


ARTICLE

The embedded *Wh- in situ* clause: French tout court? *Les spasmes musculaires incontrôlés, je sais pas c'est quoi* (France Info 20/10/21)

Gudrun Ledegen 

LLL – Université d'Orléans - PREFICS – Université Rennes 2 (France)

Email: gudrun.ledegen@univ-rennes2.fr

(Received 04 August 2022; revised 15 January 2024; accepted 30 January 2024; first published online 27 March 2024)

Abstract

Although not studied by the generative approach (Shlonsky 2017), the embedded *Wh- in situ* clause nevertheless belongs to the French language spoken in a large amount of areas, and, we believe, to “français tout court” (Blanche-Benveniste & Jeanjean 1987): until now, it has mainly been studied in Quebec (Lefebvre & Maisonneuve 1982; Blondeau & Ledegen 2021) and in Reunion Island (Ledegen 2007a, 2007b, 2007c, 2016; Ledegen & Martin 2020), but it has recently been massively attested in the Multicultural Paris French project in the suburbs of Paris (Gardner-Chloros & Secova 2018) and Strasbourg (Marchessou 2018). These new data could be read as a language contact, as a recent linguistic change, or as a long-established “popular” structure (Guiraud 1966), different analytical hypotheses that will be detailed in this study. These recent data also argue in favour of the methodology of ecological corpora, obtained within the framework of a strong acquaintanceship and located at the pole of communicative proximity (Koch & Oesterreicher 2001). The examination of various existing corpora will reveal the structural functioning of the structure and contrast the corpora following the modes of interaction, the oral or written medium, as well as on the chronological axis.

Keywords: variation; syntaxis; francophone world; communicative proximity; embedded *Wh- in situ* clause; sociolinguistics

Résumé

Bien que non étudiée par l'approche générative (Shlonsky 2017), la structure indirecte *in situ* fait pourtant partie de la langue française parlée dans un grand nombre de régions, et appartient, selon nous, au “français tout court” (Blanche-Benveniste & Jeanjean 1987) : jusqu'à présent, elle a surtout été étudiée au Québec (Lefebvre & Maisonneuve 1982 ; Blondeau & Ledegen 2021) et à La Réunion (Ledegen 2007a, 2007b, 2007c, 2016 ; Ledegen

& Martin 2020), mais elle a récemment été massivement attestée dans le cadre du projet Multicultural Paris French en banlieue parisienne (Gardner-Chloros & Secova 2018) et strasbourgeoise (Marchessou 2018). Ces nouvelles données pourraient être lues comme un contact de langue, comme un changement linguistique récent, ou comme une structure "populaire" établie de longue date (Guiraud 1966), différentes hypothèses d'analyse qui seront détaillées dans cette étude. Ces données récentes plaident également en faveur de la méthodologie des corpus écologiques, obtenus dans le cadre d'une connaissance forte et situés au pôle de la proximité communicative (Koch & Oesterreicher 2001). L'examen de différents corpus permettra de révéler le fonctionnement structurel de la structure et de contraster les corpus selon les modes d'interaction, la modalité orale ou écrite, ainsi que sur l'axe chronologique.

Mots-clés: variation; syntaxe; monde francophone; proximité communicative; structure enchâssée *Quin situ*; sociolinguistique

We will deal here with the embedded *Wh- in situ* clauses in French, an underestimated sociolinguistic and diatopic variable, yet not much studied by the generative approach (Shlonsky 2017). Embedded *Wh- in situ* clauses are usually considered as very strongly regionally marked (Shlonsky, p.c.), but they nevertheless belong to spoken French in surprisingly many areas throughout the realm of "Francophonie", and can even be held straightforwardly as typical "français tout court" (Blanche-Benveniste 1983). There are plenty of occurrences as unmarked, 'ordinary' utterances in the French spoken in Quebec (Lefebvre & Maisonneuve 1982; Blondeau & Ledegen 2021) or in Reunion Island (Ledegen 2007a, 2007b, 2007c, 2016; Ledegen & Martin 2020). Indeed, this convergence may be seen as some kind of an archaism (according to Bartoli's "lateral area/geolinguistic norm effect" (1945)), which would have remained in usage through the historical parallels between these two territories. A new type of corpus, *ecologically* conceived by the recording of people who know each other well (Koch & Oesterreicher 2001; Guerin 2017), has provided evidence of a very high rate of usage of this structure in the suburbs of Paris (Gardner-Chloros & Secova 2018) and of Strasbourg (Marchessou 2018).

French *Wh* interrogation reveals different strategies. The direct structures use *Wh* in initial position, with (a) or without (b) inversion of the subject and the verb, initial *Wh* with *est-ce que* (c), the *in situ* option (d), the clefting option (e) and initial *Wh* with non-inverted *est-ce que* (f):

- (1.a) Où tu pars ?
- (1.b) Où pars-tu ?
- (1.c) Où est-ce que tu pars ?
- (1.d) Tu pars où ?
- (1.e) C'est où que tu pars ?
- (1.f) Où c'est que tu pars ?

The embedded structures reveal initial *Wh* (g), with *est-ce que* (h) or *c'est que* (i), and the *in situ* option (j); this paper is concerned with this last option¹:

- (1.g) Il sait où tu pars.
- (1.h) Il sait où est-ce que tu pars.
- (1.i) Il sait où c'est que tu pars.
- (1.j) Il sait c'est où que tu pars.
- (1.k) Il sait tu pars où.

The new data concerning the embedded *Wh- in situ* clauses call for different analytical hypotheses: recent linguistic change, language contact, learner's variant, long-established structure of 'registre populaire' (Guiraud 1966). The focus here will be on the embedded *Wh- in situ* clause, with elements of comparison with the embedded *qu'est-ce que* interrogative clause. After examining previous studies in different French-speaking areas, and different media elements that give the impression of an increased use in France, we conduct further explorations on the embedded *Wh- in situ* clauses through the *Fonds de données linguistiques du Québec* (FDLQ): the contrast between the different corpora gathered in the FDLQ will allow us to analyse the current patterns of the embedded *Wh- in situ* clauses, according to modes of interaction, following the oral or written medium, as well as on the chronological axis (diachrony – synchrony). These latter explorations will highlight the status of the embedded *Wh- in situ* clauses in Quebec French, and reveal its presence through time.

The embedded *Wh- in situ* clauses being found in other French-speaking countries, with close sociolinguistic representations (colloquial French), the emergence of the structure will be analysed, while it feeds into our pan-French perspective (Chaudenson, Mougeon & Beniak 1993; Boutin & Gadet 2012: 19), which takes non-standard as its primary reference (Ploog 2002, 2019; Poplack & Dion 2009). Our analytical hypothesis is that the embedded *Wh- in situ* clause belongs to "français tout court" (Blanche-Benveniste 1983: 27), being part of the uses of language that form coherent systems of forms in the description. And by exploring the usages more fully, we could give a renewed place to this structure, as attested for the embedded interrogative with *est-ce que* (cf. *infra*).

Thus, this study uses different approaches, by studying multiple corpora (including normative opinions) and discussing syntactic theory, and enters by multiple angles: diachrony-synchrony, morpho-syntaxis and sociolinguistics. Our conviction is indeed that this combination can encompass the complexity of this non-standard structure.

1. IS THE *IN SITU* STRUCTURE GAINING GROUND, ALSO IN THE EMBEDDED CONTEXT, IF EVER...

Several studies posit that the *in situ* form is nowadays gaining ground in the direct interrogative construction: this claim can be found, for example, in

¹Blasco & Cappeau (1992) also find embedded structures with initial *Wh* and subject-verb-inversion but only in the specific – rather formal – journalistic situations, where hypercorrection or emphasis might be at work: « [Ali Campbell] à qui nous avons demandé en tant qu'auteur de la chanson Mandela's Day que ressentait-il » (1992: 17).

Dekhissi & Coveney (2018: 136 and 147), who draw on their own diachronic research on movies dealing with *banlieue* surroundings; Rossi-Gensane, Acosta Cordoba, Ursi & Lambert (2021) find increase of direct interrogatives *in situ* with *où* in dialogues of French novels during the twentieth century.² The present study will question the recent evolution of the embedded *in situ* structure, which may appear to be gaining ground too.

Another so-called non-standard structure has undergone a recent evolution towards greater acceptability: the embedded (*qu'*)*est-ce que* clause is analysed by Muller as colloquial:

- (2) je me demande **qu'est-ce qu'**il fabrique (Muller 1996: 198).

Also for the colloquial usage, Riegel, Pellat & Rioul see it as an “[alignment of] the embedded structure on the model of the independent direct interrogative” (2021: 839). The same is true for Defrancq, who attests that “the ‘particle’ [*est-ce que*] is well installed at the very heart of the embedding”³ (2000: 135). Finally, Blanche-Benveniste goes even further and classifies this embedded interrogative with *est-ce que* among:

les fautes qui n'en sont plus. [...] [Cette tournure] est une faute qui agace beaucoup certains puristes. [...] On la trouve [pourtant] partout, y compris chez les notables, les écrivains, les professeurs. (Blanche-Benveniste 1997a: 41)

It is also interesting to note that “until the eighteenth century, *est-ce que* et *c'est que* in embedded interrogation belonged to good usage” (Grevisse 1988: 683):

- (3) Je ne sais comment **est-ce que** je suis ensuite comme rapporté et rendu à moi-même (Fénelon, cit. Haase, 92)
 (4) Ce sont les noms des mots, et l'on doit regarder/En quoy **c'est qu'**il les faut ensemble accorder (Bélise, in Molière, Femmes savantes, II, 6)⁴

As for the embedded *Wh- in situ* clause, which does not have the word *Wh-* at the top,⁵ it is considered at best as a “structure without a subordinating element” (Defrancq 2000: 132), or even as “[not concerning] embedded interrogation as it is conceived [in his study (cf. note 3)]” (Defrancq 2000: 137). Thus, Defrancq did not take into account the 16.9% of embedded interrogatives (12/71) which present the interrogative element *in situ* in Lefebvre and Maisonneuve's study in Montréal (1982), arguing that “the classical conception of embedded interrogation as an

²See also Farmer's diachronic and filmic study (2013 and 2015, cited in Dekhissi & Coveney 2018) and Larrivée (2014, cited in Dekhissi & Coveney 2018).

³For Defrancq, the structure corresponds to the general rule of embedded structure by being characterised “par la présence d'un élément subordonnant [en l'occurrence le mot *Qu-* à l'initiale]” (“by the presence of a subordinating element [in this case the word *Wh-* on top]”) (2000: 132).

⁴Bélise, one of the ‘femmes savantes’, isn't using here hypercorrection or pedantic tours, but a construction characteristic of the classical language (Fournier 1993: 22).

⁵Also called ‘non-extraction du terme interrogatif’ (Combettes & Glikman 2020: 1356).

embedded structure [is not to be questioned by these examples]" (Defrancq 2000: 137).

In fact, we should probably not question the conception of the indirect interrogative as an embedded structure,⁶ but rather the conception of *embedding* as such, that should be made more flexible to match this kind of data. Indeed, there are other embedded structures that do not have a connector, such as the dropping of conjunction *que* attested in "ordinary" French in Reunion Island (Ledegen 2007a, 2007b):

- (5) ça fait longtemps j'ai pas vu là (VALIRUN⁷, Cafête5, 4–5)
 (6) j'ai l'impression ça va casser (VALIRUN, Cafête2, 3–4) (Ledegen 2007b: 320–321)

as well as in spoken corpora in France:

- (7) faut pas croire les élèves ils viennent avec des calibres en cours hein/ils viennent tranquille (Gadet 2003: 47)
 (8) je pense au contraire maintenant on commence à changer et à se dire bon voilà quoi (Ferreira 2022: 2)

or in Canada, where Martineau finds an omission rate of 28% in the Spoken French of Ottawa-Hull:

- (9) je pense je vais y aller (Martineau 1988: 279).

This absence of a connector doesn't mean absence of embedding or of syntactic binding and dependency: Deulofeu indeed uses the term *unasyndetic hypotaxis*⁸ beside *syndetic hypotaxis* (1989: 112).

2. INCREASED FREQUENCY OF ATTESTATION OF EMBEDDED *WH- IN SITU* CLAUSES

As it is the case for the direct interrogative, studies on more "ordinary" corpora are increasing, and the embedded *in situ* structure is mentioned more frequently: Lefebvre and Maisonneuve (1982) was the first study to mention it, from a corpus of French spoken in the Centre-Sud district of Montreal. This study showed examples of embedded interrogations whose interrogative term happened to occur *in situ* (for example: *Il y en a qui savent pas c'est quoi*) (1982: 190), which constituted 16.90% (12/71) of all embedded interrogations in this data. Defrancq considered this percentage to be "spectacular when one considered that Coveney's (1995) study of

⁶Lefebvre & Rossi-Gensane adopt the analysis "selon laquelle, de manière générale, les interrogatives indirectes sont des subordinées sans subordonnant" ('that, in general, sees the embedded interrogatives as embedded structures without conjunction') (2017: 71).

⁷Variétés linguistiques de La Réunion (Ledegen (dir.)), partly available on the platform *Cocoon* (<https://cocoon.huma-num.fr/exist/crdo/>).

⁸Ferreira (2022) presents an innovative analysis of syndetic, asyndetic and macro-syntactic structures with the verbs *penser* and *croire*. See also note 42 *infra*.

direct questioning attested only 13% of interrogative terms *in situ*” (2000: 136). In Quillard’s PhD thesis (2000), based on more “natural” and diverse corpora, the structure has been found to occur more frequently: the proportion of *in situ* structures in these direct interrogative constructions in “ordinary” spoken French made 15.57%.

In the French field of research, no occurrence appears e.g. in Defrancq’s study (2000), which is based on the CORPAIX corpus of the *Groupe Aixois de Recherche en Syntaxe* (GARS 1999), mainly based on interviews and life narratives, where few questions occur, especially by the interviewees. The very first mention appears in the study of the linguistic practices of young people in the Paris suburbs by Conein and Gadet (1998).⁹ These authors mention a single example of embedded *Wh- in situ* clause (“*il regardait pas c’était qui*”), and point out that “the absence of distinction between direct and indirect interrogation [has been recorded] for a long time” (1998: 110), and belongs to the “traits héréditaires populaires” (1998: 121). The structure is somehow attested in the research of Lefevre and Rossi-Gensane (2017) who find six forms: one in the *Corpus de français parlé parisien* (CFPP2000¹⁰) (out of 144 embedded interrogatives) and five (out of 98) in the ESLO corpus¹¹ (*Les repas*, dating from 2012).

In Reunion Island, the structure is found to belong to everyday French (Ledegen 2007a, 2007b, 2007c, 2016; Ledegen & Martin 2020), even more so than the embedded *est-ce que*. A careful analysis of six hours of recordings in the VALIRUN corpus that contain non-standard embedded interrogatives yields no less than 62 embedded interrogative forms, including 28 standard forms, 27 *in situ* and 7 with *est-ce que* (Ledegen 2007a). It should be noted that the embedded *Wh in situ* structure is also attested in the written SMS-Reunion corpus¹² compiled in the framework of the SMS4science project, in familiar (ex. 10) but also more formal contexts (ex. 11):

- (10) Soi dizan ct ce matin ke talé voir!tc enkor **c koi** ladress?Le meilleur c su la fin!lol
(Soi-disant c’était ce matin que t’allais voir! tu sais encore **c’est quoi** l’adresse?
Le meilleur c’est sur la fin!lol)
- (11) Bonjour.vous m’avez appelez.je peux savoir **c’est qui**.je n’ai pas eu temps de
repondre. (Bonjour. vous m’avez appelé. je peux savoir c’est qui.je n’ai pas eu
[le] temps de répondre.)

Research carried out as part of the *Phonology of Contemporary French project* (Durand, Laks & Lyche 2009) also reveals multiple examples for Reunion Island and New Caledonia, where usage is quite widespread, and with highly variable¹³ embedded verbs:

⁹It must be noted that there are no examples in *Le français ordinaire* (Gadet 1989) or *Le français populaire* (Gadet 1993).

¹⁰<http://cfpp2000.univ-paris3.fr/>

¹¹<http://eslo.huma-num.fr/>

¹²<https://repository.ortolang.fr/api/content/comere/v3.3/cmr-smslareunion.html>

¹³Which is in contrast with our data, cf. *infra*.

- (12) je me rappelle plus **c'est quel magasin** (PFC-Nouvelle-Calédonie)
 (13) je connais des gens mais je sais pas **ils habitent où** (PFC-Nouvelle-Calédonie)

Moreover, this same database contains numerous examples from Canada (Blondeau & Ledegen 2021), as will be discussed later in this study (cf. 4. *Exploring the Fonds de Données Linguistiques du Québec*). Other explorations show examples (4 *in situ* and 18 *qu'est-ce que*) of these structures in Switzerland (Ledegen & Martin 2020) in the OFROM corpus (Corpus oral de français de Suisse Romande), and in Brittany (France):

- (14) moi je sais pas **c'est quoi** sa personnalité je la connais pas assez bien (unine 15–101)
 (15) je sais même pas les petits points **c'est quoi** (oral presentation during a doctoral seminar, 2016, PhD student, Brittany)
 (16) je vais expliquer davantage **c'est quoi** l'interaction verbale (2017, written record, Master1-student, Brittany)

In contrast, in 2018, the Multicultural Paris French study will reveal a very remarkable attestation rate in the Paris and Strasbourg suburbs: the Gardner-Chloros and Secova study (2018) on the Paris suburbs attests 61 embedded *in situ* interrogations out of 159 structures, i.e. 38% of the structures. And Marchessou's study (2018) on the Strasbourg suburbs, a study made in the same MPF-project, shows 35 on 63 forms, i.e. 55%.

In a field similar to that of the suburbs, we analysed in 2020 the Rapcor corpus¹⁴ (2009–...), made up of French and Quebec rap songs, by searching the repeated segments *c'est qui* and *c'est quoi*, the most frequent forms of the embedded *in situ* structures: this exploration yielded 18¹⁵ embedded *in situ* structures (8 from the French corpus; 10 from the Quebec corpus). It is notable that only two statements could have been constructed for rhyming purposes:

- (17) approche de toi et te demande ton nom **c'est quoi**, et ben tu dis qu'tu sais pas.
 (MC Solar, Si On T Demande, 2007, Album: Chapitre 7)
 (18) Rappelle-moi donc **c'est quoi** Rappelle-moi donc Rappelle-moi donc **c'est quoi** (Gatineau, Back In The Days, 2008, Album: Gatineau)

These two statements also stand out from the rest of the corpus by not having the right dislocation, which is systematically present in the 16 other extracts:

- (19) moi je comprends pas **c'est quoi leur maladie** (Koma, La Peur Du Lendemain, 1999, Album: Le Réveil)
 (20) Jamais trop tard jamais trop tôt Je sais **c'est quoi tout recommencer à zéro** (Sans Pression, Pas Le Choix De Foncer, 2003, Album: Répliques Aux Offusqués)

¹⁴Since this study, we have heard this occurrence in a song by Orelsan: « Pas b'soin d'savoir **c'est quoi** l'Sénat Pour voir qu'les vieux riches font les lois » (L'odeur de l'essence, album *Civilisation*, 2021).

¹⁵Eighteen structures found in 700 songs (the state of the Rapcor corpus in 2020), i.e. 2.5% as attestation rate.

The corpus of rap songs thus confirms the presence of the embedded *Wh- in situ* clause among these authors who are often considered innovative and counter-normative (Lesacher 2015).

Moreover, the consultation of – non-scientific – language forums informs us that such a structure might have increased: for example, the forums mention the structure as early as March 2011:

- (21) Je vous le prédis, dans quinze ans, plus personne ne dira « tu sais qui c'est ? », mais « tu sais **c'est qui** ? ». Depuis déjà une dizaine d'années, j'entends mes élèves de collège de banlieue parisienne (je suis prof) parler de cette façon : « J'sais pas **c'est lequel** », « J'sais pas **c'est où qu'il a acheté ça** ». ¹⁶

After a long silence, the subject comes back, on two sites in 2020: on the Blabla 18–25 Forum on the one hand, and on the *Question orthographe* forum on the other. The latter is entitled, in a very normative way, “Tu sais c'est qui ... tu vois c'est quoi ... et autres immondices (règles ?)”¹⁷.

One can see that the explanation given by the *Question orthographe* website (cf. Table 1) shows a very strict standard for the direct interrogative, where the interrogative word should always be in the initial position ('le pronom interrogatif [...] se place devant elles comme d'ailleurs dans l'interrogative directe').

Table 1. A mention of embedded *Wh- in situ* clause on the *Question orthographe*-forum

Request for help:	Answer:
<p>Bonjour,</p> <p>J'entends souvent des phrases du type de celles que j'ai mises en titre, et sais pertinemment en mon for intérieur qu'elles sont erronées, cependant je n'arrive pas du tout à mettre le doigt sur la règle que je pourrais invoquer afin de permettre aux gens qui les disent ou les écrivent de se corriger.</p> <p>Pourriez-vous m'aider à la trouver ?</p> <p>En vous remerciant par avance pour votre aide.</p> <p>Cordialement.</p>	<p><i>Tu sais c'est qui</i> → <i>tu sais qui c'est</i></p> <p>tu sais qui est venu – tu sais comment il s'appelle – tu sais pourquoi il vient – tu sais où le trouver – tu sais lequel prendre</p> <p>Remarque les mots interrogatifs qui, comment, pourquoi, où, lequel ...</p> <p><i>Tu sais et c'est</i> sont deux propositions. elles sont reliées par « qui » ici pronom interrogatif.</p> <p>Tu sais introduit une interrogative indirecte et le pronom interrogatif est introducteur de ces paroles et se place avant elles comme d'ailleurs dans l'interrogative directe : <i>qui est-ce ? Où vas-tu ? Pourquoi vient-il ? ...</i></p>

The structure is also mentioned in several Top 10 French mistakes: in 2016 in the fourth place (*Les fautes de français qui nous horripilent*¹⁸), but in 2021 and 2022 in the

¹⁶I predict that in fifteen years' time, no one will say “tu sais qui c'est ?”, but “tu sais c'est qui ?”. For about ten years now, I've been hearing my pupils in a Parisian suburb (I'm a teacher) talk like this: “J'sais pas c'est lequel”, “J'sais pas c'est où qu'il a acheté ça”. <http://commeonparle.over-blog.com/article-tu-sais-c-est-qui-68954308.html>

¹⁷You know who ... you know what ... and other filth (rules?) <https://www.question-orthographe.fr/answer/re-tu-sais-cest-qui-tu-vois-cest-quoi-et-autres-immondices-regles/>

¹⁸French mistakes that exasperate us' : <https://petitsfrenchies.com/ces-fautes-de-francais-qui-nous-horripilent/>

first place (with the titles: *Top 15 des fautes de français qui arrachent l'oreille*¹⁹, and *13 fautes de français que tout le monde fait (et qui arrachent les oreilles)* cf. Figure 1²⁰):



Figure 1. Illustration of the site '13 fautes de français que tout le monde fait (et qui arrachent les oreilles)'.

1. "Je sais pas c'est qui":

Petite sœur du fameux "je sais pas c'est où" et de "c'est qui qui dit ça". On commence par du lourd avec une double faute.

On passe sur l'oubli de négation ("je NE sais pas") car on est à l'oral.²¹

Par contre, la construction par inversion du sujet (qui) et du verbe (est) n'a ici aucune raison d'être.

La bonne version: "Je ne sais pas qui c'est"²²

It is interesting to observe the confusion in this comment about *inversion*.²³ Finally, the structure is attested also in very different media formats:

- (22) Cartoon: On va leur montrer **c'est qui** le patron (Mune, Le gardien de la lune, 2014)

¹⁹Top 15 ear-splitting French mistakes': <https://www.topito.com/top-10-des-fautes-de-francais-qui-marrachent-loreille>

²⁰13 French mistakes that everyone makes (and that are ear-splitting)' <https://www.comment-economiser.fr/13-fautes-de-francais-les-plus-courantes-a-eviter.html> (121,994 views).

²¹It is remarkable that the negation *ne* is present in the illustration of this document (cf. Figure 1).

²²<https://www.comment-economiser.fr/13-fautes-de-francais-les-plus-courantes-a-eviter.html>

²³The other way around, this observation may explain the embedded *in situ* structure: the standard embedded structures might be seen by speakers as a complex unusual placement and assimilated to inversions: *je sais pas lequel c'est* vs *je sais pas c'est lequel*. This might explain the avoidance of the standard structure that places *Wh-* at the beginning of the embedded phrase.

- (23) Humorous YouTube channel Parole de chat: Chat 1: Tu sais **c'est quoi** ton problème ? Chat 2: Non. Chat 1: C'est que t'as le QI d'une huitre. (Les gros mots 3²⁴)
- (24) Film: tu vois **c'est qui** Martha (Ma 6-T va crack-er, 1997). An exploration of three films about banlieues (La haine, 1995; Ma 6-T va crack-er, 1997; Le ciel, les oiseaux et ta mère, 1999) reveals this unique instance. The film in question is known to have been filmed with amateur actors who had a great deal of freedom in the formulation of their dialogues, so they would be closer to their natural style.
- (25) Broadcast: Les spasmes musculaires incontrôlés, je sais pas **c'est quoi** (France Info 20/10/21). This last occurrence, produced by the political columnist Alix Bouilhaguet, takes place during a period of free interaction, of chatting, between more formal sections of the television news, during direct broadcasting, and thus reveals a usage attested in a context of communicative proximity (Koch & Oesterreicher 2001), and not an exclusively 'registre populaire'.²⁵

3. CONTRASTING HYPOTHESES

These different studies allow us to identify three major explanatory hypotheses that are attached to the rise of the embedded *Wh- in situ* structure:

- (a) the structure would come from a situation of language contact (Gardner-Chloros & Secova 2018),
- (b) it would be a recent linguistic change (for example, given its total absence in the GARS corpora (Defrancq 2005))
- (c) the structure would be a turn of phrase of 'registre populaire' (in the sense of Guiraud 1966) attested for a long time and characterised by the self-regulating tendencies of French.

The first hypothesis is combined with the second in the research by Gardner-Chloros and Secova (2018)²⁶: the authors see the structure as 'an instance of *change from below* (Labov, 2007), which seems to have emerged in the speech of young people of immigrant background' (2018: 182). The language contact would have been generated by 'bilingual speakers²⁷ who belong to major communities of immigrant origin' (2018: 182), although no specific contact languages that would have spread the *in situ* structure can be identified (neither in the French suburbs

²⁴Cat talk, swear words', humorous videos giving voice to cats https://www.youtube.com/watch?v=Cuhmj_eYt3k

²⁵For a critical analysis of this notion, in its registral use ('stigmatised non-standard use' (Gadet 1992: 27)), see Ledegen (2016).

²⁶They certainly mention the idea that "It might also, on the other hand, be a long-standing vernacular variant", however immediately brought back to their perspective: "which has re-emerged, with specific identity-related significance, in this particular group of speakers" (2018: 181).

²⁷These speakers are also seen as possibly producing this form as a 'simplification grammaticale caractéristique des variétés d'apprenant' (Secova 2017: 13).

or in Reunion Island) : neither English, Arabic, Berber, Reunionese Creole²⁸ or German (for the suburbs of Strasbourg studied by Marchessou 2018) present this structure.

Comparison with other vernacular corpora permits us to eliminate this contact hypothesis, which is frequently invoked in situations of language contact, even though the features belong to ordinary French, frequently attested for a long time. Poplack and St-Amand set out this frequent analysis for Canada:

L'étude du changement linguistique s'est butée depuis toujours à la rareté de données appropriées reflétant un stade antérieur de la langue. En effet, [...] les textes écrits ont l'inconvénient de ne pas toujours refléter la langue parlée, lieu privilégié des changements. Le manque de données historiques fiables explique, du moins en partie, la notion courante voulant que de nombreux traits saillants des parlers vernaculaires contemporains soient des innovations récentes. Cette idée est particulièrement répandue dans le cas des variétés canadiennes du français, qui comportent plusieurs traits distinctifs, souvent même stigmatisés. On attribue d'ordinaire ces traits au changement, censément causé par le contact massif que ces variétés ont subi avec l'anglais depuis la Conquête britannique du Canada (1760) [ex. Laurier 1989], et par des siècles d'éloignement par rapport au français métropolitain et son influence se voulant normalisatrice. (2009: 511)²⁹

The same is true for Reunion Island, where the vast majority of ordinary French features are considered – especially by teachers – as an interference with Creole (Ledegen, 2007b: 321); this explanation is not, however, the most plausible, as the *in situ* structure is the least used in the Creole embedded structures. Gadet and Jones attest this same phenomenon when it comes to non-standard or geographically distant structures:

grammarians of French may sometimes be overly quick to explain away as interference 'unusual' phenomena in non-standard and/or geographically peripheral varieties [to France]). (2008: 246)

Rather than considering this structure as 'un changement intralinguistique lié aux contraintes universelles de la parole spontanée (rapidité, manque de temps de planification entraînant une simplification syntaxique et des énoncés paratactiques)' (Secova 2017: 13), we link these different characteristics rather to the fact that it is a

²⁸The embedded *Wh-* *in situ* clause exists but is little used compared to the structure with *Wh-* as introducer of the embedded clause.

²⁹The study of linguistic change has always been hampered by the scarcity of appropriate data reflecting an earlier stage of the language. Indeed, [...] written texts have the disadvantage that they do not always reflect the spoken language, which is the preferred site of change. The lack of reliable historical data explains, at least in part, the common notion that many salient features of contemporary vernaculars are recent innovations. This idea is particularly prevalent in the case of Canadian varieties of French, which have several distinctive, and often stigmatised, features. These features are usually attributed to change, supposedly caused by the massive contact these varieties have undergone with English since the British Conquest of Canada (1760) [e.g. Laurier 1989], and by centuries of distance from metropolitan French and its normalising influence.

survival of a vernacular, colloquial, form ‘de registre populaire’ (Ledegen 2016), which is in line with the evolutionary trends of French (Guiraud 1966: 46; Lefeuve & Rossi-Gensane 2017). The term *populaire* is used here in the ordinary sense of ‘un usage non standard stigmatisé’ (Gadet 1992: 27), with all the criticisms³⁰ that the term deserves. It is a social practice of epilinguistic judgement which is common in society on the part of speakers from the metropolis. But to this sense we add another, that is the one used by P. Guiraud in his analysis of the term *populaire*: in the “système de la relative en français populaire” (although it is not the prerogative of working-class speakers of course (Gadet 1989: 156)), where the term *populaire* refers to ‘le français héréditaire tel qu’on le parlait et l’écrivait au seuil de la période classique [qui] était en train de se décanter et de s’organiser selon les trois grandes tendances qui modèrent l’évolution du français: réduction des déclinaisons, décumul des formes synthétiques, syntaxe séquentielle’ (Guiraud 1966: 46). So it is this link with ‘ordinary’ historical French, which is a common element in the fields examined here, that we wish to put in light.

4. EXPLORING THE FONDS DE DONNÉES LINGUISTIQUES DU QUÉBEC

We proceed here to a complementary analysis of the previous data through the *Fonds de Données Linguistiques du Québec* (FDLQ), recently put online by Centre de recherche interuniversitaire sur le français en usage au Québec (CRIFUQ) - University of Sherbrooke (2022). This exploration will allow us to highlight the structural regularities of the embedded *Wh- in situ* clause, but also contrast the different corpora gathered, according to the modes of interaction, following the oral or written medium, as well as on the chronological axis.

We have proceeded to a detailed examination of the 1498 *c’est quoi* and 3186 *qu’est-ce que/qu’* in the corpora, which gives us 183 occurrences of embedded *Wh- in situ* clause and 104 of embedded *qu’est-ce que* clause. The study is focused on *c’est quoi* which is the most frequent structure in the analyses concerning the embedded *Wh- in situ* clauses; it is here compared with the embedded *qu’est-ce que* clauses. The FDLQ-collection providing no POS-annotation,³¹ we couldn’t sort the 86.072 occurrences of the standard structure in *ce que/qu’* in the corpus, in order to propose a complete and statistical analysis (our consultation of the site was in August 2022).

The Table 2 shows the six corpora concerned, described by general information (number of documents and single words), according to whether they are in the oral or written medium, and the number of occurrences of *in situ* and of the documents containing them. Finally, a value for the number of occurrences of *in situ* in relation to the number of words in the corpus shows the extent of attestation of the structure.

The last column, revealing the attestation rate, shows very clearly that the *in situ* and *qu’est-ce que* structures are attested above all in the oral interviews: these are sociolinguistically oriented or focused on the journalistic environment (*Language and Journalism in Quebec*). The comparison of the two Montreal corpora (1984 and

³⁰Le regard social affable [cet usage non standard stigmatisé] de l’étiquette de *populaire*: tout ce qui est familier est susceptible d’être taxé de *populaire* si le locuteur s’y prête, et seuls certains traits populaires sont étrangers à l’usage familier non populaire.’ (Gadet 1992: 27)

³¹This project is in preparation.

Table 2. Embedded *Wh-* *in situ* and *qu'est-ce que* clauses in the FDLQ by corpus and medium

Medium	Name of the corpus ³²	Number of documents with embedded <i>in situ</i>	Number of occurrences of embedded <i>in situ</i>	Number of documents with embedded <i>qu'est-ce que</i>	Number of occurrences of embedded <i>qu'est-ce que</i>	Differential between occ. Embedded <i>in situ</i> and <i>qu'est-ce que</i>	Attestation rate ³³
Written	Corpus de littérature québécoise (1794–2022)	16	43	11	13	+30	10,8%
	Le français québécois dans les bulles (1958–2022)	2	2	–	–	+2	13,3%
Oral	Langue et journalisme au Québec (2011–2013)	8	21	6	10	+11	42,4%
	Corpus sociolinguistique de l'Estrie (1971–1974)	–	–	21	28	–28	21%
	Corpus Montréal 1984	12	33	–	–	+33	40%
	Corpus Montréal 1995	5	14	–	–	+14	35,7%
	PFC Québec (2009–...)	12	36	10	11	+25	32,4%
	Corpus Hochelaga-Maisonneuve (2012–2015)	20	34	21	42	–8	46,5%

³²A presentation of each corpus figures in the final bibliography about the corpora used in this article.

³³This value is calculated by making the percentage of documents with *in situ* & *qu'est-ce que* over the total number of documents.

1995) seems to show a decrease in attestation (respectively 33 et 14 occ.), but this impression is contradicted by the more recent PFC corpus (36 occ.) and the Hochelaga-Maisonneuve corpus (34 occ.), where the structures occur more frequently. It must be noted that in all these sociolinguistic corpora interviewers and interviewees produce both embedded structures, and for the PFC-corpus, they occur equally in the free as in the guided discussions (Ledegen 2021).

Another remarkable result is that the first oral corpus – Corpus sociolinguistique de l’Estrie (1971–1974) – doesn’t reveal any embedded *Wh- in situ* clause, but only *qu’est-ce que* structures; this corpus is joined by the most recent one – Hochelaga-Maisonneuve (2012–2015) – which presents more occurrences of the *qu’est-ce que* structure than the *in situ* one, as shown by the negative differential value (respectively -28 and -8). Is this a specificity of the localisations of these corpora (Estrie for example being a region at the intersection of the dialectal areas of Eastern and Western Quebec) or of their semi-guided methodology (the Hochelaga-Maisonneuve-corpus “were intended to mimic the flow of natural conversation and were loosely based on a script” (Blondeau, Tremblay, Bertrand & Michel 2021: 7))?

The occurrences in the literary and paraliterary corpora are also interesting: we find 42 embedded *in situ* in respectively 14 and two documents. The Corpus de littérature québécoise (CLIQ) includes several occurrences in quoted speech:

- (26) Les trois gars, **a poursuivi Steeve entre deux gorgées**, s’obstinent pour savoir **c’est quoi** l’animal le plus dangereux de la Gaspésie. (Bernard, 2017, *La bête creuse*. [CLIQ]³⁴)

but also in narratives:

- (27) Maria demande **c’est quoi** la différence de maman. (Turcotte, 1991, *Le bruit des choses vivantes* [CLIQ])

Figure 2 illustrates the evolution of the number of occurrences of embedded *in situ* and *qu’est-ce que* clauses, and reveals that the attestations of *qu’est-ce que* are present since 2011, but appear to be taken over by the *in situ* structure from the 2010s³⁵ onwards:

We will see later that this literary corpus contains some formal particularities, which strongly differentiate it from the oral corpus and which reproduce, for some of them, stereotypes of the oral language, revealing a literary writing imitating the oral (Ledegen 2019). Thus, these literary corpora, extending from 1794 until 2022, attest the ‘informalisation’ (Fairclough 1996), and the ‘arrival of the oral’ in written practices previously closer to the prescriptive norm (Gadet 1999: 581; Meizoz 2001³⁶):

³⁴The book *La bête creuse* (Bernard Christophe, 2017) has 17 occurrences on its own. For the other books, the highest rate is three. Full bibliographical details are listed at the end of this article.

³⁵Similarly, for the two comic strips, the years of attestation of the embedded *in situ* are 2013 and 2017.

³⁶After putting the words of working-class characters into discourse, literature oralised the narrative voice at the turn of 1920s: Il y a un ‘roman parlant’, un roman qui donne l’illusion que l’on vous parle directement à l’oreille. (‘There is a “talking novel”, a novel that gives the illusion that you are being spoken to directly in your ear’) (Meizoz 2001: 16)

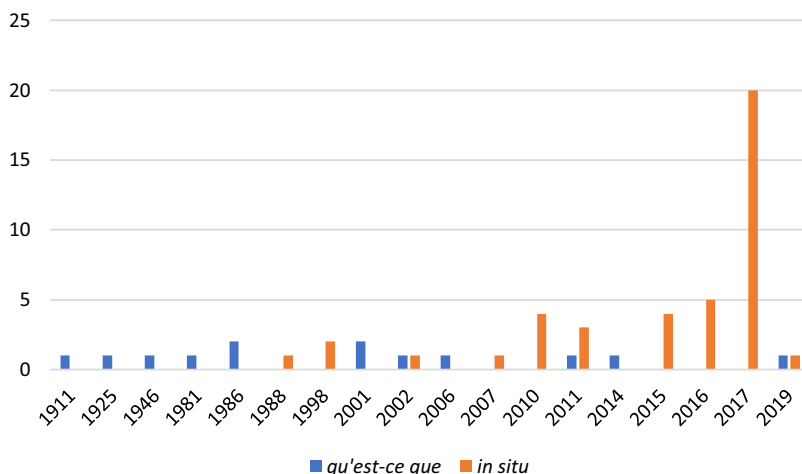


Figure 2. Number of occurrences of embedded *in situ* and *qu'est-ce que* clauses in the CLIQ corpus of the FDLQ, per year.

The engineering of informality, friendship, and even intimacy entails a crossing of borders between the public and the private, the commercial and the domestic, which is partly constituted by a **simulation of the discursive practices of everyday life, conversational discourse** (Fairclough, 1996: 7). (emphasis added)

These attestations of both structures in the Quebec literary corpus, also and above all, reveal that they are widespread in society.

A complementary corpus of the FDLQ-fund sheds an interesting light on the two structures: the *Chroniques québécoises du langage* gathers language columns published in the Quebec press from 1865 to 1996. Twelve³⁷ of them mention the embedded *qu'est-ce que* clause between 1911 and 1986: in the first mentions, we find eight (five by the same author) which are corrective:

- (28) Ne dites pas : Je ne sais pas qu'est-ce que je dois faire. (Blanchard, Étienne, *Chronique du bon langage*, La Presse, Montréal, 1er février 1919, p. 10, col. 1–5 [ChroQué])
- (29) Je ne sais pas « qu'est-ce que » je devrais faire. (Lefranc, Paul [pseudonyme of Étienne Blanchard], *Autour de la langue française*, La Presse, Montréal, 16 juin 1923, p. 46 [ChroQué])
- (30) je me demande qu'est-ce qu'il fait. – je me demande ce qu'il fait. (Rosemont, Gilles de, *Parlons correctement le français*, La Patrie, Montréal, 15 août 1954, p. 28 [ChroQué])

³⁷The mention rate is 12/7936 billets, a very low percentage (0,15%), that shows the very little interest the structure reveals in these *Chroniques*.

The more recent occurrences (all from the same author, a linguist) appear to be much more descriptive and even advocating for the structure (cf. ex. 26 & 27):

- (31) « Elle ignore ce que, ce qui... », etc., sont **systématiquement suivis dans notre tête – sélection oblige !** – du fameux « qu'est-ce que », « qu'est-ce qui » en remplacement du « ce que », « ce qui ». (Barbaud, Philippe, Parler d'ici, La Presse Plus, Montréal, 18 mai 1985, p. 16 [ChroQué]) (emphasis added)
- (32) **Disculpions-nous de tout complexe** d'avoir créé un horrible précédent quant à cet usage puisque Frédéric Deloffre constate, dans son petit livre consacré à la phrase française, que des tournures comme : « Dis-moi est-ce que tu veux venir, dis-moi qu'est-ce que tu veux » sont **depuis peu d'années de plus en plus répandues**. (Barbaud, Philippe, Parler d'ici, La Presse Plus, Montréal, 18 mai 1985, p. 16 [ChroQué]) (emphasis added)
- (33) En maintenant d'autre part son refus d'écrire **des mots que l'on dit constamment** comme, par exemple, le « TU » interrogatif, [...] et le QU'EST-CE QUE apparaissant dans les phrases suivantes : On sait-tu ce que ça veut dire ? (Barbaud, Philippe, Parler d'ici, La Presse Plus, Montréal, 11 janvier 1986, p. 16 [ChroQué]) (emphasis added)

As for the embedded *in situ* clauses, they are not mentioned in these language chronicles at all. Moreover, neither of the two structures are mentioned on a website tracking in Quebec 'la faute de français qui vous dérange le plus'³⁸: it even shows an occurrence of the *in situ* structure in a critical comment about the word *jour* showing that the structure isn't seen as 'not correct':

- (34) 'Laisser un petit jour lorsqu'on ouvre une fenêtre ou une porte (pour celui-là, je sais pas **c'est quoi** le vrai mot qu'on doit dire. Un petit jeu, peut-être?)'

A normative publication figuring on the OQLF³⁹-website about the sustained spoken French of future teachers mentions the embedded *qu'est-ce que* clause: 'j'sais qu'est-ce que tu veux' (Ostiguy, Champagne, Gervais & Lebrun 2005: 29, note 13) and the embedded *in situ* clause as colloquial variants, 'caractérisée[s] par l'intrusion d'une interrogation directe en guise de complétive indirecte':

- (35) plusieurs variantes familières: «i' m'demandent toujours **quand est-ce que j'reviens**» (quand) «elle était v'nue voir comment que j'reagissais»⁴⁰ (comment) «j'sais pas j'sais pas ça fait combien de temps j'parle» (depuis combien de temps) «j'me souviens j'me souviens **c'est comment**, une classe» (comment c'est, une classe) «je savais je savais **c'était qui**» (qui c'était) «elle leur demandait **c'était quoi** leur objectif» (quel était) «je sais pas **c'est quoi** le chemin» (quel est). (Ostiguy, Champagne, Gervais & Lebrun 2005: 47)

³⁸C'est quoi la faute de français qui vous dérange le plus ? (861 comments; 254 pages; 14213 words) (https://www.reddit.com/r/Quebec/comments/r1vrwl/cest_quoi_la_faute_de_fran%C3%A7ais_qui_vous_d%C3%A9range/)

³⁹Office québécois de la langue française: <https://www.oqlf.gouv.qc.ca/recherche/>; <https://vitrinelinguiste.que.oqlf.gouv.qc.ca/23281/la-syntaxe/linterrogation/formules-interrogatives-ce-que-et-quest-ce-que>.

⁴⁰We can see here another variant of indirect interrogations with *Wh-* followed by *que*.

These different publications reveal clearly the existence of both embedded structures and the evolution of the attitudes towards them.

As a whole, the embedded *in situ* and *qu'est-ce que* structures are clearly present in Quebec, as shown by the oral corpora, the literary sources, and more prescriptive writings. Several indications point out that the structures belong to the 'ordinary' everyday language.

The presence of the embedded *in situ* structures in particular among the speakers in Quebec and in Reunion Island, might⁴¹ thus lead to a probable common source in the seventeenth century⁴²:

Comparative reconstruction has equally yet to be fully exploited. The goal is to examine the French spoken overseas or creoles formed on the basis of spoken French taken abroad by colonizers in the seventeenth century. If common features can be identified in these varieties, especially if they occur in areas which are geographically disparate, it may be possible to hypothesize that these features are present in the common source, namely seventeenth-century spoken French (cf. Chaudenson 1973, 1994; Valdman 1979). The seventeenth century is a particularly fertile area for such investigation since French was taken to three principal areas which are geographically widely separated: North America, notably Acadia and Quebec; Central America, and especially Caribbean islands such as Guadeloupe, Martinique, Dominica, and Saint Lucia, and the Indian Ocean islands of Reunion and Madagascar. (Ayres-Bennett 2004: 32)

What seems sure, is that the *in situ* structure doesn't constitute a recent innovation of modern spoken French, but rather a case of "long-standing differences in written and spoken usage" (Ayres-Bennett 2004: 38).

We will now examine the embedded status of the structure we are analysing in this research, and then describe its characteristics in view

⁴¹In a personal discussion by e-mail, W. Ayres-Bennett compares the direct *in situ* constructions with the presence/absence of the *ne* of negation: "My conclusion there is that we have tantalising glimpses: the 'popular' construction without *ne* was already possible in the seventeenth century, but the evidence is patchy. I think the same is true of evidence for *in-situ* interrogation. There is indeed evidence for it, for example in Heroard's journal – and here it is also used by the adults as well as by the Dauphin." (personal communication 2017 W. Ayres-Bennett)

⁴²Chaudenson mentions a melting-pot of linguistic usages of colons in the seventeenth century: "La variation intrasystémique d'une langue (en l'occurrence l'idiome des colons du dix-septième siècle que je me garde pour le moment, par prudence, de nommer le « français ») peut engendrer, dans la migration de populations françaises, sans doute relativement hétérogènes, une koïnèisation, c'est-à-dire l'émergence d'une variété relativement cohérente et stable." (Chaudenson 1998 : 163).

Valdman also underlines their colloquial character: "La présence de nombreux traits partagés par le français populaire et les parlers des isolats américains démontre clairement en effet que ceux-ci procèdent du langage familier plutôt que du parler soutenu de l'élite. [...] Les traits morphosyntaxiques qui différencient les parlers des isolats du français standard se retrouvent dans les autres communautés francophones d'Amérique et dans le français populaire sur une vaste aire en France." (Valdman 1979: 195).

of the FDLQ corpus, which allows us to contrast different situations of use of the structure.

5. STRUCTURAL CHARACTERISTICS OF THE EMBEDDED *IN SITU* CLAUSE

The embedded *in situ* clause, e.g. *je me demande c'est quoi*, combines the two characteristics given by Muller (1996) of indirect interrogative: equivalence with 'cela'⁴³; paraphrasable by 'quel est'⁴⁴:

(36) *Je me demande c'est quoi/cela*, where the pronoun designates the 'identity of things'

(37) *Je me demande c'est quoi = Je me demande quelle est la chose*

Moreover, the subordination is "étanche" ('impermeable', i.e. true subordination), according to Blanche-Benveniste's criteria⁴⁵ (1982: 77–81). A final syntactic criterion concerns the "equivalence" between the different variants: as in the direct interrogative, where the variants *qu'est-ce?*, *qu'est-ce que c'est?* and *c'est quoi?* coexist, to mention only the main structures examined here, the indirect interrogative

⁴³In order to identify embedded interrogative clauses, Muller (1996) gives a series of distinctive properties: by analysing the ambiguous sentence *Je lui ai demandé ce qu'elle tenait*, he makes clear the distinction between the relative sentence, pointing to the object (which the speaker would demand to obtain), and the embedded interrogative sentence, dealing with the identity of the object (which the speaker is questioning). Thus, the subordinate can be replaced by *le/cela* in both cases, but with very different meanings:

(2) *Je mange ce que Paul a acheté/cela*: representative of a concrete term (relative clause)

(3) *Je me demande ce que Paul a acheté/cela*: designating the identity of things, *cela* being here a "demonstrative of identification" (interrogative clause) (Muller 1996 : 210).

⁴⁴Muller gives another reliable way to identify indirect Wh- interrogative clauses: the possibility of constructing a paraphrase with *quel est* that requires to specify the classifying name of the domain (*qui = quelle est la personne*):

(1) *Je sais qui vient d'entrer = Je sais quelle est la personne qui vient d'entrer* (interrogative clause)

(2) *Je remercie qui vient d'entrer = *Je remercie quelle est la personne qui vient d'entrer* (relative clause) (Muller 1996 : 198).

⁴⁵- The two verbs have a constructor status: they construct their subject and complements, each having their "valence elements";

- The second verbal construction is "a part of the valence of the first, in the same way as a pronoun" (1982: 105): as in *il sait qu'il part demain*, where "qu'il part demain" can be equated with a pronoun (*il le sait, il sait ça*), "c'est quoi" in *il sait pas c'est quoi* is equivalent to the pronoun *le* (*il le sait pas ; il sait pas ça*). The second verb belongs to a subordinate clause ruled by the first.

- The first verb is a constructor verb that freely takes on interrogative or negative modalities (Blanche-Benveniste, 1982: 79): *savez-vous qu'il part demain ? ; on ne savait pas qu'il part demain*. The same is true of *il sait pas c'est quoi ; il sait c'est quoi ; il sait c'est quoi?*

This last point concerning the interrogative modality can be discussed and rather seen as a phrase modality, and thus related to macro-syntaxis (Blanche-Benveniste et al. 1990). The further analysis of the *weak verbs* (Blanche-Benveniste 1989 ; Blanche-Benveniste & Willems 2007, 2016), i.e. parenthetical verbs (of which *verba dicendi*, verbs of knowledge...), will deepen this analysis (cf. conclusion).

presents the variants⁴⁶ *il ne sait pas ce que c'est, il ne sait pas qu'est-ce que c'est, and il ne sait pas c'est quoi*.

We can also note that the introduction verbs turn out to be classical introducers of subordination (Riegel, Pellat & Rioul 2018; Muller 1996), “dont la valeur interrogative ne saute pas aux yeux” (Grevisse 1988: 683, §411c, Remarques)⁴⁷: they are naturally verbs of interrogation (28 occurrences) (*demander*,⁴⁸ *se demander*), a subcategory of *verba dicendi* (of whom 12 other occurrences are attested) (*expliquer, dire...*), and in great majority verbs of knowledge (143 occurrences) (*savoir, montrer, comprendre, se souvenir, connaître, voir, regarder, ressentir, se rappeler, découvrir, penser, avoir une idée...*); in sum, verbs that can signify “knowledge about the truth value of their complement” (Muller 1996: 201–203). Table 3 lists the different verbs and one compound verbal according to these two categories⁴⁹.

Table 3. Introduction verbs of the embedded *in situ* clause with *quoi* in FDLQ

Verba dicendi (40), of which (28) verbs of interrogation	<i>expliquer</i> (6); <i>dire</i> (4); <i>se dire</i> (2) <i>demander</i> (22); <i>se demander</i> (6)
Verbs of knowledge (143)	<i>savoir</i> (105); <i>se souvenir</i> (11); <i>voir</i> (7); <i>comprendre</i> (4); <i>se rappeler</i> (5); <i>connaître</i> (2); <i>montrer</i> (2); <i>regarder</i> (2); <i>avoir une idée</i> (1) ⁵⁰ ; <i>chercher</i> (1); <i>découvrir</i> (1); <i>penser</i> (1); <i>ressentir</i> (1)

⁴⁶The application of the concept of variation to syntax has been problematic for a long time, in particular because of difficulties related to the very definition of the variable and the equivalence between forms, as well as the lack of occurrences in the data and the difficulty of applying quantitative analyses (Gadet 1997; Tagliamonte 2012). Furthermore, syntactic variation is, more than other areas (pronunciation, lexicon), considered as incorrect, and spontaneous oral syntax, as chaotic (Blanche-Benveniste 1997: 5–6): Knecht (1999: 37) points out that ‘l’enseignement ne tolère aucune liberté dans ce domaine, considéré plus que tout autre comme critère de correction’ (Ledegen & Martin, 2020).

⁴⁷“Le support de l’interrogation indirecte peut être un verbe de sens nettement interrogatif, comme *demander*, mais ce n’est pas toujours le cas. Si le verbe support implique l’incertitude par son sens (*ignorer*) ou par sa construction négative ou interrogative, ou si le verbe support est à l’impératif ou au futur, il s’agit de quelque chose qu’on ignore et dont on s’enquiert; la nuance interrogative est donc perceptible. [...] La nuance interrogative est peu sensible quand les conditions données ci-dessus sont absentes. C’est seulement la présence d’un mot interrogatif qui permet de reconnaître une interrogation indirecte: Je sais quelle est la réponse.” (Grevisse 1988:1693, §1102) (“The support for indirect questioning can be a verb with a clearly interrogative meaning, such as *demander*, but this is not always the case. If the support verb implies uncertainty by its meaning (*ignorer*) or by its negative or interrogative construction, or if the support verb is in the imperative mood or future tense, it is something that is not known and that is being asked about; the interrogative nuance is therefore perceptible. [...] The interrogative nuance is not very noticeable when the conditions given above are absent. It is only the presence of an interrogative word that makes it possible to recognise an indirect interrogation.”)

⁴⁸The categories are not mutually exclusive: *demander* belongs to the interrogative verbs who belong to the *verba dicendi*.

⁴⁹The same observations are made for the embedded *qu’est-ce que* interrogative: *verba dicendi* (31 occ.), of which verbs of interrogation (9 occ.); verbs of knowledge (73 occ., with 55 occ. of *savoir*).

⁵⁰The unique occurrence of ‘avoir une idée’ shows that this compound verbal form behaves like single verbs: je s/, j’ai aucune idée c’est quoi les règles euh (Côté (2009-), *Entretien guidé qcqbc1g, Montréal [PFC-Québec]*).

As for the subjects used by these introduction verbs, they are almost exclusively pronominal (93.8%), and concern essentially – and classically in interlocution (Halliday 1985) – the two persons of the interlocution *je* and *tu*, as Table 4 clearly shows:

Table 4. Subjects of the introduction verbs of the embedded *in situ* clause with *quoi* in FDLQ

Subjects	Occ.	%
<i>je</i>	79	43,2%
<i>tu</i>	34	18,6%
<i>il</i>	11	6%
<i>elle</i>	5	2,7%
<i>on</i>	12	6,6%
<i>ça</i>	1	0,5%
<i>il impersonal</i>	3	1,6%
<i>tout le monde, personne</i>	4	2,2%
<i>vous</i>	5	2,7%
<i>ils</i>	9	4,9%
NPwa	4	2,2%
PN	6	3,3%
TOTAL	183	100%

Consistent with these major trends in oral language where mainly pronominal subjects appear, we note that subjects in the form of nominal phrase (NP) or proper nouns (PN) (6.2%) are quasi exclusively found in the literary corpus (6 occ.)⁵¹ and in the metalinguistic interviews with the journalist (1 occ.):

- (38) Les trois gars, a poursuivi Steeve entre deux gorgées, s'obstinent pour savoir **c'est quoi** l'animal le plus dangereux de la Gaspésie. (Bernard, 2017, *La bête creuse* [CLIQ])

These nominal subjects reveal a certain formality of discourse; the same is true for subject proper nouns, which in ordinary speech would be completed with a pronoun (40):

- (39) Maria demande **c'est quoi** la différence de maman. (Turcotte, 1991, *Le bruit des choses vivantes* [CLIQ])
 (39') Maria elle demande **c'est quoi** la différence de maman

⁵¹Three of which are from the same book *La bête creuse*, already mentioned in note 30 for its high rate of indirect *in situ* structures.

Thus, the literary corpus shows here informalisation, by stereotyping certain features of speech, without entirely corresponding to the orally attested uses.

Then, as regards the form and modality of the introduction verbs, it should be noted that interrogation (by modality (for ex.: *As-tu pensé c'est quoi ?* (Vincent *et al.*, *Entrevue Paul G* [Montréal 1995])) or verb type (for ex.: *demander*)) and negation together form three quarters of the examples, as show in Table 5⁵².

Table 5. Modalities of the introduction verbs of the embedded *in situ* clause with *quoi* in FDLQ

	Occ.	%
Affirmative form	33	17,9%
Negative form	79	43,2%
Interrogative modality	24	13,1%
Interrogative verb	27	14,8%
Total	183	100%

Thus, the interpretation in terms of the truth value of the subordinate is above all indeterminate (Muller 1996: 202), as noted in note 37 ('il s'agit de quelque chose qu'on ignore et dont on s'enquiert ; la nuance interrogative est donc perceptible' (Grevisse 1988: 1693, §1102)): the truth value remains suspended as well because the speaker does not know which of the terms is the right one [negation], or because he knows and refuses to say [interrogation] (Muller 1996: 208). But the affirmative form is still largely attested, revealing that the truth value can be fully assumed by the speaker:

- (40) Oui je sais **c'est quoi** (Blondeau *et al.*, *Entrevue Christine* [Hochelaga-Maisonneuve 2012])

Finally, as Table 6 shows, as far as the verb present in the embedding is concerned, the structure is very strongly fixed and used for the process of identification.

Table 6. Verbs in the embedded *in situ* clause with *quoi* in FDLQ

Embedded verbs	Occ.
(c')est	136
(c')était	43
(ça) va être	1
(ça) donne	1
(ça) fait	1
(ça) voulait dire	1
Total	183

⁵²The same is true for the *qu'est-ce que* interrogative: affirmative form (32 occ.); negative form (49 occ.); interrogative modality (14 occ.); interrogative verb (9 occ.).

Almost all verbs turn out to be copulas, serving to identify: we find the form *c'est* mainly in the present indicative (136 occ.), and more rarely in the imperfect tense (43 occ.) or in the future tense, more specifically the periphrastic future (1 occ.). It should be noted that among the 43 forms in the imperfect tense, 22 (i.e. more than the half) come from the CLIQ literary corpus, which confirms that the embedded verbs show rather little formal variation. These results also confirm Secova's findings, concerning the length of the embedded clause: « la forme *in situ* [...] est favorisée dans des subordinées courtes, comme *c'est quoi, c'est qui* ou *c'est où* » (Secova 2017: 13).

A final point worth examining concerns the detachment of a nominal or verbal phrase, also known as left (before) or right (after) dislocation: Blanche-Benveniste (1997b) has proposed an analysis of direct interrogative on the inanimate subject that sheds light on the detached elements near the embedded *in situ* interrogatives. She proposes a grammatical and not a stylistic justification, as the detached turns of phrase (*c'est quoi, le N?* or *qu'est-ce que c'est, un N?*) are distributed according to the degree of referentiality of the subject: the turn of phrase *qu'est-ce que c'est* is used with noun phrases of generic value, often with an indefinite article, acting as requests for general definition (*qu'est-ce que c'est, comme sorte de chose?*), while the turn *c'est quoi* has a clearly deictic value, aiming to elucidate a specific term present in the context (*c'est quoi, cette chose?*) (Blanche-Benveniste 1997b: 143–144).

In 2007, half of our data from Reunion Island (Ledegen, 2007a) showed detachments, especially placed on the right (14 out of 25 occurrences of embedded *in situ* clause), including eight noun phrases introduced by a definite article and three with an indefinite article. The FDLQ data reveal the same organisation,⁵³ in an even more massive way, as shows Figure 3:

The majority of dislocations are placed on the right, and two thirds are with a determined noun phrase.

- (41) pluguer puis bien je veux dire peut-être qu'ils ont dit toujours pluguer mais des:⁵⁴ des mots dans ce gen: dans ce style là qu'on se sert couramment là mais que: ils deviennent heu: ça: ça devient pas-mal en: ou bien donc on sait plus **c'est quoi** le français de tel mot. (Thibault *et al.*, Entrevue Christine Q [Montréal 1984])
- (42) je sais pas si tu sais **c'est quoi** le coin là (Blondeau *et al.*, Entrevue Alex [Hochelaga-Maisonneuve 2012])

Once again, the literary corpus reveals overuse: among six left dislocations, three come from this particular corpus, with two remarkable examples cumulating two dislocations (as in ex. 43, subject of the main clause, and dislocation of the demonstrative pronoun *c'*, subject of the embedded clause) and 45 (twice dislocation of the demonstrative pronoun *c'*, subject of the embedded clause) and one example positioning the dislocation at the beginning of the embedded clause (as in ex. 44,

⁵³The embedded *qu'est-ce que* interrogative doesn't combine as frequently with dislocation: we find only four occ. before and one occ. behind, and only one of them coming from the literary corpus (*Ces jeunesses-là, ça sait-il bien qu'est-ce que ça veut.* (Cloutier, 1925, L'erreur de Pierre Giroir [CLIQ]). Confirming the analysis of Blanche-Benveniste, we find three GN indetermined, one GN determined and one infinitive.

⁵⁴In the transcriptions, « : » stands for long pronunciation.

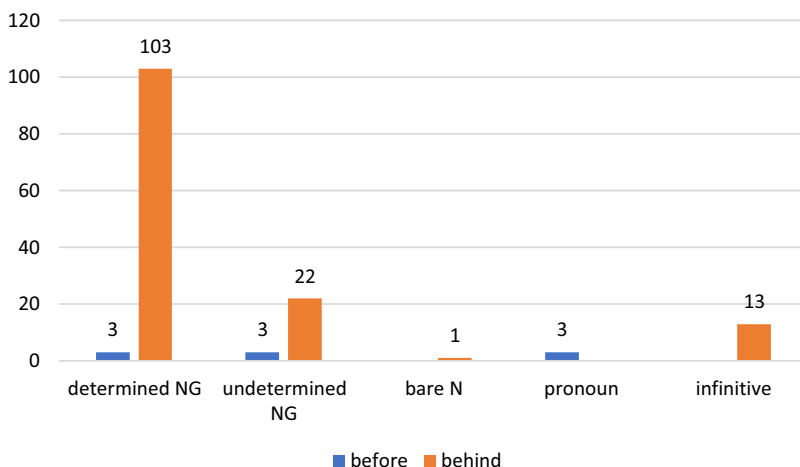


Figure 3. Dislocations in the embedded *in situ* clauses with *quoi* in FDLQ⁵⁵.

dislocation of the demonstrative pronoun *c'*, subject of the embedded clause), a sequence that is not attested in the oral corpora:

- (43) Aznavour, la misère, il sait pas trop **c'est quoi**, anyway. (Bienvenu, 2014, Chercher Sam. [CLIQ])
- (44) Tu sais le gambit, **c'est quoi** ? (Hamelin, 2010, La constellation du Lynx. [CLIQ])
- (45) Des patates pilées, ça il savait **c'était quoi**. (Bernard, 2017, La bête creuse. [CLIQ])

The literary corpus again shows features that are stereotypical of the spoken language, but not actually attested structures.

6. A NEW PIECE TO THE PUZZLE: ATTESTATION IN PRE-SCHOOL PRACTICES

The fact that the structure is practised by children before they go to school (cf. Palasis & Faure, in this volume) makes it possible to posit that the embedded *in situ* clause constitutes a feature of “primary knowledge of grammar”, i.e. the “initial knowledge mastered by children before the age of primary school” (Blanche-Benveniste 1990: 52), supplemented or replaced subsequently at school, where the “second knowledge of grammar” is learned (Blanche-Benveniste 1990: 52). Secova also finds more embedded *in situ* clauses used by the youngest speakers of the PMF-study (10–14 years old),⁵⁶ and supposes that “la forme standard est

⁵⁵It is interesting to note that 10 determined NG are completed with relative clauses (six with *que* and four with *qui*): *Je peux tU* savoir c'est quoi les voyages que tu as faits* (Thibault *et al.*, Entretien Marcel V. [Montréal1984]); *Puis si je t'avais demandé c'est quoi l'événement qui t'a le plus marqué là depuis la première fois qu'on t'a vu* (Thibault *et al.*, Entretien Marcel V. [Montréal1984]). [*The transcription *tU* indicates the interrogative particle, specific of Quebec's French (Thibault & Vincent 1988: 29).]

⁵⁶They prefer the *in situ* structures (14 occ.) rather than standard ones (2 occ.) (Secova 2017: 11).

majoritairement acquise au fur et à mesure de l'instruction scolaire" (2017: 13). The structure seems to belong to those constructions that are "continually censored and, it seems, continually alive" (Blanche-Benveniste 1995: 128).

Palasis' corpus (2009),⁵⁷ recorded in France, attests the embedded *in situ* structure⁵⁸ with preschool children, more specifically three children in the second and third year of nursery school. It should be noted that they use various introduction verbs (*se demander, dire, regarder*⁵⁹) and embedded verbs (*c'est, il y a, je + verb*):

- (46) je me demande **c'est quoi que ça s'appelle** ? (MAT, 2nd year of preschool class)⁶⁰
 (47) on disait **c'est laquelle** une fille. (MAT, 2nd year)
 (48) i(l) regarde y a **quoi**. (ELI, 2nd year)
 (49) eh ben regardez **j(e) peux faire quoi** moi. (LAN, 3^d year)
 (50) regardez **j(e) peux faire quoi** avec mes yeux. (LAN, 3^d year)

It should be noted that within this corpus, the form of the embedded clause with *est-ce que* is also attested:

- (51) poisson pigeon papillon canard et ce truc là je sais pas **qu'est-c(e)+que c'est**. (LSN, 2nd year)

and much more extensively than the *in situ* form, as shown in Table 7⁶¹.

Table 7. Non-standard embedded *in situ* and *est-ce que* clauses by pre-school children in Palasis' corpus (2009)

Embedded clauses/pre-school classes	1 st year	2 nd year	3 ^d year
<i>in situ</i>	/	3 occ.; 2 children	2 occ.; 1 child
<i>est-ce que</i>	1 occ.; 1 child	30 occ.; 7 children	17 occ.; 8 children

⁵⁷Made in France, this corpus covers a three-year period, with 22 children aged 2;5.5 to 4;0.1 from the same kindergarten class; they are all L1 French speakers except for two children (of whom Eli from the ex. (42)): <https://childes.talkbank.org/access/French/Palasis.html>.

⁵⁸The transcriptions of the examples clearly show that these are embedded structures, by indicating pauses and stops in the productions of the children.

⁵⁹The verbs *dire* and *regarder* make part of the verbs considered introductory of embedded interrogative constructions (cf. *supra*). The frontier with relatives is fuzzy as shows also the following remarks of the *Chroniques québécoises du langage-corpus*: *Cette dissidence affecte en outre les propositions relatives démonstratives du type illustré par notre exemple-vedette à un point tel que cette locution peut être modifiée par certains éléments comme dans* : « Tout qu'est-ce (que) j'ai à dire, je l'ai dit » ; « Rien de qu'est-ce (que) t'as fait est correct » (Barbaud, Philippe, « Parler d'ici », La Presse Plus, Montréal, 18 mai 1985, p. 16 [ChroQué]). Lefeuve & Rossi-Gensane indicate, in hexagonal French, the presence of *est-ce que* in relative clauses: for example, in ESLO1, produced by speakers born in 1900 or 1910 and having left school between the age of 11 and 14 (2017 : 67–68).

⁶⁰We can note the presence of a « dispositif de clivage » here (cf. *supra*) : *c'est ... que/qui* « accueille les valences et rections d'un autre verbe » (Blanche-Benveniste 2002 : 89), here *appelle*.

⁶¹They are produced by ten distinct children: five children realised occurrences during two or even three consecutive years.

These attestations with pre-school children reveal that the embedded *in situ* clause might be part of the self-regulation process (Chaudenson *et al.* 1993: 38): they put into light this point of “weakness” or “fragility” of the interrogative in French (Chaudenson *et al.* 1993: 6–7). The embedded *in situ* clause might belong to the panlectal variation of French studied by Chaudenson *et al.* who “[traquent] le français en liberté, à l’état de nature, indemne des contraintes que lui ont imposées les grammairiens du XVIIe siècle et que la tradition académique a sanctionnées”⁶² (Manessy 1993: 3).

Moreover, Foulet’s historical analysis of interrogative forms reveals that the *in situ* order became widespread from the seventeenth century onwards: “De 1350 à 1650 l’effort de la langue a consisté principalement à faire triompher l’ordre sujet-verbe-complément, en d’autres termes à se débarrasser tant bien que mal des nombreuses inversions dont elle avait hérité et qui étaient désormais contraires à son génie”⁶³ (Foulet 1921: 262). The author considers that the turn of phrase “Votre père partira quand ?” constitutes “un développement très logique et au fond très naturel, dont la force peut un jour devenir irrésistible. Il entraînerait la langue encore plus loin de l’inversion : il en ferait disparaître jusqu’au souvenir même”⁶⁴ (Foulet 1921: 348). The *in situ* structure in direct interrogative then became widespread from the seventeenth century onwards, and could be, according to Ayres-Bennett, a case of “long-standing differences in written and spoken usage” (2004: 38). If the direct *in situ* structure remains rare in French texts between the seventeenth and nineteenth centuries, it is above all because it is non-standard: Ayres-Bennett (2004: 50–58) gives several examples in theatre texts and Mathieu (2009) attests an example in Diderot’s *Le Rêve de d’Alembert* (1784): “Mademoiselle de L’Espinasse: *Et cette comparaison se fait où ? Bordeu.*” (Mathieu 2009: 59). The attestation of the embedded *Wh- in situ* structure on multiple terrains of the French-speaking world thus shows that “there is evidence for [its] usage in the French spoken in France at the time of the colonization” (Ayres-Bennett 2004: 38).

A parallel could be drawn here with Anthony Kroch’s (1978) analysis of pronunciation, where he contrasts speakers of prestige dialects with speakers of vernaculars:

Our position [...] is that prestige dialects resist phonetically motivated change and inherent variation because prestige speakers seek to mark themselves off as distinct from the common people. [...] Thus, we are claiming that there is a particular ideological motivation at the origin of social dialect variation. This ideology causes the prestige dialect user to expend more energy in speaking than does the user of the popular vernacular. (1978: 30)

⁶²[track down] French in freedom, in a state of nature, unharmed by the constraints imposed on it by the grammarians of the seventeenth century and sanctioned by the academic tradition’.

⁶³From 1350 to 1650 the effort of the language consisted mainly in making the subject-verb-complement order triumph, in other words in getting rid, as best as it could, of the numerous inversions it had inherited and which were henceforth contrary to its genius.’

⁶⁴a very logical and basically very natural development, whose force may one day become irresistible. It would take the language even further away from inversion: it would make even the memory of it disappear.’

Also in syntax, the ordinary realisations constitute the common way of speaking, from which the dominant speakers try to differentiate themselves in certain contexts.

The embedded *Wh- in situ* clause seems to belong fully to these vernacular practices: it conforms to the evolutionary trends of French, where the reduction of declensions, the decumulation of synthetic forms, and the sequential syntax are attested (Guiraud 1966; Lefevre & Rossi-Gensane 2017). This fixity of word order thus has the joint advantage of maintaining the parallel between direct and indirect structures, and of allowing the interrogative status to be highlighted by the use of *quoi* rather than the form *que*, which can be taken for the relative structure.

7. CONCLUDING REMARKS

The embedded *Wh- in situ* clause turns out to be a long-lived and attested structure in view of its presence in many areas of the French-speaking world, a large part of which has in common being an ordinary French-speaking world dating from the seventeenth century. This vernacular turn of phrase corresponds to the evolutionary tendencies of French, and is structurally rather fixed: the embedded *in situ* clause is mainly constructed with *c'est*, and completed by dislocated elements; as for the introduction verbs, they are above all made up of verbs of knowledge, frequently in the negative or interrogative form (but also in an affirmative form (25%)), or of interrogative verbs.

The structure is attested in a very wide variety of French-speaking areas, without coming from the influence of other contact languages: its presence even in current Quebec literature and paraliterature reveals its real use; French and Quebec rap songs, which may be counter-standard, attest the structure too. Finally, the presence of the structure among preschool children confirms its existence in contexts of lesser normativity, also attested in certain ‘marginal’⁶⁵ francophone contexts (Acadia, Ontario, Newfoundland) where the structure is found as well (Ledegen 2007c). These different elements not only make it possible to link the structure to “the ‘first’ knowledge of grammar” (Blanche-Benveniste 1990), but also to eliminate the categorisation of the turn of phrase as being exclusively of “register populaire”.

If the embedded *in situ* structure is more visible today, in sociolinguistic corpora, but also in the media, or in metalinguistic commentaries, it is, on the one hand, because of the informalisation of media and public practices,⁶⁶ and, on the other

⁶⁵This term, proposed by Chaudenson (1989), refers to “marginalité à la fois géographique et structurelle [...]. Les français marginaux se rencontrent en particulier en Amérique du Nord, essentiellement dans des zones où le français s’est transmis, des siècles durant, par voie exclusivement orale, indépendamment de toute superstructure culturelle, le français n’y étant pas transmis par l’école [: la Louisiane, le Missouri, Saint-Barthélemy (Antilles françaises)].” (“both geographical and structural marginality [...]. Marginal French is found in particular in North America, essentially in areas where French has been transmitted for centuries exclusively orally, independently of any cultural superstructure, since French is not taught in school [: Louisiana, Missouri, Saint-Barthélemy (French West Indies)].”) (Chaudenson 1997: 154)

⁶⁶France Martineau interestingly considers that the impression of a decline in the quality of French comes mainly from its visibility. « De nos jours, les peu-lettrés écrivent beaucoup plus, et publiquement sur Twitter ou Facebook. Le nombre croissant de communications écrites révèle davantage les points faibles des conventions orthographiques et le gros écart entre l’écrit et l’oral. Avant de juger l’orthographe, il faut considérer l’objectif. Les textos visent une communication ultrarapide plutôt qu’une démonstration de son brio. » (Martineau 2022)

hand, because the corpora studied are more varied, and closer to the axis of communicative proximity (Koch & Oesterreicher 2001): sociolinguistic interviews are more frequently organised in a free form than by following a strict questionnaire⁶⁷ that gave the respondent an exclusive responding role in a less ordinary way.

Our analysis of the previous data and the complementary analysis of the FDLQ corpora have allowed us to argue for the extended use of oral and written corpora, methodologically obtained in an ecological way, i.e. obtained within the framework of a strong inter-acquaintance and situated at the pole of communicative proximity (Koch & Oesterreicher 2001), as used in the MPF project (Gadet 2017), rather than for corpora bringing back the ‘paradox of the observer’ (Labov). We therefore argue for the adoption of a communicative approach to variation (Guerin 2017) in the light of the “sociolinguistics of language” (Gadet & Guerin 2021). These investigations might be built with different age groups, to compare youth to the other generations, in different places and departments, including Overseas Departments and the French-speaking world as a whole, and combine with epilinguistic discourses revealing the attitudes and representations of the speakers. We thus wish that a sociolinguistic inquiry combining the lessons from the MPF-Project with the investigations of the Francophone world by the project *Phonologie du français contemporain* (PFC) (Durand, Laks & Lyche 2009) may see the light soon.

In the future, in order to continue seeking for the embedded *Wh- in situ* clause, we will pursue the exploration of oral and written corpora established in contexts of communicative proximity, and take a closer look at introductory verbs, particularly *weak* verbs (Blanche-Benveniste 1989; Blanche-Benveniste & Willems 2007, 2016), from a macro-syntactic perspective; moreover, the study of the extended contexts of each attestation observed here, in its complementarity with other embedded or non-embedded *Wh-* structures (standard, all the *Wh-* words...), will offer a promising perspective. Finally, an exploration in historical sociolinguistics of old judicial corpora containing ordinary reported speech (Dourdy & Spacagno 2020; Dourdy 2021) should allow us to date the rise of direct and embedded *in situ* interrogatives more precisely, and possibly attest the existence of the *Wh- in situ* in embedded form as early as the seventeenth century.

Competing interests declaration. None.

REFERENCES

- Ayres-Bennett, Wendy, 2004, *Sociolinguistic Variation in Seventeenth-Century France: Methodology and Case Studies*, Cambridge, Cambridge University Press.
- Bartoli, Matteo, 1945, *Saggi di linguistica spaziale*, Torino, Rosenberg & Sellier.
- Blanche-Benveniste (C.) & Willems (D.), 2016, « Les verbes faibles », in *Encyclopédie Grammaticale du Français*, en ligne : <http://encyclogram.fr>
- Blanche-Benveniste, Claire, 1982, Examen de la notion de subordination, *Recherches sur le français parlé*, 4, 71–115.

⁶⁷The corpus Trois Rivières (1972) of the DFLQ for instance does not reveal our two structures by being too strictly bound by a questionnaire, giving little space for ordinary productions by the informants.

- Blanche-Benveniste, Claire**, 1983, L'importance du français parlé pour la description du 'français tout court', *Recherches sur le français parlé*, 5, 23–45.
- Blanche-Benveniste, Claire**, 1989, Constructions verbales "en incises" et rection faible des verbes, *Recherches sur le français parlé* 9, 53–73.
- Blanche-Benveniste, Claire**, 1990, Grammaire première et grammaire seconde : l'exemple de EN, *Recherches sur le français parlé*, 10, 51–73.
- Blanche-Benveniste, Claire**, 1995, Quelques faits de syntaxe, in Antoine, Géraud & Martin, Robert (dirs), *Histoire de la langue française 1914-1945*, Paris, CNRS Editions, 125–147.
- Blanche-Benveniste, Claire**, 1997a, *Approches de la langue parlée en français*, Paris, Ophrys.
- Blanche-Benveniste, Claire**, 1997b, A propos de *Qu'est-ce que c'est* et *C'est quoi...*, *Recherches sur le français parlé*, 14, 127–146.
- Blanche-Benveniste, Claire**, 2002, Auxiliaires et degrés de verbalité, *Syntaxe et sémantique*, 3, 75–97.
- Blanche-Benveniste, Claire, Bilger, Mireille, Rouget, Christine & Van den Eynde, Karel**, 1990, *Le français parlé. Etudes grammaticales*, Paris, CNRS Editions.
- Blanche-Benveniste, Claire & Jeanjean Colette**, 1987, *Le français parlé. Transcription et édition*, Paris, Didier Erudition.
- Blanche-Benveniste, Claire & Willems, Dominique**, 2007, Un nouveau regard sur les verbes faibles, *Bulletin de la Société Linguistique de Paris* 102/1, 217–254.
- Blasco, Mylène & Cappeau, Paul**, 1992, Sujet postposé et double marquage, *Recherches sur le français parlé*, 11: 11–30.
- Blondeau, Hélène & Ledegen, Gudrun**, 2021, « Chaque français » et sa pratique syntaxique ordinaire. *Personne sait c'est qui* (Les Parent (2009–2010)), in Bertin, Annie, Gadet, Françoise, Lehmann, Sabine & Moreno Kerdreux, Anaïs (éds), *Réflexions théoriques et méthodologiques autour de données variationnelles*, Chambéry, PU Savoie Mont Blanc, 169–182.
- Boutin, Béatrice & Françoise Gadet**, 2012, Comment ce que montrent les français d'Afrique s'inscrit/ ne s'inscrit pas dans les dynamiques des français, *Le français en Afrique*, 27, 19–34.
- Chaudenson, Robert**, 1973, Pour une étude comparée des créoles et parles français d'outre-mer : survivance et innovation, *Revue de Linguistique Romane*, 37, 342–371.
- Chaudenson, Robert**, 1994, Français d'Amérique du Nord et créoles français : le français parlé par les immigrants du XVIIe siècle, in *Mougeon, Raymond & Beniak, Edouard, Les origines du français québécois*, Sainte-Foy, Les Presses de l'Université de Laval, 167–180.
- Chaudenson, Robert**, 1997, Français marginaux, in Moreau, Marie-Louise (ed.), *Sociolinguistique. Concepts de base*, Bruxelles, Mardaga, 154.
- Chaudenson, Robert**, 1998, Variation, koinésation, créolisation : français d'Amérique et créoles, in Brasseur, Patrice (ed.), *Français d'Amérique. Variation, créolisation, normalisation*, Avignon, Université d'Avignon/CECAV, 163–179.
- Chaudenson, Robert, Mougeon, Raymond & Bénéak, Edmond**, 1993, *Vers une approche panlectale de la variation du français*, Paris, Didier-Erudition.
- Combettes, Bernard & Glikman, Julie**, 2020, Chapitre 36. Syntaxe de la phrase complexe, in Marchello-Nizia, Christiane, Combettes, Bernard, Prévost, Sophie & Scheer, Tobias (eds), *Grande grammaire historique du français*, volume 2, Berlin, Walter de Gruyter, 1338–1465.
- Conein, Bernard & Gadet, Françoise**, 1998, Le « français populaire » de jeunes de la banlieue parisienne, entre permanence et innovation, in Androutsopoulos, Jannis K. & Scholz, Arno (eds), *Langue des jeunes*, Paris, Peter Lang, 105–123.
- Coveney, Aidan**, 1995, The use of the QU-final interrogative structure in spoken French, *Journal of French Language Studies*, 5, 143–171.
- Defrancq, Bart**, 2000, Un aspect de la subordination en français parlé : l'interrogation indirecte, *Etudes Romanes*, 47, 131–141.
- Defrancq, Bart**, 2005, L'interrogative enchâssée, Louvain, Duculot, Coll. « Champs linguistiques ».
- Dekhissi, Laurie & Coveney, Aidan**, 2018, La variation dans l'emploi des interrogatives partielles dans le cinéma de banlieue, in Béguelin, Marie-José, Coveney, Aidan & Guryev, Alexander (eds), *L'interrogative en français*, Berne, Peter Lang, 119–151.
- Deulofeu, José**, 1989, Les couplages de constructions verbales en français parlé : effet de cohésion discursive ou syntaxe de l'énoncé, *Recherches sur le français parlé*, 9, 111–141.

- Dourdy, Laura-Mai**, 2021, Entend-on vraiment les paysans ? Ce que la variation syntaxique nous dit des représentations sociales dans la littérature des XV^e et XVI^e siècles, *Cahiers Internationaux de Sociolinguistique*, **18**, 53–82.
- Dourdy, Laura-Mai & Spacagno, Michela**, 2020, Donner la parole aux interrogés : une étude de l'oral représenté dans les comptes rendus de procès médiévaux, *Langages*, **217**, 119–132.
- Durand, Jacques, Laks, Bernard & Lyche, Chantal**, 2009, Le projet PFC : une source de données primaires structurées, in Durand, Jacques, Laks, Bernard & Lyche, Chantal (eds), *Phonologie, variation et accents du français*, Paris, Hermès, 19–61.
- Elalouf, Aurélie**, 2012, La notion de 'grammaire seconde'. Tentative de reconstruction épistémologique. CMLF 2012–3e Congrès Mondial de Linguistique Française, Lyon, France. 737–755, (10.1051/shsconf/20120100328).
- Fairclough, Norman**, 1996, Border Crossings: Discourse and Social Change in Contemporary Societies, in Coleman Hywel & Cameron Linne (eds), *Change and Language*, Clevedon, British Association for Applied Linguistics/Multilingual Matters, 3–17.
- Ferreira, Auphémie**, 2022, Etudes des constructions asyndétiques avec les verbes *croire* et *penser* en français parlé contemporain : analyse syntaxique et approche sociolinguistique, Thèse de doctorat, Université Sorbonne-Nouvelle.
- Foulet, Lucien**, 1921, Comment ont évolué les formes de l'interrogation, *Romania*, **47**, 243–384.
- Fournier, Nathalie**, 1993, La langue de Molière et le bon usage de Vaugelas dans Les femmes savantes, *L'Information Grammaticale*, **56** : 19–23.
- Gadet, Françoise**, 1989, *Le français ordinaire*, Paris, Colin.
- Gadet, Françoise**, 1992, *Le français populaire*, Paris, Presses Universitaires de France.
- Gadet, Françoise**, 1997, La variation, plus qu'une écume, *Langue française*, **115**, 5–18.
- Gadet, Françoise**, 1999, La langue française au XXe siècle. I. L'émergence de l'oral, in Chaurand, Jacques (dir.), *Nouvelle histoire de la langue française*, Paris, Seuil, 581–671.
- Gadet, Françoise**, 2003, *La variation sociale en français*, Paris, Ophrys.
- Gadet, Françoise**, 2017, *Les parlers jeunes dans l'Île-de-France multiculturelle*, Paris, Ophrys.
- Gadet, Françoise & Guerin, Emmanuelle**, 2021, Décrire le 'français tout court' : pourquoi et à quelles conditions, Colloque international Description de l'oral et méthodes d'analyse en linguistique : perspectives actuelles. Hommage à Claire Blanche-Benveniste (1935–2010), Nancy, 20–22 octobre 2021.
- Gadet, Françoise & Jones Mary C.**, 2008, Variation, Contact and Convergence in French Spoken Outside France, *Journal of Language Contact*, **2**, 238–248.
- Gardner-Chloros, Penelope & Secova, Maria**, 2018, Grammatical change in Paris French: *in situ* question words in embedded contexts, *Journal of French Language Studies*, **28**, 181–207.
- Grevisse, Maurice**, 1988, 12e édition, *Le Bon Usage. Grammaire française*, Paris, Ducolot.
- Guerin, Emmanuelle**, 2017, Eléments pour une approche communicationnelle de la variation, in Bilger, Mireille, Cappeau, Paul, Guerin, Emmanuelle & Tyne, Henry (eds), *La variation en question(s)*, Bruxelles, Peter Lang, Coll. GRAMM-R, vol. 36, 57–76.
- Guiraud, Pierre**, 1966, Le système du relatif en français populaire, *Langages*, **3**, 40–48.
- Halliday, Michael**, 1985, *Spoken and Written Language*, Oxford, Oxford University Press.
- Koch, Peter & Oesterreicher, Wulf**, 2001, Langage parlé et langage écrit, in *Lexikon der Romanistischen Linguistik, Tome 1*, Tübingen, Max Niemeyer Verlag, 584–627.
- Kroch, Anthony**, 1978, Toward a theory of social dialect variation, *Language in Society*, volume **7.1**, 17–36.
- Laurier, Michel**, 1989, Le subjonctif dans le parler franco-ontarien : un mode en voie de disparition ?, in Mougéon, Raymond & Béniak, Edmond (eds), *Le français canadien parlé hors Québec : aperçu sociolinguistique*, Québec, Presses de l'Université Laval, 105–126.
- Ledegen, Gudrun**, 2007a, L'interrogative indirecte *in situ* à la Réunion : *elle connaît elle veut quoi*, dans Abécassis, Michael, Ayosso, Laure & Vialleton, Elodie (Eds), *Le français parlé du 21^{ème} siècle : normes et variations géographiques et sociales*, Actes du Colloque à l'Université d'Oxford (23 et 24 juin 2005), Paris, L'Harmattan, 177–200.
- Ledegen, Gudrun**, 2007b, Résonance SMS « Jc c koi mé javé pa réalisé sur le coup ! », *LINX*, **57**, 101–111.
- Ledegen, Gudrun**, 2007c, Morpho-syntaxe du français « ordinaire » de La Réunion : résonance intra- et interlinguistique et éclairage phonétique, *Le français en Afrique*, **22**, 319–330.

- Ledegen, Gudrun**, 2016, Traits syntaxiques ‘populaires’ dans le français ‘régional’ de La Réunion : très populaires dans la francophonie, *Langages*. ‘Le français calédonien (Nouvelle-Calédonie), une ‘variété régionale’ de français au sein de la francophonie’, Pauleau Christine (Dir.), 203, 87–102.
- Ledegen, Gudrun** 2019, Les interrogatives béruriennes : *C’est-y un escandale, oui z’ou chose ?*, ‘Chemins de traverse’, *Les Cahiers Frédéric Dard*, 3, Editions Universitaires de Dijon, 133–146.
- Ledegen, Gudrun & Martin, Philippe**, 2020, L’interrogative indirecte in situ dans le corpus OFROM. ‘Ils posaient la question c’était quoi’, *Studia linguistica Romanica* 4, 175–194, <https://studialinguisticaromanica.org/index.php/slr/article/view/40/48>.
- Lefebvre, Claire & Maisonneuve, Huguette**, 1982, La compétence des adolescents du Centre-Sud : les structures complexes, in Lefebvre, Claire (ed.). *La syntaxe comparée du français standard et populaire : approches formelle et fonctionnelle* 1. Québec. Office de la langue française.
- Lefevre, Florence & Rossi-Gensane, Nathalie**, 2017, Les interrogatives indirectes en discours informel oral, *Langue française*, 196, 51–73.
- Lesacher, Claire**, 2015, *Le rap comme activité (s) sociale (s) : dynamiques discursives et genre à Montréal* (approche sociolinguistique), Thèse de doctorat, Université Rennes 2.
- Manessy, Gabriel**, 1993, Introduction, in Chaudenson, Robert, Mougeon, Raymond & Bėniak, Edmond, *Vers une approche panlectale de la variation du français*, Paris, Didier-Erudition, 3–4.
- Marchessou, Agnes**, 2018, Strasbourg, another setting for sociolinguistic variation in contemporary French, *Journal of French Language Studies*, 28, 265–289.
- Martineau, France**, 1988, Variable Deletion of Que in the Spoken French of Ottawa-Hull, in Montreuil, Jean-Pierre & Birdsong, David (dirs), *Advances in Romance Linguistics*, Dordrecht, Foris, 275–287.
- Martineau, France**, 2022, Non, les jeunes ne font pas plus de fautes qu’avant, in *L’actualité*, 1/6/2022, <https://lactualite.com/societe/non-les-jeunes-ne-font-pas-plus-de-fautes-quavant/>
- Mathieu, Eric**, 2009, Les questions en français : micro et macro-variation, in Martineau, France, Mougeon, Raymond, Nadasdi, Terry & Tremblay, Mireille (eds), *Le français d’ici : études linguistiques et sociolinguistiques de la variation*, Toronto, GREF, 37–66.
- Meizoz, Jérôme**, 2001, *L’Age du roman parlant (1919–1939). Ecrivains, critiques, linguistes et pédagogues en débat*, Genève, Droz.
- Muller, Claude**, 1996, *La subordination en français*, Paris, Colin.
- Nadeau, Jean-Benoît**, 2022, « Non, les jeunes ne font pas plus de fautes qu’avant », ITW de France Martineau, *L’actualité*, 1/6/2022, <https://lactualite.com/societe/non-les-jeunes-ne-font-pas-plus-de-fautes-quavant/>.
- Ostigny, Luc, Champagne, Éric, Gervais, Flore & Lebrun, Monique**, 2005, Le français oral soutenu chez des étudiants québécois en formation pour l’enseignement au secondaire. Collection « Suivi de la situation linguistique ». Etude 4, Bibliothèque nationale du Québec. https://www.oqlf.gouv.qc.ca/ressources/sociolinguistique/2005-2006/etude4_compl.pdf
- Palasis, Katerina**, 2009, *Syntaxe générative et acquisition : le sujet dans le développement du système linguistique du jeune enfant*, Thèse de doctorat, Nice.
- Ploog, Katja**, 2002, L’approche syntaxique des dynamiques langagières : non-standard et variation, *Cahiers de grammaire*, 27, 77–96.
- Ploog, Katja**, 2019, Linguistique du locuteur, linguistique de la complexité, linguistique générale : l’apport du français d’Afrique à la linguistique « tout court », *Langue française*, 202: 27–42.
- Poplack, Shana & Dion, Nathalie**, 2009, Prescription vs. praxis: The evolution of future temporal reference in French, *Language*, 85, 557–587.
- Poplack, Shana & St-Amand, Anne**, 2009, Les Récits du français québécois d’autrefois : Reflet du parler vernaculaire du XIXe siècle, *Revue canadienne de linguistique*, 54–3, 511–546.
- Quillard, Virginie**, 2000, *Interroger en français parlé. Etudes syntaxique, pragmatique et sociolinguistique*, Thèse de Doctorat, Université de Tours.
- Riegel, Martin, Pellat, Jean-Christophe & Rioul, René**, 2021, 8^e édition, *Grammaire méthodique du français*, Paris, Presses Universitaires de France.
- Rossi-Gensane, Nathalie, Acosta Cordoba, Luisa Fernanda, Ursi, Biagio & Lambert, Margot**, 2021, Les structures interrogatives directes partielles fondées sur *où* dans les dialogues de romans français du XXe siècle, *Journal of French Language Studies*, 31, 169–191.
- Sankoff, David & Cedergren, Henrietta J.**, 1976, The dimensionality of grammatical variation, *Language* 52, 163–178.

- Secova, Maria**, 2017, Changement grammatical et discursif en français multiculturel de la région parisienne : éléments de comparaison, *Linguisticæ Investigationes*, 40–2, 173–199.
- Shlonsky, Ur**, 2017, *Wh in situ* and criterial freezing, in Aboh, Enoch, Haerberli, Eric, Puskás, Genoveva & Schönenberger, Manuela (eds), *Elements of Comparative Syntax*, Berlin, De Gruyter, 349–369.
- Tagliamonte, Sali**, 2012, *Variationist Sociolinguistics: Change, Observation, Interpretation*, Chichester, Wiley-Blackwell.
- Thibault, Pierrette & Vincent, Diane**, 1988, La transcription ou la standardisation des productions orales, *LINX*, 18, 19–32.
- Valdman, Albert**, 1979, Créolisation, français populaire et le parler des isolats francophones d'Amérique du Nord, in Valdman, Albert (dir.), Chaudenson, Robert & Manessy, Gabriel (collab.), *Le français hors de France*, Paris, Champion, 181–197.

Corpora :

- CHILDES French Palasis Corpus, 2006–..., sous la direction de Palasis Katerina, <https://childes.talkbank.org/access/French/Palasis.html>
- CORPAIX, Groupe aixois de recherche en syntaxe (GARS), 1999, Corpus d'Aix-en-Provence. Non publié.
- Fonds de données linguistiques du Québec, 2022, Centre de recherche interuniversitaire sur le français en usage au Québec (CRIFUQ), Université de Sherbrooke, <https://fdlq.flish.usherbrooke.ca> (consulted on 8/2022) :

- Chroniques québécoises de langage (ChroQué), Claude Verreault, Louis Mercier & Wim Remysen (dirs) (1865 à 1996): Corpus of language columns published in the Quebec press from 1865 to 1996, dealing with language, more specifically with its good and bad uses. The corpus contains 7.936 entries by some forty columnists, 5.154.788 words.

Citations from the *Corpus ChroQué (Chroniques québécoises)*, dir. by Wim Remysen and Hélène Cajolet-Laganière. Consulted on FDLQ-site in August 2022. [fdlq.usherbrooke.ca] :

- Blanchard, Étienne, Chronique du bon langage, La Presse, Montréal, 1er février 1919, p. 10, col. 1–5
- Lefranc, Paul [pseudonyme de Étienne Blanchard], Autour de la langue française, La Presse, Montréal, 16 juin 1923, p. 46
- Rosemont, Gilles de, Parlons correctement le français, La Patrie, Montréal, 15 août 1954, p. 28
- Barbaud, Philippe, Parler d'ici, La Presse Plus, Montréal, 18 mai 1985, p. 16
- Barbaud, Philippe, Parler d'ici, La Presse Plus, Montréal, 18 mai 1985, p. 16
- Barbaud, Philippe, Parler d'ici, La Presse Plus, Montréal, 11 janvier 1986, p. 16
- Corpus de littérature québécoise (CLIQ), Wim Remysen & Hélène Cajolet-Laganière (dirs): The corpus brings together the main Quebec literary works published between 1794 and 2022: literary texts written in French (not translated) by authors living in Quebec, including Quebecers of foreign origin. The genres are varied (narrative, theatrical, poetic and argumentative) as is the language (chastened, but also popular or spontaneous). The corpus contains 250 documents and 13.481.018 words.

Citations from *Corpus CLIQ (Corpus de Littérature québécoise)*, dir. by Wim Remysen and Hélène Cajolet-Laganière. Consulted on FDLQ-site in August 2022. [fdlq.usherbrooke.ca] :

- Bernard, Christophe (2017), *La bête creuse*, Montréal, Le Quartanier.
- Turcotte, Élise (2016) [1991], *Le bruit des choses vivantes*, Montréal, Bibliothèque québécoise.
- Bienvu, Sophie (2015) [2014], *Chercher Sam*, Montréal, Le Cheval d'août.
- Hamelin, Louis (2012) [2010], *La constellation du Lynx*, Montréal, Boréal.
- Cloutier, Joseph (1925), *L'erreur de Pierre Giroir : roman*, Québec, Imprimerie « Le Soleil ».
- Ébullition : le français québécois dans les bulles, Anna Giaufret, Philippe Rioux et Wim Remysen (dirs): This corpus is a representative selection of Quebecois comics from 1904 to 2022 (serials published in the print media, albums). The comic strip is “a genre that offers a stylised and contextualised representation of spoken language that allows us to study both the practice of the

language and the representations that surround it” (presentation on fdlq.usherbrooke.ca). The corpus contains 15 documents and 72.992 words.

- *JournaLangue2013 : langue et journalisme au Québec*, Franz Meier (dir.) (2011–2013): This corpus consists of 33 semi-structured interviews conducted with 32 professionals working in the Quebec print media from 2011 to 2013, as part of a sociolinguistic study focusing on the conception of language and language norms that exist in this environment. The corpus contains 33 documents and 249.553 words.
- *Corpus sociolinguistique de l’Estrie*, Normand Beauchemin, Pierre Martel & Michel Théoret (dirs) (1971–1974): Sociolinguistic interviews conducted in twenty localities in the Eastern Townships (approximately 60 kilometres around the city of Sherbrooke), a region at the intersection of the dialectal areas of Eastern and Western Quebec. A selection of 100 interviews, 50 of which were conducted with women and 50 with men, aged between 18 and 70, whose mother tongue was French and who were born and raised in their village, just like their parents. The corpus contains 100 documents (for each respondent) and 471.581 words.
- *Corpus variationniste Hochelaga-Maisonneuve 2012*, Hélène Blondeau, France Martineau, Mireille Tremblay & Yves Frenette (dirs) (2012): This corpus contains semi-directed interviews with participants aged 18 to 89, educated in French and living for at least five years in the Hochelaga-Maisonneuve district of Montreal; this corpus, which is part of the FRAN (Français en Amérique du Nord) corpus, allows for more in-depth diachronic comparisons with the uses attested in the Montreal corpora (Montreal 1971, 1984, 1995). The corpus contains 43 documents (for each respondent) and 930.638 words.
- *Montréal 1984*, Pierrette Thibault & Diane Vincent (dirs) (1984): The corpus contains 30 semi-structured interviews with half of the witnesses in the Montreal 1971 corpus (Sankoff & Cedergren 1976) (i.e. 60 people), plus 12 young people, to study language change in real time. The corpus contains 30 documents (for each respondent) and 647.501 words.
- *Montréal 1995*, Diane Vincent, Marty Laforest & Guylaine Martel (dirs) (1995): This is the third and final part of a series of sociolinguistic surveys conducted in Montreal from the 1970s onwards, this corpus brings together interviews conducted with informants already met in 1971 and 1984 (14 people). The corpus contains 14 documents (for each respondent) and 361.366 words.
- *Phonologie du français contemporain: corpus Québec (PFC-Québec)*, Marie-Hélène Côté (dir.) (2009–): Carried out within the framework of the Phonology of Contemporary French (PFC) project (see Durand *et al.* 2002), the PFC-Quebec corpus comprises a series of linguistic surveys conducted across the different regions of Quebec between 2009 and 2017. The data collected are used to illustrate and analyse French as spoken in Quebec at the beginning of the twenty-first century, with a particular focus on the regional, generational and stylistic variation that affects it: each witness is invited to participate in two conversations, one formal and one informal, and to perform three reading tasks (reading a follow-up text and two word lists). The corpus contains 68 documents (for each respondent) and 364.119 words.

Citations from the sociolinguistic oral corpus, dir. by Wim Remysen and Hélène Cajolet-Laganière. Consulted on FDLQ-site in August 2022. [fdlq.usherbrooke.ca] :

- Vincent, Diane, Marty Laforest et Guylaine Martel (1995), « Entrevue Paul G., 2’95 ». Cité dans *Corpus Montréal 1995*. [Montréal1995]
- Thibault, Pierrette et Diane Vincent (1984), « Entrevue Christine Q, 4’84 ». Cité dans *Corpus Montréal 1984*. [Montréal1984]
- Thibault, Pierrette et Diane Vincent (1984), « Entrevue Marcel V., 46’84 ». Cité dans *Corpus Montréal 1984*. [Montréal1984]
- Blondeau, Hélène, France Martineau, Mireille Tremblay et Yves Frenette (2012), « Entrevue Christine, 003F24 ». Cité dans *Corpus Hochelaga-Maisonneuve 2012*. [HoMa2012]
- Blondeau, Hélène, France Martineau, Mireille Tremblay et Yves Frenette (2014), « Entrevue Alex, 015M18 ». Cité dans *Corpus Hochelaga-Maisonneuve 2012*. [HoMa2012]
- Côté, Marie-Hélène (2009–), « Entretien guidé cqcbc1g, Montréal ». Cité dans *Phonologie du français contemporain : corpus Québec (PFC-Québec)*. [PFC-Québec]

OFROM, 2012–2019, Corpus oral de français de Suisse romande, Avanzi, Mathieu, Béguelin, Marie-José & Diémoz, Federica (dirs), (since 2019, Corminboeuf, Gilles & Johnsen, Laure Anne (Dirs)), <http://www.unine.ch/ofrom>.

Parole de chat, Les gros mots 3, https://www.youtube.com/watch?v=Cuhmj_eYt3k

APCOR, 2009–..., Corpus français de chansons de rap, sous la direction de Policka, Alena, Université Marazyk, République Tchèque, <https://www.sketchengine.eu/rapcor-french-rap-corpus/>

VALIRUN, 2000–..., Variétés linguistiques de La Réunion, sous la direction de G. Ledegen, partly on Cocoon (<https://cocoon.huma-num.fr/exist/crdo/>).

Cite this article: Ledegen G (2024). The embedded *Wh-* in situ clause: French tout court? *Les spasmes musculaires incontrôlés, je sais pas c'est quoi* (France Info 20/10/21). *Journal of French Language Studies* 34, 39–71. <https://doi.org/10.1017/S095926952400005X>