Activation levels:
A fresh perspective on French wh in-situ

Lena Baunaz
Université de Genève
lena.baunaz@gmail.com

Caterina Bonan
University of Cambridge
caterina.bonan@outlook.com

Abstract

In the last five decades, French wh in-situ has been the center of much work in theoretical linguistics. Nonetheless, scholars still disagree on the distribution of these constructions, and on their interpretation. While whether or not wh in-situ is necessarily presuppositional has been debated for years (Cheng & Rooryck 2000, Baunaz 2011, Shlonsky 2012, a.o.), we believe this question is too narrow. Here, we investigate the ESLO 1-2 corpora of spoken French and provide a fresh understanding of in-situ questions based on the notion of 'discourse activation' (Dryer 1996, Larrivée 2019a, Garassino 2022). By demonstrating both the passage from a predominantly ex-situ system to a predominantly in-situ system, and a significant augmentation of *non*-context-bound in-situ occurrences, we redefine the conditions under which these structures are licenced in Hexagonal French, and how they have evolved from a micro-diachronic perspective (1970s-2010s).

Keywords: wh in-situ, French, wh-interrogatives, interpretation, activation.

1. Introduction

Linguists have produced numerous syntactic and semantic-pragmatic descriptions of French wh in-situ, but nonetheless still disagree on its distribution and on whether there is a specific contextual condition that licenses the structure in the language.

In this paper, we propose to solve this issue by considering two dimensions related to the diachrony of this construction: (i) the quantitative dimension, which corresponds to the evolution of the in-situ construction in the last century, and (ii) the interpretive dimension, i.e., the evolution of the semantico-pragmatics of the in-situ construction over time. For the first dimension, we will analyze abundant data that we gathered from two existing corpora of Hexagonal French (ESLO 1-2).[1] The data, following Larrivée (2019a) and Garassino's (2022) proposed classifications of in-situ, will be assessed by taking into consideration the context that precedes each occurrence of wh in-situ, both synchronically and micro-diachronically.

In a nutshell, we demonstrate an evolution from a predominantly ex-situ interrogative system to a predominantly in-situ syntax for French, as well as the progressive augmentation of in-situ occurrences in discourse-new clauses, i.e., non-context-bound sentences. We subsequently propose that the controversies found in the literature concerning the interpretation of wh in-situ are motivated by the specific time-frame investigated in each study, and a consequence of language evolution. In other words, different judgments (and thus, different claims) were made in different studies on wh-in situ depending on when the work was conducted. The confusion, we claim, was made possible because the construction evolved importantly over a short span of 40 years.

This paper is organized as follows. In §2, we outline some properties of partial wh-questions in French and introduce the wh in-situ strategy. §3 is an overview of some of the controversies commonly found in the syntactic literature on this construction. §4 presents Larrivée's (2019a) diachronic study of the licensing conditions of the French wh in-situ construction from its emergence until today. §5 outlines and discusses our diachronic study and the reasons behind it, while §6-§7 present and discuss our main results.

2. French in-situ questions – some properties

This paper is concerned with partial interrogation in French. In this type of question-formation strategy, also known as wh-question, the wh-element scopes over a constituent or part of it. The answer to a partial question is open. Hence, to the question in (1), B can answer (virtually) anything but *yes* or *no*:

---

[1]   **ESLO**, Enquête Sociolinguistique à Orléans, was elaborated by the Laboratoire Ligérien de Linguistique of the University of Orléans (France): http://eslo.huma-num.fr/index.php?option=com_content&view=article&id=1.

(1)   A.   **Où**   il   va?
              where   he   go-PRS-3SG
              'Where does he go ?'
       B.   A   Genève. /   Nulle part. /   A la maison.
              to   Geneva /   Nowhere /   Home

Polar questions, on the other hand, require a closed answer of either the 'yes/no' or the dubitative type. An example is given in (2).

(2)   A.   Tu   viens   en   cours?
              you   come   to   class
              'Are you coming to class?'
       B.   Oui / Non / Sans doute / Je sais pas
              Yes / No / Maybe   / I don't know

French displays various strategies for the formation of single, mono-clausal wh-questions (Gadet 1989, Coveney 1996/2002, 2011, Cheng & Rooryck 2000, Mathieu 2009, Elsig 2009, Baunaz 2011, Tailleur 2013, Guryev 2017, Bonan 2019, a.o.); some of these are illustrated in (3a) to (3d):[2]

(3)   a.   **Où**   va-t-il?   ex-situ, inversion
              where   go-PRS-3SG-HE
              'Where does he go*?'*
       b.   **Où**   il   va?   ex-situ, no inversion
              where   he   go-PRS-3SG
       c.   **Où**   est-ce   qu'il   va?   ex-situ, *est-ce que*
              where   *est-ce que*   that=he   go-PRS-3SG
       d.   Il   va   **où**?   in-situ
              he   go-prs-3sg   where

Leaving aside the relative position of the verb w.r.t to the subject, and focussing on the position of the wh-word for now, one immediately sees that most of the strategies in (2) involve a wh-element at the beginning of the clause (a-d). Another strategy consists in leaving the wh-element clause-internally, as in (3d). To understand the difference between the first three questions in (3) and the last one, consider the declarative sentence in (4):

---

[2]   Crucially, all question types in (3) are used as requests for information, i.e. they are information-seeking questions. These signal ignorance about a certain topic, and consequent requests for information. Answer-seeking questions must be distinguished from "special questions" (incl. rhetorical questions) which do not request an answer, and also from echo questions. The latter are requests for confirmation, or repetition, 'or a showing of politeness, or concern, or an expression of surprise or disbelief, or the like,' (Boeckx 1999:76). Echo wh-phrases involve specificity (Starke 2001) and heavy stress (Mathieu 2002), or a 'high + rising echo intonation' (Boeckx 1999: 76, Mathieu 2002, Glasbergen-Plas et al. (2021)). Echo-questions thus display specific pragmatic, semantic and prosodic properties which contrast from those of information-seeking questions.

(4)   Il   va   **à**   **Genève**.
       he   go-PRS-3SG to   Geneva
       'He goes to Geneva.'

In (4), *à Genève* follows the verb directly. To ask a question that bears on the locative of the verb 'to go', *à Genève* can be substituted by the wh-word *où* ('where'). When a wh-word appears at the end of the clause, it occupies (at least apparently) the same position as its declarative counterpart: it is *in-situ*. Conversely, when a wh-word appears in a position different from that of its non-interrogative counterpart, as in (3a-c), it is *ex-situ*.[3]

Typologically speaking, languages form their partial questions according to whether their wh-phases appear ex-situ, as in English and German in (5)-(6), or in-situ, as in Mandarin Chinese, in (7).

(5)   English
       Where does he go?

(6)   German
       Wohin geht   er?
       where goes   he
       'Where's he going?'

(7)   Mandarin (Cheng 1991: 9, (6))
       Hufei   mai-le   **sheme**?
       Hufei   buy-asp   what
       'What did Hufei buy'

Contemporary French thus constitutes a third type in this classification, as it allows both strategies.[4]

According to Tailleur (2013:52), the peaceful coexistence of several question-formation strategies in the variety of French under investigation is a result of its historical evolution, and the fact that 'certain variants are 'specialized' to certain registers, therefore taking a specific function that can be observed pragmatically. All of them have remained in competition because of this difference in usage context.'. Accordingly, scholars vastly agree that the in-situ variant is mainly a spoken phenomenon that is favored in informal registers (cf. *français familier* Valdman 1968; Behnstedt 1973; Söll 1985), or in written registers that mimic spontaneous speech (such as text messages (Guryev & Delafontaine 2015; Guryev 2017 and WhatsApp, as well as emails (Dagnac 2013) and plays (Tailleur 2013; Larrivée 2019a; Zimmermann & Kaiser 2019 a.o)).

---

[3]   See Bonan (2021) and references therein cited for empirical evidence that clause-internal wh-elements are not necessarily in-situ in all languages.

[4]   French is not the only contemporary language with a mixed system. Other Romance languages appear to have both ex-situ and in-situ constructions, i.e., Brazilian and European Portuguese (Cheng & Rooryck 2000, 2002, Kato 2012, a.o.), Spanish (Jiménez 1997, Etxepare & Uribe-Etxebarria 2005, Biezma 2018, a.o.), numerous northern Italian dialects (Munaro 1999, Poletto 2000, Manzini & Savoia 2005, Bonan 2021, a.o.). Larrivée (2016) counts 23 languages attested in WALS (https://wals.info/) that display both ex-situ and in-situ options.

3. The controversies

Studies from the last 20 years have claimed that French wh in-situ constructions are semantically, prosodically and syntactically different from their fronted counterparts (Cheng & Rooryck 2000; Mathieu 1999, 2002, 2004, 2009; Boeckx 1999, 2003; Starke 2001; Adli 2006 and Baunaz 2005, 2011, 2016; Baunaz & Patin 2011, 2012; Déprez, Syrett, & Kawahara 2013; see also Faure & Palasis 2021 for a recent comparison, a.o).

Notably, to date scholars only agree on three properties of wh in-situ. First, as the data in (3) show, it does not involve subject inversion; second, there is no *est-ce que* reinforcement. Third, it is widely acknowledged that these structures are normally excluded from embedded environments in Hexagonal French (Mathieu 2002, 2004; Baunaz 2011; Shlonsky 2012, a.o), as illustrated in the examples in (8).

(8) a. Paul   se   demande   qui_i   elle   a
         Paul   REF   wonder-3SG   whom   she   has-3SG
         invité __i.
         invited
     b. *Paul   se   demande   elle   a   invité qui.
         Paul   REF   wonder-3sg   she   has-3SG   invited who
         'Paul wonders who she invited.'

If the distribution of moved wh-elements has been described in (quite) homogeneous and consistent ways for contemporary French (Mathieu 2009, Faure & Palasis (2021) and references therein cited), the same does not apply to in-situ wh-elements, of whose distribution linguists have given confusing descriptions. In the next two subsections, we provide a non-exhaustive overview of the disagreements on the syntax and semantics of French wh in-situ found in the literature.[5;6]

3.1. The syntax of wh in-situ

Linguists agree on the fact that the distribution of wh in-situ vs. wh ex-situ is different. They do not agree, however, as to how much in-situ wh-elements are syntactically restricted.

Some authors claim that in-situ wh-elements are restricted to root clauses and cannot be moved out of infinitival CP-complements (Chang 1997, Boeckx 1999, Cheng & Rooryck 2000, Boskovic 2000, Mathieu 1999, 2002, although Mathieu 2002 describes them as grammatical with modals); they are trapped in negative islands (Chang 1997; Bošković 1998; Mathieu 1999, 2002, 2004, 2009; Cheng & Rooryck 2000; Zubizarreta 2003; Shlonsky 2012) and scope islands (mainly universal quantifiers but not only, see Chang 1997; Cheng & Rooryck 2000; Etxepare & Uribe-Etxebarria 2005, but see Mathieu 2002, for whom wh in-situ take narrow scope in scope islands).

---

[5]   Here, we only raise the main disagreement points on the topic. See Zimmermann & Kaiser (2019) for a recent, more detailed state of the art.
[6]   We do not review here the literature on the prosody of wh in-situ. On this, see Hamlaoui (2009); Baunaz (2011); Baunaz & Patin (2011, 2012); Déprez et al. (2013); see also Glasbergen-Plas et al. (2021) for a recent study, and references.

For other scholars, in-situ wh-elements in spoken French are actually productively used (i) in embedded clauses (Starke 2001; Baunaz 2005, 2011, 2016; Etxepare & Uribe-Etxebarria 2005; Oiry 2011, Dagnac 2013, Adli 2006; Shlonsky 2012, a.o.)[7], as in (9); (ii) with modals (Starke 2001; Baunaz 2005, 2011, 2016; Etxepare & Uribe-Etxebarria 2005; Dagnac 2013, and Adli 2006)), as in (10); (iii) in negative islands (Starke 2001, Baunaz 2005, 2011, 2016; Etxepare & Uribe-Etxebarria 2005; Adli 2006 a.o), as in (11); (iv) in scope islands (Zubizaretta 2003; Adli 2006, Baunaz 2005, 2011, a.o), as in (12) and (13).

(9) a.   Tu   penses qu'elle   invite qui   à   sa   fête?
            you   think that=she   invites whom   at   her   party
            'Whom do you think she'll invite at her party?'

     b.   Tu   penses qu'il   vient quand/comment/où?
            you   think that=he   comes when/how/where
            'When/How/Where do you think he'll come?'

     c.   Tu   as   décidé   de   venir quand?
            You   have   decided   to   come   when
            'When did you decide to come?'

(10) a.   Il   peut   rencontrer   qui?
            he   can   meet   who
            'Who can he meet?'   (Adli 2006: 16, (13))

     b.   Il   peut/doit   aller   où?
            he   can/must   go   where
            'Where can/must he go?'   (Adli 2006: 16, (14))

(11) a.   Elle   a   pas   mangé quoi?
            she   had   not   eaten what
            'What didn't she eat?'

     b.   Il   (ne)   peut/doit   pas   aller où?
            he   (NE)   can/must   not   go   where
            'Where can't/mustn't he go?'   (Adli 2006: 14, (9))

(12) a.   Plusieurs   personnes   ont   reconnu   qui?
            several   persons   have   recognized who
            'Who did several people recognise?'   (Adli 2006:16, (15))

     b.   Plusieurs chênes ont été coupés où/quand?
            several oaks have been cut where/when
            'Where/when were several oaks cut down?' (Adli 2006: 6, (16))

(13) [?]Tu   passes toujours   par   quel   chemin   quand
         you   go   always   by   which way   when

---

[7]   Oiry (2011) shows experimentally that both fronted wh-elements and in-situ wh-elements can appear in non-root questions. In a corpus study from literary prose texts (see fn.14), Zimmerman & Kaiser (2019) also find in-situ wh-elements in embedded clauses. Dagnac (2013) also cites an example from the PFC spoken corpus (https://www.projet-pfc.net/).

tu   rentres?
you   go.home
'Which way do you always take when you go back home?'

Concerning the distribution of wh in-situ, there exist two main groups of descriptions: the more conservative one (Chang 1997, Chang & Rooryck 2000 etc), and the more permissive one (Starke 2001, Adli 2006, Baunaz 2005, 2011 etc), with disagreements within each groups (in fact mainly in the first group). Interestingly, for the more permissive group, the syntax of wh in-situ is very similar to that of wh ex-situ (although they agree on the fact that wh in-situ constructions do not allow subject-verb inversion, *est-ce que* reinforcement and embedded indirect questions).

3.2. The semantics/pragmatics of wh in-situ

Linguists disagree on what the interpretational property of wh in-situ is, how this should be defined and whether it is categorical or not. For a property to be categorical, it must be necessarily present, as well as different from that observed in a syntactic counterpart of the structure under investigation, such as ex-situ in this case.

3.2.1. The conservative view

Starting with Chang (1997), wh in-situ has been described as involving a strong presupposed context (see also Boeckx 1999, 2003, Cheng & Rooryck 2000, Mathieu 2002; Obenauer 1994; Coverney 1996/2002; Zubizaretta 2003; Boucher 2010 a.o). The general idea is that the interpretation of wh in-situ questions elicits 'details on an already established (or presupposed) situation' (Chang 1997:45). Thus, for these authors, an in-situ question like *tu vas lui acheter quoi* 'What will you buy for him?' can only be uttered in a context where the event of buying something is presupposed by the speaker (see Chang 1997, Cheng & Rooryck 2000), like, for instance, in the context of Marie's birthday in (14)

(14)   A: C'est   l'anniversaire de   Pierre la   semaine prochaine.
              ce=is   the-birthday   of   Pierre the   week   next
              'It's Pierre's birthday next week.'
       B.   Et   tu   vas   lui   acheter   quoi?
              and   you   will   to.him buy   what
              'And what will you buy for him?' (Cheng & Rooryck 2000: 5, fn.(i))

One (crucial) property of presuppositions is that they are not cancellable. It has been proposed that a way to discriminate between presupposed and non-presupposed contexts is to test whether the answer to a question can be cancellable or not, i.e. using a negative answer. If a negative answer is possible, it means no presupposition is involved. In other words, answering *rien* ('nothing') to (14) should be impossible if presupposition is indeed involved in all wh in-situ constructions; conversely, replying *rien* should be perfectly fine in case of wh-fronting. And this is indeed what these linguists claim (see Coveney 1989, 1995; Chang 1997; Boeckx 1999, 2003; Cheng & Rooryck 2000; Zubizaretta 2003, a.o). The dialogue in (15) illustrates this idea, while (16) illustrates that *rien* is actually a possible answer for ex-situ questions.

(15)   Question: Marie a   acheté   quoi?
                     Marie has   bought   what
                     'What did Marie buy?'
        Answer: ??Rien.
                     'Nothing.'   (Cheng & Rooryck 2000: 4, (8))

(16)   Question: Qu'est-ce que   Marie   a   acheté?
                     What=*est-ce-que* Marie   has   bought
                     'What did Marie buy?'
        Answer: Rien.
                     'Nothing.'   (Cheng & Rooryck 2000: 4, (7))

An issue about the 'strong presupposition' hypothesis is that there is no clear definition in the literature of what a *strong* presupposition is supposed to be.[8] In that respect, authors differ on what they considered is presupposed, i.e. is it the event as a whole that is presupposed (as in the discussion around (14)), in which case, all questions are presupposed after all,[9] or is it the presupposed set (of answers) only (involving *existential* presupposition)?[10]

All in all, for the conservatives, there is a specific context in which these questions are licensed, and which distinguishes it from wh-fronting.

3.2.2. *The liberal view*

Liberals do not agree that there is a specific discursive context in which in-situ constructions are licensed. These authors have indeed shown by means of experiments (Oiry 2011[11]), corpus studies (Zimmerman & Kaiser 2019), and judgements tasks

---

[8]   The reader will refer to Mathieu (2002) and Baunaz (2011) for a thorough review of the old semantic/pragmatic literature on in-situ constructions.

[9]   All partial questions generate a presupposition (see Boucher 2010; Larrivée 2019a, a.o). The ex-situ question in (i) indeed involves the presupposition that the addressee is going somewhere for the holidays.

(i)   Où   vas-tu pour   les   vacances?
       where   go=you for   the   holidays
       'Where do you go for the holidays'

It is thus unclear what exactly 'presupposition' means when it comes to wh in-situ.

[10]   Aware of these difficulties, some linguists have decided to abandon the 'presupposition' account and identify the context in which in-situ appears. Tieu (2012) proposes that these questions are used in contexts of Verum Focus (i.e. they involve both Verum focus and questions). Other proposals favor accounts of the wh in-situ variant as being conditioned by prosody and information structure. In particular, Hamlaoui (2009, 2011) argues that in in-situ constructions, the wh-element is highly prominent and conveys new information (it is focussed), while the remainder of the question is given. A similar account has been put forward by Déprez et al. (2013).

[11]   In her experiment, Oiry (2011) shows that fronted wh-elements and in-situ wh-elements are perfectly grammatical in non-root questions, with non-presuppositional contexts being more frequent.

(Mathieu 2004; Adli 2006; Baunaz 2005, 2011, 2016; Baunaz & Patin 2011, 2012; Oiry 2011; Shlonsky 2012; Garassino 2022, a.o.) that all presupposition tests fail.

Today, the consensus among liberals is that answering with a negation to a wh in-situ question is perfectly acceptable. If wh in-situ constructions appear in both presupposed and non-presupposed contexts, then presupposition is not a categorical feature of the construction (Tailleur 2013, Elsig 2009 for Laurentian French[12]).[13] Interestingly, Zimmermann & Kaiser (2019) make a similar observation for the "colloquial spontaneous language spoken in Metropolitan France", in their corpus study from literary prose texts.[14] This fact suggests that in-situ wh-elements trigger similar answers in similar contexts as ex-situ wh-elements (i.e., "in a standard Hamblin framework for questions the set of possible answers is the same, [and] the semantics seems to be the same in that respect", Walner 2018). There is thus a consensus concerning the non-categorical status of presupposition with wh in-situ.

In addition, among those linguists who accept wh-elements within islands (negative islands, scope islands and/or adjunct islands), there is also a consensus today that the presuppositional status of these wh-elements involve an existential presupposition. How this is defined (does it involve *D-linking* as in Pesetsky 1987 or rather familiarity as in Enç 1991, or else?) and whether it should be subdivided into two types (specificity and partitivity as in Starke 2001; Baunaz 2005, 2011, 2016; Baunaz & Patin 2011, 2012) is still under debate.

All in all, for the liberals, there is no specific context in which these questions are licensed, and which distinguishes them from wh-fronting.

3.3. Why these descriptive differences for the wh in-situ strategy?

The confusing picture provided in the wh in-situ literature deserves an explanation. In what follows, we discuss some (potential) reasons for this.

3.3.1. The diglossic approach

One possible explanation for the contrastive descriptions of wh in-situ relies on the heavy influence of the norm on grammaticality judgements on spoken language (see also Adli 2015[15]). Bonan (2019: 292) claimed that while it is quite easy to get data from native informants of any language, the normative pressure negatively influences the genuineness of the grammaticality judgements delivered by French speakers. Accordingly, "one should always be careful while establishing what is grammatical and what is not in this language.".

---

[12]   The variety of French spoken in Québec.

[13]   Tailleur (2013) notes that present-day French speakers from France are keener to use the wh in-situ constructions than the Laurentian French speakers.

[14]   Zimmermann & Kaiser (2019) conducted a corpus study in which they extracted in-situ wh-elements from a corpus of forty-five novels and one collection of short-stories published in 1995-2018, all written by native Metropolitan France authors. This literary prose is argued to reproduce colloquial spontaneous language spoken in Metropolitan France (Zimmermann 2018 for details on the corpus).

[15]   Adli (2015) tries to understand how the different partial interrogative strategies are accepted by French speakers. However, he does not mention the variability in judgements under discussion here.

Since wh in-situ is a spoken phenomenon, we argue that it is capital to put aside standard French and focus on the contemporary spoken variety when dealing with these structures: French has (at least) two grammars, and this property needs to be taken into account both while trying to establish an empirical description of wh in-situ in this language and when trying to fit the relevant data into a solid theory (Bonan 2019: 292-293).[16]

This view is not new in French linguistics, as it was formalized in an approach called the *Diglossic* approach (Massot 2010, Zribi-Hertz 2011 for Hexagonal French[17], see also Tailleur 2013 about the Laurentian French interrogative system and, more recently, Faure & Palasis 2021 in the context of 'Metropolitan' French interrogatives)). The basic idea behind diglossia is that speakers have access to different registers, which are different enough to be considered as different grammars. One grammar that is accessible is the 'normative' stage of the language (i.e. Ferguson's 1959 *highly codified* variety, or Crystal's 1991 'High' language). This is codified, acquired passively through media and actively in school, mainly written or used in very formal situations (etc). Conversely, the other variety is acquired naturally, not used in school education, mainly used in spoken contexts (the low varieties of Crystal's 1991).

Under the diglossic approach, speakers have at least two grammars in their brains, which explains the variation. Massot 2010 speaks of the non-'étanchéité' of the grammars, which are rather in competition. Thus judgements may be unstable as a result of grammar competition (see Massot 2010 about grammar in competitions).[18]

---

[16]   In a similar vein, Baunaz (2011, ch. 2, p. 33-56) distinguishes between two varieties of European French: *Standard Colloquial* (SC) and *Non Standard Colloquial French* (NSC), building on previous works such as, non-exaustively, Baunaz (2005), Adli (2006), Starke (2001). What Baunaz shows is that in NSC, wh in-situ is a very productive question-formation strategy, often felicitous in contexts where the SC variety excludes it categorically. The contexts where wh in-situ is felicitous in NSC are (minimally) the following: (i) long distance finite and non-finite questions; (ii) in the scope of negation; (iii) in the scope of modals; (iv) construed with quantifiers; (v) construed with adverbs. The expressions *Standard Colloquial French* and *Non Standard Colloquial French* to refer to speakers using or not using the specific questions in French are, admittedly, too vague and probably not representative enough. Yet, it suggests a bi-partition existing in European French, making a distinction between speakers using the in-situ construction and those not-using it. Actually Bonan (2019) argues that this bi-partition might not be enough to represent the usage of this construction: she claims that a further subdivision into regional and socio-linguistic varieties might be necessary too. Nonetheless, the author insists that "this first division constitutes a first, capital step towards a correct description and analysis of French wh in-situ."

[17]   See also Rowlett's (2007) distinction between Modern French and Contemporary French. See also Rowlett (2011).

[18]   In addition to the diglossic approach, another approach to linguistics variation exists, i.e., the variationist approach (Labov 1996; Beeching et al.2009; Blanche-Benveniste 1997; Gadet 2007). Contrarily to the diglossic approach, for which two different grammars coexist in the brain, there is only one (unique) grammar for variationists. The hypothesis put forward is that the various variants of French depend on various social and stylistic factors. We refer to Coveney 1996; Quillard 2000; Elsig 2009; Adli 2006, see also Guryev 2017 for discussion.

3.3.2. Diatopic variation

Another possibility, related to the diglossic approach just discussed is geographical variation: different low varieties are likely to be found in different regions (see for instance Tailleur 2013 for the differences between the vernacular Laurentian French vs. contemporary French). Classifying the in-situ data into vernacular varieties is actually what people do more and more in the field of Generative Grammar (see for instance Hamlaoui 2011; Mathieu 2004 and more recently Tailleur 2013, Faure & Palasis 2021).

It follows that the subdivision into groups presented in sections 3.1-3.2 looks a bit artificial: Baunaz (2016) notes that the type of French described by Chang 1997, Boeckx 1999, Cheng & Rooryck 2000, Bošković, 2000, Mathieu 2002 (i.e., SC) is not uniform, and as such, should not be considered as merely one group, nor as the standard. There are indeed variations within this variant – and crucially concerning the availability of in-situ wh-elements in root infinitives, or with modals. Accordingly, "Judgments described by Starke 2001, Adli 2006 and Baunaz 2005, 2011 also show variations, yet, these variations are more subtle (context-dependent, e.g., with negative and scope islands)." (Baunaz 2016:132, fn 6). Importantly, the majority of Starke's and Baunaz's informants are from the same region (Geneva, Switzerland). Hence, all of these studies describing colloquial French differently from other authors are actually describing a peculiar diatopic variant of the language (Genevan French, Laurentian French, Metropolitan French etc.).

3.3.3. Language evolution

Another possible explanation for the non-harmonious descriptions of the wh in-situ phenomenon relates to language evolution. Wh in-situ is indeed often described as a relatively young phenomenon in the French language, and while linguists agree that the wh in-situ strategy is the most recent of the question-formation strategies introduced in (2) for French, they disagree on when it appeared in the language.

Wh in-situ being a colloquial 'spoken' phenomenon, one expects to find it in writings that mimic spoken interactions such as theatrical plays). Larrivée (2019a,b) claims that the in-situ construction is not attested in theater plays in the Frantext corpus before the 15th century.[19] Additionally, he only finds three total occurrences of wh in-situ for this century. Dekhissi (2013) claims that the wh in-situ strategy can be found from at least the 18th century, just like Mathieu (2009) and Canel (2012). In her study of the history of the French wh-interrogative system, Tailleur (2013:30) does not find wh in-situ constructions before the Modern French period (19th century) (see also Tuaillon 1975, Coveney 2011 and Guryev 2017).[20]

The results of Larrivée's (2019) Frantext search of wh in-situ in theatrical texts for the 18th and 19th centuries are given in Table 2.

---

[19]   The database Frantext (https://www.frantext.fr/) contains samples of French texts from different periods (9th-21st c.) and of different written genres (novels, theatrical plays, diaries, linguistic texts, recipes etc).
[20]   Tailleur's (2013) written corpus is composed of historical grammars and writings representing the spoken languages, i.e., theater plays.

Table 1. *In-situ* between 1700 et 1900 in Frantext (Larrivée 2019a: 123, Table 2)

	18th c	19th c	Total *in-situ*	Total "wh-?"
Qui?'who'	1	2	3	1409
Comment?'how'	0	1	1	1411
Combien 'how much'	0	3	3	112
Totals	1	6	7	2932[21]

It is important to note that the author exclusively looked for instances of *qui* ('who'), *comment* 'how' and *combien* ('how much') immediately followed by a question mark, therefore excluding all possible occurrences of non sentence-final wh in-situ.[22]

When it comes to spoken corpora, work has been done for both European and Quebec French. In his overview of the literature, Guryev (2017) observed the distribution of partial questions in 5 different corpora of spoken (European) French covering the period 1965-2005. He summarizes his findings in a table whose data we report below (where 'S' stands for *subject*, 'V' for *verb*, 'Q' for *question word* and 'ES' for *est-ce que*):

---

[21]   This number has been recalculated. In Larrivée's table, 1409 is indicated as total, which we reckon must be a mistake.

[22]   For a recent analysis of the emergence of the in-situ construction as being promoted by two discourse patterns ('routines discursive'), see Guryev & Larrivée 2021. Note that the period investigated is between 1100 and 1840. The data come from the Frantext database. After data-cleansing, they found 32 relevant in-situ constructions, including 26 "questions retardées" (half of them from the XVIIth c.). The interested reader is referred to this work for details.

Table 2. Constituent questions in five different corpora of spoken French covering the period 1965-2005 (adapted from Guryev 2017: 116-117, figure 3)[23]

Study	SVQ	QSV	QV-Scl	QV NP	QESV	Q=S V	Other	N=
1. Pohl (1965)	2%	3%	28%	0.5%	66%	-	–	184
2. Behnstedt (1973)	33%	46%	3%	2%	12%	-	4%	446
3. Coveney (1996)	15.6%	23.8%	6.6%	2.5%	48.4%	3.3%	–	118
4. Quillard (2000: 96)	41.6%[24]	16.2%	5.2%	9.7%	22.4%	2.7%	2.2%	670
5. Adli (2015)	57.6%[25]	15.2%	3.7%	6.4%	16.7%	?	0.4%	1680

The numbers in table 2 are difficult to handle as they are virtually impossible to compare: each corpus contains data from a different region (Belgian for Pohl, Paris for Adli, French Picardie for Coveney), different backgrounds (only middle class for Behnstedt 1973, speakers from different backgrounds for Quillard 2000), non uniform age ranges (80-70 for Pohl 1965; 17-37 years old for Coveney; 19-49 for Adli 2015) etc. Also, the number of speakers varied drastically (from 2 in Pohl 1965, to 30 for Coveney 1996, and 101 in Adli 2015). Additionally, the way the data is classified varies across authors: Behnstedt (1973) includes *c'est Q* and *c'est Q que* in what Guryev calls QESV, while Adli treats these as separate categories (but see fn 25), and Coveney (1996) does not count 'c'est Q?' or 'C'est Q que…', because these are absent from his corpus. Also studies may involve different methodologies: there are methodological variations as to how data were collected. For instance, Quillard (2000) uses a corpus of spoken French with recordings of different types of interactions, while Adli (2015) is based on recordings of interviewer and interviewees in one very specific setting designed by the author.

---

[23]   Pohl's (1965) study was based on the variety of French spoken by a Belgian elderly couple, while Coveney (1996) discusses recordings of spoken French from the 1980s, Quillard (2000: 96) French Spoken interactions of various types from the Corpus du Français, and Adli (2015) spoken French recorded in 2015. Behnstedt's (1973) study, cited by Coveney (1996), covers middle class spoken French.

[24]   Guryev (2017) includes structures like *c'est wh* (*C'est quoi ce truc?*) and of wh in-situ clefts like *C'est quoi que tu fais?*, which Quillard treats separately.

[25]   Guryev (2017) includes regular wh in-situ (*tu vois qui devant la fenêtre?* (Adli 2015: 178, (1b)) and in-situ clefts (*c'est quand que tu l'as fait le dessin?*), as well as non-final wh in-situ (*tu le fait quand le dessin?*), which are counted separately in Adli (2015: 181). Wh in-situ is the most frequent interrogative strategy in his corpus (56.2%), while in-situ clefts constitute only 1.0%.

One can nonetheless observe a trend whereby the most recent corpora (4 and 5 in Table 2) show a high rate of in-situ questions compared to all other question-formation strategies. Additionally, Quillard (2001) notes a correlation between age and the usage of wh in-situ: this structure is mainly used, in her corpus, by speakers under 35 years old.

For Quebec French, it has been shown that the evolution of wh in-situ structures has gone rather fast since the 2000. Tailleur (2022) highlights this evolution by comparing three different corpora of spoken Quebec French from closely related regions (Ottawa, Sherbrook, Montreal) within three different periods: early 80s, 1971-1974 and 2012-2014. For this, Tailleur used data from the Ottawa-Hull French corpus (Poplack 1989), the *Corpus sociolinguistique de l'Estrie* (Beauchemin, Martel & Théoret 1971-1974) and the Corpus FRAN (Martineau & Séguin 2016), respectively. The author's claim is that the wh in-situ strategy shows a generalized growth over the years.[26]

In the next section, we come back to the debate between linguists arguing in favor of wh-in-situ being always presupposed and those arguing that the presupposition criterion is not categorical. Following an idea initially advocated for in Larrivée (2019a), we shall claim that the presuppositional status of wh in-situ was categorical when it emerged in the language, but became lost over time.

4. The wh in-situ construction across time

In §3.2, we saw that linguists disagree on whether there is a categorical feature that licenses in-situ constructions in French. Larrivée (2019a) addressed the issue from the point of view of historical pragmatics[27], for the period that goes from the 15th c. until today. We present and discuss his findings in §4.1-2.

4.1. On discourse activation and lack thereof (Larrivée 2016, 2019a)

To identify the pragmatics of wh in-situ in French, Larrivée investigates the relationship of these constructions with the preceding discourse. To do this, he relies

---

[26]   We refer the reader to the published version of this talk, Tailleur (in prep.) for the exact figures.

[27]   Meillet (1912) claims that new grammatical expressions appear to convey specific pragmatic values.

> 'Languages undergo [...] a sort of spiral development: they add extra words to obtain an intensified expression; those words weaken, wear out and are reduced to the level of simple grammatical tools; new or different words are added for expressive purposes; the weakening process begins anew, and so on without end.' (translation by Mosegaard Hansen 2013: 52)

Larrivée proposes to elucidate i) how the "intensity" of new grammatical variables is linguistically instantiated, ii) why there exists a pragmatic value for new grammatical variables and iii) at which moment in time this value is lost. What he discovers is that there is an activation condition to the emergence of these new grammatical structures, according to which "rare emerging grammatical variables representing less than 1% of uses in agrammatical category are characterized by a pragmatic value of explicit activation" (Larrivée 2019a: 127).

on the notion of *Discourse Activation* introduced by Dryer (1996). Accordingly, the author distinguishes two levels of activation: *explicit activation* (henceforth EACT) and *non-activation* (N-ACT).

To be explicitly activated means "that the propositional content of the question has already been mentioned in the discourse" (Garassino 2022: X), i.e., it is "discourse-old information explicitly primed by antecedent context" (Larrivée 2019a: 118). Conversely, in the context of *non-activation* the question literally introduces a new topic into the discourse. These notions are reported in Table 3.

Table 3. Discourse activation levels as in Larrivée (2019a)

Activation level	**Propositional content**
Explicitly activated (**EACT**)	Discourse old
Non-activated (**N-ACT**)	Discourse New

Larrivée utilizes the examples in (17) and (18) to illustrate the contextual difference between explicitly activated and non-activated wh in-situ.

(17)   **EACT (discourse old)**

OW26   dans   les   jeux   antiques   euh   *ils   se*
              in   the   games old   *euh* they   refl
              dopaient   quand même   avec   des   méthodes   un
              dope.past   even   with   part   methods   a
              peu   bizarres   mais [...]
              little weird   but
              'In the old (olympic) games they'd dope themselves with weird methods…'

ch_PP6   *ils se dopaient* comment ?
              they   refl   dope.past   how
              'How did they dope themselves?'

OW26   ils   prenaient   euh   des   plantes
              They   took   *euh*   part   plants
              'They took plants.'
                              (ESLO2_ENT_1026, Larrivée 2019a: 120, (13))

(18)   **N-ACT (discourse new)**

finalement   tu   trouves **comment**   la   vie   à Orléans*?*
So   you   find   how   the   life   in Orleans
'So, how do you find life in Orléans?'
                              (ESLO2_ENT_1022, Larrivée 2019a: 120, (14))

In (17), what is old information (underlined) has already been explicitly activated in the discourse, while no such contextual anchor is present in (18). It is especially important to understand that in (17), no echo-question effect is triggered (no confirmation-request, no repetition-request, no expression of surprise and the like),

and speaker ch_PP6 is genuinely asking for information. Conversely, in (18) the speaker introduces a new discourse topic, which is not activated by the preceding discourse. This type of sentence appears either at the very beginning of a conversation (see Adli 2006: 184, see also Garassino 2022), or is used by the speaker to express a change of discourse topic.

4.2. The emergence of new grammatical constructions (Larrivée 2016, 2019a)

Larrivée (2019a) claims that from the 15th century to the first half of the 19th century, the rare occurrences of in-situ needed a particular discursive context to license their presence, that of explicit activation. In contemporary French, this is no longer the case, to the effect that the use of wh in-situ is more diversified: wh in-situ is used both in contexts of explicit activation and of non-activation.

To reach this conclusion, Larrivée conducted two corpus studies, one based on (non-)vernacular literary texts and theatrical plays spanning from the 15th c. to the early 20th c. (from a corpus of his own, and Frantext Search, Larrivée 2019b for details) and another study based on two corpora of spoken French: ESLO1 and ESLO2. [28]

Following his spoken French corpus study, Larrivée claims that i) overall, the wh in-situ strategy is under-represented (0.91% in ESLO1[29,30], from which 30.8% are explicitly activated[31] and 46.2% are non-activated[32]; 6.8% in ESLO2[33], from which 2.8% are explicitly activated[34] and 85.7% are non-activated[35]) ; ii) explicit activation was much more in ESLO 1 (30.8%) than in ESLO 2 (2.8%); iii) the rate of new information in-situ is 10 times greater in ESLO 2 than in ESLO 1. Larrivée thus concludes that explicit activation is a categorical feature of rare, emerging wh in-situ (less than 1%), which gets lost once the rate for the structure gets higher than 1%.

Larrivée's work makes interesting conclusions, and supports one of our hypotheses on the confusion around the interpretation of wh in-situ found in the literature. In particular, his conclusion suggests that the wh in-situ strategy encoded the categorical feature of discourse oldness until the end of the 18th c., which subsequently got lost in the 20th c. once the construction became more frequent, opening the structure to the increased interpretational flexibility that we observe in this study.

---

[28]   See fn.1. for details.
[29]   13 out of 1434 occurrences.
[30]   Larrivée notes que "the percentages of *wh in-situ* are indicative; the total number of *comment* refers to all the tokens and include uses other than interrogative ones." (Larrivée 2019a: 120, fn.7).
[31]   4 out of 13 wh in-situ.
[32]   6 out of the 13 wh in-situ.
[33]   70 out of 1113 occurrences.
[34]   2 out of 70 wh in-situ.
[35]   60 out of the 70 in-situ.

5. The recent evolution of wh in-situ

Despite the interesting conclusions, Larrivée's work was mainly pioneering. We wanted to test and expand Larrivée's study to several other bare wh-words. The author's spoken corpus study is indeed limited to only one wh-word, namely *comment* 'how'.[36] Given that the historical literature has demonstrated that wh-elements evolve at different paces (Tailleur 2013 and references therein cited) and that today's *comment* 'how' seems to be an almost exclusively fronted wh-word (see Guryev 2017 for references), it seems interesting to reduplicate the study and determine whether other wh-words have evolved similarly. Another reason that motivated our decision to reduplicate his study is the fact that Larrivée's dataset included non-interrogative uses of *comment* 'how' (see fn. 30), making it difficult to sort out what the specific syntactic properties of interrogatives are. Finally, the numbers at the author's disposal suggest that he only investigated part of the whole corpus, which could (but does not necessarily) constitute a sampling error. We thus decided to conduct a study that expands Larrivée's work to all non-lexically restricted wh-words that alternate between the ex-situ and the in-situ positions, and investigates the entirety of the ESLO 1-2 corpus.

       The aim of our study was twofold: to solve the controversy on the interpretive conditions that successfully license wh in-situ, and to propose a formalization of the different types of wh in-situ which we base on the notion of 'activation level'.

       Accordingly, our research questions are the following:

    A. To what extent does the overall proportion between wh in-situ and wh ex-situ vary over the chosen time frame, and to what extent does the overall proportion of <u>each</u> in-situ wh-word vary over time (Q1)?

    B. To what extent does the proportion of activation levels vary over time globally, and to what extent does the proportion of activation levels vary over time for each of the considered wh-words (Q2)?

To answer the questions in (A)-(B), we conducted a micro-diachronic corpus study spanning over the 1970s-2010s period that expands and implements Larrivée's (2019a) work. We introduce and discuss the study in §5.1.

5.1. The chosen corpus of spoken French

Just like the most recent part of Larrivée's work on contemporary French (20th-21st c.), our study is based on the ESLO corpus, 'Enquêtes Sociolinguistiques à Orléans'. ESLO is a corpus of spoken French composed of two sub-corpora, ESLO 1 and ESLO 2, which document how French was spoken in Orléans, France, in the 1970s (1969-1974) and 2010s, respectively. It was created by the University of Orléans and the CNRS, and supported by the French Ministry of Culture, and 'Centre of France regional government'. The recordings are transcribed and the two corpora, which are

---

[36]   Larrivée is aware that only focussing on *comment* 'how' and not to look at other wh-words could be seen as a shortcoming. Yet he claims that "[t]he reasoning is that either the relation between the item and a pragmatic value is consistent, to support the claims in the current literature, or it is not, supporting the counterclaims, inviting for further investigation in either case". (Larrivée 2019a: 118, fn.3)

collections of spoken interactions on different topics recorded in different situations,[37] are freely accessible.

While Larrivée (2019a) based his study on one type of interaction, i.e., semi-directed interviews, we used the whole of the recordings to reduce sample biases to a minimum.[38;39]

5.2. Data collection[40]

Because not all wh-elements alternate freely between the ex-situ and the in-situ position in French, we focused our work exclusively on those non-lexically restricted wh-words that are able to surface either ex-situ or in-situ. These are:

a. *comment* ('how')[41];
b. *quand* ('when');
c. *où* ('where');
d. *qui*-object ('who');
e. *quoi*-indirect object ('what').

Indeed, it is widely acknowledged that French *que* ('what') and quiS ('whoS') can only be fronted, as illustrated in (19) and (20), while *quoi* ('what') can only surface ex-situ if part of a PP, as in (21).

(19)   a.   Que   veux-tu?
               Que   want=you
       b.   *Tu   veux   que?
               You   want   *que*
               'What do you want?'

(20)   a.   Qui   as   réveillé   le   bébé?
               quiS has   woken.up   the   baby
       b.   *(Il)   a   réveillé   le   bébé   qui?
               it   has   woken.up   the   baby   *whoS*
               'Who woke up the baby?'

---

[37]   The recorded interactions include, non exclusively: questionnaire-based interviews, unplanned interactions in different contexts (markets, shops, etc.), phone calls, planned interviews with public personalities and researchers, conferences/debates, cinema dialogues, school interactions, children's books, etc.
[38]   For a socio-linguistic study of simple direct partial questions in ESLO1 and ESLO2, see Thiberge *et al*. 2021.
[39]   Further studies will be needed to confirm that our sampling was indeed unbiased, including data from other French towns but also from other countries where French is spoken, i.e., Switzerland or Quebec. For the time being, we aim at producing the preliminary results that will constitute the basis of our future work on the topic.
[40]   The repository dedicated to this study is publicly available at the following link: https://github.com/CaterinaBi/interrogatives-corpus-work.
[41]   *Combien* was only kept in its bare form, meaning 'how', while all partitive usages (*combien de*+N, 'how many N') were discarded.

(21) a.   À quoi   dois-tu   ton   succès?
              QuoiIO   owe=you   your   success
              'What do you owe your success to?'
      b.   ??Quoi   veux-tu   manger?
              QuoiDO   want=you   eat
              'What do you wish to eat?'

     Moreover, lexical restrictions have been shown to play a role in the distribution of wh-elements (Pesetsky 2000 and related works), whence our choice to exclude all wh-phrases such as *quel(le)*+N (*quel professeur*, 'which professor').

     We searched each of the words in (a)-(e) in the corpus using the platform's search tool, making 'exact word' (*mot exact*) queries. These gave .csv files in return. Since all the items in (a)-(e) have both interrogative and non-interrogative uses (cf. *La ville où j'habite* 'the city where I live' where *où* is a relative pronoun), we sorted the .csv files using a Python script that matched sentences containing a question mark. Since the ESLO corpus has been meticulously transcribed, this automatic sorting method was trustworthy, as confirmed by subsequent manual checks. The latter were especially meant to get rid of those interrogative sentences containing one of the words in (a)-(e) that did not bear on the wh-word itself, such as that in (22):

(22) **Quand** vous écrivez   à   vos   amis,   est-ce que
      When   you   write   to   your   friends   *est-ce que*
      vous   faites un   brouillon?
      you   make   a   draft
      'Do you make a draft when you write to your friends?'

                                                    (ESLO1_ENT_001_C)

     Questions like (22), despite containing one of the wh-words under investigation in this paper, had to be eliminated from our data pool because the wh-word, here *quand* ('when'), had a non-interrogative role in them. Here, for instance, *quand* introduces the temporal reference for the polar ('yes/no') question that follows immediately (*est-ce que vous faites un brouillon?*).

     All remaining interrogatives then underwent additional triage, which was done by in two rounds during which both the transcript and the audio file were assessed manually, as prosody is important to distinguish between different types of interrogative, i.e., wh in-situ in echo questions vs wh in-situ information-seeking questions. During these, we only selected simple direct partial information-seeking questions (making a distinction between ex-situ, in-situ in non-final position and in-situ in final position, cf. §5.3 for additional info). This means that we systematically discarded from our pool of interrogatives:
  (i)   long distance partial wh-questions, as (23);
  (ii)   embedded partial questions, as (24);
  (iii) clefts, both complete like (25) and elliptical like (26);
  (iv)   infinitivals, as (27);
  (v)   multiple wh-phrase constructions, such as (28);[42]

---

[42]   Note that this question is also an echo question, i.e. the teacher is not asking a genuine question but rather a request for repetition.

(vi)   non-information seeking questions such as rhetorical questions, as (29), introspective questions, as (30), echo questions, also (28), quiz questions, as (31), and fragments questions (*Quoi?, Comment?, Où?*);

(vii)   all questions uttered by non-native speakers of French.

(23)   Combien   de   livres croyez-vous   que   vous   lisez
         How.many   of   books   believe=you   that   you   read
         dans   une   année?
         in   one   year
         'How many books a year do you think you read?'   (ESLO1_ENT_046)

(24)   Je   sais   pas   comment   vous   posez   la   question…
         I   know   neg   how   you   ask   the   question
         'I have no idea how you'd ask the question…'   (ESLO1_ENT_028_C)

(25)   Qui   c'est   qui   t'a   dit   ça?
         Who   *ce*=is   who   you=has   told   this
         'Who's told you that?'
         Lit: 'Who is it that told you that?'   (ESLO2_REPAS_1254)

(26)   C'est   qui?
         Ce=is who
         'Who's that?'   (ESLO2_REPAS_1260)

(27)   [...]   ils   vont   venir   sur   Orléans   pour   faire quoi?
                 they   will   come   to   Orleans   to   do what
         'They're coming to Orleans to do what?'   (ESLO2_DIA_1226)

(28)   Context: A pupil tells something to the teacher. The teacher does not understand, and asks:
         Qui   est-ce   qui   t'a   demandé   quoi?
         Who   is=*ce*   who   you=has   asked   what
         'Who did what?'   (ESLO2_ECOLE_1281)

(29)   Comment   voulez-vous   qu'il   s'en   sorte?
         How   want=you   tha=he   refl=part   make.it
         'How can you expect him to make it?'   (ESLO1_ENT_022_C)

(30)   Un   un   comment   dirais-je?   Un   responsable
         A   a   how   would.say=I   A   responsible
         euh   assez   haut   placé
         euh   quite   high   placed
         'A, a…how would I say it? Someone in a high-end responsibility job.'
                                                          (ESLO1_ENT_001_C)

(31)   Context: A teacher is reading a book to her pupils. Occasionally, she stops to ask questions about the story.

Elle	part	loin,	elle	part	avec	qui,	là?
She	leaves	far	she	leaves	with	whom	there

Lit: 'She's leaving, she's going with whom?'

The reason behind the cleaning in (i)-(viii) was the desire to only compare sentences that are equivalent both syntactically and interpretationally, whence our choice to only study interrogatives of the partial (wh-) type, and only those who contained one single wh-element and were matrix questions. Indeed, biclausal structures such as clefts and indirect questions have been widely acknowledged in the literature to have different behaviors with respect to mono-clausal structures (Shlonsky 2012, Belletti 2009 2015 refer to Haegeman *et al*. 2014 for arguments in favor of the bi-clausality of clefts), and so have non-information seeking questions (Baunaz 2005, Mathieu 2009 and references therein cited). The choice to exclude non-native speakers of French (by resorting to the metadata provided by the platform), even when they appeared to master the language fully, was an extra precautionary measure to ensure that our database did not contain sample errors or noise.

5.3. Data classification

Once our dataset was clean, we classified each entry along the criteria in Table 4.

Table 4. Classification criteria originally adopted for the dataset used in this study

Criteria	Choices
Type	Ex-situ, in-situ (final), in-situ (non final)
Structure	VS, SV, (est-ce) que
Activation level (in-situ only)	EACT, N-ACT

The data was once again classified by us, manually, over a total of two rounds. It must be noted that the distinction between clause-final wh in-situ, as that in (32), and non-final wh in-situ, as in (33), was originally made because we were expecting this alternation to carry interpretational meaning.

(32) Vous partez **quand**?
      You   leave   when
      'When are you leaving?'   (ESLO1_TEL_371)

(33)	Euh	vous	êtes	arrivé	**quand**	à	Orléans?
	Euh	you	are	arrived	when	in	Orleans

'When did you arrive in Orleans?'
Lit: 'You arrived when in Orleans?'   (ESLO1_ENT_148_C)

Unfortunately, our calculations showed that no significant difference could be detected between the two structures, therefore we shall not discuss it here.

5.3.1. More than just EACT and N-ACT

Despite the fact that we originally intended to merely reproduce Larrivée's (2019a) on a bigger dataset cleaned on the basis of the syntax-informed criteria in (i)-(ix), during the first round of data classification we quickly realized that there existed a third type of context in which wh in-situ was used productively. In this peculiar context, wh in-situ was neither explicitly activated, nor non-activated, but nonetheless 'discourse old', i.e., already established in the discourse. We thus decided to add a third activation level to our study of wh in-situ, previously introduced in Garassino (2022): inferred (INF) wh in-situ.

Garassino (2022), a quantitative work on partial rhetorical and information request questions in Italian and French,[43] follows Dryer (1996) and proposes that when the propositional content of the question is not 'explicitly mentioned in the conversation' but nonetheless 'easily accessible thanks to our world knowledge', in-situ is inferable. The author provides the example in (34) to illustrate the inferable activation level:

(34)   NAT: Et   qu'est-ce que   tu   as   acheté d'autre
               and   what=*est-ce que*   you   have   bought of=other
               alors?
               then
               'And so, what else did you buy?'
        MAI: Et   ben   on   a   acheté &euh la   table   avec
               And   well   we   have   bought &euh the   table   with
               les   quatre chaises /#   sept-cent   balles//#
               the   four   chairs /#   seven hundred   euros
               'Well, we bought a table with four chairs, 700 euros'
        JOS: *Pour   mettre où ? #*
               for   to.put where
               'Where are you going to put them?'
                                    (C-ORAL-ROM, ffamcv05, Garassino 2022 : 10, (12))

In (34), the fact that the new table and chairs will have to be placed somewhere after buying has not been explicitly mentioned in preceding discourse, but is accessible thanks to our word knowledge (it is quite straightforward that once new bulky objects of the sort are bought, they will have to be placed somewhere in the room they are meant for).

Garassino (2022) attributes an elaboration function to French wh in-situ in both explicitly activated and inferable contexts, as these are 'used to clarify or to add further details to a discourse topic raised in the previous context.'. Table 5 completes the data in Table 4 by including the three discourse activation levels used to classify our data, and their function.

---

[43]   Unlike Larrivée 2019, Garassino 2022 does not study the evolution of the in-situ construction through time, but he focuses on its (pragmatic) licensing conditions in contemporary French. He performed a corpus search in the French section of the C-ORAL-ROM corpus (Cresti & Moneglia 2005), which contains data of spontaneous interactions (face-to-face conversations and phone calls) of different registers of French, from the regional variety of Aix-en-Provence.

Table 5. Discourse Activation levels as in Larrivée (2019a) and Garassino (2022)

Activation level	Propositional content	Function (Garassino 2022)
Explicitly activated (**EACT**)	Discourse Old (previously established in the discourse)	Elaborates on a previously introduced topic
Inferable (**INF**)	Discourse Old (**not** explicitly mentioned in the discourse, but easily accessible)	Elaborates on a previously introduced topic
Non-activated (**N-ACT**)	Discourse New (not established in the discourse)	Introduces new topic

In §6, we present our results by organizing them into the four research questions presented in §5 in (A)-(E).

6. Overview of the results

Table 6 shows the raw occurrences for each investigated wh-word left after sorting and cleaning our dataset as detailed in §6. The numbers combine all occurrences of wh ex-situ and wh in-situ, independently of their activation.

Table 6. Raw occurrences of wh-words across the two corpora under consideration

Wh-word	ESLO 1	ESLO 2	tot
comment	562	472	1034
où	190	293	483
quand	65	61	126
quiO	24	27	51
quoiO	276	408	684
tot	1117	1261	2378

It should be noted that out of a total of 2378 partial questions, we have a total of 988 in-situ, including 208 in ESLO 1 and 780 in ESLO 2. Also, the overall representation for each of the ESLO corpora is well balanced, with 1117 total interrogatives in ESLO 1, and 1261 for ESLO 2. We present the distribution of wh ex-situ and wh in-situ in the two corpora more in detail in §6.1.

6.1. Q1: To what extent does the overall proportion between wh in-situ and wh ex-situ vary over time (i.e. between ESLO1 and ESLO2)?

Figure 1 shows the evolution in the distribution of all occurrences of wh in-situ and wh ex-situ in micro-diachrony, i.e., between ESLO 1 and ESLO 2.

Figure 1. Distribution of in-situ vs ex-situ constructions in micro-diachrony
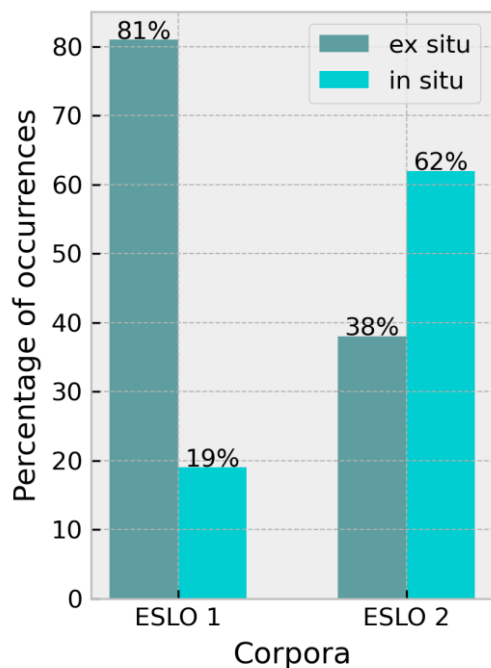
Figure 1 illustrates that the overall proportion of wh ex-situ and wh in-situ changes over time, with wh ex-situ constituting 81% of all occurrences in ESLO 1 (913 in total) and only 38% in ESLO 2 (481). This change is significant, as confirmed by our calculations ($\chi^2 = p < 0.05$).[44] Therefore, while ex-situ is still the more prominent structure in ESLO1, i.e., in the 1970s, in-situ had already become more prominent by the time ESLO2 was established.

The individual growth of each of the wh-words under consideration follows a trend similar to the general one, with wh in-situ becoming dominant over wh ex-situ. Observe the raw data in Table 7.

---

[44]   We performed χ^2 tests to determine whether our proportions were significantly different. Note that our tests were performed on raw numbers, not percentages.

Table 7. Raw occurrences of ex-situ vs in-situ wh-words across the two corpora

	ESLO 1		ESLO 2	
Wh-word	**ex-situ**	**in-situ**	**ex-situ**	**in-situ**
comment	539	23	333	139
où	141	49	84	209
quand	54	15	32	29
quiO	21	3	8	19
quoiO	158	118	24	384

Figure 2 illustrates the overall proportion of wh ex-situ and wh in-situ in ESLO 1 for each of the five wh-words investigated here.

Figure 2. Distribution of specific wh-words in ESLO1 (1970s)
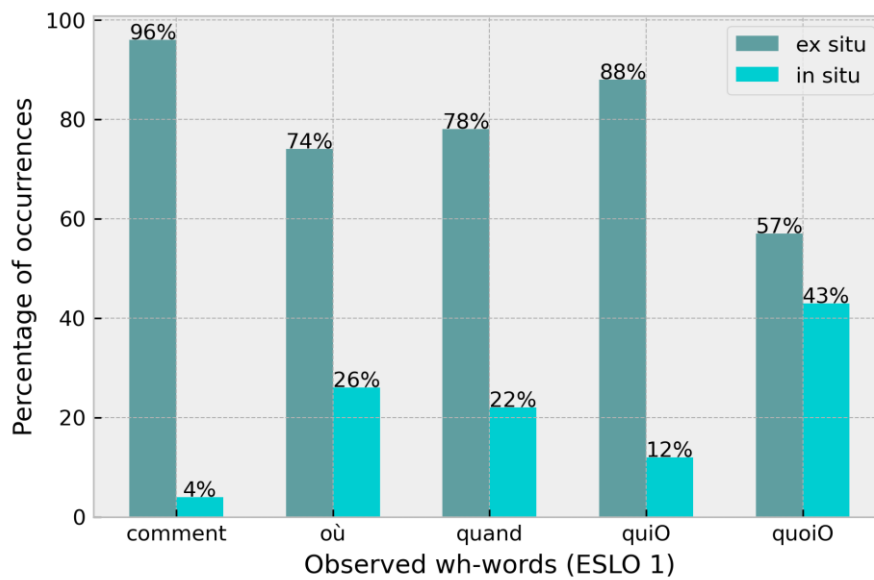
In the 1970s (ESLO 1), wh ex-situ was thus still the predominant construction for each of the wh-words, with proportions that span from 96% for comment ('how') to 57% for quoiO ('what). The same is however not true for the 2010s (ESLO 2), as illustrated in Figure 3.

Figure 3. Distribution of specific wh-words in ESLO2 (2010s)
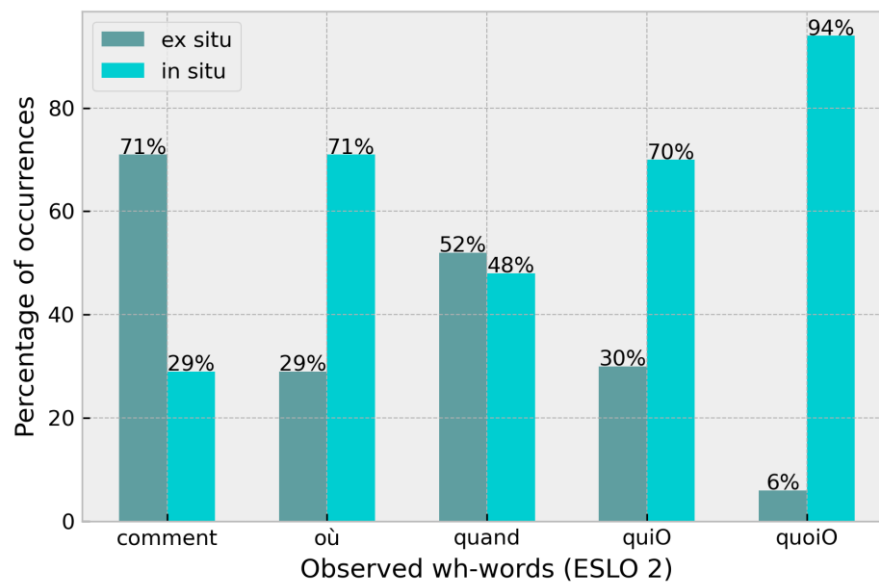
The data from ESLO 2 in Figure 3 demonstrates an important augmentation of the overall proportion of wh in-situ for each wh-word, as confirmed by our calculations which show $\chi^2 = p < 0.05$. This means that, by the 2010s, wh in-situ had already become the most productive question-formation strategy for answer-seeking, matrix single wh-questions in French.

However, it should be noted that the individual growth of each of the wh-words in Figure 3 is not equivalent, i.e., while *comment* ('how') in-situ increases from 4% to 29%, wh in-situ already constituted 43% of all occurrences of *quoi*-IO in the 1970s (Figure 2) and had reached an impressive 94% by the 2010s. Moretheless, it is noteworthy that some wh-words are more frequent than others, as illustrated in Figure 4, which shows the proportion for each of the wh-words under consideration in our pools of wh in-situ questions for each of the corpora.

Figure 4. Wh in-situ occurrences in micro-diachrony (1970s-2010s): distributions for each of the wh-words under consideration
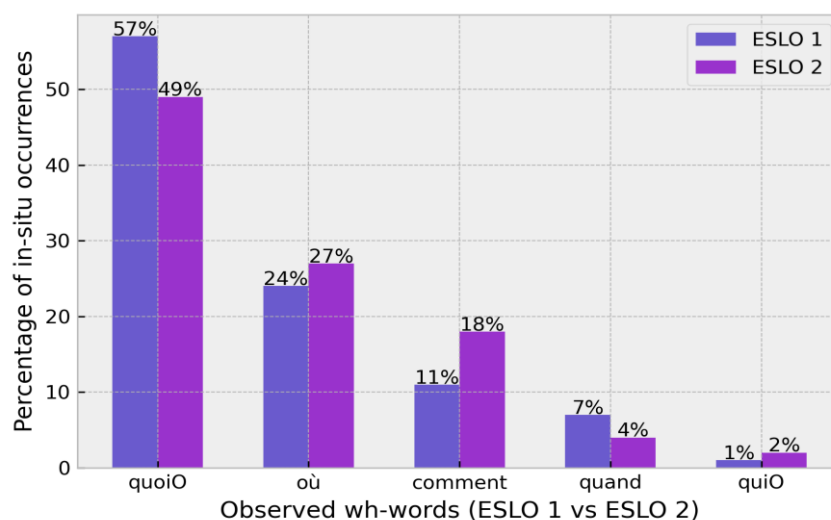
The dominance of *quoi*-IO in-situ questions is striking in both corpora, as the 118 occurrences thereof in ESLO 1 and 384 in ESLO 2 correspond to 57% and 49% of all in-situ occurrences for each corpus. Conversely, *où* and *comment* constitute a smaller part of, respectively, 24% and 11% for ESLO 1 and 27% and 18% for ESLO 2. *Quand* and *qui* in-situ, unfortunately, constitute such a small part of all in-situ occurrences in our dataset that we are unable to make significant discussions and predictions on their evolution, as our calculations confirmed an undesirable $\chi^2 = p > 0.05$ in all fit tests. Note that our results are in line with Garassino's (2022), Adli's (2015) and Elsig's (2009) studies, all of which register a majority of *quoi* and a minority of *qui*-object in-situ.

One might rightfully wonder why *qui*-O is very scarce in our corpus (and in Adli's 2015, Elsig's 2009 and Garassino's 2022). A possible answer is that the availability of *qui* in-situ is restricted to animate object, making the occurrence of this wh-word in-situ rather rare. Conversely, *quoi* often appears in constructions headed by some of the most frequent verbs of the language, such as *faire* 'make, do' and *être* 'be'.

Figure 4 suggests that the nature of wh-words impacts their ability to adapt to change. As the in-situ construction becomes more productive, arguments seem to be licensed in it first and quicker (*quoi*-IO here), only later followed by non arguments (*où*, *comment*). This pattern has been attested in other studies on the evolution of wh-interrogatives: the wh-*est-ce que* construction, for instance, started out only with *que/qui* in the 12th c., then with *comment* later in the 12th c., followed by *où* in the 15th c. and *quand* at the end of the 16th c. (Tailleur 2013, Rouquier 2002, a.o).

6.2. Q2: To what extent does the proportion between the three activation levels vary over time?

In §6.1, we observed the overall proportion of wh ex-situ and wh in-situ changed significantly between the 1970s (ESLO 1) and the 2010s (ESLO 2). Accordingly, wh in-situ had already become the predominant structure for single matrix wh-interrogatives by the 2010s, crucially evolving from constituting only 19% of all interrogatives in ESLO 1 (208 in total) to an impressive 62% in ESLO 2 (680).

Interestingly, the micro-diachrony of wh in-situ shows not only an increase of the percentage of occurrences of this structure but also significant changes in the interpretational contexts under which the structure itself is licensed. We provide a general picture of the observed changes in §6.2.1, and fine grained observations on each wh-word in §6.2.2.

6.2.1. Overall changes

We adopted a tripartition of wh in-situ based on its activation with respect to the preceding context (§5.3), whereby all occurrences are either classified as EACT (when the propositional content has already been mentioned in the preceding discourse), N-ACT (when wh in-situ is uttered 'out-of-the-blue') or INF (when the propositional content can be recovered from one's knowledge of the world).

The raw numbers resulting from our classification are given in Table 8.

Table 8. Raw occurrences of in-situ wh-words across the two corpora under consideration, divided into three activation levels[45]

ESLO 1			ESLO 2		
EACT	**N-ACT**	**INF**	**EACT**	**N-ACT**	**INF**
38	18	152	122	261	394

While the raw numbers in Table 8 fail to show what our classification demonstrates about the interpretive contexts in which in-situ is licensed, Figure 5 illustrates the changes in distribution observed for each of the three activation levels in ESLO 1 vs ESLO 2 in a straightforward way.

Figure 5. Proportions of wh in-situ according to their level of activation in micro-diachrony
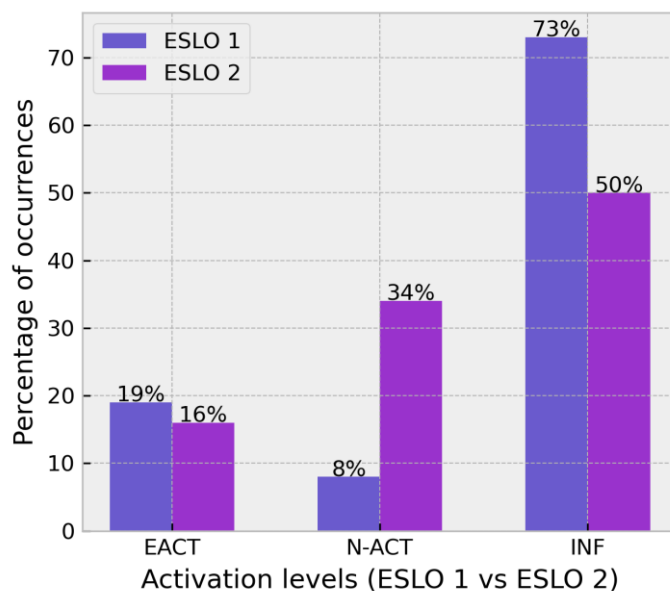
In Figure 5, the iris columns illustrate the distribution of each activation level in ESLO 1, while the purple columns are for ESLO 2. The percentages for EACT structures are very close between the the two corpora (19% and 16% for a total of 38 and 132 raw occurrences, respectively), while N-ACT and INF structures show different trends, with the former increasing from 8% to 34% over time (18 to 261 occurrences in our datasets) and the latter decreasing from representing 73% (152) of all occurrences to 50% (394).

We ran χ^2 tests to understand the graph in Figure 5 and our calculations showed that the observed decrease from 19% to 16% of explicitly activated in-situ questions between ESLO 1 and ESLO 2 is not significant ($\chi^2 = p > 0.05$), to the effect that explicit activation cannot be considered the variable that drives change in these constructions.

---

[45] Note that henceforth, as briefly mentioned in §6.1, we exclude the *qui*-O and the *quand* data from our discussion, as our tests of fitness returned $\chi^2 = p > 0.05$ for both, meaning that our data is not significant. This is not surprising, given the small number of occurrences that we collected for each, as previously illustrated in Tables 6 and 7.

Conversely, the changes in overall proportions observed between ESLO 1 and ESLO 2 for N-ACT and INF structures are statistically significant as they returned a value of $p < 0.05$.

This suggests that while the weight of explicit activation did not change between the two periods under investigation here, inferable contexts did become less prominent, while N-ACT contexts gained in productivity. According to Larrivée's (2019b) claim, in-situ questions started out in exclusively EACT contexts and then progressively lost their categorical status and became compatible with different licensing contexts too. That the evolution of EACT froze and the overall proportion of INF decreased while N-ACT gained importance between the 1970s and the 2010s suggests that by the 2010s French in-situ wh-questions had already gained independence from the linguistic context, becoming the default construction to ask questions in Hexagonal French (see Baunaz 2011, but also Faure & Palasis 2021).

Noteworthily, INF structures, overlooked in Larrivée's (2019b) discussion, constitute half (73% in ESLO) to more than half (50%) of all occurrences of wh in-situ in our dataset. These data confirm the importance of the present study for the understanding of the licensing criteria for matrix single wh in-situ in French, as well as the descriptive power of the tripartition of wh in-situ proposed in Garassino (2022) and adopted here.

6.2.2. Changes per wh-word

Let us now turn to the evolution of single wh-words, remembering that *qui-O* and *quand* have been excluded from the discussion because the numbers in our possession were not significant. Table 9 shows the raw numbers in both corpora for each of the activation levels under consideration.

Table 9. Raw occurrences of in-situ wh-words across the two corpora under consideration, divided into three activation levels

	ESLO 1			ESLO 2		
Wh-word	**EACT**	**N-ACT**	**INF**	**EACT**	**N-ACT**	**INF**
comment	5	2	16	19	56	64
où	10	10	29	29	83	97
quoi-IO	20	4	94	66	114	204

Figures 6 and 7 illustrate the different representation of the three activation levels for each of the wh-words into consideration (*quoi-IO*, *où* and *comment*) in ESLO 1 and ESLO 2, respectively.

Figure 6. Activation levels for each wh-word (ESLO1, 1970s)
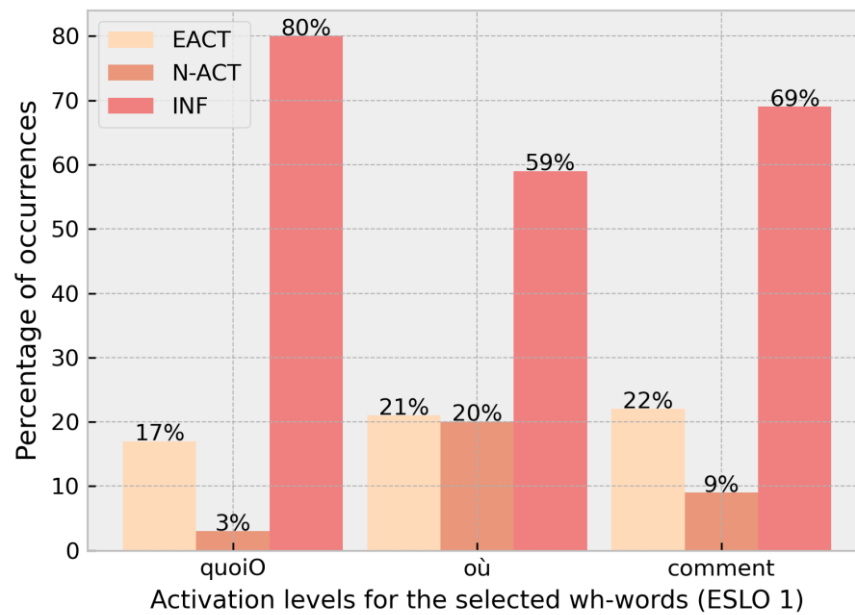
Figure 7. Activation levels for each wh-word (ESLO2, 2010s)
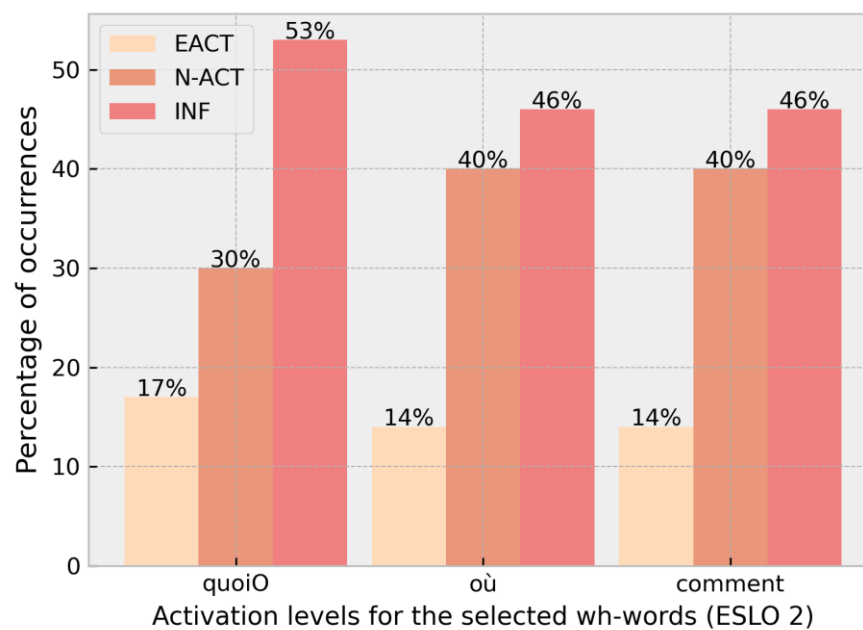
Globally speaking, by merely looking at the percentages in Figures 6 and 7 alone, the trend observed in §6.2.1 seems confirmed, as we visually acknowledge little to no variations in the overall proportions of EACT contexts, a generalized raise in the percentage of N-ACT in-situ (N-ACT goes from 3% to 30% for *quoi-IO*, 20% to 40% for *où*, and 9% to 40% in the case of *comment*), and drops in the distribution of INF in-situ.

However, this apparent trend was not confirmed for every wh-word by our calculations: while the evolution of *quoi-IO* reproduces that of the global trend observed in Figure 5 in §6.2.1 (with the changes for N-ACT and INF scoring $p < 0.05$

in our χ^2 tests, as well as the non significant change ($\chi^2 = p > 0.05$) registered for EACT in-situ), that of *où* and *comment* do not. The significant drop observed in INF in-situ globally and for *quoi-IO* is indeed not observed in either *où* or *comment*, with both scoring $p > 0.05$ in our χ^2 tests. Conversely, the change from 9% to 40% of total N-ACTs observed for *comment* is indeed significant ($\chi^2 = p < 0.05$), thus confirming the evolution of the structure as increasingly context-free, while the same cannot (yet) be said of *où*, whose passage from a total of 20% to a total of 40% N-ACT in-situ cannot be considered significantly different ($\chi^2 = p > 0.05$).

Therefore, while *quoi-IO* saw both an increase in the distribution of N-ACT occurrences and a decrease of INF structures in the passage from ESLO 1 to ESLO 2, *comment* only underwent the first change, and *où* none. This further confirms the pattern discussed in §6.1 whereby arguments like *quoi-IO* are commonly found in new constructions (such as N-ACT in-situ) before adjuncts like *où* and *comment*. Remember Tailleur's (2013) and Rouquier's (2002) claim that *comment* was already found in *est-ce que* interrogatives in the 12th c., while *où* only entered these constructions later, in the 15th c. What we observe in the passage from ESLO 1 to ESLO 2 is a very similar evolution, with N-ACT in-situ starting to become more prominent for *comment*, but not yet for *où*.

It follows that the global evolution observed in Figure 5 (§6.2.1) does not apply to all wh-words observed in this study yet, but nonetheless represents the direction in which all wh-words are likely evolving. The reason why the global picture corresponds exactly to that of *quoi*-IO undoubtedly resides in the fact that *quoi-IO* in-situ is by far the most common wh-element and represents more than 60% of all in-situ occurrences in our database.

7. Discussion

7.1. The evolution of in-situ vs. ex-situ questions from the 15th c. to today

Let's review what we know about the evolution of wh in-situ construction vis-à-vis ex-situ from the 15th c. until today. In §3.3.3, we saw that in the corpus of literary texts and theatrical plays studied by Larrivée (2019b) wh in-situ only constitutes less than 1% of all partial interrogatives. During this period, in which in-situ was an emerging and still rare construction, it was only licensed under a unique and specific pragmatic condition, i.e., that of 'explicit activation'. At the beginning of the 20th c., this categorial property was already lost, as wh in-situ then already constituted a proportion of more than 2% of the total pool of questions (Larrivée 2019b). While we did not verify Larrivée's claims for the beginning of the 20th c. and preceding stages, our study confirmed the author's claim that wh in-situ had already lost its categorial status by the 1970s, as in-situ interrogatives constitute 19% of all partial interrogatives in the 1970s (ESLO 1), and then 62% in the 2010s (ESLO 2). Moreover, we have seen that in both corpora investigated here in-situ questions were not only licensed in EACT contexts, but also in N-ACTs and INFs.

Over the relatively short span of four decades, we thus observe a striking evolution within the system of partial interrogatives in French. This corresponds to noteworthily increased rates of use of wh in-situ, which by the first decade of the 21st c. already constituted the most represented partial interrogative strategy of the

language (as already demonstrated in previous quantitative studies such as Huková 2006, Adli 2015, and Guryev 2017 about Swiss French text messages, a.o.).

One might wonder why the change from predominant wh ex-situ to predominant wh in-situ has happened so fast in French, while language change is widely acknowledged to happen slowly (see Marchello-Nizia 1999, 2003; Tailleur 2013; Guryev 2017 about the history of the French interrogative system). In this respect, it is important to understand that for the in-situ construction to be rare in the literature before the 20th c. (§3.3.3) does not entail that the structure was absent also from everyday speech at that stage of the history of the language. The construction has indeed often been characterized as a phenomenon of the spoken (familiar or vernacular) varieties of the French language, so it seems reasonable to hypothesize that the structure simply was not utilized in written texts at the time.[46;47;48]

The question should thus be reformulated as to why the change from predominant wh ex-situ to predominant wh in-situ has been happening increasingly faster. A common cause of language change has been identified as language contact (Trudgill 2020, Gadet 2021, a.o.). As suggested by D. Sportiche and U. Shlonsky (p.c), we wish to tentatively propose that the change that we have observed in French interrogatives is related to a combination of the overall relaxation of social norms following the 1969 revolutions and the widespread exposure to the internet that western societies have known since the 2000s. Accordingly, McCulloch (2019) argues that "The internet, then, makes language change faster because it leads to more weakties… [and] you can get to know people who you never would have met otherwise." Since at least the 2000s, the use of social networks has been in constant evolution and is accessible to virtually everyone in the Western World. People use applications and tools that mimic or even adopt spontaneous speech on a daily basis: instant messaging (Facebook Messenger, WhatsApp, Telegram, etc.), e-mails, internet forums, social media like Facebook and Instagram etc.), video platforms like YouTube and TikTok, and much more. All these applications, in which the adopted writing and speaking style is often relaxed (if not neglected), greatly contribute to the progressive and unprecedentedly fast abandonment of the Standard norm.

That the speed of linguistic change is not constant has actually been attested by various linguists, including Trudgill (2020), who claims that the

---

[46]   On wh-words not being well represented in written texts, but fairly in the spoken language, see also Foulet (1921).

[47]   Using written means to mimic spontaneous speech is not widespread in the 15th-19th c. French literature. Theater plays did exist, but were still mainly characterized by the use of high-register French. For a collection of French plays from the 16th-20th centuries, see https://github.com/CaterinaBi/parameters-corpus-work/tree/main/raw_data_theatre/data.

[48]   It could very well be the case that the licensing of wh in situ in French is multi-factorial, i.e. not completely free, but constrained by a number of different (grammatical) pressures. In that respect, one issue that we haven't mentioned yet, but which has been raised by one of the reviewers, is the influence of prosody. During its development, the French prosody system changed drastically after the loss of V2 (V2 started declining gradually around 1100-1500) and the configuration of intonational phrases were modified (see Adams 1987). Focus became associated with accent rather than with leftward positions. This could explain (or at least be one of the factors pressing for) the development of the wh in-situ construction. See Mathieu (2016), a.o.

'speed of linguistic change is not constant: it differs between different languages and dialects, and between different chronological periods. These differences are, at least to some extent, conditioned by social parameters. Two major social factors are involved in producing these different rates of linguistic change. There is, first, the role of the relative degree of contact versus isolation which speech communities have experienced: a good example is provided by the contrast between Faroese and Icelandic as opposed to the continental Scandinavian languages. There is, secondly, an important role for relative social stability versus social instability in the histories of communities.'
(Trudgill 2020: 1)

The impact of both the relaxation of social norms and the widespread access to and use of the Internet on the way people speak in France and more generally in the West should therefore never be overlooked. It is therefore quite possible that no real *grammatical* change has taken place in the partial interrogatives of French in the past century, and that the increased rates of use of the in-situ strategy has more to do with social factors than the core properties of the language itself. Conversely, though, in §7.2 we shall argue that the same does not apply to the interpretation and contexts of licensing of the structure, which have rather clearly been the subject of an on-going evolution that was already striking for *quoi-IO* in the passage from the way French was spoken in the 1970s (ESLO 1) to the 2010s (ESLO 2).

While the ex-situ strategy is still alive and constitutes 38% of all partial questions available in ESLO2, its impressive decrease observed between the 1970s and the 2010s and the increase of N-ACT structures observed for *quoi-IO* suggest that the process that we observed here is still undergoing and is unlikely to stop soon. Accordingly, ex-situ constructions will either cease to exist when the on-going process is complete, or acquire new meaning and start encoding a specialized categorical feature (see Faure & Palasis 2021 for a claim that this is actually already the case in Metropolitan French). We leave this question open for future investigation - a follow-up study of the contexts in which ex-situ constructions were licensed in the 1970s vs the 2010s is definitely in order.

7.2. The evolution of the interpretation of in-situ questions between the 15th c. and today

While we have attributed the increase in the overall proportion of in-situ structures between the ESLO 1 and the ESLO 2 corpora to causes beyond grammar alone (relaxation of social norms after the 1970s, widespread use of the Internet from the 2010s), what we observed about the contexts in which in-situ questions are licensed constitutes the true change undergone by these structures in the last four decades.

If the wh in-situ strategy mainly appears in restricted interpretational contexts between the 15th and the 19th centuries (Larrivée 2019b), our study demonstrates that the contexts that license wh in-situ diversified and evolved significantly in the second half of the 20th c. Accordingly, we singled out three different activation levels that license in-situ, thus combining previous classifications proposed in Larrivée (2019b) and Garassino (2021): those of explicit activation and non-activation, and that of inferrable wh in-situ. None of these are categorical in the corpora that we investigated,

and noteworthy trends can be observed in the passage from the way in-situ is licensed in ESLO 1 to ESLO 2, especially by looking at single wh-words separately.

If wh in-situ appears in a large quantity of explicitly activated and inferable (and thus 'old discourse') contexts during the two periods under consideration here, we nonetheless observed a significant, major rise of non-activated in-situ both for *quoi-IO* and *comment* (+25%), and a significant decrease of INF in-situ for *quoi-IO*. As for explicitly activated contexts, we saw no significant variation over time for either of the wh-words under consideration. The fact that non-activation is becoming increasingly productive in constructions with *quoi-IO* and *comment* crucially suggests that in-situ has lost its strictly context-bound status that it had when it emerged as a question-formation strategy. While no significant change was observed in the evolution of *où*, as the proportion of N-ACT in-situ in constructions bearing on this wh-word did not change significantly between ESLO 1 and ESLO 2, we have claimed that this not makes our claim invalid, as it has been demonstrated that *où* is more resistant to change than *quoi-IO* and *comment*. We thus forecasted two future changes to be expected: a significant decline of INF in-situ for *comment* and later *où*, and an increase in the distribution of N-ACT in-situ for *où*.

If we abide by Hooper's (1987) famous claim whereby '[s]ubsystems are either innovating and spreading out from an earlier more restricted usage, or are contracting and being abandoned from an earlier wider use …' (Hooper 1987: 144), the significant rise of in-situ in non-activated context observed for *quoi-IO* and *comment* suggests that the change is entering the grammar of French and we are in the process of switching to a new interrogative system: a wh-in situ system in which ex-situ occurrences will either be highly specialized or extremely scarce.

7.3. What our data say about the controversy

In §3, we introduced a controversy that exists among scholars whereby it is unclear whether or not French wh in-situ involves a presupposition. Conservative scholars, we claimed, argue that there is a unique and particular contextual condition licensing wh in-situ. Conversely, liberals argue that there is no particular contextual condition licensing wh in-situ.

Interestingly, the work of the former dates back to the mid-1990s, while that of liberals takes into consideration data collected between the second half of the 2000s until today. As a consequence of what we observed in the present paper, it is possible to claim that both factions are right, only for different periods. Indeed, conservative scholars were describing a stage of French in which wh in-situ was still relatively context-bound (as N-ACT occurrences were still very scarce at the time, as in ESLO 1), while the liberals are basing their discussion on varieties of French in which, like in ESLO 2, wh in-situ is very productive, increasingly free from the surrounding context, and more varied interpretively.

Our work therefore highlights the superiority of the tripartition of French wh in-situ proposed in Garassino (2022) over the traditional understanding of wh in-situ based on the poorly-defined notion of 'presupposition', but also the importance of considering French and all languages not as monolites but as living creatures. Especially when dealing with spontaneous speech and phenomena that display (apparent) optionality, the approach that seems to yield the most reliable results is that of micro-diachrony.

8. Conclusions

Our study duplicated Larrivee's work, which was based on a portion of the ESLO 1-2 corpora, and extended it to the totality of the recordings. Additionally, we did not limit our investigation to *comment* alone, but rather opened it to all those non-lexically restricted wh-words that can alternate between the in-situ and ex-situ position: *comment* 'how', *quand* 'when', *où* 'where', *qui-O* 'who' and *quoi-IO* 'what'. Then, our data cleaning methods were significantly different with respect to Larrivée's. Finally, we decided to classify our in-situ data following Garassino's (2022) claim that a third level of discursive activation exists which was not considered by Larrivée, namely that of inferable in-situ.

      Our data demonstrated that by the 2010s, Hexagonal French had already moved to a predominantly in-situ interrogative system, and in-situ constructions had gained independence from the surrounding linguistic context, as shown by a remarkable increase in the overall proportion of N-ACT in-situ. Conversely, the overall representation of EACT structures did not change between the two corpora under consideration, and the percentage of INF in-situ only decreased with *quoi-IO*. Our results are therefore different from those of Larrivée's, with the notable exception that we both found a significant increase of N-ACTs. They also argue that the controversy that existed in the literature about the pragmatic licensing conditions of in-situ was linked to the time frame in which the different authors collected their data, with conservatives describing a system that roughly resembled that of ESLO 1, and liberals basing their work on varieties of French similar to that represented by ESLO 2.

      At a broader level, this paper highlights the importance of systematically considering languages as dynamic systems. Changes often begin slowly, then go through an accelerated transition during which the change becomes observable, as demonstrated for in-situ between the 1970s and the 2010s in our study. Our take-home message is that when dealing with spontaneous speech data, it is crucial to focus on short periods of time that may help see potential changes in progress, understand and predict variability, and sometimes even solve forty-year long controversies.

References

Adams, Marianne. 1987. *Old French, null subjects and verb second phenomena*. PhD thesis, University of California, Los Angeles.

Adli Aria. 2004. Y a-t-il des morphèmes intonatifs impliqués dans la syntaxe interrogative du Français ? Le cas du *qu-in situ* . In M. Trudel & M. Selig (eds.), *Nouveaux départs en phonologie: les conceptions sub- et supra-segmentales*. 199-215. Tübingen : Narr.

Adli, Aria. 2006. French wh-in-situ Questions and Syntactic Optionality: Evidence from three data types. In *Zeitschrift für Sprachwissenschaft* 25: 2. 163-203. https://doi.org/10.1515/ZFS.2006.007

Adli, Aria. 2015. What you like is not what you do: Acceptability and frequency in syntactic variation. In Adli, A., García García, M. & Kaufmann, G. (eds.). *Variation in language: System- and Usage-based Approaches*. 173-200. Berlin/Boston: Walter De Gruyter. https://doi.org/10.1515/9783110346855-008

Baunaz, Lena. 2005. The syntax and semantics of wh in-situ and existentials: the case of French. *Leiden Working Papers in Linguistics* 2/2: 1-27.

Baunaz, Lena. 2011. *The Grammar of French Quantification*, Dordrecht, Springer. https://doi.org/10.1007/978-94-007-0621-7

Baunaz, Lena. 2016. Wh-phrases in-situ: interface strategies. *Acta Linguistica* 63 (2): 125-168. https://doi.org/10.1556/064.2016.63.2.1

Baunaz, Lena & Cédric Patin. 2011. Prosody refers to semantic factors: evidence from French wh-words. In H.-Y. Yoo, E. Delais-Roussarie (eds.), *Actes d'Interface Discours & Prosodie*, 93-107. http://makino.linguist.jussieu.fr/idp09/actes_en.html.

Baunaz, Lena & Cédric Patin. 2012. Quand la prosodie et la sémantique vont de pair, le cas des mots-*qu* en français. In de Saussure L. & Rihs A. (eds). *Etudes de sémantique et pragmatique françaises*. 357-378. Berne: Lang.

Beauchemin, Normand, Pierre Martel & Michel Théoret. 1971-1974. *Corpus sociolinguistique de l'Estrie*. Université de Sherbrooke.

Beeching, Kate, Nigel Armstrong, & Françoise Gadet. 2009. *Sociolinguistic variation in contemporary French*. Amsterdam: Benjamins. https://doi.org/10.1075/impact.26

Belletti, A. 2009. *Structures and Strategies*. Routledge Leading Linguists 16. New York.

Belletti, A. 2015. 'The Focus map of clefts: Extraposition and Predication'. In U. Shlonsky (ed.) *Beyond Functional Sequence. The Cartography of Syntactic Structures*

10. Oxford Studies in Comparative Syntax. Oxford University Press. 42-59. https://doi.org/10.1093/acprof:oso/9780190210588.003.0003.

Behnstedt, Peter. 1973, Viens-tu ? Est-ce que tu viens ? Tu viens ? *Formen und strukturen des direkten Fragesatzes im Französischen.* Tübingen, Narr.

Biezma, M. 2018. Givenness and the difference between wh-fronted and wh-in-situ questions in Spanish. *Romance Languages and Linguistic Theory: Selected papers for Going Romance* 29. John Benjamins. 21-39. doi:10.1075/rllt.13.03bie.

Blanche-Benveniste, Claire. 1997. A propos de Qu'est-ce que c'est et c'est quoi. *Recherches sur le français parlé* 14 : 127-146.

Boeckx, Cédric. 1999. Decomposing French questions. *University of Pennsylvania Working Papers in Linguistics* 6/1: 69-80 .

Boeckx, Cédric. 2003. French Wh-in-situ Interrogatives as (C)overt Clefts, ms. Harvard University.

Bonan, Caterina. 2019. *On clause-internally moved wh-phrases: Wh-to-foc, nominative clitics, and the theory of Northern Italian wh-in situ*. Ph.D. thesis, Université de Genève.

Bonan, Caterina. 2021. Romance interrogative syntax. Formal and typological dimensions of variation. Amsterdam: Benjamins. https://doi.org/10.1075/la.266

Bošković, Željko. 1998. LF movement and the Minimalist Program. In N. T.Pius /K. Kusumoto (eds.), *Proceedings of NELS 28*. 43-57. Amherst, MA, GLSA.

Bošković, Željko. 2000. Sometimes in [Spec, CP], sometimes in situ. In R. Martin, D. Michaels, J. Uriagereka (eds.), *Step by Step. Essays on Minimalist Syntax in Honor of Howard Lasnik*. 53-88. Cambridge, MA, MIT Press.

Boucher, Paul. 2010. *Wh*-questions in French and English. Mapping syntax to information structure. In C. Breul & E. Göbbel (eds.), *Comparative and contrastive studies of information structure*, 101-137. Amsterdam: Benjamins. https://doi.org/10.1075/la.165.05bou

Canel, Angela. 2012. *Les interrogatives in situ en français : une étude syntaxique.* Tesi di Laurea. Università Ca' Foscari Venezia.

Chang, Lisa. 1997. *Wh-in-situ phenomena in French*, MA thesis, University of British Columbia.

Cheng, Lisa Lai-Shen 1991. On the Typology of Wh-questions. Ph.D. thesis, MIT. Distributed by MIT Working Papers in Linguistics.

Cheng, Lisa Lai-Shen & Rooryck, Johan. 2000. Licensing wh-in-situ. *Syntax* 3/1, 1-19. https://doi.org/10.1111/1467-9612.00022

Cheng, L. L.-S. & Rooryck, J. 2002. Types of Wh-in-situ. Unpublished MS. Leiden University.

Coveney, Aidan. 1989. Pragmatic constraints on interrogatives in spoken French. *York Papers in Linguistics* 13: 89-99. https://doi.org/10.1017/S0959269500002738

Coveney, Aidan. 1995. The use of the QU-final interrogative structure in spoken French. *Journal of French Langage Studies* 5: 143-171. https://doi.org/10.1017/S0959269500002738

Coveney, Aidan. 1996 (2002). *Variability in Spoken French: interrogation and negation.* Bristol: Intellect Books.

Coveney, Aidan 2011. L'interrogation directe. *Travaux de linguistique* 63, 112-145. https://doi.org/10.3917/tl.063.0112

Cresti, Emanuela & Massimo Moneglia. 2005. *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam-Philadelphia, John Benjamins. https://doi.org/10.1075/scl.15
Crystal, David. 1991. *A dictionary of linguistics and phonetics.* 3rd ed. Cambridge, MA: Basil Blackwell.

Dagnac, Anne. 2013. La variation des interrogatives en français. hal-00988751v2

Dekhissi, Laurie. 2013. *Variation syntaxique dans le français multiculturel du cinéma de banlieue*. Ph.D. thesis, University of Exeter.

Déprez, Viviane, Kristen Syrett, and Shigeto Kawahara. 2013. The interaction of syntax, proso- dy, and discourse in licensing French wh-in-situ questions. *Lingua* 124 (1): 4-19. https://doi.org/10.1016/j.lingua.2012.03.002

Dryer, Matthew S. 1996. Focus, Pragmatic Presupposition, and Activated Propositions. *Journal of Pragmatics* 26: 475-523. https://doi.org/10.1016/0378-2166(95)00059-3

Elsig, Martin. 2009. *Grammatical Variation across Space and Time: The French interrogative system*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/silv.3

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22. 1-25.

Etxepare, Ricardo & Myriam Uribe-Etxebarria. 2005. In-situ wh-phrases in Spanish: locality and quantification. *Recherches linguistiques de Vincennes* 33, 9-34. https://doi.org/10.4000/rlv.1238

Faure, Richard & Katerina Palasis 2021. Exclusivity! *Wh*-fronting is not optional *wh*-movement in Colloquial French. *Natural Language and Linguistic Theory* 39, 57-95. https://doi.org/10.1007/s11049-020-09476-w

Ferguson, Charles A. 1959. Diglossia. *Word* 15, pp. 325-340. https://doi.org/10.1080/00437956.1959.11659702

Foulet, Lucien. 1921. Comment ont évolué les formes de l'interrogation? *Romania 47*, pp. 243-348.

Gadet, Françoise. 1989. Le français ordinaire. Paris: Colin.

Gadet, Françoise. 2007. *La variation sociale en français*. 2e éd. Paris: Ophrys. https://doi.org/10.4000/praxematique.847

Gadet, Françoise. 2021. Le changement linguistique. *Langage et société* (Hors série) : 41-46. https://doi.org/10.3917/ls.hs01.0042

Garassino, Davide. 2022. A contrastive perspective on French and Italian wh-in situ questions : A discourse-pragmatic approach. In *Functions of Language* 22 (1): 25-57. https://doi.org/10.1075/fol.00037.gar

Glasbergen-Plas, Aliza, Stella Gyllia & Jenny Doetjes. 2021. The prosody of French wh-in-situ questions: echo vs. non-echo. *Journal of Linguistics* 57: 569-603. doi:10.1017/S0022226720000158

Guryev, Alexander. 2017. La forme des interrogatives dans le Corpus suisse de SMS en français. Étude multidimensionnelle. Ph.D. thesis. Université de Neuchâtel and Université Paris 3 Sorbonne Nouvelle..

Guryev, Alexander & François Delafontaine. 2015. La variabilité formelle des questions dans les écrits SMS. *Travaux neuchâtelois de linguistiqu*e 63: 129-152. https://doi.org/10.26034/tranel.2015.2973

Guryev, Alexander & Pierre Larrivée. Routines discursives comme contextes d'émergence de l'interrogative partielle in situ en diachronie. *Langue française 212 (4))*, 75-90.

Haegeman, Liliane, André Meinunger & Aleksandra Vercauteren. 2014. The architecture of it clefts. *Journal of Linguistics 50(2)*: 269-296.

Hamlaoui, Fatima. 2009. *La focalisation à l'interface de la syntaxe et de la phonologie: le cas du français dans une perspective typologique*. Ph.D. thesis, Université Paris III Sorbonne Nouvelle.

Hamlaoui, Fatima. 2011. On the role of phonology and discourse in Francilian French wh-questions. *Journal of Linguistics 47*: 129-162.

Hopper, Paul. 1987. Emergent grammar. In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society* 13. 139-157. https://doi.org/10.3765/bls.v13i0.1834

Huková, Lubomíra. 2006. *La variation syntaxique des interrogatives directes en français parlé.* Master thesis, Universizita Karolva v Praze.

Jiménez, M. L. 1997. Semantic and pragmatic Conditions on Word Order in Spanish. Ph.D. thesis. Georgetown University.

Kato, M. 2013. Deriving "wh-in-situ" through mouvement in Brazilian Portuguese. In V. Camacho- Taboada, Á. L. Jiménez Fernández, J. Martín-González & M. Reyes-Tejedor (eds) *Information Structure and Agreement*. Linguistics Today/Linguistik Aktuell 197. The John Benjamins Publishing Company. 175-192. doi:10.1075/la.197.06kat.

Labov, William. 1996. When intuitions fail. In McNair, L. et al. (eds.), *Papers from the Parasession on Theory and Data in Linguistics* 32, 77-106

Larrivée, Pierre. 2016. Les interrogatives in-situ sont-elles pragmatiquement marquées en français vernaculaire ? Investigation synchronique et historique. Paper presented at the Kolloquium Konstanz, June 2016.

Larrivée, Pierre. 2019a. Historical pragmatics, explicit activation and wh in situ in French. In I. Feldhausen, M. Elsig, I. Kuchenbrandt & M. Neuhaus (eds.), *Romance languages and linguistic theory 15. Selected papers from 'Going Romance' 30, Frankfurt,* 114-132. Amsterdam: Benjamins. https://doi.org/10.1075/rllt.15.06lar

Larrivée, Pierre. 2019b. Contextes promoteurs et émergence des questions in situ en français In A. Dufter, K. Grübl & T. Scharinger(eds). *Des parlers d'oïl à la francophonie: Contact, variation et changement linguistiques*. 97-116. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110541816-005

Manzini, Maria Rita & Savoia, Leonardo Maria. 2005. *I dialetti italiani e romanci. Morfosintassi generativa* I-III. Edizioni dell'Orso.

Marchello-Nizia, Christiane. 1999. Le français en diachronie: douze siècles d'évolution. Éditions Ophrys.

Marchello-Nizia, Christiane. 2003. Le français dans l'histoire. In Yagello, M. (eed.), Le grand livre de la langue française, 11-88.

Martineau, France & Marie-Claude Séguin. 2016. Le Corpus FRAN : réseaux et maillages en Amérique française *, Corpus 15*. https://doi.org/10.4000/corpus.2925

Massot, Benjamin. 2010. Le patron diglossique de variation grammaticale en français. *Langue française* 168.4 : 87-106.

Mathieu, Eric. 1999. WH in situ and the intervention effect. *UCL Working Papers in Linguistics* 11: 441-472.

Mathieu, Eric. 2002. *The syntax of non-canonical quantification: a comparative study*. Ph.D. thesis, University College London.

Mathieu, Eric. 2004. The mapping of form and interpretation: the case of optional WH-movement in French. *Lingua* 114: 1090-1132. doi:10.1016/j.lingua.2003.07.002

Mathieu, Eric. 2009. Les questions en français: micro- et macro-variation. In F. Martineau, R. Mougeon, T. Nadasdi, M. Tremblay (eds.), *Le français d'ici: études linguistiques et sociolinguistiques sur la variation du français au Québec et en Ontario*, 37-66. Toronto : GREF.

Mathieu, E. 2016. The Wh parameter and radical externalization. In L. Eguren, O. Fernández Soriano & A. Mendikoetxea (eds.), *Formal grammar and syntactic variation: Rethinking parameters*, 252-290. Oxford: Oxford University Press.https://doi.org/10.1093/acprof:oso/9780190461737.003.0009

Mcculloch, Gretchen. 2019. Because Internet: Understanding the New Rules of Language. New York, NY: Riverhead Book. 10.1386/eme_00039_5

Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia* 12: 384-400. Mosegaard Hansen Maj-Britt. 2013. Negation in the history of French. In D. Willis, C. Lucas & A. Breitbarth. (eds.), *The history of negation in the languages of Europe*, 51-76, Volume 1. Oxford: Oxford University Press.

Munaro, N. 1999. *Sintagmi interrogativi nei dialetti italiani settentrionali*. Rivista di Grammatica Generativa. Monografie. Unipress.

Obenauer, Hans-Georg. 1994. *Aspect de la syntaxe A-barre – Effets d'intervention et mouvement des quantifieurs*. Ph.D. thesis, Université de Paris VIII.

Oiry, Magda. 2011. A case of true optionality: Wh in situ patterns like long movement in French. *Linguistic Analysis* 37(1-2) : 115-142.

Pesetsky, D. 2000. *Phrasal movement and its kin*. Linguistic Inquiry Monograph 37. The MIT Press.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In E. Reuland & A. ter Meulen (eds.), *The Representation of (In)definites*. 98-129. Cambridge, MA: MIT Press.

Pohl, Jacques. 1965. Observations sur les formes d'interrogation dans la langue parlée et dans la langue écrite non littéraire. *Actes du Xe Congrès International de Linguistique et de Philologie Romanes*. Tome 2. Paris: Klincksieck.

Poletto, Cecilia. 2000. *The Higher Functional Field. Evidence from Northern Italian Dialects*. Oxford Studies in Comparative Syntax. Oxford University Press.

Poplack, Shana. 1989. The care and handling of a mega-corpus. R. W. Fasold & D. Schiffrin (eds.). *Language Change and Variation*, 411-451. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.52.25pop

Quillard, Virginie. 2000. Interroger en français parlé: études syntaxique, pragmatique et sociolinguistique. Ph.D. thesis,. Université de Tours.

Quillard, Virginie. 2001. La diversité des formes interrogatives : comment l'interpréter ? *Langage et société 95*, 57-72. https://doi.org/10.3917/ls.095.0057

Rouquier, Magali. 2002. Les interrogatives en 'qui/qu'est-ce qui/que' en ancien français et en moyen français. *Cahiers de Grammaire* 27 : 97-120.

Rowlett, Paul. 2007. The Syntax of French. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511618642

Rowlett, Paul. 2011. Syntactic variation and diglossia in French. *Salford Working Papers in Linguistics and Applied Linguistics* 1: 13-26.

Shlonsky, Ur. 2012. Notes on wh in situ in French. In L. Brugé, A. Cardinaletti, G. Giusti, N. Munaro, C. Poletto (eds.), Functional Heads. The Cartography of Syntactic Structures. Volume 7, 242-252. Oxford: Oxford University Press.

Söll, Ludwig. 1985. Gesprochenes und geschriebenes Französisch, Berlin, Schmidt.

Starke, Michal. 2001. *Move dissolves into merge: a theory of locality*. Ph.D. thesis, University of Geneva.

Tailleur, Sandrine. 2013. *The French Wh Interrogative System: Est-ce que, Clefting?* Ph.D. thesis, University of Toronto.

Tailleur, Sandrine. 2022. Historical data for a contemporary debate: the use of wh-elements in 19th-20th c. Quebec French. Talk given at the Journée d'études 'L'interrogative in situ : aspects formels, pragmatiques et variationnels'.

Tailleur, Sandrine. in prep. Partial interrogatives in Quebec French: a diachronic look at a spreading variant. In Dekhissi et al. *Journal of French Language Studies. Special issue*.

Tieu, Lyn. 2012. Semantic-pragmatic conditions on wh-in-situ in English. Poster presented at the 86th Annual Meeting of the Linguistic Society of America, Portland.

Trudgill, Peter. 2020. Sociolinguistic typology and the speed of linguistic change. *Journal of Historical Sociolinguistics* 6(2). https://doi.org/10.1515/jhsl-2019-0015

Tuaillon, Gaston. 1975. Analyse syntaxique d'une carte linguistique : ALF 25 : « Où vas-tu? ». Revue de linguistique romane 38 : 79-96.

Valdman, Albert .1968. Normes pédagogiques: Les structures interrogatives du français, *International Review of Applied Linguistics in Language Teaching* 5, 3-10.

Walner, Ramona. 2018. French wh-in-situ questions: Why 'giveness' doesn't explain all. Talk given at GenWH 2018, 18.01.2018. University of Geneva.

Zimmermann, Michael. 2018. A Corpus of WH-in-Situ Interrogatives from Contemporary Hexagonal French Literary Prose Writing, Konstanz, Fachbereich Linguistik der Universität Konstanz.

Zimmermann, Michael & Katharina Kaiser. 2019. Refining current insights into the *wh*-in-situ interrogative construction in French: the case of Contemporary Hexagonal French. *Romanistisches Jahrbuch* 70 (1): 123-157. https://doi.org/10.1515/roja-2019-0005

Zribi-Hertz, Anne. 2011. Pour un modèle diglossique de description du français : quelques implications théoriques, didactiques et méthodologiques. *Journal of French Language studies* 21.2 : 231-256.

Zubizarreta, María Luisa. 2003. Intervention effects in the French wh-in-situ construction. Syntax or interpretation?. In: R. Núñez-Cedeño, L. López, R. Cameron (eds.), *A Romance Perspective on Language Knowledge and Use. Selected Papers from the 31th Linguistic Symposium on Romance Languages (LSRL), Chicago, 19-22 April 2001.* 359-379. Amsterdam: Benjamins. https://doi.org/10.1075/cilt.238.26zub