

- b. *il* *lui* *le* *montre*
 he.NOM:3.Sg:Masc him.DAT:3.Sg:Masc it.ACC:3.Sg:Masc
 show.PRES:3.Sg
- (2) a. ***demande*** ***-le*** ***-lui***
 ask.IMP:2.Sg it.ACC:3.Sg:Masc him.DAT:3.Sg:Masc
- b. *demande* *-lui* *-le*
 ask.IMP:2.Sg him.DAT:3.Sg:Masc it.ACC:3.Sg:Masc
- c. ***rends*** ***-le*** ***-moi***
 give-back.IMP:2.Sg it.ACC:3.Sg:Masc me.DAT:1.Sg
- d. *rends* *-moi* *-le*
 give-back.IMP:2.Sg me.DAT:1.Sg it.ACC:3.Sg:Masc
- (3) a. ***rends*** ***-le*** ***-moi (=2c)***
 give-back.IMP:2.Sg it.ACC:3.Sg:Masc me.DAT:1.Sg
- b. *rends* *-me* *-le*
 give-back.IMP:2.Sg me.DAT:1.Sg it.ACC:3.Sg:Masc

In (1b), the 3SG.DAT clitic pronoun *lui* precedes the 3MSG.ACC clitic pronoun *le*, contrary to what is expected in the standard variant (example (1a)). In examples (2b) and (2d), in postverbal position (only possible in standard French in positive imperatives, see Heap *et al.* forthcoming, section 3.2), 1SG.DAT and 3SG.DAT clitics precede the 3MSG.ACC *le*, again in opposition to the standard French order (examples (2a) and (2c)). Finally, example (3b) illustrates the clitic morphophonological variant *me* of the 1SG.DAT/ACC weak pronoun form *moi* (see Cardinaletti/Starke 1999, Heap *et al.* forthcoming: section 1), the regular one in preverbal position, but not found in standard French in enclitic position (example (3a)). Except for this last observation, case seems to be at stake here, i.e. a general preference for consistent DAT-ACC ordering in French clitic clusters in non-standard varieties.

As for the first variant (example (1b)), *lui le* instead of *le lui* (3SG.DAT – 3MSG.ACC vs. 3MSG.ACC – 3SG.DAT) in preverbal, proclitic position, Miller/Monachesi (2003: 109) label it "regional French", Grévisse/Goosse (2008: §682) localize it more precisely ("Southern France, Bretagne"), Violin-Wigent (2010) even more so ("Briançon"), and

Tuaille (1983: 234) reports that it is found in "[...] toute la basse vallée de l'Isère, la région de Valence, une bonne partie de l'Ardèche [...]"¹. The second variant (example (2b)), *lui le* instead of *le lui* (3SG.DAT – 3MSG.ACC vs. 3MSG.ACC – 3SG.DAT) in postverbal, enclitic position, seems to be clearly confined regionally: it is said to be found in Bretagne, Ardèche and Isère (cf. Dagnac 2012: 6), as well as in Gallo (i.e. the Eastern part of Bretagne) and Saint-Etienne regional French (Morin 1979a: 308). Locations are given in Figure 1.

Insert Figure 1 about here

Figure 1. Areas where the non-standard variants (1b) and (2b) are attested according to the existing literature.

The third variant (example (2d)), *moi-le* instead of *le-moi* (1SG.DAT – 3MSG.ACC vs. 3MSG.ACC - 1SG.DAT) in enclitic position, with the weak pronoun *moi* instead of the clitic *me*, represents an exception to the preceding ones, as this variation is more often considered to be stylistic (see e.g. De Kok 1985: 378; Miller/Monachesi 2003: 39; Grévisse & Goosse 2008: §683) than geographical. The descriptive grammar of Riegel et al. (1994: 204) labels it as a variant of colloquial French ("langue familière"), as does Laenzlinger (1994: 88), while Dauzat (1930: 432) even considered it as "vulgaire".

Finally, the last variant, with the same order and context but different morphophonological material to the third one (example (3b)), *me-le* instead of standard *le-moi*, 1SG.DAT – 3MSG.ACC vs. 3MSG.ACC - 1SG.DAT), is again classified as stylistically marked (see e.g. Larthomas 1979: 22; Laenzlinger 1994: 88) or regionally confined, even very clearly attributed to the Swiss French variety (see e.g. Singy 1996; Bürgi 1999: 159).

To sum up, the following variationist classifications of our four variants were found in the literature:

- a) Variants that may be regionally distributed ((1b), (2b), (3b));

¹ Linguistically, the latter 4 areas belong to the Southern regional French area.

- b) One variant that is not regionally distributed ('colloquial French': (2d)).

Systematic empirical evidence to assess the scientific validity of these classifications is lacking however, as insufficient data are available for serious quantitative analyses of the distribution of the four variants in the present-day usage of speakers of European French. The aim of this paper is therefore to present a new approach to variation in contemporary European French (Belgium, France and Switzerland), more specifically to geographic variation in object clitic clusters, based on crowdsourced data.

The paper is structured as follows: after this short introduction setting the scene, a brief state-of-the-art of the abundant literature on French clitic clusters (section 2) reviews the potential explanations for their internal structure, which will lead to the formulation of three hypotheses (section 3). The method (crowdsourcing survey, cartography and statistics) will be presented in section 4, and the quantitative results in section 5. We will then discuss the results in section 6 in the light of preceding research and our hypotheses, and conclude in section 7.

2. State-of-the-art

2.1 Accounts of standard orders in French object clitic clusters

In order to describe what happens in standard clitic clusters (examples (1a), (2a) and (3a)), at least three different observations have been presented in the literature.

Firstly, the ban on combinations such as the one in (4), unattested in any variety of French and in many, if not all languages of the world, is described as the "Person Case Constraint" (PCC):

- (4) **Il* *me* *lui* *montre*
 he.NOM:3.Sg:Masc me.ACC:1.Sg: him.DAT:3.Sg:Masc
 show.PRES:3.Sg

This constraint can be formulated as follows:

"In a ditransitive, where both internal arguments are realized as phonologically weak elements, the direct object must be third person" (Adger/Harbour 2007: 4).

It is discussed, e.g., by Perlmutter (1971), Laenzlinger (1993), Miller/Monachesi (2003) for French and Romance, and by Adger/Harbour (2007) in general, and has been shown convincingly by Rezac (2010) to be *syntactic* in nature, as it is not the morphophonological material that inhibits combinations such as *me lui*, but the syntactic function of the dative argument in question – with ethical datives being freely permitted in *me lui* etc. combinations (Rezac 2010: 153f.).

Secondly, preverbal proclitic clusters (example (1a)) are usually described by the template approach from Perlmutter (1970: 226):

Class I	Class II	Class III	Class IV	Class V
me	le	lui	y	en
te	la	leur		
se	les			
nous				
vous				

Table 1. Order of representation of the standard French clitic sequences according to Perlmutter (1970: 226)

The basic descriptive generalization is:

Dative_{1st+2nd pers.} – Accusative

Accusative – Dative_{3rd pers.}

which contains inconsistent ordering for third person clitics, in contrast to most other standard Romance languages (Italian, Spanish) with consistent dative–accusative ordering (Miller/Monachesi 2003), and also differs from older stages of the Romance languages, including French (until the 16th century, cf. de Kok 1985: 366ff.).

Thirdly, postverbal enclitic object clitics (examples (2a) and (3a)) show the following ordering:

Accusative – Dative

(+weak pronouns for 1st and 2nd pers. Sing.: *moi* - not *me*, *toi* - not *te*).

This means that pre- vs. postverbal clitic clusters in standard French do not have the same order (mainly DAT-ACC in preverbal position, but mainly ACC-DAT in postverbal position), and that standard French differs in that again from most standard Romance languages, Old French and many Gallo-Romance dialects (see sections 2 and 6.1).

In order to explain the observed and additional ordering restructurations in standard French clitic clusters besides the PCC, two principled types of analyses have been proposed: in very few cases, a syntactic one (Laenzlinger 1993, 1994; Anagnostopoulou 2003), and in most cases a postsyntactic, i.e. morphophonological one, based on morphological processes and phonological incompatibilities / repairs (see e.g. Bonet 1991, 1995 for Catalan, Miller/Sag 1997, Miller/Monachesi 2003, Bonami/Boyé 2007, Goldbach 2007, Pescarini 2014, to name but a few). Rezac (2010: 153, and section 3 on the variation illustrated in our examples (1)) provides sufficient evidence for the non-syntactic nature of many restrictions and variations within French clitic clusters, as opposed to the PCC, which is based on the crucial syntactic difference between internal verbal arguments (**me lui*) and extra-argumental datives (*me lui* possible). The fact that most recent analyses of ordering issues and variation in French clitic clusters consider them as post-syntactic will be of importance in the discussion of our hypotheses, to be developed in section 3.

2.2 Dialectal variation in Gallo-Romance

For reasons of space, we cannot give a comprehensive overview of dialectal variation in Gallo-Romance object clitic clusters. We will thus present the most important facts and observations in the literature for the variation observed in our data (see section 5) and

refer the interested reader to Dalbera (1986), Heap/Kaminskaïa (2001), and Brun-Trigaud (2013) for more information. For the most recent pan-Romance picture, the interested reader is referred to Heap et al. (forthcoming: section 3).

We speak of *Gallo-Romance* in what follows as a cover term for three different groups of dialects, belonging to three different Romance languages which for historical reasons all belong (completely or partially) to French-speaking territories. Northern dialects in France and Belgium represent the first group, the *Oïl* dialects (cf. Smith 2016: 292), against the *Oc* or Occitan dialects of Southern France (Olivieri/Sauzet 2016; *oïl* and *oc* being the medieval particles for ‘yes’ in the respective dialects) as the second group, and Francoprovençal varieties in South-Western France and French-speaking Switzerland as the third group (plus the Val d’Aosta in North-Western Italy, cf. Kristol 2016):

Insert Figure 2 about here

Figure 2. The division of Gallo-Romance dialects. The dots represent the survey points of the *Atlas Linguistique de France* (Gillieron and Edmont 1902-1910).

These three groups of dialects developed directly from Latin, and the first group gave birth to standard modern French, a roof language for all three groups of dialects nowadays.

We have already seen in the introduction that three out of four variants ((1b), (2b), (3b)) are frequently considered to be a regional phenomenon, and that DAT-ACC ordering ((1b), (2b)) as well as consistency in pre- and postverbal object clitic clusters ((2d)) is typical for standard Romance languages other than French. This observation also holds for Gallo-Romance dialects. Thus, Morin (1979a: 307) reported that varieties with DAT-ACC ordering and consistency in pre- and postverbal clitic clusters are found in Northern (*Oïl*) dialects²:

² All examples taken from Morin (1979a); spelling has been normalized to standard French forms by the authors.

- (5) a. *Lorrain dialects* (cf. Aub-Büscher, 1962)
 il me le donne donne-me-le
- b. *Vendéen dialect* (cf. Rézeau, 1976)
 il me le donne donne-me-le
 il lui le donne donne-lui-le
- c. *Walloon dialect of la Gleize* (cf. Remacle, 1937)
 il me le donne donne-moi-le
- d. *Normand dialect of Saire* (cf. Lepelley, 1974)
 il me le donne donne-me-le
 il lui le donne donne-lui-le
- e. *Gallo dialects of Plouguenast* (cf. Hervé, 1973) and Pléchatel (cf. Dottin et Langouet 1901)
 il me le donne donne-me-le
 il lui le donne donne-lui-le

In Belgium (Wallonie), Western (Vendée) and Eastern dialects (Lorraine), we thus find the variant (3b), with the clitic *me* instead of standard *moi* in postverbal position.

However, in Gallo-Romance dialects, DAT-ACC as well as ACC-DAT can be found in object clitic clusters (cf. Heap et al. forthcoming, section 3.3). This is shown for Occitan varieties in examples (6a) and (6b) below and for Francoprovençal varieties in (6c) and (6d):³

- (6) a. *Provençal dialects* (cf. Ronjat 1937)
 il me le donne donne-me-le
 il le lui donne donne-le-lui
- b. *Gascon dialects* (cf. Séguy 1973)
 il le me donne donne-le-me
 il lui le donne donne-lui-le
- c. *Vaud (Switzerland) dialect* (cf. Reymond and Bossard 1979)
 il le me donne donne-le-me

³ Spelling has been normalized to the standard French forms in order to increase readability.

il le lui donne	donne-le-lui
d. <i>Bellegarde (Ain) dialect</i> (cf. Clédat 1887)	
il me le donne	donne-me-le
il le lui donne	donne-le-lui

There is, nevertheless, again a tendency towards consistency of clitic placement in pro- and enclitic position for clusters with identical person and number feature values. It should be pointed out, however, for all dialectal observations as well as for our results (sections 5 and 6) that the observed variation may be person/number sensitive, i.e. orders observed for 1st-3rd-person singular clusters are not necessarily the same for clusters of clitics with different person-number feature values.

Turning back to section 2.1, where we saw a clear preference in the formal literature for explaining most variation in French clitic clusters postsyntactically, a dialectal finding from Violin-Wigent (2010) on the variation seen in examples (1a) vs. (1b) (*le lui* vs. *lui le*) in Briançon (cf. Figure 1 *supra*) is relevant here. The author found that the most important factor for choosing one variant over the other is the phonetic nature of the sound following the clitic cluster (vowel or consonant, see Violin-Wigent 2010: 2007), and that there are differences according to different clitic forms rather than person or case features – which, again, points to postsyntactic regularities. This is in line with quite early observations from Foulet (1924: 89f., quoting Clédat 1925), who considered the variants under (7) and (8) acceptable, as opposed to (2d):

- | | | | |
|-----|----------------|--------------|-------------------------|
| (7) | <i>donne</i> | <i>-moi</i> | <i>-la</i> |
| | give.IMP:2.Sg | me.DAT:1.Sg | it.ACC:3.Sg: Fem |
| (8) | <i>apporte</i> | <i>-nous</i> | <i>-les</i> |
| | bring.IMP:2.Sg | us.DAT:1.PI | them.ACC:3. PI |

Vowel quality (full vowels [a] and [e] under stress in *la* and *les* as opposed to schwa in *le*) and prosody seem to play a role: *le* is clearly dispreferred in final (stressed) position of clitic clusters, but fine in (9):

(9) *donne* *-moi* *-le* *donc*
 give.IMP:2.Sg me.DAT:1.Sg it.ACC:3.Sg:Masc MODPART

Following this brief state-of-the art, we will now formulate three hypotheses that will guide our interpretation of the data presented in section 5.

3. Hypotheses

We saw in sections 1 and 2 that the non-standard variants in (1) and (2) are typologically consistent in a pan-Romance picture (DAT-ACC in most Romance languages and varieties), and that the variation observed in (2) as opposed to (1) is also diachronically explainable (ACC-DAT being the older word order, which has become fixed in affirmative imperatives plus lengthening and diphthongization of the final <e> in *me/te* under stress, leading to the weak pronoun forms *moi/toi*, cf. Foulet 1924: 61; Goldbach 2007: 73f.). The variation under (3) is, however, not explainable on these grounds, as, once the diphthongization *me/te* > *moi/toi* had taken place (from the 13th c.), the weak pronouns are always found in positive imperatives, except in areas where this diphthongization did not occur (see e.g. some rare attestations of *dis-me* in Picard, Northern France, cf. Foulet 1958: 121).

In front of the rich variation observed in the different dialectological studies, we would like to formulate two hypotheses to account for the variation observed in (1)-(3) in regional French, which for centuries has been in contact with the original dialects:

Hypothesis 1: the observed variation in examples (1) to (3) is due to language contact, more precisely dialectal substrate influence, as diverging orders in clitic clusters are attested in many Occitan, Northern French and Francoprovençal dialects in close contact to the varieties of regional French in question.

Especially when (in)consistency in pre- and postverbal object clitic clusters in Standard French is concerned, one may also assume a general tendency of analogical leveling to explain consistent variants such as (2d) against (2c), often at work in non-standardized varieties:

Hypothesis 2: the ordering variants in (1b), (2b and d), and (3b)⁴ **are due to analogical leveling** inside the template towards a more consistent ordering (see also the above-mentioned pan-Romance picture).

The fact that in dialectological studies on variation within clitic clusters with stable person-number features phonetic features such as vocalic or consonantal onset in the following context, vowel quality or stress also play a role leads us to our **Hypothesis 3**, concerning the nature of the observed variation: **variation in examples (1) to (3) takes place on the level of the phonetic form and is not linked to syntactic variation.**

These three hypotheses are not, of course, mutually exclusive; the dialectal substrate influence can enhance general analogical leveling tendencies, and both are possible in the system of the regional French varieties observed because of the postsyntactic nature of the variation at issue, i.e. they do not violate existing rules or introduce foreign syntactic rules into the respective systems.

4. Method

It has been shown repeatedly that spoken corpora are not useful when dealing with complex syntactic phenomena in French (see e.g. Blanche-Benveniste 2010). Accordingly, we found very few occurrences of clitic clusters in several databases of spoken French recorded in different areas of French-speaking Europe, and all displayed standard order.⁵

Therefore, in order to assess "who" uses the structures under scope, we decided to rely on the results of two questionnaire-based surveys that were broadcasted through mailing lists and social media. These surveys were specifically designed to investigate the vitality and the extension of some alleged regional expressions in European French on the levels of the lexicon, morphology and syntax (Belgium, France and Switzerland, see

⁴ See also Laenzlinger (1994: 88), referring back to Foulet (1924: 88-90).

⁵ For Swiss French, we investigated the OFROM corpus (Avanzi et al. 2016), for Belgian French the VALIBEL corpus (Bolly et al. 2016). For metropolitan French, we conducted research in the CFPP corpus (Lefevre & Branca 2016) as well as in the CLAPI database (Groupe ICOR 2016).

AUTHOR-1 for more details on the design and content of the surveys). The two surveys (named Euro-1 and Euro-2) were conducted between June 2015 and September 2016. Participants were invited to take part via email or social media and by activating a link included in the message, reached an online platform for crowdsourcing. They then had, first, to answer a few questions on their sociolinguistic background (sex, age, country, places where they had spent most of their life and where they currently lived, etc.), and were, second, asked to answer a set of 40 questions in a row. For each question, a brief context explaining a virtual situation was given, accompanied by a picture. Participants had to read the description and then to select the expressions or sentences corresponding best to what they would have said if in the situation described (see examples in section 5). They were asked to answer as naturally as possible, as if they were engaged in an informal conversation with relatives or peers from the place where they had spent most of their childhood.

In total, slightly more than 20,000 European French speakers took part. Table 2 provides the basic demographic information regarding participants, namely age and country where the participants declared to have spent most of their lives:

Survey	Country	Npart.	Age		
			mean	min.-max.	sd
Euro-1	Belgium	443	34.2	15-84	14.1
	France	8,296	34.5	11-90	12.9
	Switzerland	2,366	38.6	12-88	15.4
	Total	11,105	35.3	11-90	13.6
Euro-2	Belgium	133	39.5	18-82	16.8
	France	5,329	36.9	11-92	14.4
	Switzerland	3,892	43.5	10-87	17.3
	Total	9,354	39.7	10-92	16.1

Table 2. Basic demographics of participants for each survey. Age values are given in years.

The *ggmap* R package was used subsequently to map the results of the surveys (Kahle and Wickham 2013, R development core Team 2016). Stamen[®] maps were used to plot

the background. As for spatial coordinates (longitudes, latitudes and polygons), we used the coordinate files provided by ArcGis®. On the maps presented in the following section, each symbol refers to a ZIP-code (the postal code of the area where participants declared they had spent most of their lives), while polygons define the borders of a department in France, a province in Belgium and a canton in Switzerland where participants said they had spent most of their lives. The shade of polygons varies depending on the percentage of respondents who gave a particular answer: the higher the percentage, the darker the color, and *vice versa*.

5. Results

In order to address the vitality and the distribution of clusters such as (1b), a first question was taken from the Euro-1 survey. The context given for this question was the following: ‘Paul came home last weekend. He has just bought new skiing equipment, which he is very proud of. He shows it to his father. This means that:’. The following two options were then proposed:

- (1) a. *il* *le* *lui* *montre*
 he.NOM:3.Sg:Masc it.ACC:3.Sg:Masc him.DAT:3.Sg:Masc
 show.PRES:3.Sg
- b. *il* *lui* *le* *montre*
 he.NOM:3.Sg:Masc him.DAT:3.Sg:Masc it.ACC:3.Sg:Masc
 show.PRES:3.Sg

The map in figure 3 shows the percentage of participants who indicated that they would use the non-standard variant (1b):

Insert Figure 3 about here.

Figure 3. Percentage of participants who declared using the non-standard variant (1b) as a function of the zip code/area where they spent most of their lives (Euro-1).

First, it can be seen that the use of the non-standard variant in (1b) is quite rare in contemporary French, since it was chosen by only 631 participants (out of 11,105 participants, which amounts to 5.6%). Second, the cartography shows that there is a considerable variation within French-speaking Europe. In Belgium, the percentages vary between 2.9% (Liège) and 7.1% (Brussels) with a mean of 4.8%; in Switzerland, they vary between 5.3% (Jura) and 14.5% (Vaud and Fribourg). In France, the situation is more complex: only one department shows 20% of answers with the non-standard variant (21.8% for Haute-Loire), and 6 departments have percentages above 10% (Hautes-Alpes and Alpes-de-Haute-Provence in the South-East with 11.5% in average; Ardèche and Loire, which are adjacent to Haute-Loire, reach 18.2%; the Meuse department (in the North-East) and the Ille-et-Vilaine department in the North-West, 13.4%). The other departments present percentages between 0% and 10%, but the distribution of the non-standard variant does not follow any cohesive trend.

Concerning non-standard object clitic clusters such as (2b), one question was taken from the Euro-2 survey. The given context was the following: ‘You brought one of your work colleagues to have a drink with your best friends. The colleague tells you he has a crush on one of the people in the group, but that he does not dare to ask for his/her phone number. You tell him...?’. The following two options were then proposed:

- | | | | |
|--------|----------------|-------------------|-------------------|
| (2) a. | <i>demande</i> | <i>-le</i> | <i>-lui</i> |
| | ask.IMP:2.Sg | it.ACC:3.Sg:Masc | him.DAT:3.Sg:Masc |
| b. | <i>demande</i> | <i>-lui</i> | <i>-le</i> |
| | ask.IMP:2.Sg | him.DAT:3.Sg:Masc | it.ACC:3.Sg:Masc |

The map in figure 4 shows the percentage of participants who claimed to use the non-standard variant (2b):

Insert Figure 4 about here

Figure 4. Percentage of participants who declared using the non-standard variant (2b) as a function of the zip code/area where they spent most of their lives (Euro-2).

Again, the use of the structure is quite rare in French, since the non-standard variant was chosen by only 543 participants (out of 9,394 participants, 5.7%), and again, there is considerable variation within French-speaking Europe. In Belgium, none of the participants declared using the construction (this can however be explained by the small number of persons who took part in this survey). In Switzerland, percentages varied between 1.5% (Valais) and 5.8% (Vaud and Neuchâtel). Interestingly, in France, the situation is again more complex: 4 departments present percentages above 20% (Loire and Haute-Loire reach 29.1%; Haute-Marne and Meuse in the North reach 22.5%). The other departments present percentages between 0% and 20%, but here again, the geographical distribution does not follow any cohesive trend.

As for the third question we will present, it was designed to address the vitality and the distribution of clitic clusters such as (2d) and (3b). It was taken from the Euro-1 survey. The context given for this question was the following: "You lent your cellphone to your younger brother, but he does not want to give it back to you. A little bit upset, you order him:". The following three options were then proposed:

- | | | | |
|--------|--------------------|------------------|------------------|
| (2) c. | <i>rends</i> | <i>-le</i> | <i>-moi</i> |
| | give-back.IMP:2.Sg | it.ACC:3.Sg:Masc | me.DAT:1.Sg |
| d. | <i>rends</i> | <i>-moi</i> | <i>-le</i> |
| | give-back.IMP:2.Sg | me.DAT:1.Sg | it.ACC:3.Sg:Masc |
| (3) b. | <i>rends</i> | <i>-me</i> | <i>-le</i> |
| | give-back.IMP:2.Sg | me.DAT:1.Sg | it.ACC:3.Sg:Masc |

Figure 5 shows the map with the percentage of participants who claimed to use the non-standard variant (2d), *rends-moi-le*:

Insert Figure 5 about here

Figure 5. Percentage of participants who declared using the non-standard variant (2d) as a function of the zip code/area where they spent most of their lives (Euro-1).

First, we can see that the use of the structure is not as rare as the two preceding ones. In fact, the non-standard variant (2d) was chosen by 1657 participants (out of 11,105 participants, which represents 14.9%). As can be seen on the map above, the percentages are relatively low in the French-speaking parts of Belgium and of Switzerland. In Belgium, the percentages vary between 0% (Luxemburg) and 5.1% (Hainault) with a mean of 2.6%; in Switzerland, they vary between 3.9% (Vaud) and 19.7% (Jura) with a mean of 7.3%. In France, there are some pronounced differences between the departments: we can see that the structure is more frequently used in a large crescent covering all the departments of the West and the South of the territory, with some peaks in the departments of the former Franche-Comté region (including the Jura, Haute-Saône, Haute-Marne and Doubs, which present percentages reaching 44.3% in average), in the former Rhône-Alpes region (including Savoie, Haute-Savoie and Loire, with percentages reaching 35.5% in average) and in isolated departments such as Lozère (in the Middle-South, 61.5%) and Ariège (on the Spanish-French border, 37.1%). We also find some randomly distributed peaks in departments located in the Northern part of France (the Orne department in Normandy reaches 38.1%, the Ardennes department in the North-West reaches 37.7%). On the contrary, the lowest values are found in Bretagne (Côtes d'Armor, Finistère and Morbihan cumulate 6.68% in average) and in the South-West (Aquitaine and Landes, 6.1%).

Figure 6 shows the percentage of participants who claimed to use the non-standard variant (3b):

Insert Figure 6 about here

Figure 6. Percentage of participants who declared using the non-standard variant (3b) as a function of the zip code/area where they spent most of their lives (Euro-1).

Contrary to what was observed for the other three non-standard clitic clusters, the variant (3b) is quite clearly confined in an areal way. In fact, the structure is attested in Switzerland (more precisely in the Bern and Neuchâtel cantons, with 27.1% and 15% of

participants who declared using this variant; none of the other cantons had a score higher than 10%). In France, the clitic cluster *rends-me-le* is weakly attested in the Francoprovençal area, especially in the Haute-Savoie department (4.1% of the participants), and more frequently in the South-Western part of France, in a territory that corresponds to the historical region of Gascony (which includes the departments of Gers, Gironde, Landes, Lot-et-Garonne; Hautes-Pyrénées and Pyrénées-Atlantiques, 10.4% in average). Otherwise only some scattered occurrences of the cluster are found in the North of France.

6. Discussion

The three non-standard variants where order in object clitic clusters is concerned ((1b), (2b) and (2d)) do not seem to be regionally well-distributed, i.e. the observed distribution does not correspond to any cohesive area: (2d) is attested everywhere in European French (and its use ranges from 0 to 60%); (1b) and (2b) are also attested everywhere, but they are much less frequent (uses range from 0 to 25% and 0 to 20%). Second, only one variant, (3b), where the selection of the form (*me* vs. *moi*) is at issue, seems to be regionally confined.

Two main explanations, which are closely interwoven, come to mind when one has to account for the existence of the non-standard variants in (1) to (3). The non-standard examples may be the consequence of language contact (= our H1). For centuries, the French spoken in Belgium, France and Switzerland has been in contact with the underlying Gallo-Romance dialects. Variants such as (2b), (2d) and (3b) may also be due to analogical leveling yielding consistency in pre- and postverbal clitic clusters (= H2).

Unfortunately, the areas where (1b) is found do not coincide with the areas where e.g. Morin (1979a) or Ronjat (1937; see our examples (6)) reported the order *lui-le*, except for Gallo. For (2b), *lui-le* in enclitic position, we have a partial overlap with Gallo dialects and the respective regional French in our study. This variant is also found in Vendée, close to Bretagne, but with higher percentages in our study. However, we do not find it with particular frequency in Gascony (cf. example (6b), following Séguy 1973). (2d) seems to be preferred in the Eastern parts of France, which does not

correspond at all to the dialectal observations made in Morin (1979a), examples under (5), which localized it in Northern dialects (Normandie and Gallo, Bretagne). Finally, (3b), the third variant to be considered as regional in the literature, is attested in dialects from Wallonie, Lorraine, Vendée, also in Southern France (Provence) and Ain (see respective examples under (5) and (6)), but is very clearly confined to Gascony and Switzerland (and Nord-Pas-de-Calais) in our data. All in all, direct correspondences between dialectal orders in object clitic clusters and the geographical distribution of the respective orders in regional French are very rare. But even for the scarce cases where we can find overlapping areas, one major fact sheds some doubt on direct dialectal substrate influence as the sole explanation for our data.

As we saw in section 5, the non-standard variants (1b), (2b) and (2d) do not form cohesive areas in their regional distribution. And the same holds for Gallo-Romance dialects: there are no well-delimited or cohesive areas that present a certain order and not the opposite. The following map (Figure 7) gives an overview of the order of object clitics in Gallo-Romance clitic clusters, that correspond to our examples (2c)/(2d)/(3b):

Insert Figure 7 about here

Figure 7. Synthetic map of the variants of the standard French sentence *dis-le-moi* in Gallo-Romance dialects, according to ALF (map 310)⁶.

Given this situation, one should be very wary of explaining contemporary regional variation in object clitic clusters as an instance of the influence of dialectal substrate – we simply do not know and cannot decide, neither on the basis of dialectal nor of modern crowd-sourcing data, as the spatial patterns they both show are too diffuse.

The variation seen in clusters such as (2c) against (2d) can be accounted for by the fact that there is inconsistent ordering in standard French pre- and postverbal clitic clusters, because of the maintenance of the archaic order (ACC-DAT) in positive imperatives (cf. Foulet 1924, Hirschbühler and Labelle 2001, Goldbach 2007). Indirect evidence for the

⁶ Some data could not be morphologically analyzed because of complex morphophonological processes in the respective varieties or complete fusion of the clitics, which explains why some points are left blank on the map.

partially unnatural development in standard French is the vivid debate that took place among grammarians during the classical period (17th c.) regarding the question of the ‘correct’ order in P1-P3 positive imperatives: DAT-ACC or ACC-DAT (cf. De Kok 1985: 376ff.). Since the classical period at the latest, in proclitic position, the dative precedes the accusative (*il me le donne*), except for the case of two third person clitics, while it is the contrary in enclitic position (*donne-le-moi*). This inconsistency may trigger some variation, by favoring analogical orders in order to produce symmetrical clusters, such as (2d).

As for the clusters in (1b) and (2b), they are less frequent than the cluster in (2d). As has been shown in previous studies, the combination *le-lui* is not euphonic, and tends to be avoided by speakers (Morin 1979b), usually by dropping one clitic. Since in the surveys the participants had to choose between two possibilities, neither of which corresponded to their actual use, the fact that the non-standard clitic clusters were not chosen very often might not reflect what speakers would actually have produced in everyday life. But still, and despite the fact that the question to assess the order in proclitic clusters was extracted from a different survey than the question to test the order in enclitic clusters, it appeared that the peaks of percentages for both non-standard variants (*lui-le*, see figure 3 and 4 above) were located in the same departments, confirming a preference for consistent ordering, which speaks strongly in favor of analogical leveling, and which also explains the preference for (2d) over (2c).

Thus, the situation we observe in regional French is similar to the one we observe in Gallo-Romance dialects: variation in order in object clitic clusters is not associated with a specific dialectal area, different ordering variants tend to be found everywhere, and orders may vary according to person and number features. We consider that this provides enough evidence to claim an important role for analogy, which can be, of course, enhanced by a corresponding dialectal substrate order, without dialectal substrates being only or overwhelmingly responsible for the observed variation in space. Based on this reasoning, both hypotheses, 1 (language contact) and 2 (analogical leveling) may be confirmed, with a preference for 2 over 1.

As for clitic selection in regional French object clitic clusters, the finding that stands out the most clearly is that its distribution and vitality appear to be regionally constrained: (3b) is attested mostly in French-Speaking Switzerland and in two peripheral areas of France, namely Gascony and Nord-Pas-de-Calais. The most likely hypothesis to explain the origin of this structure is, again, that the non-diphthongized form *me* in enclitic position is the result of an interference with substrate dialects. However, the non-diphthongized form exists all over the Gallo-Romance area, except in the Oïl dialects spoken around Paris (see Figure 7 above, the survey points where diphthongization took place are indicated by a star symbol), where diphthongization took place under stress: *me* > *moi* (cf. Foulet 1924). So how do we account for the clear geographical delimitation of the variant *rends-me-le* (3b)? Obviously, the form *me* in French postverbal object clitic clusters such as (3b) is a form which comes from the substrate dialects. The order *me-le* is, however, not attested for dialects in Switzerland, rather the opposite one (see example (6c) above; Séguy reports *le-me* also for Gascony in (6b)). The question thus remains as to why the variant *rends-me-le* (3b) is found only in some specific areas of Europe and not elsewhere. An external explanation may lie in the fact that the peripheral areas where *me* was maintained in clitic clusters in the respective regional varieties of French are linguistically conservative and more resistant to innovative influences of the (Parisian) standard than other areas, also for political reasons (Gascony and French-speaking Switzerland show in fact many archaic features when compared with other regions, cf. e.g. Séguy 1950). For the Nord-Pas-de-Calais result, we know that the structure is present in the dialect and not transposed to non-dialectal non-standard usage (i.e. it is only employed by dialect speakers). And for Gascony and French-speaking Switzerland, we saw in examples (6b) and (6c) above (see also figure 6) that their substrate dialectal pattern (*le-me*) is the marked one (much less frequently attested than *me-le*), with ACC-DAT order also in preverbal clitic clusters (against the pan-Romance picture). The structural difference between their dialects and the standard or more frequent *me-le* ordering in preverbal clitic clusters may thus have caused linguistic insecurity, which led to the otherwise infrequent pattern *rends-me-le* as an amalgam of their dialectal form *me* and an – erroneous – analogy, almost like a case of hypercorrection, of the learned standard preverbal order *me le* towards the postverbal order *me-le* as well.

7. Conclusion

The aim of this paper was to shed some new light on object clitic cluster variation in French, with a focus on internal ordering variation (DAT-ACC vs ACC-DAT) and internal ordering variation plus clitic selection:

- | | | | | |
|--------|--------------------|-------------------|-------------------|---------------|
| (1) a. | <i>il</i> | <i>le</i> | <i>lui</i> | <i>montre</i> |
| | he.NOM:3.Sg:Masc | it.ACC:3.Sg:Masc | him.DAT:3.Sg:Masc | |
| | show.PRES:3.Sg | | | |
| b. | <i>il</i> | <i>lui</i> | <i>le</i> | <i>montre</i> |
| | he.NOM:3.Sg:Masc | him.DAT:3.Sg:Masc | it.ACC:3.Sg:Masc | |
| | show.PRES:3.Sg | | | |
| (2) a. | <i>demande</i> | <i>-le</i> | <i>-lui</i> | |
| | ask.IMP:2.Sg | it.ACC:3.Sg:Masc | him.DAT:3.Sg:Masc | |
| b. | <i>demande</i> | <i>-lui</i> | <i>-le</i> | |
| | ask.IMP:2.Sg | him.DAT:3.Sg:Masc | it.ACC:3.Sg:Masc | |
| c. | <i>rends</i> | <i>-le</i> | <i>-moi</i> | |
| | give-back.IMP:2.Sg | it.ACC:3.Sg:Masc | me.DAT:1.Sg | |
| d. | <i>rends</i> | <i>-moi</i> | <i>-le</i> | |
| | give-back.IMP:2.Sg | me.DAT:1.Sg | it.ACC:3.Sg:Masc | |
| (3) a. | <i>rends</i> | <i>-le</i> | <i>-moi (=2c)</i> | |
| | give-back.IMP:2.Sg | it.ACC:3.Sg:Masc | me.DAT:1.Sg | |
| b. | <i>rends</i> | <i>-me</i> | <i>-le</i> | |
| | give-back.IMP:2.Sg | me.DAT:1.Sg | it.ACC:3.Sg:Masc | |

In the literature, examples such as (1b), (2b) and (3b), displaying consistent dative-accusative ordering, have frequently been claimed to be specific of some regional varieties of French. In contrast to this, the form (2d) is often considered as “colloquial”.

Our quantitative results, stemming from two crowd-sourcing surveys with more than 20,000 francophone participants from all over French-speaking Europe, reveal that the three non-standard variants where order in object clitic clusters is concerned ((1b), (2b) and (2d)) are not regionally well-distributed, i.e. the observed distribution does not correspond to any cohesive area and very often not to corresponding dialectal variants in the respective regions. In contrast, only one variant (3b)), where the selection of the form (*me* vs. *moi*) is at issue, seems to be regionally confined: it is found in French-speaking Switzerland, in Gascony, and we also found some rare attestations of it in the North of France.

In order to see whether the observed variation is driven by language contact (H1), we have described the situation in Gallo-Romance dialects, which are substrates for the regional French varieties in question. Based on previous studies (Morin 1979; Brun-Trigaud 2013; Heap et al. forthcoming) and thanks to a new look at existing material (Clédat 1887, 1925; Gilliéron and Edmont 1902-1910, Séguy 1973), we are able to show that in Gallo-Romance dialects, we find either (rarely) ACC-DAT or DAT-ACC ('pan-Romance' order) in object clitic clusters, but not necessarily in coinciding areas. What we do observe in the dialectal data, however, is the fact that the order is always consistent within the same cluster in pro- or enclitic position, contrary to standard French. As inconsistencies of a rigid learned norm tend to be ignored in spontaneous language use by analogical leveling, our second hypothesis (H2), analogical leveling as an overall source of the observed variation, seems to be confirmed. This phenomenon may be enhanced in some cases by parallel systems in the dialectal substrate varieties (H1), cf. already Morin (1979a: 308). To account for the regionally clearly confined variant in (3b), a look at the underlying dialectal systems shows that it is only in these areas that we find an overall ordering ACC-DAT in preverbal clitic clusters. We also find there the non-diphthongized form *me*1.SG. instead of the standard weak pronoun *moi*1.SG in enclitic position, but this non-diphthongized form exists all over the Gallo-Romance area (except in the Oïl dialects spoken around Paris, where diphthongization took place under stress). Thus, we argue that it is the linguistic distance between the dialectal substrates and the standard combined with the extra-linguistic fact that French-speaking Switzerland and South-Western France (Gascony) are very conservative and historically and politically quite isolated linguistic areas, more resistant to

innovations/influences of the (Parisian) standard (cf. e.g. Séguy 1950), which explains the observed regional distribution of variant (3b) as, maybe, a case of hypercorrection. As we have observed a rather accidental and also geographically random distribution of the four variants taken into consideration here, we finally consider this kind of variation as taking place at the postsyntactic, morphophonological level (H3). Further work is needed, however, to corroborate this impression, by taking systematically into account other clitic clusters (with different person/number feature values) and argumental vs. extra-argumental clitics in order to further confirm our hypotheses or uncover alternative regularities and explanations.

References

- Adger, David, and Daniel Harbour 2007. "Syntax and Syncretisms of the person case constraint". *Syntax* 10-1: 2-37.
- Anagnostopoulou, Elena. 2003. *The Syntax of Ditransitives. Evidence from Clitics*. Berlin/New York, Mouton de Gruyter.
- Aub-Büscher, Gertrud. (1962). *Le parler rural de Ranrupt*. Paris, Klincksieck.
- Avanzi, Mathieu, Marie-José Béguelin, and Federica Diémoz. 2016. "De l'archive de parole au corpus de référence: la base de données orales du français de Suisse romande (OFRON)". *Corpus* 15: 309-342.
- Blanche-Benveniste, Claire. 2010. *Le français. Usages de la langue parlée*. Leuven/Paris, Peeters.
- Bolly, Catherine, Georges Christodoulides, and Anne Catherine Simon. 2016. "Disfluences et vieillissement langagier. De la base de données VALIBEL aux corpus outillés en français parlé". *Corpus* 15: 195-212.
- Bonami, Olivier, and Gilles Boyé. 2007. "French pronominal clitics and the design of Paradigm Function Morphology". *Proceedings of the 5th Mediterranean Morphology Meeting*, Fréjus, 291-322.
- Bonet, Eulalia. 1991. *Morphology after syntax: Pronominal clitics in Romance*. Unpublished PhD dissertation, Massachusetts Institute of Technology.
- Bonet, Eulalia. 1995. "Feature structure of Romance clitics". *Natural Language and Linguistic Theory* 13: 607-647.

- Brun-Trigaud, Guylaine 2013. "Les chaînes de clitiques : traitements cartographiques". In *Mémoires du terrain : enquêtes, matériaux, traitement des données*, ed. by Francis Manzano, 139-155. Lyon, CEL/Lyon III.
- Bürgi, Anne. 1999. "Le pronom 'ça' en français vaudois", *Vox Romanica* 58: 149-171.
- Cardinaletti, Anna, and Michal Starke 1999. "The typology of structural deficiency: A case study of three classes of pronouns", in *Clitics in the languages of Europe*, ed. by Henk van Riemsdijk, 145-233. Berlin: Mouton de Gruyter.
- Clédat, Léon. 1887. "Les patois de la région lyonnaise". *Revue des patois* 1: 1-10.
- Clédat, Léon. 1925. "Les pronoms personnels avant et après le verbe". *Revue de philologie française* 37: 43-48.
- Dagnac, Anne. 2012. "La variation régionale des proformes faibles en français". Ms, Université de Toulouse.
- Dalbera, Jean-Philippe. 1986. "Les pronoms personnels atones dans les parlers des Alpes-Maritimes. Champ et mécanismes de variation". In *Actes du XVIIIe Congrès International de Linguistique et de Philologie romane*, ed. by Dieter Kremer, vol. 3, 599-611. Tübingen: Max Niemeyer.
- Dauzat, Albert. 1930. *Histoire de la langue française*. Paris, Payot.
- De Kok, Ans. 1985. *La place du pronom personnel régime conjoint en français: une étude diachronique*. Amsterdam: Rodopi Bv Editions.
- Dottin, Georges & Jean Langouet (1901). *Glossaire du parler de Pléchatel*. Geneva, Slatkine.
- Foulet, Lucien. 1924. "L'accent tonique et l'ordre des mots: formes faibles du pronom personnel après le verbe". *Romania* 50: 54-93.
- Foulet, Lucien. 1958³. *Petite syntaxe de l'ancien français*. Paris: Honoré Champion.
- Gilliéron, Jules, and Edmond Edmont. 1902-1910. *Atlas linguistique de la France*. Paris: Champion.
- Groupe ICOR. 2016. "CLAPI, une base de données multimodale pour la parole en interaction: apports et dilemmes". *Corpus* 15: 165-194.
- Grévisse, Maurice, and André Goosse. 2008¹⁵. *Le bon usage*. Bruxelles, Duculot.
- Goldbach, Maria. 2007. "The Distributed Morphology of Object Clitics in Modern French". *Zeitschrift für Sprachwissenschaft* 26: 41-81.

- Heap, David, and Sevtlana Kaminskaïa (2001). "Variable Clitic Sequences in Nonstandard French: Feature Geometry or Optimality?". *University of Pennsylvania Working Papers in Linguistics* 7/3: 101-114.
- Heap, David et al. (forthcoming). "Clitic pronouns". In *Manual of Romance linguistics: Morphosyntax and Syntax*, ed. by Andreas Dufter, and Elisabeth Stark. Berlin/New York: de Gruyter.
- Hervé, Bernard. 1973. *Le parler de Plouguenast*. Mémoire de maîtrise, Université de Haute Bretagne
- Hirschbühler, Paul, and Marie Labelle. 2001. "La position des clitiques par rapport au verbe à l'impératif dans l'évolution du français". *Recherches Linguistiques de Vincennes* 30: 13-37.
- Kahle, David, and Hadley Wickham (2013). "ggmap: Spatial Visualization with ggplot2". *The R Journal* 5/1: 144-161.
- Kristol, Andrès 2016. "Francoprovençal". In *Oxford Guide to the Romance Languages*. Adam Ledgeway, and Martin Maiden, 350-362. Oxford: Oxford University Press.
- Laenzlinger, Christopher. 1993. "A syntactic view of Romance pronominal sequences". *Probus* 5/3: 242-270.
- Laenzlinger, Christopher. 1994. "Enclitic clustering: The Case of French Positive Imperatives". *Rivista di Grammatica Generativa* 19: 71-104.
- Lefevre, Florence, and Sonia Branca 2016. "Le CFPP2000: constitution, outils et analyses. Le cas des interrogatives indirectes". *Corpus* 15: 265-285.
- Lepelley, René. 1974. *Le parler normand du Val de Saire*. Caen: Musée de Normandie.
- Miller, Philipp, and Ivan A. Sag (1997). "French Clitic Movement without Clitics or Movement". *Natural Language and Linguistic Theory* 15/3: 573-639.
- Miller, Philipp, and Paola Monachesi (2003). "Les pronoms clitiques dans les langues romanes". In *Les langues romanes, problèmes de la phrase simple*, ed. by Danièle Godard, 67-123. Paris, Editions du CNRS.
- Morin, Y. (1979a). "More remarks on French clitic order". *Linguistic Analysis* 5/3: 293-312.
- Morin, Y. (1979b). "La morphophonologie des clitiques en français populaire". *Cahiers de linguistique* 9, 1-36.

- Olivieri, Michèle, and Patric Sauzet 2016. "Southern Gallo-Romance (Occitan)". *Oxford Guide to the Romance languages*, ed. by Adam Ledgeway, and Martin Maiden, 319-349. Oxford: Oxford University Press.
- Perlmutter, David, M. 1970. "Surface structure constraints in syntax". *Linguistic Inquiry* 1: 187-255.
- Perlmutter, David, M. 1971. *Deep and Surface Structure Constraints in Syntax*. New York et al.: Holt, Rinehart and Winston.
- Pescarini, Diego 2014. "Prosodic restructuring and morphological opacity. The evolution of Italo-Romance clitic clusters". In *Diachrony and Dialects: Grammatical Change in the Dialects of Italy*, ed. By Paola Benincà, Adam Ledgeway, and Nigel Vincent, 155-176. New York/Oxford: Oxford University Press.
- R development core Team. 2016. "R: A language and environment for statistical computing. R Foundation for Statistical Computing".
- Remacle, Louis. 1937. *Le parler de la Gleize*. Liège: H. Vaillant-Carmanne.
- Reymond, Jules, and Maurice Bossard. 1979. *Le patois vaudois. Grammaire et vocabulaire*. Lausanne: Cabédita Editions.
- Rezac, Milan. 2010. "Ineffability through modularity: Gaps in French clitic clusters. In Defective paradigms: Missing forms and what they tell us". In *Proceedings of the British Academy* 163, ed. by Matthew Baerman, Greville G. Corbett, and Dunstan Brown, 151-180. Oxford: Oxford University Press.
- Rézeau, Pierre. 1976. *Un patois de Vendée: le parler rural de Vouvant*. Paris: Klincksieck.
- Riegel, Martin, Jean-Christophe Pellat, and René Rioul (1994). *Grammaire méthodique du français*. Paris: PUF.
- Ronjat, Jules. 1937. *Essai de syntaxe des parlers provençaux modernes*. Mâcon: imp. Protat.
- Séguy, Jean. 1950. *Le français parlé à Toulouse*. Toulouse: Privat.
- Séguy, Jean. 1973. *Atlas linguistique et ethnographique de la Gascogne* (vol. 5). Paris: Éditions du CNRS.
- Singy, Pascal 1996. *L'image du français en Suisse romande. Une enquête sociolinguistique en Pays de Vaud*. Paris, L'Harmattan.

- Smith, John C. 2016. "French and northern Gallo-Romance". In *Oxford Guide to the Romance languages*, ed. by Adam Ledgeway, and Martin Maiden, 292-318. Oxford: Oxford University Press.
- Tuailon, Gaston. 1983. "Régionalismes grammaticaux". *Recherches sur le français parlé* 5: 227-240.
- Violin-Wigent, Anne. 2010. "Tu lui les as donnés ? Variation dans l'ordre des pronoms multiples en phrases déclaratives". In *Actes du 2^{ème} Congrès Mondial de la Langue Française*, ed. by Franck Neveu et al., 2001-2014, Nouvelle-Orléans.