [mot] in the French variety under study). The mid vowel in *dossier* should be realized as [o] according to both phonotactics (due

to the *loi de position* before CGV) and paradigm uniformity (because the base *dos* is pronounced as [do] in the relevant French variety). Therefore, the fact that these words are morphologically complex is probably harmless: they should have the same phonology whether their phonology is realized according to phonotactics or paradigm uniformity. Other neighbours were morphologically simple verbs, common names or proper names.

Word-position is known to affect the realization of mid vowels: non-final vowels are shorter than word-final vowels and this causes F1 undershoot (Gendrot and Adda-Decker, 2005; Storme, 2017a). This may make it hard to detect an effect of paradigm uniformity in base-derivative pairs. For instance, assuming paradigm uniformity of the suffixed word *fête-ra* with the base [fɛt], the underlying /ɛ/ in /fɛtʁa/ could raise close to [e] due to undershoot in non-final syllables. To avoid this confound, the target syllable was put under contrastive focus for all words: in this context, vowels should be longer and therefore less reduced. In order to elicit contrastive focus on the first syllable of target words, participants were told that they should imagine they are correcting the pronunciation of a non-native French speaker. The target word *Y* was included in a carrier sentence *On dit pas X, on dit Y* 'one does not say X, but Y', where *X* was a faulty pronunciation of *Y* with the mid vowel in the initial syllable of *Y* being incorrectly pronounced as a high or low vowel. For instance, X was *fitard* and Y *fêtard*. Each word was repeated three times. The stimuli were presented in three lists where the sentences appeared in pseudo-random order.

Speakers were recorded in a sound attenuated booth with a head-mounted Shure SM35-XLR connected to a computer via a Shure X2u XLR-to-USB signal adapter. The recordings were made using the Audacity software, with 44 kHz/16 bit sampling. The distance (approx. 5 cm) of the participants to the microphone was held constant across all recording sessions.

3.3. Acoustic analyses

Vowels were segmented and labelled in Praat (Boersma and Weenink, 2019). The segmentation was done using voicing and changes in formant trajectories as indicators. Formant measurements (F1, F2 at vowel midpoint; in Hertz) were done automatically with a Praat script. Manual corrections were performed when a formant value was obviously wrong (in a handful of cases, F2 was misidentified as F3). F1 and F2 were z-scored.

As a measure of mid-vowel quality, the distance to the center of the F1/F2 space was chosen. Close-mid vowels are expected to be further away from this center than open-mid vowels. The center of the F1/F2 space was calculated as follows. For each speaker, three averages were calculated for each of the first two formants: the average realization for target front mid vowels, the average realization for target back mid vowels, and the average realization for 10 occurrences of [a] (these occurences came from the negation *pas*). Then, these three numbers were averaged to obtain the center of the F1/F2 space. The average was computed in two steps rather than in a single step because the sample was not balanced between front and back: due to the absence of [e] in consonant-final bases, there were more back vowels than front vowels in the stimuli.

3.4. Statistical analyses

The models of mid-vowel realization discussed in section 2.2 were implemented as mixed effects models in R using lme4. The dependent variable was the distance to the F1/F2 center. Fixed effects included the vowel in the base (as established by pronunciation dictionaries, i.e. [ɛ o ɔ]), the vowel's broad segmental context (_C#, _CV, _CGV, _CLV, and _CCV), and the interaction of these two factors. Random effects included a random intercept by speaker and a random intercept by word. Models with more complex random structure did not converge.

What differed between the different models of mid-vowel quality in suffixed stems (see sections 2.2 and 2.3.2 for a presentation of these models) was the way mid vowels were assigned to their segmental context in derivatives. When a given derivative was predicted to behave according to

paradigm uniformity with the base, the segmental context was the same as in the base (e.g. the segmental context for the stem mid-vowel in *fêt-ard* is treated as _C#, as in the base *fête*). When it was predicted to behave according to phonotactics, the segmental context was the same as in the phonotactic baseline (e.g. the segmental context for the stem mid-vowel in *fêt-ard* was treated as _CV, as in the neighbour *feta*).

To determine which of the base or the derivative was more frequent, frequencies based on French subtitles were used (Lexique 3.80; New et al 2007). The relative frequency of the base with respect to the derivative was coded as a binary variable (see Base saliency in the appendix), taking value “yes” if the base was more frequent and “no” otherwise. It should be noted that the derivative was more frequent than the base in only five of the 42 pairs used in the study.

Although phonology, frequency, and morphology could all play a role in determining a derivative’s pronunciation, only a single factor at a time was considered. The reason is that the statistical analysis was set up so that derivatives could only be assigned to a single segmental context at a time. As a consequence, the analysis did not determine whether the relevant factors (i.e. suffix’s phonological shape, suffix’s morphological function, frequency of the base with respect to the derivative) significantly affected a derivative’s pronunciation but rather what was the single best predictor among them.

Model selection was done using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The model with the lowest AIC or BIC was preferred. Likelihood ratio tests could not be used because the models under consideration were not nested. AIC and BIC evaluate both model fit and complexity: there is a larger penalty for models with more parameters. Because of the way suffixed words were assigned to segmental contexts, the different hypotheses all had the same number of parameters and therefore the same complexity. Therefore, AIC and BIC actually only compared model fit.

4. Results

Table 1 shows the AIC and BIC values for all six models discussed in section 2.2. The model in the first row of the table is the model proposed in this paper, where (underlyingly) vowel- and glide-initial suffixes are cohering and other suffixes are non-cohering. The phonotactic model corresponds to the hypothesis that all suffixed words are realized according to the regular phonotactics of the language. The paradigm uniformity model corresponds to the hypothesis that all suffixed words are realized like the base. The resyllabification model corresponds to the hypothesis according to which resyllabifying suffixes are cohering whereas non-resyllabifying suffixes are non-cohering. The morphological model corresponds to the hypothesis according to which derivational suffixes are cohering and inflectional suffixes non-cohering. The frequency model corresponds to the hypothesis according to which suffixed words with salient bases (the base is more frequent) are non-cohering and suffixed words with non-salient bases (the base is less frequent) are cohering.

	AIC	BIC
Proposed model (=H1)	**-1031**	**-919**
Phonotactic model	-991	-879
Paradigm uniformity model	-985	-873
Resyllabification model	-1006	-894
Morphological model	-979	-867
Frequency model	-984	-871

Table 1: Model comparison

The model proposed in this paper has the smallest AIC and BIC. This means that the phonological shape of the suffix as vowel/glide-initial or consonant-initial is the single best predictor of suffix cohesion among the different factors considered. These results corroborate the hypothesis developed in section 2 based on [ə]-[ɛ] alternations.

Figure 1 compares the realization of mid vowels in the derivative with its realization in the base and in the phonotactic neighbour as a function of the base mid vowel ([ɛ o ɔ]) and the suffix-initial segment (vowels, glides, liquids, and other consonants). Hypothesis (H1) predicts that mid-vowels in derivatives should be closer to the corresponding mid-vowels in neighbours with vowel- and glide-initial suffixes (see the six leftmost panels) but closer to the corresponding mid-vowel in bases with other consonant-initial suffixes (see the six rightmost panels). Note that, for [o], for all V-, G-, and L-initial, the same vowel quality is expected for derivatives, their bases, and their neighbours. Indeed, phonotactics and paradigm uniformity predict a close-mid quality in these three cases.

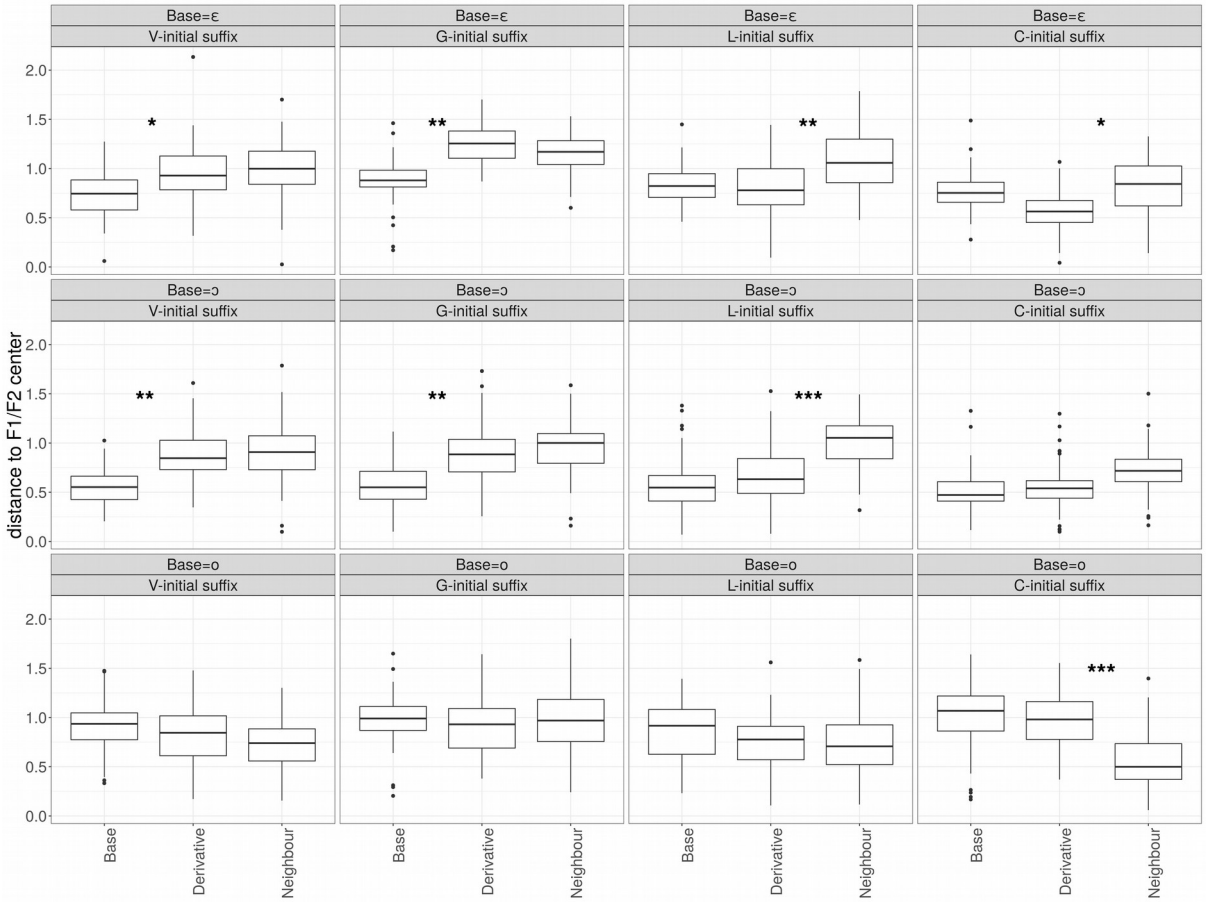

Figure 1. Mid vowel realization in the study (F1 and F2 values z-scored).

The results are largely compatible with the predictions of the analysis proposed in this paper. With vowel-initial and glide-initial suffixes, suffixed words do not behave significantly differently from their neighbours, i.e. they behave according to the regular phonotactics of the language. With liquid-initial suffixes and other consonant-initial suffixes, suffixed words do not behave significantly differently from their bases, i.e. they behave according to paradigm uniformity. The results are described in more details below, focusing first on bases in [ɛ] and [ɔ] (the first two rows in Figure 1) and then on bases with [o] (the last row in Figure 1).

4.1. Bases in [ɛ] and [ɔ]

For bases with open-mid vowels [ɛ ɔ] (see the first two rows in Figure 1), stem mid vowels are significantly more peripheral in the words suffixed with vowel- and glide-initial suffixes than in the corresponding bases (see columns 1 and 2 in Figure 1). This is compatible with the hypothesis that the quality of stem mid vowels is determined by the *loi de position* before these suffixes. Indeed, the *loi de position* requires close-mid vowels [e o] before CV and CGV morpheme-internally, and [e o] are more peripheral than [ɛ ɔ].

Stem mid vowels in derivatives built with liquid-initial suffixes are significantly more central than in the corresponding neighbours (see column 3 in Figure 1). This is also compatible with the current hypothesis. Indeed, for bases with open-mid vowels [ɛ ɔ], paradigm uniformity before liquid-initial suffixes predicts [ɛ ɔ] in the corresponding derivatives and the *loi de position* predicts more central realizations (namely [e o]) in neighbours before CLV.

Finally, the current hypothesis predicts an open-mid vowel quality [ɛ ɔ] for derivatives suffixed with other consonant-initial suffixes and for their neighbours. For derivatives, this is because mid-vowel quality is copied from bases with [ɛ ɔ]. For neighbours, this is because the *loi de position* predicts an open-mid vowel quality [ɛ ɔ] in non-final closed syllables. However, contrary to this prediction, mid-vowel quality was found to be slightly less peripheral in neighbours than in derivatives (see column 4 in Figure 4; note that this effects reaches significance only for [ɛ]). How can we explain this difference between derivatives and neighbours? Open-mid vowels are expected to be more peripheral acoustically word-finally than non-finally, due to vowel undershoot in nonfinal syllables (see Storme, 2017a). If paradigm uniformity effects are sensitive to this level of fine grained phonetic detail (see Steriade, 2000), then we actually do expect open-mid vowels [ɛ ɔ] to be more peripheral in derivatives than in the corresponding neighbour: under Steriade's hypothesis that paradigm uniformity cares about the base's detailed phonetic realization, the derivative is expected to copy the vowel quality of the base without undershoot and not just the phonological label [ɛ ɔ]. Note that the difference between derivatives and neighbours cannot be attributed to stem mid vowels in derivatives also copying the long duration of base mid vowels in addition with their quality. Indeed, although mid vowel duration is longer in bases than in neighbours, it is not longer in derivatives than in neighbours in the relevant conditions, as shown in Figure 2.

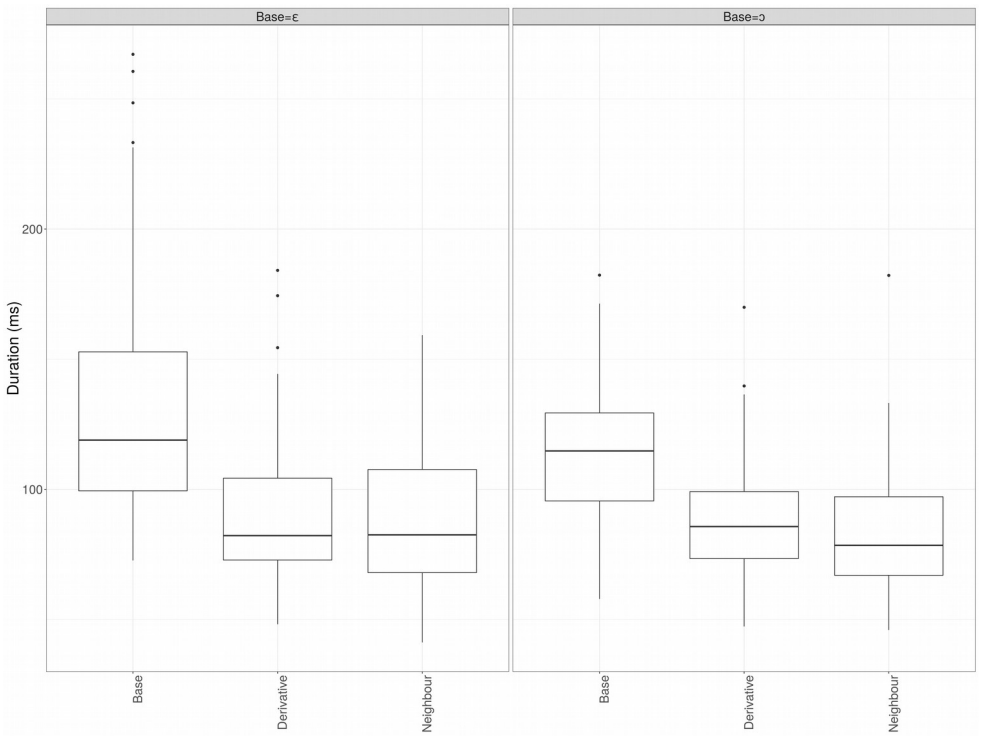

Figure 2. Mid vowel

duration in derivatives suffixed with C-initial suffixes, in the corresponding bases, and in the corresponding neighbours.

4.2. Bases in [o]

For bases in [o] (see the last row in Figure 1), stem mid vowels in derivatives suffixed with V-, G-, and L-initial suffixes are not significantly different from the corresponding mid vowels in bases or neighbours (see columns 1 to 3 in Figure 1). This is expected under the present analysis. In neighbours, we expect a close-mid quality [o], due to the *loi de position* before CV, CGV, and CLV. In derivatives, we also expect a close-mid quality [o] both with V- and G-initial suffixes (due the *loi de position*) and with L-initial suffixes (due to paradigm uniformity). Before other C-initial suffixes, mid vowels are more peripheral in derivatives than in neighbours (see column 4 in Figure 1). This is expected under the present analysis. Indeed, paradigm uniformity predicts a close-mid quality [o] for derivatives suffixed with C-initial suffixes while, in neighbours where the mid vowel appears in a closed syllable, the *loi de position* predicts a more central, open-mid quality [ɔ].

5. Discussion

The hypothesis that suffix cohesion is conditioned by the suffix-initial segment in French was corroborated by acoustic evidence on [e]-[ɛ] and [o]-[ɔ] alternations. Section 5.1 highlights some limits of this study and makes suggestions for future research. Section 5.2 focuses on explanations for the split between vowel-/glide-initial suffixes and other suffixes.

5.1. Limits of the study

It was difficult to find neighbours that perfectly matched the corresponding derivatives segmentally. For example, the neighbour of *fêt-ard* was *féta*, without final rhotic. In some cases, the discrepancy

between the neighbour and the derivative was greater: for instance, the neighbour of *cloche-ton* was *pochtron*, where the target mid vowel is preceded by one consonant instead of two, followed by three instead of two, and the word-initial cluster/consonant have different places of articulation ([kl] vs. [p]). In a future study, the segmental context should be controlled for more carefully. This could be done by considering nonce words as neighbours (e.g. fake proper names like *Fétar*). As mentioned earlier, the C-initial suffix *-ment* was represented in several suffixed words because it was difficult to find actual examples of words with other suffixes. To overcome this problem, it might also be necessary to consider nonce words built with other consonant-initial suffixes.

The acoustic study only considered a small pool of Parisian French speakers in their twenties. Therefore, caution should be used before generalizing, in particular to other French varieties. French varieties notoriously differ in their patterns of distribution of mid vowels. For instance, in Southern French, there are no contrasts between close-mid and open-mid vowels (Coquillon and Turcsan, 2012). In particular, in contrast with Parisian French, all mid vowels in Southern French are open-mid before word-final consonants (e.g. *côte* is pronounced as [kɔt] instead of [kot]). This means that suffix cohesion cannot be diagnosed by contrast maintenance in suffixed stems in this variety. As a consequence, the protocol used in this study should be slightly adapted to take this type of differences into account when dealing with other French varieties.

5.2. Interpretation

Suffix cohesion cannot be straightforwardly captured in terms of resyllabification of the stem-final consonant, as discussed in section 2.1 and as further shown in section 4. What is the motivation then? In this section, I summarized the key points of the analysis I proposed in Storme (2017c), where the split is argued to be phonetically motivated. The analysis makes a central use of non-derived environment blocking, a concept developed mainly by Kiparsky (1993). Non-derived environment

blocking describes a pattern where a phonological process P applies only if accompanied by another phonological process P'.

In Storme (2017c), I proposed that suffix cohesion in French depends on how the suffix affects the phonetic similarity between the suffixed stem and the independent base (e.g. the similarity between the base *fête* and the stem in *fêt-iez* vs. *fête-rez*), with greater phonetic similarity entailing further phonological similarity, in a "rich get richer" effect. In a nutshell, the proposal is that a phonetic change P affecting the release cues of the stem-final consonant (and induced by the suffix-initial segment) feeds a further phonological change P' affecting the mid vowel in the final syllable of the stem, according to the logic of non-derived environment blocking.

The basic split between vowel-initial and stop-initial suffixes is explained as follows. Vowel-initial suffixes modify the release cues of the stem-final consonant by adding release formant transitions that were absent in the isolated base. This change then licenses a further change in the stem, according to the logic of non-derived environment blocking. This further change is motivated by the language's phonotactics, namely the *loi de position*. Contrary to vowels, stops do not have formant structure: as a consequence, stop-initial suffixes do not affect as dramatically the release cues of the stem-final consonant. This absence of change blocks any further change from happening in the stem, again according to the logic of non-derived environment blocking.

The difference between glide- and liquid-initial suffixes is particularly problematic for an account in terms of resyllabification, as both suffixes should trigger resyllabification and therefore behave as cohering suffixes. How is this difference accounted for in the phonetically based analysis? The difference between glide-initial and liquid-initial suffixes is expected if glides affect the release properties of a preceding consonant more than liquids do. Both glides and liquids have formant structure and therefore provide release transitions to a preceding consonant (Wright, 2004). However,

glides are acoustically more similar to vowels than liquids are, and therefore potentially provide more salient release transitions to a preceding consonant. Also, Storme (2017c) provides preliminary experimental evidence that word-final consonants are perceptually more similar to their pre-liquid counterparts than to their pre-glide counterparts. In a nutshell, a [t] extracted from *fête* is judged as more similar to a [t] extracted from *fête-rez* than to a [t] extracted from *fêt-iez*. This difference then underlies the different patternings of glide-initial and liquids-initial: glide-initial suffixes induce a sufficiently salient change to allow a further change to happen in the stem whereas liquid-initial suffixes do not and therefore block the *loi de position* in stems.

How is the behavior of stems followed by epenthetic schwa explained (see *jett-e-riez*)? In Storme (2017c), I proposed the following explanation. Schwa epenthesis at morpheme boundary does not license a further change in the stem because schwa can also be epenthesized at the end of a base. For instance, *fête* is usually realized as [fɛt] but can also be realized as [fɛtə], with a final epenthetic schwa. As a consequence, the specific release transitions that are added to the stem-final consonant in case of schwa epenthesis before a suffix (e.g. *fêteriez* [fɛtəʁje]) can also be present in the base-final consonant, in case of word-final schwa epenthesis (e.g. *fête* [fɛtə]). This makes the stem and the base sufficiently similar to block a further phonotactically motivated change from /ɛ/ to [e] before Cə.

6. Conclusion

This paper has provided converging evidence for a previously undescribed pattern of suffix cohesion in French, where words suffixed with vowel- and glide-initial suffixes behave phonologically like simple units and words built with other suffixes behave phonologically like complex units. This finding is relevant for theories of suffix cohesion because it suggests that other factors beyond resyllabification, morphological function, and frequency play a role in determining whether a suffix is cohering or not. It was suggested that suffix cohesion in French can be interpreted in terms of how the

suffix-initial segment affects the phonetic similarity between the suffixed stem and the corresponding base.

The results are also relevant for theories of paradigm uniformity more generally. Indeed, some of the results in section 4.1 suggest that paradigm uniformity is sensitive to fine-grained phonetic detail, as hypothesized by Steriade (2000). However, these results are very preliminary and should be evaluated further in a future study.

Finally, the results are also relevant for transcription conventions for French. In the dictionary *le Petit Robert* (Robert, 2013), suffixed words with vowel-initial or glide-initial suffixes are not consistently transcribed according to the language's regular phonotactics. For instance, *fêt-ard* is transcribed according to paradigm uniformity as [fɛtaʁ] and differently from monomorphemic *feta* 'feta cheese', transcribed as [feta]. The findings in this paper suggest that words with vowel- and glide-initial suffixes should be more consistently transcribed as phonotactically regular, at least in the French variety studied here. Further studies should be carried out to assess whether the results reported here extend to other French varieties.

Appendix

Base V	Base	Derivative	Neighbour	Context	Suffix	Base saliency	Inflection
ɛ	bête	bêt-ise	Betty	_CV	V-initial	yes	no
ɛ	rêve	rêv-ais	réveil	_CV	V-initial	yes	yes
ɛ	fête	fêt-ard	féta	_CV	V-initial	yes	no
ɛ	mètre	métr-age	étrave	_CLV	V-initial	yes	no
ɛ	maigre	maigr-e-let	écrevisse	_CLV	L-initial	yes	no
ɛ	aile	ail-ier	bélier	_CGV	G-initial	yes	no
ɛ	lègue	léguiez	Séguier	_CGV	G-initial	yes	yes
ɛ	mette	mett-iez	métier	_CGV	G-initial	yes	yes
ɛ	guette	guette-rie	Guétry	_CLV	L-initial	yes	no

ɛ	pète	pète-ra	Pétra	_CLV	L-initial	yes	yes
ɛ	tète	tète-ra	tétra	_CLV	L-initial	yes	yes
ɛ	saine	saine-ment	ennemi	_CCV	C-initial	yes	no
ɛ	chère	chère-ment	ferment	_CCV	C-initial	yes	no
ɛ	vête	vête-ment	Etna	_CCV	C-initial	no	no
o	Paule	Paul-ine	police	_CV	V-initial	no	no
o	cause	caus-ais	Cosette	_CV	V-initial	yes	yes
o	faute	faut-eur	fauteuil	_CV	V-initial	yes	no
o	autre	autr-ui	autruche	_CLV	G-initial	yes	no
o	gaule	gaul-ois	Olloix	_CGV	G-initial	no	no
o	fausse	fauss-iez	dossier	_CGV	G-initial	yes	yes
o	taule	taul-ière	Molière	_CGV	G-initial	yes	no
o	chauffe	chauffe-rie	Sofri	_CLV	L-initial	yes	no
o	saute	saute-rie	Kotri	_CLV	L-initial	yes	no
o	côte	côte-lé	potelé	_CCV	L-initial	yes	no
o	miaule	miaule-ment	dolmen	_CCV	C-initial	no	no
o	haute	haute-ment	Otmani	_CCV	C-initial	yes	no
o	frôle	frôle-ment	colmater	_CCV	C-initial	yes	no
o	pauvre	pauvr-e-té	hôtelier	_CV	C-initial	yes	no
ɔ	folle	fol-ie	folique	_CV	V-initial	yes	no
ɔ	ogre	ogr-esse	maugrée	_CV	V-initial	yes	no
ɔ	bloque	bloqu-é	loquet	_CV	V-initial	no	yes
ɔ	colle	coll-age	cola	_CV	V-initial	yes	no
ɔ	moque	moqu-iez	Mauquier	_CGV	G-initial	yes	yes
ɔ	botte	bott-ier	Bauthier	_CGV	G-initial	yes	no
ɔ	vol	vol-ière	Vaulière	_CGV	G-initial	yes	no
ɔ	gobe	gobe-ra	cobra	_CLV	L-initial	yes	yes
ɔ	coq	coque-let	mot-clé	_CLV	L-initial	yes	no
ɔ	os	osse-let	Oslo	_CLV	L-initial	yes	no
ɔ	cote	cote-rie	Autriche	_CLV	L-initial	yes	no
ɔ	molle	molle-ton	poltron	_CCV	C-initial	yes	no
ɔ	cloche	cloche-ton	pochtron	_CCV	C-initial	yes	no
ɔ	sobre	sobr-e-ment	saugrenu	_CLV	C-initial	yes	no

Number of words: about 8,600 (among which 350 come from the appendix)

References

Boersma, P. and Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.56, retrieved 20 June 2019 from http://www.praat.org/.

Boomershine, A., Hall, K. C., Hume, E., and Johnson, K. (2008). The impact of allophony versus contrast on speech perception. *Contrast in phonology: Theory, perception, acquisition*, *13*, 145-172.

Bürki, A., Fougeron, C., Veaux, C., and Frauenfelder, U. H. (2009). How similar are clusters resulting from schwa deletion in French to identical underlying clusters? In *INTERSPEECH-2009*, 2271-2274.

Cohen, C. (2014). Probabilistic reduction and probabilistic enhancement. *Morphology*, *24*(4), 291-323.

Coquillon, A., and Turcsan, G. (2012). An overview of the phonological and phonetic properties of Southern French. *Phonological variation in French: Illustrations from three continents*, 105-127.

Dell, F. (1978). Certains corrélats de la distinction entre morphologie dérivationnelle et morphologie flexionnelle dans la phonologie du français. *Etudes linguistiques sur les langues romanes. Montreal working papers in Linguistics* 10:1-10.

Dell, F. 1985. *Les règles et les sons.* Paris: Hermann.

Dixon, R. M. W. 1977. Some phonological rules in Yidiny. *Linguistic Inquiry* 8:1-34.

Eychenne, J. (2014). Schwa and the *loi de position* in Southern French. *French Language Studies* 24:223-253.

Fougeron, C., Gendrot C., and Bürki, A. (2007). On the phonetic identity of French schwa compared to /ø/ and /œ/. Paper read at *5emes Journées d'Etudes Linguistiques* (JEL), at Nantes, France.

Fougeron, C. and Steriade, D. (1997). Does deletion of French schwa lead to neutralization of lexical distinctions? In EUROSPEECH 1997, volume 2, 943-946.

Gendrot, C. and Adda-Decker, M. (2005). Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. In *Proceedings of Interspeech'2005 - Eurospeech: 9th European Conference on Speech Communication and Technology*, 2453-2456. Lisbon.

Goslin, J. and Frauenfelder, U. H. (2000). A comparison of theoretical and human syllabification. *Language and Speech* 44:409-436.

Haspelmath, M. and Sims, A. (2010). *Understanding morphology.* London: Routledge.

Hay, J. (2001). Lexical frequency in morphology: Is everything relative?. *Linguistics*, 39(6), 1041-1070.

Kiparsky, P. (1993). Blocking in nonderived environments. In *Studies in lexical phonology,* 277-313. Academic Press.

Malécot, A. and Chollet, G. (1977). The acoustic status of mute-e in French. *Phonetica* 34:19-30.

Martinet, A. and Walter, H. (1973). Dictionnaire de la prononciation française dans son usage réel. Paris: France-Expansion.

New, B., Brysbaert, M., Veronis, J. and Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics* 28:661-677.

Nguyen, N. and Fagyal, Z. (2008). Acoustic aspects of vowel harmony in French. *Journal of Phonetics* 36:1-27.

van Oostendorp, M. (2004). Crossing morpheme boundaries in dutch. *Lingua* 114:1367-1400.

Plag, I. 2018. *Word-formation in English.* Cambridge: Cambridge University Press.

Raffelsiefen, R. (2015). Phonological restrictions on English word-formation. *Word-Formation: An international handbook of the languages of Europe*, Vol. 2., 894-917. Berlin: De Gruyter. URL http://nbn-resolving.de/urn:nbn:de:bsz:mh39-42112.

Robert, P. (Ed.). (2013). *Le Petit Robert : dictionnaire alphabétique et analogique de la langue française (nouvelle édition)*. Paris: Le Robert.

Schriefers, H., Friederici, A., & Graetz, P. (1992). Inflectional and derivational morphology in the mental lexicon: Symmetries and asymmetries in repetition priming. *The Quarterly Journal of Experimental Psychology*, *44*(2), 373-390.

Shepard, R. N. 1972. Psychological representation of speech sounds. In *Human communication: A unified view*, ed. Edward David and Peter Denes. New York, NY: McGraw-Hill.

Steriade, D. (2000). Paradigm uniformity and the phonetics-phonology boundary. *Papers in laboratory phonology V: Acquisition and the lexicon*, *3*, 13-334.

Storme, B. (2017a). The effect of schwa duration on pre-schwa mid-vowel lowering in French. In Jesney, K., O'Hara, C., Smith, C. and Walker, R. (eds.) *Supplemental Proceedings of the 2016 Annual Meeting on Phonology.* http://dx.doi.org/10.3765/amp.v4i0.3984.

Storme, Benjamin (2017b). The *loi de position* and the acoustics of French mid vowels. *Glossa 2*(1):64. https://doi.org/10.5334/gjgl.300.

Storme, B. (2017c). *Perceptual sources for closed-syllable vowel laxing and derived-environment effects.* PhD dissertation, MIT.

Tranel, B. (1987). *The sounds of French: An introduction.* Cambridge: Cambridge University Press.

Walker, D. C. (1993). Schwa and /oe/ in French. *Canadian Journal of Linguistics* 38:43-64.

Wright, R. (2004). A review of perceptual cues and cue robustness. *Phonetically based phonology*, 34-57.

Zuraw, K., and Peperkamp, S. (2015). Aspiration and the gradient structure of English prefixed words. In *ICPhS*.