Je ll'ai vu: Perception-driven allomorphic optimization of French *l'*[1]

Joaquim Brandão de Carvalho
Université Paris 8 / CNRS UMR 7023 SFL
jbrandao@univ-paris8.fr

Abstract
This article aims to explain the optional gemination in the elided form of the French 3rd p. object pronouns *le* and *la*, i.e. before vowel, as in [ʒəllɛvy] for *je l'ai vu(e)* 'I saw him/her/it'. This geminate, which cannot be accounted for in purely phonological terms, is shown to follow from a boundary shift within the morphological sequence /il+lə/la+V/, providing the 3rd p. object pronouns with a new geminate allomorph before vowel; thereby, /ll/ can spread to the entire paradigm. It is argued (a) that the resulting allomorphy is the strategy found by speakers to eliminate the irregular allomorphy of the 3rd p. subject pronoun *il* before consonant; (b) that a perception grammar is needed to capture the reasons for the new allomorphy.

Keywords: allomorphy, pronominal allomorphy, allomorphic choice, allomorphic ranking, recutting, perception grammar, French morpho-phonology.

A well-known feature of many varieties of French, albeit seldom reported by grammars, is that the singular third person object pronouns /lə/ *le* (masculine) and /la/ *la* (feminine) are often geminated *when their vowels are elided*, that is before vowel, as shown in (1).

(1) a. /…lə+a+di…/   [õladikõpetã]   *on l'a dit compétent*
                       [õlladikõpetã]   'he was said to be competent'

     b. /…la+a+di…/   [õladikõpetãt]   *on l'a dit compétente*
                       [õlladikõpetãt]   'she was said to be competent'

     c. /…la+di…/   [õladikõpetãt]   *on la dit compétente*
                       *[õlladikõpetãt]   'she is said to be competent'

In this (non-standard) usage, geminate and single variants of *le* and *la* freely alternate in the speech of the same speakers, even word-initially, as exemplified in (2) (pers. obs.), without any clear correlation between *l'*-gemination and focus position or emphasis.

(2) a. [ʒ(ə)lɛvy] b. [ʒəllɛvy] *je l'ai vu(e)* 'I saw / have seen him/her/it'

[tylasy] [tyllasy] *tu l'as su* 'you knew / have known it'

[õladi] [õlladi] *on l'a dit* 'it was / has been said'

[nulavõfɛ] [nullavõfɛ] *nous l'avons fait* 'we did / have done it'

[ʒ(ə)lɛm] [ʒəllɛm] *je l'aime* 'I love/like him/her/it'

[kilekut] [killekut] *...qui l'écoute* '…who listens to him/her'

[latyvy] [llatyvy] *l'as-tu vu(e)?* 'did you see him/her/it?'

From a historical point of view, only the non-geminate forms are expected: no regular phonetic change explains *l'*-gemination; like other Western Romance languages, French has eliminated all geminates inherited from Latin. Only in modern times, under the influence of spelling, did a small number of (optional) geminate sonorants emerge in such words as *illustre* 'illustrious', *collègue* 'colleague', *sommet* 'top, summit', *inné* 'innate', which shows that the constraint forbidding lexical geminates is no longer active for sonorants in modern French.[2] Yet, *no elided morpheme other than the 3ʳᵈ p. object pronoun* undergoes such a process: neither the elided form of the similar definite articles *le* and *la*, like in *de l'homme*, *à l'homme* 'of the man', 'to the man', nor the elided variant of the 1ˢᵗ p. object pronoun, which also consists of a single sonorant geminate: *tu m'as vu*

---

[2] 3ʳᵈ p. *l'*-gemination is attested since the beginning of the XVIIth century: see Morin (2007: Section 4). In XVIIth c. texts, it is generally rendered as in *je le lai* for *je l'ai* 'I have it'. This pronunciation is regarded as "*bourgeoise*", as opposed to the Court's usage, by François de Callières, in his *Mots à la mode* (1693) quoted by Elias (1977:182).

'you saw me' is realized [tymavy], never *[tymmavy]. How, then, can the forms in (2b) be explained?

This article is organized in five sections. First, I argue that the allomorphy of the 3rd p. object pronoun is closely linked to the distributional irregularity of the 3rd p. *subject* pronoun *il* (Section 1). The need both for a regular allomorphy of *il* and for a more salient exponent of *le*/*la* before vowel is then shown to have favoured a boundary shift between *il* and *l'*, providing *le* with a prevocalic geminate allomorph in colloquial registers, and thereby allowing analogical levelling (Section 2). I next show how this proposal avoids the problems met by previous accounts (Section 3), and summarize its implications for a variationist view of the data, which involves both constraint ranking and allomorphic ordering (Section 4). The conclusion (Section 5) emphasizes the two main results of the proposed explanation for morphological theory: on the one hand, making a new allomorphy may be an optimal strategy if this allomorphy replaces an older one that is less optimal in terms of distributional simplicity; on the other hand, this involves perception and listener-based reanalysis, that is taking the point of view of the hearer, and not only of the speaker.

1 Constraints on allomorphic choice

Why is lengthening restricted to a specific melody, a specific pronoun, and a specific context? As noticed by a large number of scholars (see Section 3), *l*-gemination is a natural result when the 3rd p. object pronouns follow the 3rd p. subject pronouns *il(s)* and *elle(s)*, as shown in (3).

(3) /il+lə+a+di/   [illadi]   *il l'a dit*   'he said it'
     /ɛl+lə+a+di/   [ɛlladi]   *elle l'a dit*   'she said it'

However, this does not explain why gemination spreads to *je ll'ai vu*, *tu ll'as su*, *on ll'a dit*, etc., and occurs even word-initially, as in *ll'as-tu vu?* 'did you see him?' Paradigmatic levelling, if any, typically involves suppletion of one allomorph by another.[3] Hence, as [illadi] is the realization of /il+lə+a+di/, a heteromorphemic cluster cannot expand to other persons. Thus, if the geminate can spread, this means that it *belongs* to the object pronoun, *il* being thus reduced to [i]. Interestingly, one fact partly supports this analysis.

It is well-known that *il(s)* and, to a lesser extent, *elle(s)* show a widespread and very ancient [i] – [il] ([ɛ] – [ɛl]) allomorphy, as exemplified under (4), where both [i] and [il] can be found before consonant, while only [il] occurs before vowel.

(4) a. [isɛm]   b. [ilsɛm]   *il sème*   'he sows'
       [idi]   [ildi]   *il dit*   'he says'
       [kɛskidi]   [kɛskildi]   *qu'est-ce qu'il dit?* 'what does he say?'

     c. [ilɛm]   *il aime*   'he likes / loves'
       [iladi]   *il a dit*   'he said'
       [kɛskiladi]   *qu'est-ce qu'il a dit?* 'what did he say?'

It should be noted that [i] is the most common variant before consonant in colloquial speech, while [il] is felt to be more suitable for formal registers. Historically, realizations like [ilsɛm], [ildi], etc. are clearly "learned" pronunciations, widely dictated by spelling, and relatively recent: Gilles Vaudelin, who proposed in 1713 a quite "phonological"

---

[3] For OT-based views of paradigmatic levelling, see Kenstowicz (1996), Benua (1997), Burzio (2005) and McCarthy (2005).

orthography for French, systematically notes *il* before vowel and *i* before consonant in his *Instructions cretiennes mises en ortografe naturelle* (1715) (Kawaguchi 2009).[4]

The 3rd p. subject pronoun in (4a,c) involves a "light" syllabic pattern that is shared by most French proclitics, whose last consonant, if any, is always mute before consonant, and resyllabified before vowel: *je* ('I'), *tu* ('you, sg.'), *nous* ('we'), *vous* ('you, pl.'), *me* ('me'), *te* ('you, obj.'), *se* ('3rd p. refl.'), *on* ('one, impersonal'), *le* ('him/it'), *la* ('her, obj.'), *les* ('the, pl.; them'), *lui* ('3rd p. dat.'), *un* ('a(n)'), *des* ('indef. pl.'), *ce(t)* ('this, masc.'), *ces* ('these'), *mon* ('my, masc.'), *ton* ('your, masc.'), *son* ('his/her, masc.'), *ma* ('my, fem.'), *ta* ('your, fem.'), *sa* ('his/her, fem.'), *mes* ('my, pl.'), *tes* ('your, pl.'), *ses* ('his/her, pl.'), *nos* ('our, pl.'), *vos* ('your, pl.'), etc. The "heavy" (C)VC pattern conveyed by the learned form in (4b), on the other hand, characterizes a much smaller set of morphemes, such as *cette* ('this, fem.'), *notre* ('our, sg.'), *votre* ('your, sg.'), *leur* ('their; them, dat.'), etc. Interestingly, while all heavy clitics are morphologically marked (feminine demonstrative, plural possessives and dative), all morphologically unmarked clitics are phonologically light. Based on these observations, I propose the constraint, or rather the constraint family, in (5), in line with the approach by Keine and Müller (2015) who propose a formal account to capture correlations between morphosyntactic markedness and phonological length.[5]

---

[4] As opposed to the learned character of [il]+C, some regional varieties show [(i)j] before vowel, which is a low variant effectively repressed by the norm in France, but widespread in Canadian French (Morin 2007) (see Section 3).

[5] In functional-cognitive linguistics, (5a,b) are particular cases of a general iconic principle of isomorphism, the so-called "quantity principle": conceptual complexity corresponds to formal complexity (Givón 1994:49). Note, however, that this bi-uniqueness is not complete: morphological markedness does not necessarily involve phonological markedness, nor does phonological unmarkedness necessarily involve morphological unmarkedness.

(5) CONGRUENCE:

    a. Marked exponents should be assigned to morphologically marked clitics.

    b. Morphologically unmarked clitics should be assigned unmarked exponents.

It follows from (5b) that the masculine singular 3rd p. subject pronoun *il* should have an unmarked exponent.

Two points are worth noting about CONGRUENCE. First, morphological (un)markedness does not refer to a single feature, but to a set of specifications: for example, though plural is the marked term of a privative contrast, *nous* and *vous* show unmarked exponents, but crucially not *notre* and *votre*, which, having an additional possessive feature, satisfy (5a); similarly, despite being possessives, *mon*, *ton*, *son* are phonologically unmarked, but not *notre*, *votre* and *leur*, which are plural as well.[6]

Secondly, CONGRUENCE is supposed to be context-sensitive: [il] does not violate (5b) because it ends in a consonant instead of a vowel, but only when its consonant generates a closed syllable. Thus, CONGRUENCE is satisfied when [il] occurs before vowel, as in *il a dit*, since resyllabification does not yield a marked exponent (i.e. with a coda); it is violated, however, when [il] occurs before consonant, as in *il dit*. Note that all the light proclitics listed above can be said to have one single underlying form containing a monosyllabic CV skeleton and a floating consonant which undergoes *liaison* before vowel (Encrevé 1988). *Il* (along with *elle*) is, thus, different in that the final consonant

---

[6] Like many other constraints, CONGRUENCE may be statistically true although it can be violated in some cases that should require further examination. For example, as opposed to *notre*, *votre*, the plural forms *nos*, *vos*, being light, are phonologically unmarked despite having an additional specification. "Templatic levelling" due to *mes*, *tes*, *ses* (which is historically true by the way) might explain this anomaly, assuming that analogical pressure is not only a diachronically observable effect, but also a synchronic feature of the morphophonological grammar.

may also surface before consonant; in other words, among the morphologically unmarked proclitics, only *il* and *elle* can violate CONGRUENCE.

Clearly, despite its learned origin, the allomorph [il], which also surfaces before vowel, has some kind of priority: assuming with Mascaró (2007; see also Bonet, Lloret and Mascaró 2007) that there is a lexical ranking of allomorphs that determines allomorphic choice when phonological effects are insufficient, [il] will be viewed as the allomorph favoured by (6).

(6)   PRIORITY: Respect lexical priority (ordering) of allomorphs (Mascaró 2007).

PRIORITY and CONGRUENCE have contradictory effects, viz. [il] *vs* [i] for the 3rd p. subject pronoun before consonant. As these allomorphs are in free variation, and characterize formal *vs* colloquial registers of speech, I will assume that the mutual ranking of PRIORITY and CONGRUENCE is also variable, depending on the sociolectal variety.[7] This is illustrated in the tableaux under (7).

(7)   a. 3rd p. subj. / _C (formal speech)

/{il>i}+di/	PRIORITY	CONGRUENCE
[idi]	*!	
☞ [ildi]		*

     b. 3rd p. subj. / _C (colloquial speech)

/{il>i}+di/	CONGRUENCE	PRIORITY
☞ [idi]		*
[ildi]	*!	

---

[7] This is independently supported by the sociolectal variation affecting *liaison*, and the historical restoration of many formerly mute consonants in French (cf. Carvalho and Klein 2010). As to how sociolectal variation and style levels can be accounted for by factorial typologies of constraint rankings, see e.g. Anttila (1997, 2002) and van Oostendorp (1997).

As regards /{il>i}+a+di/, PRIORITY, whatever its ranking, concurs with ONSET in selecting [iladi] instead of *[iadi], CONGRUENCE being satisfied in both formal and colloquial styles.

However, albeit marked and much less frequent in colloquial speech, the allomorph [il] has an interesting, and problematic, effect on the realization of the 3rd p. object pronoun. Let us compare the forms in (8, 9) with those in (10).

(8) /il+di/ *Il dit* 'he says'   /il+mə+di/ *Il me dit* 'he says to me'
     a. [idi]   c. [imədi]   e. [imdi]
     b. [ildi]   d. [ilmədi]

(9) /ty+di/ *Tu dis* 'you say'   /ty+lə+di/ *Tu le dis* 'you say it'
     a. [tydi]   b. [tylədi]   c. [tyldi]

(10) /il+di/ *Il dit* 'he says'   /il+lə+di/ *Il le dit* 'he says it'
     a. [idi]   c. [ilədi]   e. ?[ildi]
     b. [ildi]   d. [illədi]

The interesting point here is the interpretation of [ildi]. In certain cases, this might be a possible phonetic rendition of *il le dit*, namely, as pointed out by one reviewer, in sentences involving left dislocation of direct object NP's (e.g., [saildisɑ̃fwapaʁʒuʁ] *Ça, il le dit cent fois par jour* 'he says that a hundred times a day'). In most cases, however, especially out of context, a native speaker will preferentially interpret [ildi] as *il dit*, even in colloquial registers, in which this sequence is realized as [idi], not as [ildi]. It follows that the 3rd p. object pronoun, when it is preceded by the 3rd p. subject pronoun, *and only in this case,* cannot be satisfactorily perceived unless it preserves not only its consonant,

but also its vowel, in such a way that the entire underlying form of the morpheme surfaces as in (10c,d). Two points are worth highlighting.

Firstly, (10e) does not violate any high-ranked markedness constraint, as shown by (10b). Clearly, what generally prevents [ildi] from being the realization of /il+lə+di/ is (10b) [ildi] for /il+di/, where [l] is assigned to the subject pronoun *il*, not to *le*.

Secondly, not only does (10b) prevent [ildi], which does not violate any high-ranked markedness constraint, from being the preferred realization of /il+lə+di/, but it also generally compels /ə/ to surface in (10c) [ilədi], although this pronunciation *does* violate the well-known constraint of (non-southern varieties of) French under (11).

(11) *VCəCV: No schwa within a two-consonant cluster.[8]

By virtue of *VCəCV, (8e) [imdi] and (9c) [tyldi] are generally preferred to (8c) [imədi] and (9b) [tylədi], schwa being only obligatory in French when it helps to avoid heavy consonant clusters (as in [tymlədi], [tyməldi] *tu me le dis* 'you say it to me').[9] Thus, *VCəCV is a low-ranked constraint: not only is it generally violated by the 3rd p. object pronoun in (10c), but also, albeit less systematically, elsewhere, as in (8c) and (9b).

In sum, [ildi] and [i(l)lədi] for /il+di/ and /il+lə+di/ respectively suggest that some sort of faithfulness is at stake here, although it is not directly based on phonological objects

---

[8] Actually, *VCəCV is either an epiphenomenon of other constraint interaction, as observed by one reviewer, or simply the effect of the Empty Category Principle, in line with the Government Phonology approach: an empty nucleus does not surface if it is properly governed by the following nucleus (see, for example, Charette 1991).

[9] CCC-clusters are allowed in French only if they contain branching (*muta cum liquida*) onsets (e.g., *marbré*) and/or the fricative /s/ (*austral*, *e*[ks]*tase*).

like segments or moras. What the output must be faithful to is the pronoun itself: /il+lə+di/ should not be realized as [ildi], because this output is preferably perceived as the realization of /il+di/.

Similarly, [iladi] may be, under certain circumstances, a possible rendition of *il l'a dit*, as in (12c) – namely, once again, in cases of very close antecedents (e.g. [siiladi] *Si, il l'a dit!* 'Yes, he did say it!', [salɛtʁilaɑ̃vwajediʁɛktəmɑ̃upa] *Sa lettre, il l'a envoyée directement ou pas?* '[as to] his letter, did he send it directly or not?'). Nevertheless, native speakers will generally interpret [iladi] as *il a dit*, as in (12a).

(12) /il+a+di/   *Il a dit* 'he said'   /il+lə+a+di/   *Il l'a dit* 'he said it'
    a. [iladi]   b. [illadi]   c. ?[iladi]

Since hiatuses containing [əV] (or identical vowels like [aa]) are disfavoured in French, the geminate *l* in (12b), although it violates CONGRUENCE, appears as the only available means for the object pronouns *le* and *la* to be perceived before vowel in both formal and colloquial registers.

Let us, thus, assume the two constraints in (13).

(13) a. MAX_{MPH}:   Every morpheme of the input has a correspondent in the output (Bonet, Lloret and Mascaró 2015).

    b. NOHIATUS: No hiatus containing either identical vowels or əV is tolerated.[10]

---

[10] As is well known (see e.g. Encrevé 1988), the only apparent violations of this constraint involve the so-called '*h* aspiré' (as in [ləibu] *le hibou* 'the owl', [laaʃ] *la hache* 'the axe'). Otherwise, different means are employed to avoid hiatuses (cf. *ma fille* 'my daughter' vs *mon amie* 'my friend, fem.').

2 From realization to interpretation

The constraints that have hitherto been defined are ranked as in the tableaux under (14, 15), in which pronominal morphemes are coindexed with their inputs.[11] MAXMPH outranks PRIORITY and CONGRUENCE, as can be deduced from (15); so does NOHIATUS. The ranking of *VCəCV cannot be explicitly demonstrated; it is arbitrarily placed at the lowest rank in (14), since it is often optionally violated, as in (8d, 9b, 10c).

(14) a. 3rd p. subj. + 3rd p. obj. / _C (formal speech)

/{il>i}$_1$+lə$_2$+di/	MAXMPH	PRIORITY	CONGRUENCE	*VCəCV
[i$_1$di]	*!	*		
[il$_1$di]	*!		*	
[i$_1$l$_2$di]		*!		
[i$_1$lə$_2$di]		*!		*
☞ [il$_1$lə$_2$di]			*	

  b. 3rd p. subj. + 3rd p. obj. / _C (colloquial speech)

/{il>i}$_1$+lə$_2$+di/	MAXMPH	CONGRUENCE	PRIORITY	*VCəCV
[i$_1$di]	*!		*	
[il$_1$di]	*!	*		
☞ [i$_1$l$_2$di]			*	
☹ [i$_1$lə$_2$di]			*	*!
[il$_1$lə$_2$di]		*!		

(15) a. 3rd p. subj. + 3rd p. obj. / _V (formal speech)

/{il>i}$_1$+lə$_2$+a+di/	MAXMPH	NOHIATUS	PRIORITY	CONGRUENCE
[il$_1$adi]	*!			
☞ [il$_1$l$_2$adi]				*
[i$_1$l$_2$adi]			*!	
[i$_1$lə$_2$adi]		*!	*	
[il$_1$lə$_2$adi]		*!		*

---

[11] I leave aside [il$_1$l$_2$di] in (14), as it is ruled out by a high-ranked *CCC constraint which forbids most complex clusters in French (cf. note 9).

b. 3rd p. subj. + 3rd p. obj. / _V (colloquial speech)

/{il>i}$_1$+lə$_2$+a+di/	MAXMPH	NOHIATUS	CONGRUENCE	PRIORITY
[il$_1$adi]	*!			
☹ [il$_1$l$_2$adi]			*!	
☞ [i$_1$l$_2$adi]				*
[i$_1$lə$_2$adi]		*!		*
[il$_1$lə$_2$adi]		*!	*	

As can be seen, constraint computation does not predict (all) the expected results for colloquial registers: it gives [i$_1$l$_2$di] and [i$_1$l$_2$adi], but not [i$_1$lə$_2$di] and [il$_1$l$_2$adi]. In other words, it predicts that *il le dit* and *il l'a dit* are respectively rendered as (10e) [ildi] and (12c) [iladi], which have been said to be preferably perceived as the realization of *il dit* and *il a dit*, even in colloquial registers, in which the former sequence is pronounced [idi], not *[ildi]: cf. (7b). This is because MAXMPH does not achieve what it is supposed to do: both [i$_1$l$_2$di] and [i$_1$l$_2$adi] do satisfy MAXMPH, as every input morpheme can be said to be *realized* by an output allomorph. We thus need not a constraint on morpheme realization, but a constraint on interpretation by the listener, as in Functional Phonology (Boersma 1998, 1999, 2009), or in Bidirectional Optimality Theory (cf. Benz and Mattausch 2011). In sum, what is at stake here is a constraint on *underlying representations*, as is required by a perception grammar.

Why, then, are the two winning candidates [i$_1$l$_2$di] and [i$_1$l$_2$adi] in (14b, 15b) bad? For one basic reason: because their inputs violate the OCP in a perception grammar (cf. Boersma 2000), as they involve two identical contiguous melodies (/l+l/). This is all the more crucial in the case of the masculine object pronoun as the melodic material of /lə/ is entirely enclosed within that of /il/, schwa being an empty (i.e. featureless) nucleus, as shown in (16).

(16)   O N
         |
         l

But the feminine object pronoun is exposed to the same risk with a vowel-initial verb, as [ila(di)] may refer either to /il+la+a/ (*il l'a dit*) or, more probably, to /il+a/ (*il a dit*).

Thus, the underlying representations /il+lə-la+di/ and /il+lə-la+a+di/ violate a constraint that could be formulated as in (17), and whose evaluation will be made clear in (22), once the candidates have been defined within a perception grammar.

(17) NoOVERLAP: Adjacent morphemes should have distinct melodic content.

Like CONGRUENCE (cf. note 5), this constraint may be motivated by the general iconic principle of isomorphism: "a bi-unique correspondence tends to be established between signans and signatum" (Haiman 1980) in the speech chain; the sequential order is optimal if there is no overlap.

NoOVERLAP explains that speakers have readjusted the underlying representations of the 3rd p. object pronoun. From a historical point of view, they appear to have done so by two different means – before consonant and before vowel – that show an interesting common feature nevertheless. Let us first consider the case of consonant-initial verbs, as in (14b). I have assumed that [lə] and [l] (in all contexts) derive from a single underlying representation /lə/. However, the colloquial form [ilədi] for *il le dit*, where schwa is exceptionally favoured in a potential two-consonant cluster (as opposed to [tyldi] for *tu le dis*), challenges this hypothesis, and may bring evidence – together with word-final realizations like [dilø] *dis-le* 'say it!' contra Old French [dilᵊ], and *ceci* 'this' = *ceux-ci* 'these ones' = [søsi] in modern Parisian usage – for assuming that schwa is merging with

the "normal", fully specified, vowel phonemes /ø, œ/, namely in clitics (Hansen 1994, Walker 1996). In sum, the strategy found by speakers to prevent *le* from being totally masked by *il* before consonant consists in replacing the schwa of *le* with /ø, œ/, which, unlike the empty vowel in (16), cannot undergo deletion, as they contain the elements I, U and A assumed by unarist models (or the corresponding features). For the sake of brevity, however, I will leave aside the case of consonant-initial verbs, which involves the intricate problem of the phonological status of French schwa.

With vowel-initial verbs, the boundary shift under (18) will be assumed to have occurred within the morphemic sequence, giving rise to a new allomorph of *le* before vowel – /ll/ – in colloquial registers.

(18) /il+lə+V/ → /i+ll+V/

Both strategies – schwa becoming a full vowel and recutting (cf. Diertani 2011: 212-226) – enhance the prosodic autonomy of the 3rd p. object pronoun by bringing (at least) one additional mora (μ) to the verb, as is illustrated, before vowel, in (19), where the examples are repeated from (2), and the resulting heavy syllable is underlined.

(19) a. [ʒ(ə)lɛvy] C(V).CV.CV   = 2/3 μ   b. [ʒəllɛvy] <u>CVC</u>.CV.CV   = 4 μ

   [tylasy]   CV.CV.CV   = 3 μ   [tyllasy]   <u>CVC</u>.CV.CV   = 4 μ

   [õladi]   V.CV.CV   = 3 μ   [õlladi]   <u>VC</u>.CV.CV   = 4 μ

   [nulavõfɛ] CV.CV.CV.CV = 4 μ   [nullavõfɛ] <u>CVC</u>.CV.CV.CV = 5 μ

   [ʒ(ə)lɛm] C(V).CVC   = 2/3 μ   [ʒəllɛm]   <u>CVC</u>.CVC   = 4 μ

   [kilekut] CV.CV.CVC   = 4 μ   [killekut] <u>CVC</u>.CV.CVC   = 5 μ

   [latyvy]   CV.CV.CV   = 3 μ   [llatyvy]   <u>VC</u>.CV.CV.CV   = 4 μ

Let us turn back to the synchronic analysis of what is presumably internalized by native speakers. Phonologically, the recutting in (18) can be analysed as involving a morpheme-specific compensatory lengthening: the loss of the schwa of /lə/ before vowel is compensated by geminating the preceding consonant, that is, according to a Strict CV approach, by /l/ spreading onto the empty onset of the verb, as shown in (20). To paraphrase Encrevé's (1988) *liaison sans enchaînement*, this could be treated as a kind of *élision sans enchaînement*.[12]

(20)   /lə+a/ → /ll+a/   *(tu) l'as (dit)*

```
O   N   O   N
|‸‸‸‸‸‸‸    |
l           a
```
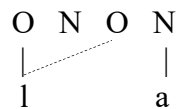
This approach is all the more interesting as it motivates the constraint under (21), which bans the geminated allomorph before consonant, and therefore candidates like *[ill(ə)di] for *il le dit*: consonant-initial verbs provide no empty O-slot for /l/ propagation.

(21)   *llC: No geminate before consonant.

The emergence of the new prevocalic allomorph /ll/ of the 3rd p. object pronoun in colloquial styles has five interesting consequences. First, we no longer need to assume that *il* has two lexicalized allomorphs in those registers: a unique form /i(l)/ with a floating

---

[12] Whatever formalism may be adopted, in theories on exceptional morphophonological patterns such as Pater's (2009) constraint indexation approach, it would be expected that processes like these can be restricted to single morphemes or morpheme combinations.

consonant underlies [il] before vowel, as in [iladi] for *il a dit*, and [i] before consonant, as in both [idi] *il dit* and [illadi] *il l'a dit*; *il* now behaves like most French proclitics.[13]

Secondly, although the recutting in (18) may seem to cause some complexity, the resulting allomorphy is distributionally regular as well: /lə/ (and /la/) before consonant, /ll/ before vowel.

Thirdly, it follows that the competition between the old and the new underlying representations of *il l'* is easily won by the latter, as shown in (22).

(22) Perceptual computation of the object pronoun (colloquial registers)

[illadi]	NoOverlap
/il+lə+a+di/	*!
/il+la+a+di/	*!*
☞ /i+ll+a+di/	

Note that the OCP is not violated by the third candidate, as /ll/ is a shorthand notation for one single melody associated with two slots, as in (20).

Fourthly, the resulting production grammar gives the expected results in colloquial registers, as the new geminated allomorph of the 3rd p. object pronoun, which involves an empty nucleus in (20), satisfies *before vowel* the constraint MAX_μ in (23), which captures the generalization illustrated in (19).

---

[13] It remains to be explained, however, why *il* keeps its /l/ word-finally (cf. *le fait*-[il]? 'does he make it?)', whereas floating consonants remain silent at the pause (cf. [vulevu] *voulez-vous?* 'if you please', [dɔnle] *donne-les* 'give them', [ilepti] *il est petit* 'he is small'). This might be due to the following bias: all cases of pre-pausal *il* belong to a highly formal syntax; as seen in Section 1, in formal styles *il* has an allomorph [il] in all contexts, that is one in which /l/ is *not* a floating consonant (and which, thus, violate CONGRUENCE).

(23) MAX$_\mu$:   The exponent of a morpheme must preserve the weight of its input.

This is shown in (24), where CONGRUENCE now concerns the allomorphy of *le*/*la*. As can be seen, owing to the existence of a new geminated allomorph, MAX$_\mu$ achieves what MAX$_{\text{MPH}}$ in (15b) was unable to do: selecting [illadi] instead of [iladi].

(24)   3rd p. subj. + 3rd p. obj. / _V (colloquial speech)

/i(l)$_1$+{lə-la,ll}$_2$+a+di/	MAX$_\mu$	NoHiatus	Congruence
☞ [i$_1$ll$_2$adi]			*
[i$_1$l$_2$adi]	*!		
[i$_1$lə$_2$adi]		*!	
[i$_1$la$_2$adi]		*!	

Fifthly, and this is what independently demonstrates my analysis, the recutting in (18) allowed analogical levelling; otherwise, it is impossible to explain neither why the geminate can spread to all persons, as in *je ll'ai vu, tu ll'as su, on ll'a dit*, etc., nor why this only takes place before vowel.[14]

In sum, making the geminate tautomorphemic appears as a strategy found by speakers for eliminating the irregular *il*-allomorphy in colloquial speech, since [il] no longer occurs before consonant. Geminate lexicalization is thus an instance of "allomorphic

---

[14] The same has probably occurred with some regional (especially Picard and Canadian) pronunciations of the anaphoric pronoun *en*: [ʒənne] *j'en ai* 'I have some', [inna] *il en a* 'he has some', [nnaty] *en as-tu?* 'have you got any?' (Morin 2007: Section 1, p.c.). Assuming that *en* [ã, ãn] (< Old Fr. *en(t)* < Lat. *ĭnde*) has a dialectal *ne* variant (as in Provençal or Italian), gemination also affects the elided allomorph (*je n'ai, i(l) n'a, n'as-tu?*). In this case, the triggering context of the geminate allomorph could easily be found in such pairs as *on a vu* [onavy] with *n*-liaison 'one/we saw' / *on n'a vu* (= *on en a vu*) [onnavy] 'one/we saw some'. Note that the subject pronoun *on* is all the more frequent as it is commonly used instead of *nous* 'we' in colloquial speech.

optimization".[15] Furthermore, capturing this idea requires that optimization be computed also on the basis of interpretation by the listener (see Blutner 2000), and not only of realization by the speaker; in other words, both production and perception grammars are needed within a bidirectional perspective.

3 Solving problems met by previous accounts

That the extension of *ll'* to all persons should result from analogical levelling, caused by the 3rd p. subject pronoun and such problematic distinctions as *il a dit* / *il l'a dit*, is not a new thesis. Morin (2007: Section 5.2), who provides the most comprehensive survey of 3rd p. *l'*-gemination and of the explanations that have been proposed for over a century, quotes a number of scholars who defended this point of view. However, as Morin justly observes, these works suffer from several inadequacies. In what follows, I will briefly mention three arguments advanced by Morin against those previous accounts, and show why they do not apply to the explanation proposed here.

A first problem with some of the studies quoted by Morin is that they fail to provide any analogical model. However, this is due to the heteromorphemic status of the geminate, which disallows levelling; the problem disappears if the change in (18), yielding a tautomorphemic *ll'*, is assumed.

---

[15] Of course, the elimination of the *il*-allomorphy may entail, as noted above, the emergence of a new irregular allomorph of *le*: /lø/ between *il* and consonant. Note, however, that (i) unlike /il/ before *l'*, /lø/ satisfies CONGRUENCE; (ii) it remains to be seen if there is an ongoing levelling whereby /lø/ spreads to the other persons at the expense of /lə/, which means that pronunciations like [tylødi] for *tu le dis* 'you say it' are becoming more and more frequent at the expense of [tyldi]; this might well be the case (cf. Hansen 1994, Walker 1996)

A second inadequacy of previous works is that they predict geminate overgeneralization, that is, broadly expressed, *il la voit* 'he sees her' → **tu lla vois* 'you see her'. However, the analysis proposed above explains why gemination is restricted to elided pronouns: recutting requires an empty onset. Hence, there is no such thing as */llə, lla/ in the input.

A third and seemingly more serious difficulty arises from dialects like Canadian French where *il* gave [i] even before vowel, *and* where *l'* is nevertheless pronounced [ll] (Pupier and Pelchat 1972; Pupier and Légaré 1973; Picard 1977, 1990; Bougaïeff and Cardinal 1980; Morin 2007: Section 5.3). Clearly, if *il* has lost its consonant, the expected realization of *il a dit* / *il l'a dit* is [iadi] / *[iladi], as there is no motivation for gemination. The analysis proposed in this article, however, allows a simple explanation for this. Historically, in Canadian French not only did the change in (18) take place, which eliminated *il*-allomorphy, as in European French colloquial registers, but *il* eventually lost its floating consonant later on.[16] Otherwise, the resulting grammar, under (25), is similar to that of (colloquial) European French in (24).

(25)   3rd p. subj. + 3rd p. obj. / _V (Canadian French)

/i$_1$+{lə-la,ll}$_2$+a+di/	MAX$_\mu$	NoHiatus	Congruence
☞ [i$_1$ll$_2$adi]			*
[i$_1$l$_2$adi]	*!		
[i$_1$lə$_2$adi]		*!	
[i$_1$la$_2$adi]		*!	

---

[16] Note that French final liquids either fell out (e.g., [ʃãte] *chanter* 'sing', [ky] *cul* 'ass', [fyzi] *fusil* 'gun') or became "fixed" ([finiʁ] *finir* 'finish', [fil] *fil* 'thread').

Morin's (1979: 24-25) own solution, like my proposal, is based on lexical gemination. However, his explanation is highly unlikely, as it supposes that the geminate dates back to Lat. *ille*, *illa*, and was preserved until Modern French in these sole morphemes.

4 Summary: allomorphy, change and variation

From the constraint-based account proposed above it follows that there are two main grammars which govern the interaction between the subject and object pronouns. In the first, PRIORITY outranks CONGRUENCE (cf. [ildi] for *il dit*), and:

(26) a. /il/ has a marked allomorph (where /l/ is anchored to the skeleton).

   b. /lə/ and /la/ are the underlying representations of the 3rd p. object pronouns.

This grammar is exemplified in the tableaux in (27), repeated from (14a, 15a).

(27) a. 3rd p. subj. + 3rd p. obj. / _C (formal speech)

/{il>i}₁+lə-la ₂+di/	MAXMPH	PRIORITY	CONGRUENCE	*VCəCV
[i₁di]	*!	*		
[il₁di]	*!		*	
[i₁l₂di]		*!		
[i₁lə₂di]		*!		*
☞ [il₁lə₂di]			*	

   b. 3rd p. subj. + 3rd p. obj. / _V (formal speech)

/{il>i}₁+lə-la ₂+a+di/	MAXMPH	NOHIATUS	PRIORITY	CONGRUENCE
[il₁adi]	*!			
☞ [il₁l₂adi]				*
[i₁l₂adi]			*!	
[i₁lə₂adi]		*!	*	
[il₁lə₂adi]		*!		*

In the second grammar, which characterizes colloquial speech, CONGRUENCE outranks PRIORITY (cf. [idi] for *il dit*). As was seen in Section 2, (26a,b) then become problematic, which led speakers to create new underlying representations: /i(l)/ (with a floating /l/), /lø/ (before consonant) and /ll/ (before vowel).[17] This grammar is laid out in the tableaux under (28).

(28) a. 3rd p. subj. + 3rd p. obj. / _C (colloquial speech)

/i(l)$_1$+{lø-la,ll}$_2$+di/	MAX$_{IO}$	MAX$_\mu$	*llC	CONGRUENCE
☞ [i$_1$lø-la$_2$di]				
[il$_1$lø-la$_2$di]				*!
[i$_1$l$_2$di]	*!			
[i$_1$ll$_2$di]			*!	

b. 3rd p. subj. + 3rd p. obj. / _V (colloquial speech)

/i(l)$_1$+{lø-la,ll}$_2$+a+di/	MAX$_{IO}$	MAX$_\mu$	NOHIATUS	CONGRUENCE
[i$_1$lø-la$_2$adi]			*!	
[il$_1$lø-la$_2$adi]			*!	*
[i$_1$l$_2$adi]		*!		
☞ [i$_1$ll$_2$adi]				*

Part of French speakers has eventually generalized /ll/ to the entire paradigm, since *l'*-gemination at all persons remains optional for many speakers (see Morin 1979), geminate forms often coexisting with non-geminate ones, as exemplified in (2). Further research is needed to determine the geographical extension and/or the sociolinguistic distribution of this analogical levelling.

---

[17] Which were added to /lə/. Let us recall that *tu le dis*, for example, can always be pronounced as [tyldi] in all styles.

5 Conclusion

We can learn several lessons from the rather complex behaviour of the 3rd p. pronouns sketched above. For one thing, the emergence of the /lə/-/ll/ allomorphy clearly demonstrates that sonorant gemination, as the one that optionally appears in such words as *collègue*, *sommet*, has been phonologized in modern French. Also, it shows the relevance of Mascaró's notion of allomorphic ranking, which, just like constraint hierarchies, may capture change and variation. Another point concerns the scope of what Elsman & Holt (2009) call "phonological compensation" to morphological reduction (here elision): just as in the old Leonese contractions analysed by these authors "the burden of morphological representation shifts from the segmental to the featural level in order to prevent the complete loss of surface forms that are already short", so may it shift from the segmental to the prosodic level as well, only the "moraic" content of morphemes being preserved, that is, as shown by the compensatory lengthening in (20), only their skeletal basis.

But the most interesting outcomes concern why and how allomorphies appear. On the one hand, as is well-known, lexical allomorphy, involving distinct, listed, and weakly suppletive allomorphs (as opposed to the *go*/*went* type), are often "accidents" due to phonologization of previously allophonic alternations. The origin of the /lə/-/ll/ allomorphy, however, does not fall into this widespread category: actually, phonology – taken as the module that is responsible for the "emergence of the unmarked" (McCarthy and Prince 1994) – plays no active role in it: geminates are marked, especially in French. As a result, the emergence of allomorphies may be phonology-free under certain conditions, which runs counter to the neogrammarian view that, without support from phonology, no allomorphy is produced.

On the other hand, since the allomorphy of *le* turns out to be an indirect effect of that of *il*, which is eliminated before consonant by the former, it follows that allomorphies may be more or less optimal. Our French example clearly shows that not only is an allomorph (/i/) selected at the expense of another (/il/), but also that speakers may replace an old problematic allomorphy (/il/-/i/) with a new one (/lə/-/ll/), which outclasses the former in terms of distributional simplicity, rather than phonetic naturalness.

Lastly, as the emergence of allomorphy supposes phonologization giving rise to a new underlying form, such processes are partly based on how the listener interprets the phonetic string in terms of categories. This is particularly clear in the case of recutting, as was shown above. Therefore, a bidirectional perspective is needed to capture the emergence and optimization of allomorphs, in which the crucial step – interpretation – is assigned to the perception grammar.

References

Anttila, Arto. 1997. Deriving variation from grammar. In Frans Hinskens, Roeland van Hout & Leo Wetzels (eds.), *Variation, change and phonological Theory*, 35-68. Amsterdam: John Benjamins.

Anttila, Arto. 2002. Variation and Phonological Theory. In Jack Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*, 206-243. Oxford: Blackwell.

Benua, Laura. 1997. *Transderivational identity: phonological relations between words*. University of Massachusetts Amherst doctoral dissertation.

Benz, Anton & Jason Mattausch. 2011. *Bidirectional Optimality theory*. Amsterdam: John Benjamins.

Blutner, Reinhard. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17. 189-216.

Boersma, Paul. 1998. *Functional phonology: Formalizing the interactions between articulatory and perceptual drives* (LOT International Series 11). The Hague: Holland Academic Graphics.

Boersma, Paul. 1999. On the need for a separate perception grammar. *Rutgers Optimality Archive* 358. http://www.fon.hum.uva.nl/paul/papers/percgram.pdf (accessed 12 April 2017).

Boersma, Paul. 2000. The OCP in the perception grammar. *Rutgers Optimality Archive* 435. http://www.fon.hum.uva.nl/paul/papers/ocp_perc.pdf (accessed 12 April 2017).

Boersma, Paul. 2009. Cue constraints and their interactions in phonological perception and production. In Paul Boersma & Silke Hamann (eds.), *Phonology in perception*, 55–110. Berlin: Mouton de Gruyter.

Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró. 2007. Allomorph selection and lexical preferences: Two case studies. *Lingua* 117. 903-927.

Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró. 2015. The prenominal allomorphy syndrome. In Eulàlia Bonet, Maria-Rosa Lloret & Joan Mascaró (eds.), *Understanding allomorphy. Perspectives from Optimality theory*, 5-44. London: Equinox.

Bougaïeff, André & Pierre Cardinal. 1980. La chute du /l/ dans le français populaire du Québec. *La Linguistique* 16. 91-102.

Burzio, Luigi. 2005. Sources of paradigm uniformity. In Laura J. Downing, T. Alan Hall & Renate Raffelsiefen (eds.), *Paradigms in phonological theory*, 65-106. Oxford: Oxford University Press.

Carvalho, Joaquim Brandão & Marc Klein. 2010. Opacité phonologique et liaison en français : De la sous-détermination de la variable à la motivation des variantes. *Langue Française* 168. 53-70.

Charette, Monik. 1991. *Conditions on phonological government.* Cambridge: Cambridge University Press.

Diertani, Ariel. 2011. *Morpheme boundaries and structural change: Affixes running Amok*. Philadelphia: University of Pennsylvania dissertation.

Elias, Norbert. 1977. *La civilisation des mœurs*. Paris: Calmann-Lévy.

Elsman, Minta & D. Eric Holt. 2009. When small words collide: Morphological reduction and phonological compensation in old Leonese contractions. In Ronald Leow, Héctor Campos & Donna Lardiere (eds.), *Little words: Their history, phonology, syntax, semantics, pragmatics, and acquisition*, 21-33. Washington, D.C.: Georgetown University Press.

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

Givón, Thomas. 1994. Isomorphism in the grammatical code: Cognitive and biological considerations. In Raffaele Simone (ed.), *Iconicity in language*, 47-76. Amsterdam: John Benjamins.

Haiman, John. 1980. Iconicity and grammar: Isomorphism and motivation. *Language* 56. 515-540.

Hansen, Anita Berit. 1994. Etude du E caduc : stabilisation en cours et variations lexicales. *Journal of French Language Studies* 4. 25-54.

Kawaguchi, Yuji. 2009. Liaison au XVIIIème siècle. Analyse préliminaire des textes de Gile Vaudelin. Paper presented at the Colloque PFC [Phonologie du Français Contemporain], Paris, 10-12 December.

Keine, Stefan & Gereon Müller. 2015. Differential argument encoding by impoverishment. In Ina Bornkessel-Schlesewsky, Andrej Malchukov & Marc D. Richards (eds.), *Scales and hierarchies: A cross-disciplinary perspective*, 75-130. Berlin: Mouton de Gruyter.

Kenstowicz, Michael. 1996. Base identity and uniform exponence: Alternative to cyclicity. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*, 363- 393. Paris-X and Salford: University of Salford Publications.

Mascaró, Joan. 2007. External allomorphy and lexical representation. *Linguistic Inquiry* 38. 715-735.

McCarthy. John J. 2005. Optimal paradigms. In Laura J. Downing, T. Alan Hall & Renate Raffelsiefen (eds.), *Paradigms in phonological Theory*, 170-210. Oxford: Oxford University Press.

McCarthy, John J. & Alan Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. *Papers from the Annual Meeting of the North East Linguistic Society* 24. 333-379.

Morin, Yves Charles. 1979. La morphophonologie des pronoms clitiques en français populaire. *Cahier de Linguistique* 9. 1-36.

Morin, Yves Charles. 2007. *Tu ll'as vu?* Les sources de la gémination du pronom clitique objet élidé en français. Paper presented at the XXV$^{\text{ème}}$ Congrès International de Linguistique et de Philologie Romanes, Innsbruck, 3-8 September.

van Oostendorp, Marc. 1997. Style levels in conflict resolution. In Frans Hinskens, Roeland van Hout & Leo Wetzels (eds.), *Variation, change and phonological theory*, 207–229. Amsterdam: John Benjamins.

Pater, Joe. 2009. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.), *Phonological argumentation: Essays on evidence and motivation*, 123-154. London: Equinox.

Picard, Marc. 1977. Les règles morphophonologiques en diachronie. *Recherches Linguistiques à Montréal* 8. 129-135.

Picard, Marc. 1990. On morphologically-conditioned sound change: The deletion and gemination of *l* in Canadian French. *Probus* 2. 103-111.

Pupier, Paul & Roland Pelchat. 1972. Observations sur la phonologie des pronoms personnels du français de Montréal. *Lingua* 29. 326-346.

Pupier, Paul & Luc Légaré. 1973. L'effacement de /l/ dans les articles définis et les clitiques en français de Montréal. *Glossa* 7. 69-80.

Walker, Douglas C. 1996. The new stability of unstable *-e* in French. *Journal of French language studies* 6. 211-229.