

RESEARCH

SyMiLa and the *Atlas linguistique de la France*: A tool for the study of Gallo-Romance syntax

Anne Dagnac

CLLE (CNRS, Université de Toulouse 2 Jean Jaurès) 5 allée Antonio Machado, 31058 Toulouse Cedex, FR
dagnac@univ-tlse2.fr

This paper presents different ways in which the Gallo-Romance data of the *Atlas linguistique de la France* (ALF) can contribute to syntactic description, typology and analysis. After briefly presenting the SyMiLa project, which aims at making such data accessible, it presents four case-studies. The first one focuses on *wh*-questions, which provide a horizontal perspective over the whole Gallo-Romance varieties. The complete data set shows how the dominant *wh*-structures distribute over the Gallo-Romance domain, and how the rarer ones argue for an articulated CP where *que* ‘that’ stands in Fin, and for a compositional nature of *est-ce que*. The next section turns to a vertical perspective on the data: it focuses on a very local phenomenon, complementizer doubling, and shows how the ALF reveals an optional but consistent behavior, which adds evidence to a head-movement analysis and suggests that *complementizer doubling* and *recomplementation* correspond to two different configurations. The last section deals with isolated data, that could not *per se* lead to any clear syntactic claim. Anecdotal and unsystematic as they may seem in the atlas, these data signal structures that, as further research shows, can enrich and/or challenge previous generalizations on interrogative clause typing and Negative Concord, respectively. In the latter case, they call for typologies and analyses of Negative Concord that include both optional concord and partial concord.

Keywords: microvariation; syntax; atlases; interrogative clauses; Negative Concord; CP-field; Picard

1 Introduction

The contribution of Gallo-Romance dialect syntax to syntactic theory, typology or, as concerns the Oïl dialects, even the diachrony of French syntax, is almost null – mostly for lack of documentation. Dialectal grammars or monographs on Gallo-Romance specific dialects focus on phonology and the lexicon, sometimes (verbal) morphology. They either eschew syntax or boil it down to a few descriptive lines.¹ This is all the more regrettable since France has an old dialectological tradition, and has published various dialect atlases. Some of them, like the *Atlas linguistique de la France* (ALF), contain syntactic data, although this is still a little-known fact. The SyMiLa project aims at making these data accessible to the community, at analyzing them and at conducting new syntax-specific dialectal surveys. Such data, indeed, are of great value for dialectology, diachronic linguistics, and linguistic analysis. Dialectal areas are mostly defined on phonological and lexical properties: how syntax fits in is still an open question. Another question is how the diachrony of standard languages interferes with or can be clarified by the diachrony of its kin varieties. In this paper, I will focus on the interest these data may present for

¹ This is not as true for Occitan. Vasseur (1996) for Vimeu Picard and Remacle (1952–60) for Walloon are two outstanding exceptions for the Oïl area.

comparative syntax. My first aim is to show that, in spite of their shortcomings, atlases constitute an important tool to investigate Gallo-Romance syntactic micro-variation, and that they can contribute to syntactic typology and theoretical debates in several ways. I first briefly present the ALF and the SyMiLa project. I then pick up a few examples to show how these data can be used as a starting point for different syntactic investigations. In the third section, I show how studying a syntactic phenomenon, here *wh*-questions, in a horizontal perspective, i.e. across the whole Gallo-Romance area, can help us test or refine known generalizations and analyses. In section 4, I illustrate another way to use the ALF data, which I call a vertical perspective. I exemplify it with the study of complementizer doubling, a “rare” phenomenon which surfaces only in a small sub-area of the Picard dialect, though it can be traced back to the Middle Ages. In section 5, I focus on interrogative particles and Negative Concord: I show how sporadic occurrences in the atlas may pinpoint larger patterns that challenge usual syntactic typologies and leave way for further theoretical investigations.

2 French dialect atlases and the SyMiLa project

France has known two waves of large-scale dialectal surveys that have resulted in the publication of linguistic atlases. The first wave of systematic surveys targeting all Gallo-romance dialects (Édmont and Gilliéron 1902–1910) took place at the turn of the 20th century and gave birth to the *Atlas linguistique de la France* (ALF). A second wave, known as the *Nouvel Atlas linguistique de la France*, was launched in the middle of the 20th century: it consists in a series of partially unrelated surveys targeting different dialectal areas, with a strong lexical and ethnological content, and usually devoid of any syntactic dimension.² I will focus here on the *Atlas linguistique de la France*.

Each of the 1920 maps in the ALF features the dialectal translation of a French word or sequence of words, noted in an ad hoc phonetic alphabet. The 1421 first maps display 639 different locations (each represented by one or, sometimes, a few speakers). The 499 next ones concentrate on the Southern or South-Eastern part of France, with a few additional locations in the North. These locations spread across the Gallo-Romance dialect speaking territories of France except Corsica,³ and extend to small areas outside of France: the Picard, Walloon and Gaumais speaking parts of Belgium and Luxemburg, the Occitan speaking area of Northern Italy, and the Romance speaking part of Switzerland. On a practical basis, this area can be defined as the places where a questionnaire could be handled in French by the informants, and yield data in some version of Occitan, Oil languages, Catalan or Franco-Provençal. This atlas has served consistently for dialectological and diachronic works. However, they concentrate mainly on phonology and lexicology, as does the synthetic commented maps published by Brun-Trigaud et al. in 2005 or the dialectometric studies based on these fields (cf. Goebel 2003).

An ill-known or forgotten fact (cf. Tuailon 1975) is that the ALF survey, unlike most of the later ones, had a syntactic component. This was obliterated when the sentences were segmented for publication as lexical maps. For instance, the first volume starts as illustrated in Table 1, pointing to sheer lexical maps.

However, as the field notebooks show, the data for map 3, *abreuvoir* (‘trough’), have originally been elicited thanks to the translation of a full sentence: *Il mène les boeufs à l’abreuvoir* ‘He leads the oxen to the water trough’. The data gathered for *les boeufs* (‘the oxen’) correspond to map 141 and for *il mène* (‘he leads’) to map 835. Same thing for map 6 and 7, originally from *Je vais acheter deux chevaux à la foire et vendre les deux que j’ai*

² A few of them feature some syntactic phenomena. In other cases, some syntax was also included in the survey but is still unpublished, as is the case for the Picard domain.

³ Corsica was devoted a separate survey and a separate volume.

Table 1: First maps in the ALF.

Map Number	French Map Title
1	<i>abeille</i> ('bee')
2	<i>aboyer</i> ('to bark')
3	<i>abreuvoir</i> ('trough')
4	<i>à l'abri de la pluie, à l'abri du vent</i> ('sheltered from rain, sheltered from wind')
5	<i>absinthe</i> ('absinth')
6	<i>acheter</i> ('to buy')
7	<i>achetés</i> ('bought', past participle with masc. pl. agreement)
8	<i>acier</i> ('steel')
9	<i>quel âge</i> ('which age')

achetés l'an dernier 'I am going to buy two horses at the fair and sell the two that I bought last year' and map 9 *quel âge*, originally from *Quel âge as-tu?* 'How old are you'.

The SyMiLa project (<http://blogs.univ-tlse2.fr/symila/>)⁴ was set up, among other things, to make these syntactic data accessible. The elicited sentences were reconstructed by Guylaine Brun-Trigaud from the unpublished notebooks and the hints present in the published atlas. They amount to 180 full sentences, some of them grouped in coherent sequences. The corresponding data were then retyped using the IPA,⁵ and entered into an ad hoc database. One aim of the database is to patch together the relevant parts of each sentence for each location, respecting the right word order. This entails two technical difficulty: first, some maps feature discontinuous parts of a source sentence: for instance, map 201 features *ne...ni...ni* ('PRT...neither...nor'), which comes from *Je ne pouvais ni avancer ni reculer* 'I could neither go forward nor go back'; second, word order in the answers may differ among dialects. For now, 44 sentences out of 180 have been entered into the database (see Appendix 4). The sentences and each word in the French source sentence (called *stimulus* in the project) are tagged with morpho-syntactic properties, to enable targeted research in the database. The tagging of their dialectal versions is under way. Additionally, variants of Occitan words are linked to a standard version. The database, in its provisional state, with very basic research functions, can be accessed there: <http://symila.univ-tlse2.fr/ALF>.

What are the aims of making such data available? Since the syntax of Gallo-Romance dialects and its microvariation is, for the most part outside Occitan, unknown, the first aim is descriptive. The methodology for eliciting the data, *i.e.* recording the spontaneous translation of a stimulus in the standard official language, has well-known shortcomings (cf. Bucheli & Glaser 2002; Cornips 2002). However, I suggest, a first approximation is still better than nothing: it can trigger and guide further investigations. Not even mentioning the picture it can sketch of microvariation in such a large territory, it provides a first hint of structures that had gone completely unnoticed so far, and thus, positive data. These data can be of great import to syntactic typology, to the study of syntactic evolution in diachrony, and, to the extent that some other access to the dialects is provided, syntactic theory. Since the initial survey was not specifically designed for syntactic inquiry, the insight into syntactic micro-variation the ALF can provide may seem

⁴ SyMiLa is the current name of the project. It stands for Syntactic Microvariation in the Romance Languages of France.

⁵ Another ANR-funded project, ECLATS, (2015–2020), https://eclats.imag.fr/wordpress/?page_id=8, aims (among other things) at developing OCR tools able to retrieve and organize data from the maps, with each answer linked to the right location number. If it succeeds, it will help enrich and check the SyMiLa database.

modest. Nevertheless, the research conducted so far shows that it can have immediate contributions, that pave the way for deeper investigations. It was in particular crucial in the planning of the second part of the SyMiLa project, i.e. the conduct of new surveys, based on a fully syntax-oriented questionnaire, to further our insight into Gallo-Romance dialect syntax. The latter is centered on 5 syntactic targets, for which the ALF data have already shown they have non-trivial implications for syntactic typology and analysis: negation and Negative Concord, subjects, clitics, the structure of CP and agreement within DPs. The methodology is, to a large extent, kept similar to the ALF one (translation of a stimulus sentence), in order to get data more comparable to the ALF ones. The new questionnaire, however, provides contexts missing in the ALF to assess minimal pairs and control various parts of the potential analyses, as well as judgement tasks for limited crucial points. The latter can be optionally extended while passing the questionnaire. Unlike the first surveys, the new ones have been recorded: the sound files will be segmented and accessible in the database. About 50 new surveys have already been conducted,⁶ most of them transcribed and segmented, and they will be added to the same database.

In the next sections, I will focus on some findings provided or fostered by the ALF data. It will illustrate several ways in which they can contribute to syntactic studies, refine typologies, add independent ground to theoretical proposals and challenge others.

3 A horizontal perspective: Testing theoretical proposals

The surveys on which the ALF was based have interesting characteristics, especially compared to the later regional atlases. They were conducted, in a small time-span (less than 4 years), by the same interviewer, with the same questionnaire (only expanded for the Occitan area) and the same methodology, on more than 600 locations. As a consequence, lots of comparative data are available in a single place. The method, i.e. the spontaneous translation of a stimulus sentence phrased in the normative standard variety, has flaws. A major risk, especially given the *very* strong stigmatization of dialects in France, is that informants may produce calques of the stimulus. However, when a speaker, or, even more clearly, several speakers from a given series of neighboring locations, produce translations consistently differing from the stimulus language, there is little doubt that the data are reliable: on a regular basis, corpora and later surveys have confirmed it so far. When a given syntactic phenomenon is well-implemented in the atlas, then, we can get a pretty good image of its geographical extension, and of its variation.

Such a case is provided by direct *wh*-questions in the atlas. The ALF indeed features six *wh*-questions, ranging over the 639 survey points.⁷ They are given in Table 2; the data for the first three have been entered in the SyMiLa database, and can be checked on-line.

Note that the stimulus sentences only resorted to the French most standard interrogative structure at that time: *wh*-words immediately followed by Subject Clitic Inversion (henceforth SCLI). The data corresponding to the 3834 answers show that, while a calque or other influence effect of the stimulus on the translation cannot be discarded, it seems far from prevailing. The answers indeed display much variation, both in the CP and the TP domain. An individual speaker can produce consistently one structure for TP and CP throughout the six questions, or up to 6 different patterns (the latter case most frequently in the transition zones between known dialects).

⁶ Some of them are partial surveys: only the part of the questionnaire on negation and Negative Concord has been used.

⁷ A 7th one, *Combien cela peut-il bien valoir?* ‘How much may this cost?’ (map 1514) corresponds to a partial map (Occitan only), and is not taken into account here.

Table 2: Direct *wh*-questions in the ALF.

French map title	Translation	Relevant maps
Où vas-tu ?	'Where are you going?'	map 25
Quel âge as-tu ?	'How old are you' (lit. 'Which age do you have?')	maps 9, 86
Pourquoi ne vous mariez-vous pas ?	'Why don't you marry?'	maps 817 A&B
Quel temps fait-il ?	'How is the weather?' (lit. 'Which weather does it make?')	maps 1291 A&B
Qui veux-tu que ce soit ?	'Who do you want that it be?'	maps 1416, 517
Comment crie-t-il ?	'How does it shriek?'	maps 314, 355

Within the TP domain, three different patterns can be found: SVO order, like in declaratives; SCLI; or prodrop, for which SVO/VSO cannot be directly inferred from the plain data. Within the CP-domain, several patterns may also arise, according to three parameters:

- the presence (or absence) of Doubly Filled Comp (henceforth DFC) interrogative clauses, in which the *wh*-phrase is followed by a complementizer, the dialectal version of *que* 'that' – I will refer to it as *que* for all dialects.⁸
- the presence (or absence) of a dialectal version of *est-ce* 'is this', which can be reduced to /s/. In standard French, *est-ce* is always followed by *que*, and has been given various analyses. The two main lines are (a) considering *est-ce que* as a grammaticalized interrogative particle in Modern French both in polar and *wh*-question (cf. a.o. Detges & Waltereit 2008), in which *est*, *ce*, and *que* are not, or no longer, in free combination, or (b) considering that it still is some kind of cleft, its various elements moving to different sites inside the CP (see Munaro & Pollock 2005).
- the presence or absence of a dialectal version of *c'est que*, the usual pattern for building clefts in French and more generally in the Oil languages. Again, I will use French to refer to all dialectal occurrences. Note that, unlike what may be the case in corpora, the ALF doesn't provide examples of combinations of *est-ce que* and *c'est que*, or multiple instances of *c'est que*. I assume that such combinations involve actual clefting, which requires more pragmatic context than the translation of unrelated sentences can provide and that uncombined *c'est* in questions is not an instance of actual clefting, but a particle.

By a simple rule of thumb, if the choice of the CP structure and its combination with a TP one was unconstrained, we could expect up to 30 overall different combinations of similar frequency. Only a subset of the potential combinations of the CP and TP patterns are actually attested: they are given in Table (3), with one example for each pattern,⁹ followed (or preceded) by the ALF number of the relevant location.

This overall view can give rise to several questions. The first one is: are these patterns geographically constrained? To a large extent, the answer is yes. The second one is: are these patterns structurally constrained? To some extent, the answer is also yes. A third

⁸ For simplicity, I will use the French equivalent to refer to all dialectal comparable forms, unless a dialectal variant is clearly dominant.

⁹ The examples follow the SyMiLa conventions: they are given in the IPA, but the stress (when indicated) immediately precedes the vowel; besides, unless mentioned in the database, they faithfully duplicate the ALF original segmentation into words, indicated by a space, even when it may not correspond to my own analysis.

Table 3: Attested patterns for wh-questions in the ALF.

	SVO	SCLI	Prodrop
[1] WH	[u: ty va:] 110 where you go 'Where are you going?'	[jʊ v'a tu] 40 where go you 'Where are you going?'	[ʼum bas] 699 where go.2SG 'Where are you going?'
[2] WH que	[pʉrkw'ɛ k vu s mɛrje p'a:] 6 REFL=marry not 'Why don't you marry?'		[ke 'adʒe k a] 889 what age that have.2SG 'How old are you?'
[3] WH est -ce/s que	[ʼʉs kə ty v'ɛ] 23 [u'ɛs kə t ve] 49 where.PRT that you go 'Where are you going?'		889 ¹⁰ [kumɛ s kə kr'ia] how PRT that shrieks 'How does it shriek?'
[4] WH c'est que	[kūm t o k i gœlə] 931 [kmā t ɛ k e kri] 957 how PRT that it shrieks 'How does it shriek?'		963 [komɛ t i kə keze] how PRT that shrieks 'How does it shriek?' 470 [e'ʉr t u kə va:] where PRT that go 'Where are you going?'
[5] WH est- ce Ø	[uɛ:s ty va:] 158 [w'ɛs tə va:] 906 where. PRT you go 'Where are you going?'	104 [k ɛ s vœ t y] who PRT want you 'Who do you want (it to be)?' 183 [kɔmɛ s kri t i:] how PRT shrieks it 'How does it shriek?'	191, 193, 196 ¹¹ [wi:s va: s] where.PRT go.2SG 'Where are you going?'
[6] WH c'est Ø			
[7] WH que est-ce Ø			
[8] WH que c'est Ø		273 [ki k ʃe vœ ty k ʃa sœʃ] who that PRT want you that this be.SUBJ 'Who do you want it to be?'	
[9] WH que est-ce que			
[10] WH que c'est que	284B [dʉ k ʃ'e k te vo:] wherethat PRT that you go 'Where are you going?'		[parka k e ke vu marje pas] 827 REFL marry not 'Why don't you marry?'

question is: are they equally productive? The answer is no. I will first consider questions 1 and 3, then question 2.

¹⁰ I assume that /s/ is the reduced form of *est-ce*.

¹¹ Such Walloon constructions, which occur at the second person, can be analyzed either as partial prodrop, /s/ being construed as a verbal person inflexion, or as SCLI, /s/ being an allomorph of the 2nd person clitic *tu/te* when enclitic.

Three patterns are clearly dominant among Gallo-Romance dialects: [1] plus SCLI, corresponding to 1082 answers, [1] plus prodrop with 1118 answers and [2] plus declarative order with 1243 answers. They are also geographically constrained: WH + prodrop occurs, unsurprisingly, in the areas that are generally prodrop in declarative clauses (roughly, the Occitan areas), and in Walloon second person questions. DFC with declarative order appears in a large Northern half of France, SCLI in its Western and Eastern central parts and in Walloon (3rd person), as shown in Figure 1: the darker the dot, the more consistent the usage.

The presence of *est-ce que* (199 answers) and *c'est que* (12 answers) is also geographically marked, though not as clearly for *est-ce que*, as Figure 2 shows.

However, not all variants are geographically-driven. So-called “intonative questions” (74 answers), in which the WH-P is immediately followed by a declarative order, can be found all over the place, though only sporadically in the prodrop zone, cf. Figure 3.

Another way to look at the data is to ask whether we can draw any structural insights from them. A first generalization is that a *que* can either precede or follow *c'est* (or both), cf. [4], [8] and [10], repeated in (1).

- (1) a. [4] WH *c'est que*: [kũm t o k i gœlə] 931, [kmã t ε ke kri] 957; *; [komẽ t i kə keze] 963
 b. [8] WH *que c'est* Ø: *; [ki k ʃe vœ ty k ʃa sœʃ] 273; *
 c. [10] WH *que c'est que*: [dũ k ʃ 'e k te vo:] 284B; *; [parka k e ke vu marje pas] 827

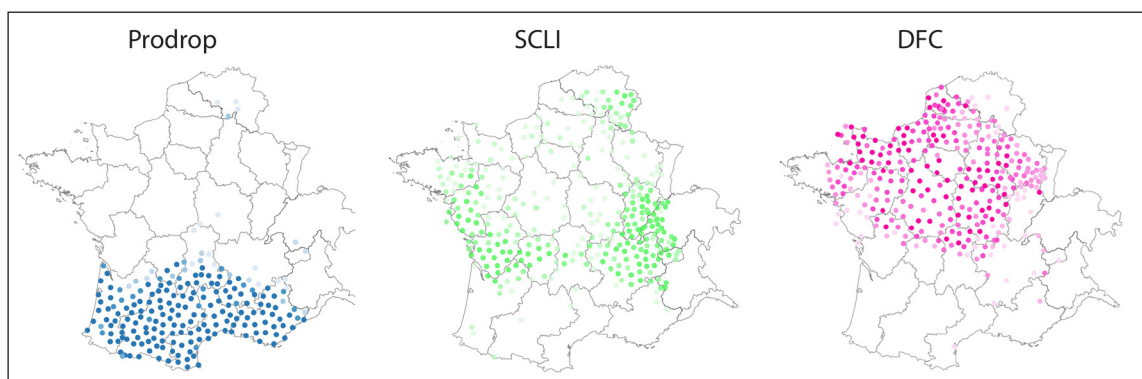


Figure 1: Dominant patterns for wh-questions in the ALF.

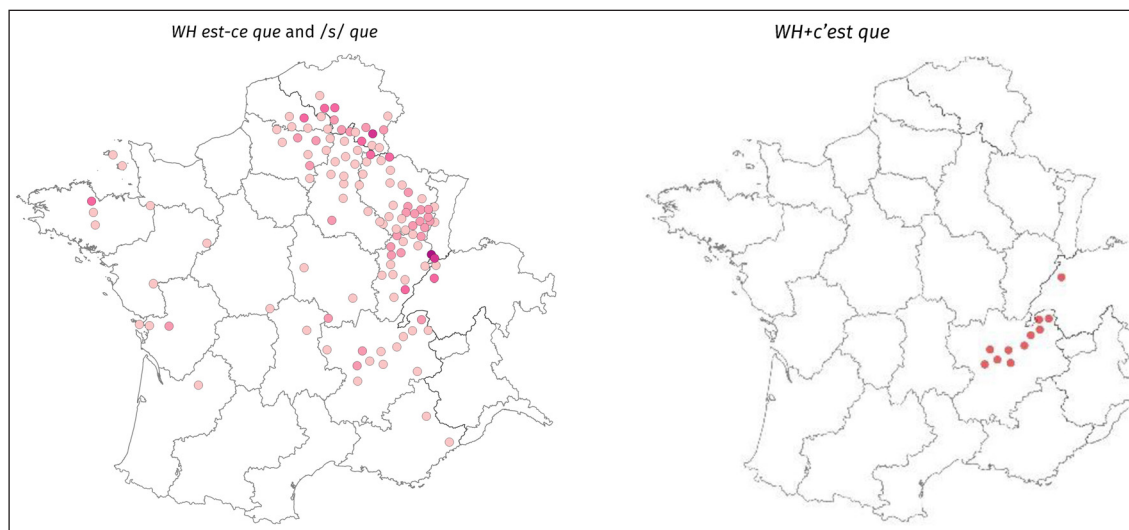


Figure 2: *Est-ce que* and *c'est que* in ALF wh-questions.

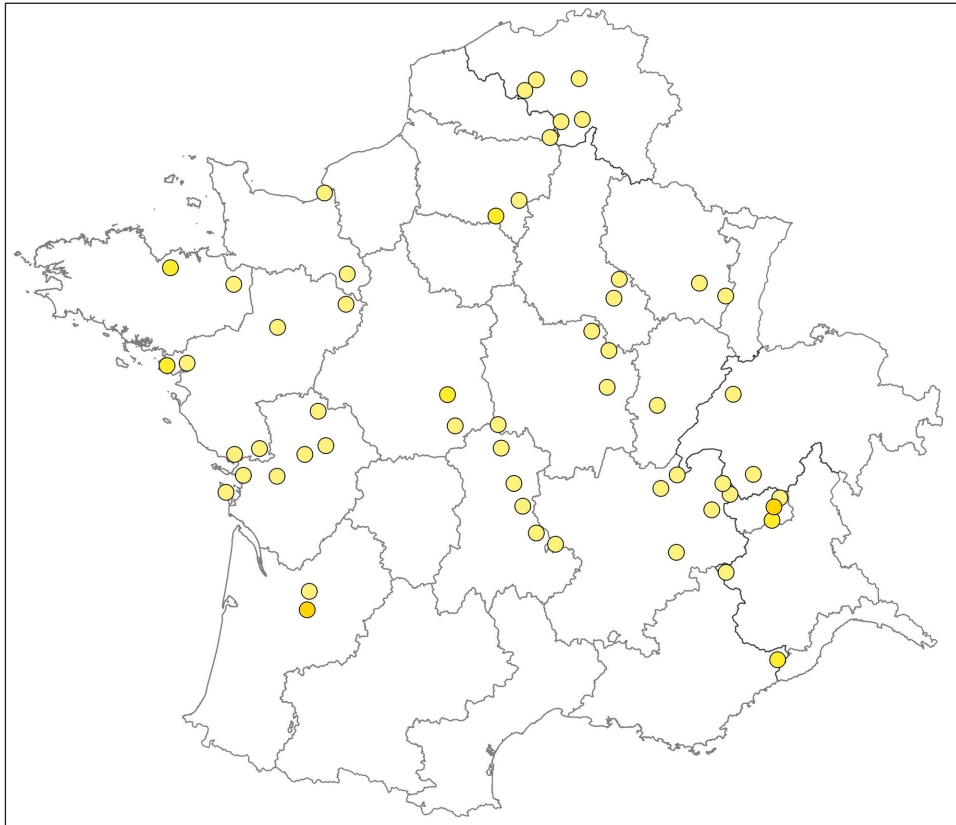


Figure 3: “Intonative wh-questions” in the ALF.

But it can only follow *est-ce*, cf. [3] vs [7] and [9] – though it must not necessarily do, as [5] shows –, cf. (2):

- (2)
- a. [3] WH *est –ce/s que* :
[¹ʁs kə ty v¹ɛ] 20, [u¹ɛs kə t ve] 49; *; [kumē s kə kr¹ia] 889
 - b. [7] WH *qu’est-ce Ø* : unattested
 - c. [9] WH *que est-ce que* : unattested
 - d. [5] WH *est-ce Ø*:
[ue:s ty va:] 158, [k ɛ s vœ ty] 104; [kəmē s kri t i:] 183, [wi:s vɑ: s] 191, 193, 196

If we consider these data, irrespective of their dialectal location and their frequency,¹² a first clear result is that SCLI only happens when no *que* is present, corresponding to lines [1], [3] and [8] in Table 4, while a *que*-construction, whether it seems to proceed from a Doubly Filled Comp construction as in [2], a *est-ce* construction as in [3] or a *c’est* one as in [4] and [10], is compatible both with declarative order and prodrop, but not with SCLI. This is materialized in Table 4 by the dark grey filling.

Though this is only positive data, the massive distribution gives additional support to Rizzi and Roberts (1989) proposal, based on standard languages, that SCLI proceeds from V/T to C movement. In a cartographic version of this idea, such as proposed by Poletto & Pollock (2004), the *que* in all these cases would stand in Fin°, where V/T raises to in its absence. Picard massively resorts to DFC; the Picartext database of Picard written texts

¹² These structures don’t have the same frequency, and while some can be found in written corpora for the given area, a lot of checking work is still needed: constituting corpora for each dialectal area, in order to check positive data, possible contexts and frequency, and fieldwork to check negative data, in particular for structures that are unattested in the ALF.

Table 4: Attested patterns in ALF wh-questions.

	SVO (a)	SCLI (b)	Prodrop (c)
[1] WH	[ʰ: ty va:] 110	[jʊ· v'a tu] 40	[ʰm bas] 699
[2] WH que	[pʰrkw'ɛ k vu s mɛrje p'a:] 6		[ke 'adʒe k as] 889
[3] WH est –ce/s que	[ʰs kə ty v'ɛ] 23 [u'ɛs kə t ve] 49		[kumɛ s kə kr'ia] 889 ¹³
[4] WH c'est que	[kūm t o k i gœlə] 931 [kmā t ɛ ke kri] 957		[komɛ t i kə keze] 963 ¹⁴ [e'ʰr t u k ə va:] 470
[5] WH est-ce Ø	[uɛ:s ty va:] 158 [w'ɛs tə va:] 906	[k ɛ s vœ ty] 104 [kɔmɛ s kri t i:] 183	[wi:s vɑ: s] 191, 193, 196 ¹⁵
[8] WH que c'est Ø		[ki k Je vœ ty k ʃa sœʃ] 273	
[10] WH que c'est que	[dʰ k ʃ'e k te vo:] 284B		[parka k e ke vu marje pas] 827

(<https://www.u-picardie.fr/LESCLaP/PICARTEXT/Public/index.php>) shows, for example, that infinitive wh-questions never display *que* after *doù* 'where', cf. (3), confirming that *que* is linked to the finite form of the verb:

- (3) A doù qu' ch'étoait? Doù s'acoufter? A doù s'éjuer ?
 where that it was where shelter.INF where play.INF
 'Where was it? Where to take shelter? Where to play?' (Collectif, Ponthieu-Vimeu)

Note that when *est-ce* is not followed by a *que*, cf. patterns in line [5], SCLI is possible in some dialects, a fact that, to my knowledge, had not been illustrated yet. It confirms the *que*/SCLI complementary distribution. We give the relevant examples in (4):

- (4) a. [k ɛ s vœ ty] 104
 who.PRT want you
 'Who do you want (it to be)?'
 b. [kɔmɛ s kri t i:] 183¹⁶
 how PRT shrieks it
 'How does it shriek?'

This strongly suggests that *est-ce que* is not a particle in the corresponding dialects, and that in (4) *est-ce* is not followed by a silent version of *que*: T raises to Fin in the absence of *que*. As for pattern [8], it allows SCLI after *c'est*, while *c'est* is preceded by *que*. The example given is the only one to illustrate the pattern in the ALF, I repeat it in (5):

¹³ I assume that /s/ is the reduced form of *est-ce* (/ɛs/).

¹⁴ Note that while the TP is prodrop, the demonstrative precedes [i] – *est* 'is' in the CP.

¹⁵ Such Walloon constructions, which occur at the second person, can be analyzed either as partial prodrop, /s/ being construed as a verbal person inflexion, or as SCLI, /s/ being an allomorph of the 2_{nd} person clitic *tu/te* when enclitic.

¹⁶ Note that the neighboring locations realize *crie* 'shriek' as [kri], not as [sekri], which can be found in other dialects. Besides, they resort to *est-ce que*, as is shown in (i) for location 184 and 185. The [s] here is then a realization of *est-ce*, not a part of the verb.

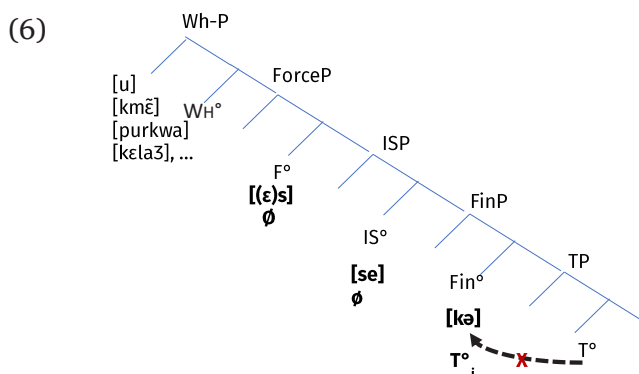
- (i) a. [komɛ s k i kri:] 184
 [kɔmɛ s k i kri:] 185
 how PRT that it shrieks

- (5) [ki k ʃe vœ ty k ʃa sœʃ] 273
 who that PRT want you that this be.SUBJ
 ‘Who do you want it to be?’

The *que* preceding *c’est*, then, shouldn’t proceed from the (Fin°) position V/T moves to. Actually, this example, as well as pattern [10a], comes from a small part of the Picard area (Artois/Ternois), in which the CP-domain has several particular properties (see section 3). The *que* could receive several analyses, for instance *qu’ ch’est* being the grammaticalized particle, or *que* being the head of the Wh-P or the ForceP. Note also that this is a long-distance question. I leave these single pattern aside for now.

A single analysis could capture almost all the ALF data despite their diversity: I propose wh-questions in Gallo-Romance dialects all share the structure in (6), where IS stands for Informational Structure – but nothing prevents, at least in the dialects illustrated in Figure 2, to consider it a mere variant of *est-ce*, an interrogative particle.

Under this view, dialects vary as whether *que* can/must be present or spelled out in Fin, whether the interrogative particle *est-ce* must/can be present or spelled out in Force to type it as a question, and whether V/T to C must apply in the absence of *que*. In this analysis, the term “Doubly Filled Comp” structures is not well suited: I assume that *que* is not the head of the Wh-P, but stands in another projection. This seems confirmed by the fact that, as some answers in the ALF show, the wh-word and *que* can be separated by discourse particles, cf. (7–9) just as *est-ce* and *que*, cf. (10). Besides, patterns [7] and [9], WH *que est-ce* ∅ and WH *que est-ce que* respectively, are then rightly expected to be unattested.¹⁷



- (7) a. [kumã dɔ̃ k a brøj] 1
 b. [kɔ̃m dɔ̃ k ɛ kri] 16
 how PRT that it shrieks
 ‘How does it shriek?’

- (8) [pukʰe: dɔ̃ k vu n vu mɛrje pʰɔ̃] 17
 why PRT that you PRT REFL marry not
 ‘Why don’t you marry?’

¹⁷ One anonymous reviewer objects to a Wh-P higher than Force. This indeed departs from Rizzi’s (1997; 2001) proposal for the left-periphery of questions. Along his lines, Wh-P s are expected to stand in FocP, as they are focal in some way; alternatively, they could raise to ForceP to type the sentence. However, these two assumptions are far from consensual cross-linguistically (see a.o. Aboh and Pfau 2011; Frascarelli 2012). For French, since Kayne and Pollock (2001), a Wh-P above ForceP has often been proposed, for different purposes (see e.g. Poletto & Pollock 2004; Munaro & Pollock 2005; Obenauer 2005). Of course, the consequences of this proposal need to be carefully checked for each dialect.

- (9) [u d'òk ty v'a:] 107
 where PRT you go
 'Where are you going?'
- (10) [kel a:ʃ ε s bẽ ky t e] 183
 which age PART PART that you have
 'How old are you?'

Though many things remain to be investigated more closely (in particular, among each given dialect, how the patterns above distribute and which predictions are made), the various structures corresponding to the elicited *wh*-questions in the atlas (up to 16, including *est-ce que* and *c'est que* questions) as well as the unattested ones argue for an articulated CP-domain. They also give additional evidence to the idea that the complementizer in “Doubly Filled Comp” constructions and the verb in Subject Clitic Inversion cannot co-occur since they target the same position – they provide more evidence that this position must be restated as *Fin*^o; they also argue for restricting the interrogative particle to *est-ce*: the *que* that follows it is not part of it, and has the same properties as *que* in DFC.

4 A vertical view: Small scale data and complementizer doubling

Another way to use the ALF data is to focus on a given small subset of neighboring locations, that may appear to form a dialectal unit. Thus, we can study the whole productions of a few individual speakers, that display phenomena that had gone unnoticed. This is the case for a small Picard zone in Artois, around Saint-Pol-sur-Ternoise (ALF 284). Most dialects of Picard display systematic (or frequent) Subject Doubling (see Auger 2003a; b; Appendix 3): any subject DP is doubled by a subject clitic. The ALF has several temporal clauses, which in this area, like in some other dialects (see Appendix 1), are introduced by a doubly filled complementizer, *quand que* ‘when that’. Two of the ALF temporal clause have a DP subject – I give them in (11). (11a) corresponds to maps 143, 458, 548A, 548B, 867, 1109 and 1405, (11b) to maps 573, 514, 469 and 972.

- (11) a. Il buvait moins quand sa femme vivait encore.
 he drank less when his wife lived still
 'He used to drink less when his wife was still alive.'
- b. Quand mon fils sera grand, je l'enverrai à Paris.
 when my son be.FUT big I him = send.FUT to Paris
 'When my son will be older, I'll send him to Paris'

The ALF also has a partial map corresponding to a relative clause with DP subject, (*la charrette*) *que le domestique a chargée* ‘(the cart) that the servant loaded’ (map 1537), and an embedded exclamative with DP subject, (*regarde*) *comme les arbres en étaient chargés* ‘(look) how the trees were loaded [with apples]’ (maps 310, 52, 513, 240).

In the small zone we are considering, these may be realized in two fashions: the *que* usually following *quand* ‘when’ can be delayed after the subject DP, as in (12), suggesting that, as for questions (cf. section 2), it stands in *Fin*, the subject DP having moved to the CP-domain:

- (12) [kã mε garʃõ k i sra grã] (287)
 when my son that he be.FUT older
 'When my son will be older (I will send him to Paris.)'

However, it can also be repeated: in this case, both instances of *que* frame the subject DP, as in (13).

- (13) [kã k mɛ fju k i sro grã] (286)
 when that my son that he be.FUT older
 ‘When my son will be older (I will send him to Paris.)’

The extension and rate of this phenomenon are given in Figure 4.

Since the two sentences are not enough to understand what is going on, and since this construction is not mentioned in the (scarce) Picard grammars and monographs, these ALF data just provide an impulse to investigate further. Corpus research confirmed that these structures are attested (though optional) in this area and, to a lesser extent, in neighboring areas, and are present in any kind of finite embedded clauses (see Dagnac 2012). It also showed that they are not limited to doubled subjects: in any kind of finite embedded clause, any left peripheral item, whatever its phonological weight or Information Structure status, can trigger (12) or (13), and a same author can produce both. In fact, the data from Ternois Picard mirror data from Southern Italian dialects (cf. Ledgeway 2005; D’Alessandro & Ledgeway 2010) and are fully consistent with their analysis: the lowest *que* is not the head of an Informational Structure phrase (vs Paoli 2007; Mascarenhas 2007; and – with qualifications – Uriegareka 1995; Demonte & Fernandez Soriano 2009; Villa-García 2010). It arguably stands in *Fin*^o, can undergo head-movement up to ForceP to check the subordinate specification of Force. In this process, both ends of the chain can be spelled out. Two elements add evidence to this proposal in Picard. First, unlike *que*, the infinitive complementizer *de* cannot be doubled, and can only occur in *Fin*^o – which

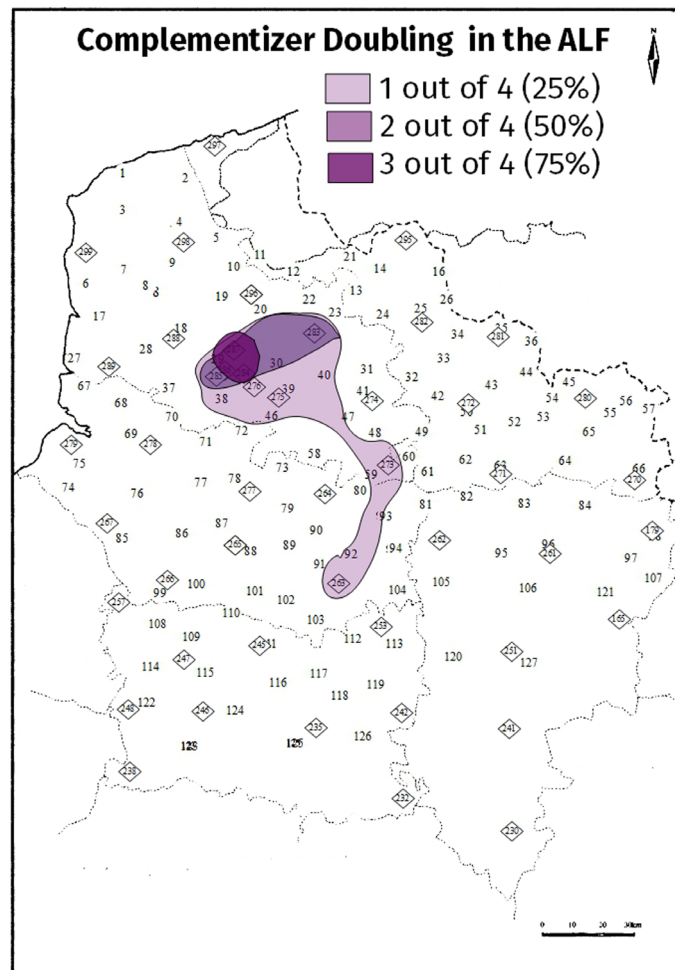


Figure 4: Complementizer doubling in the ALF.

is expected under Rizzi's (1997) analysis of the extended CP: infinitives lack the higher C position (identified as Force here), so *de* would have no higher position to go to. Second, there is an exception to complementizer doubling in embedded clauses: the relative pronoun *qui* is not followed by *que*, either single or doubled. Relative *qui* is nominative hence [+Fin], but it is also plausibly [+Sub] since it can only head embedded clauses in this dialect.¹⁸ It is then able to check both the Fin and the Force specification through wh-movement: no *que* is needed, since *qui* checks the tense feature itself on Fin, and the Subordinate feature on Force.

This phenomenon also raises non-trivial questions for diachronic perspectives within the Oïl dialects, including French. It is indeed attested, though sporadically, by Flutre (1970) in the first half of the 17th century, for Picard in a slightly larger region, and can be traced back at least to 14th century charters of the Douai/Arras region, cf. (14):

(14) Giraud et al. (2010)

[...] Et si ont enconvent li ditte Margherite, Sainte etc..., à le ditte
 and PRT have agreed the said Marguerite, Holy etc., with the said
 Evain [...], **ke** dedens les XL jours [...], **ke** il li feront,
 Evain **that** within the 40 days that they to.her have.FUT
 werpir et avoir enconvent à aquiter les dittes maisons [...].
 leave and have agreed to return the said houses
 'And Marguerite and her children formally agreed with Evain that in a 40 day
 delay [after her daughter is of age] they will return her the said houses.'

Besides, Wanner (1995) notes that complementizer doubling is attested in Old French, and Foulet (1919: 266–267, §402) presents it as frequent, though he illustrates the fact with four citations, three of which are from Picard literary texts, and one (*La Chastelaine de Vergi*) has been expertized as coming from the Paris region but with dialectal features from Picardy and Burgundy. Same thing for Buridant (2007), who quotes one Picard literary text from the XIIIth c., and one Anglo-Normand text. Given that Picard was, after Anglo-Normand, one of the first Oïl dialect to have a written tradition (Lusignan: 2011; 2012), these facts suggest that the geography and diachrony of French and Oïl dialects would greatly benefit from being studied in parallel. It is also attested for French in the 16th (cf. Gougenheim 1984: 149) and early 17th centuries (Haase 1898: 136.B), then seems to disappear, at least from grammars. Complementizer doubling is also attested in Old and Middle Occitan (Lafont 1967). Such a syntactic feature could be a Romance feature that got lost in French and most Oïl dialects but kept being transmitted in a sub-area of Picard; or it could be the case that literary French inherited a feature from dominant written models from Artois & Occitan and lost it since. In any case, considering the dialectal syntactic data from the ALF as seriously as the lexical and phonological data should provide an impulse for diachronic studies.

The ALF can then also act as a small scale tool to identify infrequent, local, but theoretically important phenomena: in this case, it adds independent evidence that the complementizer can undergo head-movement within an articulated CP-domain, and confirms that, within the Romance languages, the typology of complementizer doubling (Southern Italian dialects, Ternois Picard) and recomplementation (Spanish, Portuguese) relies on different syntactic structures, correlated to different

¹⁸ The interrogative animate pronoun, which is homophonous to the nominative relative pronoun in French, is morphologically distinct in this dialect.

distributional properties, while bringing to light new questions to investigate within diachronic studies.

5 Refining typologies and evaluating competing analyzes

Another way in which the ALF data can contribute syntactic analysis is by enriching or refining typological generalizations, and, plausibly, the theoretical analyzes they may inform. The previous sections have shown how they may argue for an at least semi-free combination analysis of *est-ce que*, or for a sub-typology of complementizer doubling. In this section, I will deal with two more cases: the first one deals with the link between polar and wh-questions interrogatives, the second one with negation and Negative Concord (henceforth NC).

5.1 *Wh-movement and yes-no particles*

Cheng (1991) proposes that the typological distinctions among languages with respect to the formation of wh-questions are directly linked to the presence/absence of interrogative particles: the presence of particles correlates with the lack of wh-movement. In particular, she states and argues: “No language has yes-no particles and also wh-movement” (Cheng 1991: 37). Her generalization is twofold (Cheng 1991: 21–22 – her (3) and (4)):

- (i) In situ languages have special markings in yes-no questions.
- (ii) Languages with special marking in wh-questions are in-situ languages.

Data from the very same dialect we considered for complementizer doubling, in the Ternois region, show that it is a wh-movement dialect. Neither the ALF nor the corresponding written corpora display *in-situ* questions, and the present speakers we could test disallow them. Yet, it has a specific marking in yes-no questions. The ALF offers a few yes-no questions, whose stimuli are listed in (15). (15a) corresponds to maps 358, 1298, 1298B, (15b) to map 575, whose data are incomplete, (15c) to maps 1417, 30, 470, 470B, 1118, (15d) to maps 85 and 775, (15e) to half-maps 1569, 1545, 1454 with some answers in the relevant area.

- (15)
- a. Crois-tu qu'elle tienne ?
'Do you think it will hold?'
 - b. Finiras-tu ?
'Will you stop?'
 - c. Voulez-vous que j'aïlle ou que j'envoie quelqu'un ?
'Do you want me to go or to send someone?'
 - d. L'as-tu lu ?
'Did you read it?'
 - e. Votre fillette est-elle déjà baptisée ?
'Is your little girl already baptized?'

These maps reveal sporadic occurrences, in this small sub-area of the Picard-speaking domain (locations 284, 285, 286, 287, 288 – see Appendix 2), of a propositional interrogative particle, *jou* (lit. 'me') in direct yes/no questions, as illustrated in (16–17):

- (16) [ʒu k te krwɛ k al taro] (287)
PRT that you believe that it hold.FUT
'Do you believe it will hold?'

- (17) [ʒu k t ell o li:] (285)
 PRT that you it=has read
 ‘Did you read it?’

A confirmation that *jou* is an interrogative particle can be found in map 659 *Est-ce que tu viendras?* ‘Will you come?’ of the Picard regional atlas, ALPIC (Carton & Lebègue 1998), and in written corpora: they all show that *jou* can also act as a verbal interrogative particle, cf. (18). It competes with a *-ti* or *-tu* verbal particle as it exists in Québec French (Vecchiato 2000; Vinet 2000; Morin 2009) and in Picard:

- (18) Il a jou cryi ? (Picartext: Devraine, Vermandois)
 he has PRT shouted
 ‘Did he shout?’

Note that *jou* is, and can only be, found in *direct* questions (unlike English *whether*, for instance) and thus qualifies as a typing particle as defined by Cheng. Besides, Picard doesn’t license multiple wh-fronting (vs Polish) though it allows multiple questions (vs Italian): along Cheng’s lines, it then appears to be an actual counterexample to (ii), of a sort that is not discussed in her thesis. The *in situ* languages she examines can realize the same or a different interrogative particle in direct wh-questions too, or display a covert one in these contexts. As pointed out by Emrik (1966/69: 4), based on ancient written texts from the surrounding Picard areas, the typing particle *jou* could optionally combine with several wh-words in previous stages of Picard, as shown in (19).¹⁹

- (19) a. Voyon èn peu c’ment jou qu’ os frons ? (18th c., Corbrie)
 see.SUBJ.P4 a bit how PRT that we do.FUT
 ‘Let’s see how we will proceed.’
 b. **Quand jou k’** ô vîz avon vu sans lôji ? (19th c., Cambrai)
 when PRT that we you have seen without home
 ‘When did we see you homeless?’
 c. **Quoi jou qu’ os** foesiez lo ? (20th c., G. Leroux d’Etaves-Bocquion)
 what PRT that you do here
 ‘What are you doing here?’

However, since this WH- + *jou* construction does not appear in the ALF, nor in the corpora we checked for the Ternois region for the same period onwards, we may assume that Ternois Picard currently checks [+Q] with the typing particle *jou* in yes-no questions and

¹⁹ Current texts in these areas may also show sporadic use of *jou* in wh-questions, but, as far as I can tell, always in association with *quand*, cf. (i-ii), and also in indirect questions, cf. (iii), though I couldn’t find it in temporal clauses: this points to the grammaticalization of *quanjou* as the interrogative form (vs *quand* in temporal clauses). Further investigation is needed on this point.

- (i) Quanjou qu’tu viens ?
 when=PRT that you come
 ‘When will you come?’ (Calais, Amiénois)
 (ii) Quanjou, pi d’où qu’il est mort ? D’où qu’o l-l’o intèrrè ?
 when=PRT and where that he is dead where that one him=has buried
 ‘When and where did he die? Where was he buried?’ (Collectif, Ponthieu-Vimeu)
 (iii) O n’savoait point quanjou qu’ él darin.ne candelle al alloait s’èteinde...
 one PRT=knew not when=PRT that the last candle she was going to go out
 ‘We didn’t know when the last candle would go out...’ (Calais, Amiénois)

by a covert particle in *wh*-questions. As suggested by the analysis in section 2, *est-ce* may alternate with this covert typing particle in *wh*-questions.

If obligatory *wh*-movement in *wh*-questions of one (and only one) *Wh*-P is assumed to be triggered by Clause Typing, in order to check an interrogative feature, Cheng's generalization and analysis don't hold. However, under a cartographic approach, it has been assumed that *wh*-movement may happen to check both a clause-typing feature, [+Q], in ForceP and a *wh*-feature in *Wh*-P (Poletto & Pollock 2004). The Ternois Picard data support the claim that *wh*-items must only check a [+WH] feature, [+Q] being checked in diachrony by *jou*, and in synchrony either by a silent version of *jou* or by *est-ce*. Cheng's generalization can then be maintained under the strict reading in which "*in situ* languages" correspond to languages in which *wh*-words don't check the [+Q] feature. This can apply to "surface" *in-situ* languages or to languages where obligatory *wh*-movement applies (only) for other reasons – a proposal already made by Cheng for some *wh*-fronting languages.

5.2 Negation and Negative Concord

All dialects of Picard also challenge the typology of Negative Concord languages. The ALF includes 16 maps with propositional negation. Very sporadically, they suggest that Picard speakers may have more than one negative marker. This is in fact largely confirmed by the ALPIC data and corpus work: many speakers have (at least)²⁰ two markers for propositional negation (*cf.* Dagnac 2015). I will lemmatize one as *point* (though it may have several phonetic and graphic realizations), the other as *mie*. The first one is way more frequent and can be found in all usual syntactic contexts; the second one seems to be restrained to tensed clauses. Both are *m*-negations in the sense of Garzonio & Poletto (2009; 2010), *i.e.* they have both been grammaticalized from former minimizers, like French *pas*.²¹

- (20) N'te casse point t'tête ! A n'va mie dureu.
 PRT = you.dat break not your = head it PRT = goes not last
 'Don't worry! It won't last.' (Vasseur, Vimeu)²²

When speakers have *mie*, unlike French *pas* (Corblin et al. 2004), it enters Negative Concord (henceforth NC) with all *n*-words.²³ I show this in (21) with a single author from the Vimeu region. As (22) shows, the *n*-word can also be preverbal, unlike what happens in Quebec French, where preverbal *n*-words preclude Negative Concord (*cf.* e.g. Burnett et al. 2015):

²⁰ Some speakers also have *pas*, which may in some cases be constrained to some syntactic contexts; when it is unconstrained, it usually works like French *pas*, and is excluded from Negative Concord. See Dagnac (2015) for details.

²¹ In Old French, *mie* and *point* also competed with *pas*. Note that Picard *mie* differs on crucial aspects from Old Italian and dialectal *mica* (Garzonio & Poletto 2012: 139–141): for instance, it cannot be clause-initial and it can appear without the preverbal particle *ne*. It also differs from Lorraine *mie* (Burnett 2015).

²² Unless otherwise specified, all examples in this section are from the Picard written corpus, Picartext. The reference gives the writer's name, followed by the geographic area.

²³ All dialects of Picard also have negative spread, in which multiple occurrences of *n*-words yield a single semantic negation. Besides, their *n*-words have a negative content of their own. They also seem to have double negation, in the right pragmatic conditions, and speakers seem reluctant to use them in NPI contexts (if-clauses or questions, for instance): fieldwork is under way to check these last two points. Note that, though Picard retains more preverbal *ne* than other Oil dialects (Auger & Villeneuve 2008), *ne* cannot be negative on its own and the presence of *ne* is not compulsory for NC or negative spread to occur, *cf.* (i) and (ii) respectively, and Dagnac (2015).

- (i) [..] j'ai point peu rien rapporter dé l' foère.
 I = have not can.PTCP nothing bring from the fair
 'I couldn't bring anything back from the fair.' (Dumont, Vimeu)
- (ii) J'm'in vos tchitter la terre, person.ne i m'voéro pu !
 I = am going to leave the earth nobody he = me = see.FUT no longer
 'I am going to die, nobody will see me any longer!' (Calais, Amiénois)

- (21) a. O n'a **mie jamoais** vu o **neune pèrt** [...].
 one PRT=has not never seen this nowhere
 'This has never been seen anywhere.' (Vasseur, Vimeu)
- b. [...] o n'a **mie pu** bsoin d' parsonne dins chés fermes.
 one PRT=has not no longer need of nobody in the farms
 'There is no longer any need of anybody in farms.' (Vasseur, Vimeu)
- c. D'abord j' n'avoés **mie rien** à dire ;
 first I PRT=had not nothing to say
 'First, I had nothing to say;' (Vasseur, Vimeu)
- d. I n'éront **mie rconnu** **parsonne**.
 they PRT= have.FUT not recognized nobody
 'They won't have recognized anybody.' (Vasseur, Vimeu)
- (22) Autrémint, **parsonne** én' n'érouot **mie** volu d' chu
 otherwise, nobody PRT=of.it=have.COND not wanted of this
 rmimbrémint.
 reparcelling
 'Otherwise, nobody would have accepted this reparcelling.' (Vasseur, Vimeu)

However, unlike what happens in some standard languages, whether they are non-strict or strict Negative Concord languages (Giannakidou 1998), i.e. whether the presence of the negative marker depends on the syntactic position of the n-word or not, Negative Concord is never compulsory in Picard. Dagnac and Burnett (2016) show that, for a same speaker, and putting aside the cases where more than one n-word appears in the sentence, it can range from 3% to 90% of the occurrences involving a n-word, according to the n-word involved. In this respect, Picard patterns with Quebec French: together with other languages and dialects like varieties of Occitan or Catalan, they show that the strict vs non-strict concord languages typology needs refining, yielding at least a four-way distinction: "compulsory strict NC" (Romanian), "optional strict NC" (Picard), "compulsory non strict NC" (Italian), "optional non strict NC" (Québec French). Besides, analyses of NC must account for this optionality. But Picard argues for an even finer-grained approach to optionality, as shown by Dagnac (2015).

The above generalization was indeed based on the behavior of the less frequent negative marker, *mie*. When it comes to the main negation, *point*, things are different. For a subgroup of speakers, *point* (vs *mie*) cannot enter NC at all: this gives us a second instance of negative markers that are unable to enter NC, as is the case with French *pas* (and Picard, for those speakers that also have it). It appears in line with theories of NC that, some way or other, place the responsibility for NC in the semantic or syntactic features of the negative marker, rather than in those of the n-words (see e.g. De Swart & Sag 2002 for an implementation). But some Picard speakers present an even more challenging use of NC. While they still have (optional) concord with *mie* and any n-word, they have (optional) selective concord with *point*. By this, I mean that *point* can enter NC only with a subset of n-words: *point* is allowed to co-occur with the nominal n-words like *parsonne* 'nobody' or *rien* 'nothing', but it is disallowed with the temporal adverbial ones *jamais* 'never' and *pu* 'no longer'.²⁴ This is hardly compatible with analyses in which NC relies on some syntactic

²⁴ As far as we can tell, *point* is compatible with *neune pèrt* 'nowhere', cf. (i) at least for one speaker. But we have very few occurrences of *neune pèrt*, so its behavior for all partial NC speakers remains to be checked.

- (i) O n'érlante **point** neune pèrt !
 one PRT=replants not nowhere
 'One doesn't replant anywhere!' (Vasseur, Vimeu)

or semantic version of resumptive quantification (cf. a.o. Zanuttini 1997; De Swart & Sag 2002): in their present state, these analyses do not predict that NC could be sensitive to the particular lexical properties of specific n-words. Conversely, analyses in which NC proceeds from the properties of n-words, for instance by assuming that they are ambiguous between negative quantifiers and Negative Polarity Items (see e.g. Herburger 2001) could account for this behavior of *point*: it could be the case that *parsonne* and *rien* are ambiguous, while *jamais* and *pu* are not: the latter, being only negative quantifiers, are predicted to yield double negation when combined with the negative marker. However, such an approach wrongly predicts that the same pattern would occur with *mie*.²⁵

Interestingly, Garzonio and Poletto (2012) have shown that Old Italian *niente* presents a different Negative Concord pattern according to its argumental vs non-argumental usage, a pattern reminiscent of the Picard opposition between, on the one side, argumental *parsonne* and *rien*, and on the other side non argumental *jamais* and *pu*.²⁶ In its non-argumental use, *niente* triggers obligatory Negative Concord; in its argumental use, *niente* triggers NC only optionally, as does *mai* ‘never/ever’. The authors explain this pattern in structural terms: only positions occupied by the adverb (between AspP and VoiceP) can yield NC, while VP internal positions don’t yield NC. The merge position then bans NC with argumental *niente*. Seemingly optional NC with argument *niente* only occurs when it can move to the adverbial position. However, this analysis cannot straightforwardly be extended to Picard, since we get the inverse generalization: in Picard, with *point*, NC is optional with arguments and excluded with adverbials. Besides, in Picard, *rien* can move to the complex T domain above the past participle verb, as in (23). However, it can also stay within an argument infinitive, as in (24), or even in a VP internal position, as in (25). As (23b), (24) and (25b) show, its position does not influence its ability to enter NC – in particular, cf. (25b), NC obtains even when it is in a merge argument position.

- (23) a. Jé n’i ai **rien** répondu, bien intindu.
I PRT = to.him = have nothing answered of course
‘I didn’t answer anything to him, of course.’ (Vasseur, Vimeu)
- b. O n’n a poé **rien** dit à parsonne.
we PRT = of.it = has not nothing told to nobody
‘We didn’t tell anything to anyone about it.’ (Vasseur, Vimeu)
- (24) a. [...] j’ai point peu **rien** rapporter dé l’foère.
I = have not can.PTCP nothing bring.back from the fair
‘I couldn’t bring anything back from the fair.’ (Dumont, Vimeu)
- b. [...] i y o point dzoin dé **rien** nous dire.
there is not need to nothing us = tell
‘we don’t need to say anything to each other.’ (Calais, Amiénois)
- (25) a. O n’a jamoais foait **rien** d’ si réussi [...].
one PRT = has never done anything of so beautiful [...]
‘We never did anything as beautiful since we started painting.’
(Vasseur, Vimeu)
- b. O n’a poé bzoin d’ cangeu **rien**.
one PRT = has not need to change nothing
‘We don’t need to change anything.’ (Vasseur, Vimeu)

²⁵ An option would be to consider *mie* itself as ambiguous between a n-word and a NPI, *point* being the only actual negative marker. In this case, *mie* would enter negative spread rather than negative concord. However, so far, there is no independent evidence that *mie* can act as a NPI.

²⁶ In its argumental uses, *niente* can be translated as ‘nothing/anything’, while in its non-argumental uses it corresponds to ‘at all’.

Very sporadic ALF data can then be used to detect an unexpected syntactic pattern or variation. In this case, the main contribution of the ALF was to suggest lexical variation: some Picard speakers use *mie* instead of *point*. But coupled with other tools of investigation, such as localized corpora, they contributed to the uncovering that these two markers have contrasted usages: one of them, *mie* ‘not’, allows (optional) Negative Concord, while the other, *point* ‘not’ either rules it out, or allows it only with a subset of n-words. This partial concord, though in some ways converging with Quebec French Negative Concord, differs from Old Italian partial concord – thus adding to the growing bulk of evidence (Déprez & Martineau 2004; 2006; Burnett et al. 2015; Dagnac & Burnett 2016; Franco et al. 2016) that the typology and analysis of Negative Concord must be refined to deal with both optional and partial concord, even in the Romance languages.

6 Conclusion

Though the ALF was not specifically conceived to study dialectal syntax, and though the methodology used to collect the data is not perfect for theoretical analysis, I have shown that it can still be an invaluable tool in a domain that dramatically lacks literature. There are many ways to use this tool. First, answers to the same stimuli over a territory and/or a given syntactic structure, produce consistent data, concerning for instance Doubly Filled Comps in Wh-questions, prodrop or subject doubling: they can give us a first rough sketch of the geographical domains where a phenomenon is expected to happen. This kind of large-scale, “horizontal” view, which the SyMiLa group will keep on establishing, can be capitalized upon at least in two manners. First, it can signal where we should go and investigate in a finer-grained way in priority: given the extent of the work to be done in order to describe and analyze the syntax of Gallo-Romance dialects, this *per se* is welcome. As I showed in section 2 for wh-questions and 4.2 for polar questions, it can also help draw and test generalizations on the structure of grammar, confirming the need for an articulated CP, restating what “Doubly Filled Comp” or *est-ce* may be, severing clause typing from wh-words in Romance. The ALF can also be used as a small scale, “vertical” tool, to spot infrequent phenomena such as complementizer doubling or the presence of alternate lexical forms such as *mie*, which are very isolated in the atlas; trusting the informants that produced them can then be an incentive to try and understand why these forms surface and start setting up corpora and specific surveys, at least while there are some speakers left. Doing so on these two cases lead us to add evidence that the typology of complementizer doubling must be revised, one type relying on head-movement of a Fin head, the other one on the presence of Informational Structure heads, and that the typology and analysis of Negative Concord should encompass both optional and partial concord.

Lastly, these first results point the fact that the ALF can serve at least two long-term objectives. First, it invites us to reassess, as a few have already started doing, the diachrony of French syntax by taking into account its potential links with the syntax of the dialects that co-existed with it for centuries, especially its next of kin. Second, it can, to some extent, enable us to assess how syntax connects (or not) with the better known phonological and lexical features that have been used to define dialects and sub-dialects, a non-trivial question both for dialectology and for the structure of grammars.

Abbreviations

FUT = future, COND = conditional, DET = determiner, PRT = particle, PTCP = (past) participle, SUBJ = subjunctive, ALF = Atlas linguistique de la France, ALPIC = Atlas linguistique et ethnographique picard, CP = complementizer phrase, DFC = Doubly Filled Comp, FinP = Finiteness Phrase, ISP = Information Structure Phrase, NC = Negative Concord, NPI = Negative Polarity Item, SVO = subject-verb-object, SCLI = Subject

Clitic Inversion, SyMiLa = Syntactic Microvariation in the Romance Languages of France.
VP = Verb Phrase, TP = Tense Phrase

Additional File

The additional file for this article can be found as follows:

- **Appendix.** The appendix provides additional maps alluded to in the paper (DFC with *quand* ‘when’, interrogative particles and DP subject realization), as well as the full list of sentences initially elicited during the ALF surveys. DOI: <https://doi.org/10.5334/gjgl.543.s1>

Acknowledgements

This research was carried out as part of the SyMiLa project (ANR-12-CORP-0014), funded by the Agence Nationale de la Recherche. I am grateful to the reviewers for their helpful comments on previous versions of this paper, to other members or collaborators of the SyMiLa project, and, most of all, to Guylaine Brun-Trigaud, who introduced me to the ALF.

Competing Interests

The author has no competing interests to declare.

References

- Aboh, Enoch O. & Roland Pfau. 2011. What’s a *wh*-word got to do with it? In Paula Benincà & Nicola Munaro (eds.), *Mapping the left periphery: The cartography of syntactic structures*, 91–124. New York: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199740376.003.0004>
- Auger, Julie. 2003a. Les pronoms clitiques sujets en picard: Une analyse au confluent de la phonologie, de la morphologie et de la syntaxe. *Journal of French Language Studies* 13 (1). 1–22. DOI: <https://doi.org/10.1017/S0959269503001066>
- Auger, Julie. 2003b. Le redoublement des sujets en picard. *Journal of French Language Studies* 13(3). 381–404. DOI: <https://doi.org/10.1017/S0959269503001200>
- Auger, Julie & Anne-José Villeneuve. 2008. Ne deletion in Picard and in regional French: Evidence for distinct grammars. In Miriam Meyerhoff & Naomi Nagy (eds.), *Social lives in language: Sociolinguistics and multilingual speech communities: Celebrating the work of Gillian Sankoff*, 223–247. Amsterdam: John Benjamins.
- Brun-Trigaud, Guylaine, Yves Le Berre & Jean Le Dù. 2005. *Lectures de l’Atlas linguistique de la France de Gilliéron et Edmont: Du temps dans l’espace*. Paris: Comité des Travaux Historiques et Scientifiques.
- Bucheli, Claudia & Elvira Glaser. 2002. The Syntactic Atlas of Swiss German Dialects: Empirical and methodological problems. In Sjef Barbiers, Leonie Cornips & Susanne van der Kleij (eds.), *Syntactic Microvariation*, 41–74. Amsterdam: Meertens Institute. [On-line: <http://www.meertens.knaw.nl/books/synmic/>].
- Buridant, Claude. 2007 [2000]. *Grammaire nouvelle de l’ancien français*. Paris: SEDES.
- Burnett, Heather. 2015. La négation dans les langues romanes de France: Données de l’Atlas linguistique de la France. Talk at the *Déterminants et Inférences Workshop*. Université Paris-Sorbonne, June 2015.
- Burnett, Heather, Mireille Tremblay & Hélène Blondeau. 2015. The variable grammar of Montréal French Negative Concord. In Sabriya Fischer (ed.), *Selected papers from NWAV 43* (Penn Working Papers in Linguistics 21 (2)), 10–20. Philadelphia: University of Pennsylvania.
- Carton, Fernand & Maurice Lebègue. 1998. *Atlas linguistique et ethnographique picard*, 2. Paris: Ed. CNRS.

- Corblin, Francis, Viviane Déprez, Henriëtte de Swart & Lucia Tovena. 2004. Negative concord. In Francis Corblin & Henriëtte de Swart (eds.), *Handbook of French semantics*, 417–452. Stanford, CSLI Publications.
- Cornips, Leonie. 2002. Variation between the infinitival complementizers om/voor in spontaneous speech data compared to elicitation data. In Sjef Barbiers, Leonie Cornips & Susanne van der Kleij (eds.), *Syntactic Microvariation*, 75–96. Amsterdam: Meertens Institute. [On-line: <http://www.meertens.knaw.nl/books/synmic/>.]
- Dagnac, Anne. 2012. How do you double your C? Evidence from an oil dialect. *Empirical Issues in Syntax and Semantics* 9. 77–94. [On-line: <http://www.cssp.cnrs.fr/eiss9/index.html>].
- Dagnac, Anne. 2015. Pas, mie, point et autres riens: de la négation en picard. In Jan Goes & Mariana Pitar (eds.), *La négation: études linguistiques, pragmatiques et didactiques*, 129–152. Arras: P.U. Artois (Artois Presse Université).
- Dagnac, Anne & Heather Burnett. 2016. Concordance négative optionnelle: Contrastes forts et faibles entre picard et québécois. In Franck Neveu, Gabriel Bergounioux, Marie-Hélène Côté, Jean-Michel Fournier, Linda Hriba & Sophie Prévost (eds.), *5e Congrès Mondial de Linguistique Française* (SHS Web of Conferences 27), 1–12. DOI: <https://doi.org/10.1051/shsconf/20162713003>
- D'Alessandro, Roberta & Adam Ledgeway. 2010. At the C–T boundary: Investigating Abruzzese complementation. *Lingua* 120. 2040–2060. DOI: <https://doi.org/10.1016/j.lingua.2010.02.003>
- Déprez, Viviane & France Martineau. 2004. Micro-parametric variation and Negative Concord. In Julie Auger, J. Clancy Clements & Barbara Vance (eds.), *Contemporary approaches to Romance linguistics* (Current Issues in Linguistic Theory 258), 139–158. Amsterdam/Philadelphia: John Benjamins. DOI: <https://doi.org/10.1075/cilt.258.08dep>
- Déprez, Viviane & France Martineau. 2006. Pour une approche micro-paramétrique de la concordance négative francophone. In Francis Corblin, Sylvie Ferrando & Lucien Kupferman, *Indéfini et Prédication*, 217–233. Paris: Presses Universitaires de Paris-Sorbonne.
- de Swart, Henriëtte & Ivan Sag. 2002. Negative concord in Romance. *Linguistics and Philosophy* 25. 373–417. DOI: <https://doi.org/10.1023/A:1020823106639>
- Detges, Ulrich & Richard Waltereit. 2008. Syntactic change from within and from without syntax: A usage-based analysis. In Ulrich Detges & Richard Waltereit (eds.), *The paradox of grammatical change: Perspectives from Romance* (Current Issues in Linguistic Theory), 13–30. Amsterdam & Philadelphia: John Benjamins.
- Édmont, Edmond & Jules Gilléron. 1902–1910. *Atlas linguistique de la France*. Paris: Champion.
- Flutre, Louis-Ferdinand. 1970. *Le Moyen-Picard d'après les textes littéraires du temps (1560–1660)*. Amiens: Musée de la Picardie.
- Foulet, Lucien. 1919. *Petite syntaxe de l'ancien français*. Paris: Champion.
- Franco, Irene, Olga Kellert, Guido Mensching & Cecilia Poletto. 2016. *On (negative) indefinites in Old Italian*. Frankfurt: Goethe University manuscript.
- Frascarelli, Mara. 2012. The interpretation of discourse categories. Cartography for a crash-proof syntax. In Valentina Bianchi & Cristiano Chesì (eds.), *Enjoy Linguistics! Papers offered to Luigi Rizzi on the occasion of his 60th birthday*, 180–191. Siena: CISCL Press.
- Garzonio, Jacopo & Cecilia Poletto. 2010. Quantifiers as negative markers in Italian dialects. In Jeroen van Craenenbroeck (ed.), *Linguistic Variation Yearbook 2009*

- (Linguistic Variation Yearbook 9), 127–152. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/livy.9.04gar>
- Garzonio, Jacopo & Cecilia Poletto. 2012. On *niente*: Optional Negative Concord in Old Italian. *Linguistische Berichte* 230. 131–153.
- Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/la.23>
- Giraud, Cédric, Jean-Baptiste Renault & Benoît-Michel Tock. (eds.) 2010. *Les espaces de la charte*. Nancy: Centre de Médiévisique Jean Schneider/Electronic publishing, Orléans: Institut de Recherche et d'Histoire des Textes. [On-line] <http://www.cn-telma.fr/chartae-galliae/charte213765/> Consulted on June, 30th, 2017.
- Goebel, Hans. 2003. Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): Relations quantitatives et structures de profondeur. *Estudis Romànics* XXV. 59–121.
- Gougenheim, Georges. 1984. *Grammaire de la langue française du seizième siècle*. Paris: Picard.
- Haase, Alfons. 1898. *Syntaxe française du XVIIe siècle*. Paris: A. Picard et fils.
- Herburger, Elena. 2001. The Negative Concord puzzle revisited. *Natural Language Semantics* 9(3). 289–333. DOI: <https://doi.org/10.1023/A:1014205526722>
- Lafont, Robert. 1967. *La phrase occitane – essai d'analyse systématique*. Paris: PUF.
- Ledgeway, Adam. 2005. Moving through the left–periphery: The dual complementiser system in the dialects of Southern Italy. *Transactions of the Philological Society* 103(3). 339–396. DOI: <https://doi.org/10.1111/j.1467-968X.2005.00157.x>
- Lusignan, Serge. 2011. Le français médiéval: Perspectives historiques sur une langue plurielle. In Serge Lusignan, France Martineau, Yves Charles Morin & Paul Cohen, *L'introuvable unité du français. Contact et variations linguistiques en Europe et en Amérique* (XIIe–XVIIIe siècle), 5–107. Laval: Presses de l'Université.
- Lusignan, Serge. 2012. *Essai d'histoire sociolinguistique: Le français picard au Moyen Âge*. Paris: Classiques Garnier.
- Mascarenhas, Salvador. 2007. Complementizer doubling in European Portuguese. New York: New York University manuscript. Online: <http://homepages.nyu.edu/~sdm330/>.
- Morin, Annick. 2009. On the Quebec French interrogative particle *-tu*. In Enoch O. Aboh, Elisabeth van der Linden, Josep Quer & Petra Sleeman, *Romance Languages and Linguistic Theory: Selected papers from "Going Romance" Amsterdam 2007* (Romance Languages and Linguistic Theory 1), 201–222. Amsterdam: John Benjamins.
- Munaro, Nicola & Jean-Yves Pollock. 2005. Qu'est-ce que (qu)-est-ce que? A case study in comparative Romance interrogative syntax. In Guglielmo Cinque & Richard Kayne (eds.), *The Oxford handbook of comparative syntax*, 542–606. Oxford: Oxford University Press.
- Obenauer, Hans-Georg. 2005. La syntaxe des questions non standard: Les questions de surprise en Bellunese, *Recherches Linguistiques de Vincennes* 33. 55–90. DOI: <https://doi.org/10.4000/rlv.1267>
- Paoli, Sandra. 2007. The fine structure of the left periphery: COMPs and subjects; evidence from Romance. *Lingua* 117. 1057–1079.
- Poletto, Cecilia & Jean-Yves Pollock. 2004. On the left periphery of some Romance questions. In Luigi Rizzi (ed.), *The structure of CP and IP*, 250–296. New York: Oxford University Press.
- Remacle, Louis. 1952–1960. *Syntaxe du parler wallon de La Gleize*. 3 volumes. Paris: Les Belles Lettres.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer. DOI: https://doi.org/10.1007/978-94-011-5420-8_7

- Rizzi, Luigi & Ian Roberts. 1989. Complex inversion in French. *Probus* 1(1). 1–30. DOI: <https://doi.org/10.1515/prbs.1989.1.1.1>
- Tuaille, Gaston. 1975. Analyse syntaxique d'une carte linguistique: ALF 25, 'Où vas-tu?'. *Revue de Linguistique Romane* 39(1). 79–96.
- Uriagereka, Juan. 1995. An F position in Western Romance. In Katalin E. Kiss (ed.), *Discourse configurational languages*, 153–175. Oxford: Oxford University Press.
- Vasseur, Gaston. 1996. *Grammaire des parlers picards du Vimeu, Somme: Avec considération spéciale du dialecte de Nibas*. Abbeville: Paillart.
- Vecchiato, Sara. 2000. The TI/TU interrogative morpheme in Québec French. *GG@G* 1. 141–163.
- Villa-García, Julio. 2010. Recomplementation and locality of movement in Spanish. Ms., University of Connecticut.
- Vinet, Marie-Thérèse. 2000. La polarité Pos/Nég, -tu (pas) et les questions oui/non. *Revue Québécoise de Linguistique* 28(1). 137–149. DOI: <https://doi.org/10.7202/603191ar>
- Wanner, Dieter. 1995. Les subordonnées à double complémentateur en roman médiéval. In Giovanni Ruffino (ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza (18–24 settembre 1995), sezione 1*. 421–433. Tübingen: Niemeyer.
- Zanuttini, Raffaella. 1997. *Negation and clausal structure: A comparative study of Romance languages*. New York: Oxford University Press.

How to cite this article: Dagnac, Anne. 2018. SyMiLa and the *Atlas linguistique de la France*: A tool for the study of Gallo-Romance syntax. *Glossa: a journal of general linguistics* 3(1): 85.1–23, DOI: <https://doi.org/10.5334/gjgl.543>

Submitted: 09 October 2017 **Accepted:** 30 March 2018 **Published:** 06 August 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.

