

Probing the impact of register, familiarity and information state on the form of French questions

Viviane Déprez¹ & Megan Gotowski²

(¹Rutgers University, ²MIT)

1. Introduction

As is well known, French speakers can employ multiple forms to ask *wh*- questions, as the example in (1) shows:

- | | | |
|-----|--------------------------------------|------------------|
| (1) | a. Pour qui vas-tu voter ? | inverted ex-situ |
| | for who go-you vote | |
| | ‘Who are you going to vote for ?’ | |
| | b. Pour qui tu vas voter? | ex-situ |
| | c. Tu vas voter pour qui ? | in-situ |
| | d. Pour qui est-ce que tu vas voter? | inverted cleft |
| | e. Pour qui c’est que tu vas voter? | cleft |
| | f. C’est pour qui que tu vas voter? | in-situ cleft |

These questions are not all equivalent and the problem of determining which factors influence the choice of one particular form over another is far from evident. This paper focuses on forms (b) and (c), the fronted (ex-situ) and in situ interrogatives (henceforth ex-situ and in-situ-Q, respectively) which differ in terms of whether overt movement of the *wh*-phrase has taken place or not, and abstracts away from difficulties that come with specific *wh*-phrases like *quoi*, whose moved-form differs, or *pourquoi*, for which the in-situ variant is mostly excluded¹. While discussions on the syntax of these constructions have centrally focused on the issue of how the apparent “optionality” of movement could be accounted for (see e.g. Cheng & Rooryck 2000; among others), the factors conditioning the choice of these question forms by speakers nevertheless remain poorly understood despite the numerous distinct proposals offered in the

¹ But see Glasbergen (2021) for a discussion of this point. Rizzi (2001) suggested that *wh*-words for *why* are cross-linguistically in a different position than other *wh*-words. This may be at the origin of the different behavior of *pourquoi* with respect to. in situ in French.

literature. The majority of recent research focused on the role of a variety of *discourse factors* in influencing this choice (see e.g. Mathieu 2004; Pires & Taylor 2004; Myers 2007; among others). The goal of this paper is to further explore some of these factors, first empirically through a comparative corpus study, and then experimentally via a forced-choice task. We aim to contribute to this discussion, by further probing (i) the importance of the discourse situation, in terms of register with respect to the formal vs informal social context of the interaction, (ii) the speaker/addressee relation (familiar vs non-familiar interlocutors), and (iii) the information state of the addressee (expertise vs non-expertise) with respect to the question information content.

The paper is organized as follows: in Section 2, we summarize previous accounts of *wh-in situ* in French. While syntax-based accounts have attempted to explain the restrictions on the distribution of the *in-situ* and the *ex-situ* forms, pragmatics-based accounts focused on discourse factors taken to influence the choice of one form or the other. In Section 3, we detail the results of our corpus-based analysis with the ESLO database (Abouda & Baude 2005), before turning to a discussion of the experiment in Section 4. Lastly, in Section 5, we discuss the implications of our findings and consider remaining questions for further research. As a preview our results, we will argue that discourse context has an influence on the frequency of *wh-in situ*, but we find no evidence that different registers of French are correlated with distinct grammars, such that *wh-in situ* is the default form in either one. Additionally, our experimental results support the view that familiarity between interlocutors, not just discourse register, increases the use of *in-situ*. However, addressee knowledge, i.e. answerability (Myers 2007), is not tied to *in-situ* preference, but is rather correlated with an increase in *ex-situ*. We suggest accommodating these results with a novel approach that attempts to merge syntactic considerations with pragmatic ones, but which necessitates further empirical testing.

2. Background

2.1 *Framing the Problem*

In generative frameworks, the interrogative syntax of partial questions is modelled as a relation between the complementizer projection, i.e., the seat of the illocutionary force of a sentence, and the *wh*-phrase. Concerning the French *in-situ* and *ex-situ* alternation, the central question that has been debated is whether this relation involves true optional movement or only a syntactically restricted conditioned alternation explainable in a grammar where movement is not optional, but is either enforced or blocked.

A partial answer to this question comes from the observation that in comparison to their moved counterpart in (2a), in situ-questions such as (2b) are subject to syntactic constraints restricting their distributions in a number of environments. First, while ex-situ questions are compatible with certain forms of inversion, such as clitic inversion (3) and so-called complex inversion (4), in-situ questions are not compatible with either. Similarly, while ex-situ questions are compatible with the grammaticalized clefted *est-ce-que* form, in-situ questions are not (5).² Finally, while ex-situ forms are found as embedded questions, in-situ questions are not (6).² Note furthermore that this main vs. embedded clause restriction on in-situ questions remains even if the question itself is not directly embedded under a selecting verb but is instead linked to a pronominal clitic that represents the proposition, as in (7).

- (2) a. OÙ tu vas?
 where you go
 ‘Where are you going?’
 b. Tu vas où?
 You go where
 (lit.) ‘You are going where?’
- (3) a. OÙ vas-tu?
 b. *Vas-tu où?
- (4) a. OÙ Marie-a-t-elle mis mes affaires?
 where Marie has-she put my things
 ‘Where has Marie put my things?’
 b. *Marie a-t-elle mis mes affaires où?
- (5) a. OÙ est-ce-que tu vas?
 where ESK you go
 ‘Where are you going?’
 b. * Est-ce-que tu vas où?

² The constraint against embedded in situ questions may not be observed in all dialects of French or all languages allowing optional in situ-wh questions. For instance, in-situ forms such as (i) are apparently acceptable for speakers of Quebecois French as well as for younger French speakers, although they are not grammatical for the speakers we consulted, and do not appear in our corpus.

(i) Je sais c'est quoi tu vas faire, toi, avec ta vie. Tu vas rester ici dans cette cité pourrie.

Note that such forms, however, are quite restricted. They are most commonly formed with the verb *be* and embedded under the verb *know*, often with negation. (je sais pas c'est quoi). See Ledegen forthcoming for a discussion of these forms.

- (6) a. Je me demande où tu vas.
 I self ask this that-he does
 ‘I wonder what he is doing.’
 b. *Je me demande tu vas où.
- (7) Je me le demande, où tu vas
 *Je me le demande, tu va où

If we abstract away from the cleft, the ungrammaticality of (5b) parallels that of (3b) and (4b), since all involve a form of inversion. Indeed, the non-inverted cleft form is compatible with a version of in situ, as is shown in (8a) though not with (8b).

- (8) a. C’est où que tu vas?
 it is where that you go
 ‘Where are you going?’
 b. *C’est que tu vas où?

In sum, the simple fact that the syntactic contexts where in situ questions are possible are a subset of those where ex-situ forms can occur suggests that the latter is the default form for asking questions in French. Moreover, the in-situ form appears to be a relatively recent innovation. According to Larrivée (2019), although a few examples have been attested since the 15th century, the in-situ form did not start to appear until the 18th century (or perhaps even later according to Barra-Jover (2004) who notes that they remained uncommon even in the popular idiolects of the early 19th century). Given these observations, it is no surprise that in the generative framework, as in other approaches, the ex-situ form is taken as the main form of French interrogative and the in-situ one as the secondary one. From this viewpoint, the existence of the alternation raises the question of what could license the in-situ form rather than the other way around.

2.2 *Syntax and Intonation*

An oft cited analysis of wh-in situ, whose goal was to account for the syntactic restrictions observed above, is that of Cheng & Rooryck (2000). Starting from the assumption that the wh-in-situ requires particular licensing, they argued that in similar to in-situ languages like Korean or Japanese, which exhibit a special sentence final particle to mark interrogability, French uses

a final rising tonal morpheme \uparrow . This morpheme is said to be comparable to the tonal interrogative morpheme of declarative polar questions (9).

(9) Is it raining \uparrow ?

According to this view, the tonal particle \uparrow is also pragmatically associated with a strong presupposition. This strong presupposition is claimed to render negative answers infelicitous, as in their example reproduced here in (10).

- (10) A: Marie a acheté quoi?
Marie has bought
(lit.) ‘You bought what?’
B: ??Rien.
nothing
‘Nothing.’

Both claims, however, have since been challenged in the literature (see Boeckx 2000; Mathieu 2004; Déprez et al. 2013, among others). As Mathieu (2004) has illustrated, there are corpus examples with negative answers to in-situ questions that clearly demonstrate that negative responses are fully acceptable for French speakers. As for the proposal that in-situ questions are licensed by a rising final tonal morpheme, it has been shown that the syntax/prosody link is far more complex since many in-situ questions do not manifest a final rise. In a production study, Déprez et al. (2013) in fact found a bimodal distribution, in which some participants exhibited a final rise in their in situ questions, but others did not. This speaker variability is incompatible with strict syntactic licensing. In addition, they observe that the sentence-final rise of in wh-in situ questions differs from that of declarative polar questions, with the latter manifesting a more constant and considerably higher final rise. Thus, while Déprez et al. find support for Cheng & Rooryck’s claim that wh-in situ questions tend to have rising intonation, they paint a more nuanced picture, such that (i) minimally, the Q-morpheme in yes/no and in situ questions must differ, and (ii) while the final rise could be a characteristic of wh-in situ, it cannot be its defining feature. Other syntactic accounts (e.g. Boskovic 2000, Shlonsky 2012) involve more technical solutions to the distributional asymmetry of the two options that all ride on specific reformulations of the relation between the complementizer and wh-phrase. While

the distinct syntactic proposals and their success differ, they all share the common assumption that both in-situ and ex-situ questions have largely the same interrogative syntax and the same semantics. Most are concerned with licensing the in-situ option in a grammar that is characterized as triggering movement, and not with the factors that could bring a speaker to choose one form over the other. This choice is seemingly left to the pragmatics.

2.3 *Information Structure and Pragmatics*

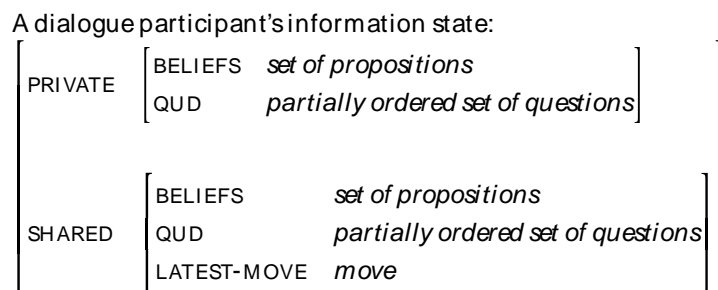
A number of pragmatic approaches to wh-in situ have focused on the role of information structure. For example, Pires & Taylor (2007) argue that wh-in situ is favored when the answers to what is being asked about are already in the Common Ground (GC) and hence shared by the speaker and the addressee. They claim that there is both an existential presupposition of the content of the question (*What did you buy?* presupposes $\exists x$. you bought x) and that the set of possible answers to the question must already be in the CG. Other approaches have taken on similar assumptions—namely that wh-in situ is associated with backgrounded material (see Mathieu 2004), or given material (Hamlaoui 2008, 2009). Déprez et al. (2012) propose to associate Givenness with fronting of the TP (i.e., the non-wh portion of the in-situ question) to Spec. CP following Pollock & Poletto (2004). On this view, in-situ questions involve Givenness-motivated fronting (Kučerová 2007) of the remnant proposition, and not of the wh-phrase. There is thus a consensus among researchers that information structure should matter for licensing in-situ, but specific aspects of the information structure often remain ill-defined (e.g., what counts as a 'strong' presupposition), and/or do not predict when wh-in situ may be used by speaker as opposed to the ex-situ form. There is little connection between the pragmatics and the syntactic realization.

Moreover, the CG-based approaches, particularly that of Pires & Taylor (2007) stand in conflict with the general assumption in the literature that French wh-in situ questions are information-seeking questions. Indeed, if the possible answers to the question are already in the CG, and hence known by the speaker and the addressee, it is unclear what information is being queried. Furthermore, claiming that there is a (set of) possible answers for an wh-in situ question does not by itself adequately distinguish in-situ from ex-situ.

Offering a different approach, Engdahl (2006) and Myers (2007) argued that speaker-addressee pragmatics contributes to the choice of a wh-in situ in French. For Myers (2007), *answerability*, defined in terms of the strength of a speaker's belief in the addressee's capacity to (easily) answer the question, has a crucial role in favoring in situ. Engdahl (2006) further refines this intuition; focusing on how information is “packaged” across discourse participants

in questions, Engdahl models the mental states of each of the participants in the conversation at a given time with dialogue boxes as in Figure (1).

Figure 1. Engdahl (2006) Information State Model



Within such a model, Engdahl (2006) proposes to represent Myers (2007) answerability proposal by claiming that the in-situ option is selected whenever the speaker is confident that the addressee has the answer. The proposal differs from Pires & Taylor (2007), because here the answer is not assumed by the Speaker to be in the shared common ground, but is rather part of the Addressee's set of beliefs. Engdahl's model hence makes clear where the answerability factor lies; to put it more clearly, in situ should be favored whenever the Speaker perceives the Addressee as an *expert* that possesses the knowledge to answer the question. This model also makes clear predictions about the Addressee's expertise: the more expertise an addressee is perceived to have by the speaker, the more likely in-situ is to be the preferred form.

2.4 Variation in Wh-In Situ Rates

Previous research hence suggests that the choice of a question form involves a rather complex interaction between syntax, prosody and discourse factors, including the relationship between the Speaker and the Addressee. Another component, often mentioned by French speakers, is the importance of the social discourse setting, or register that is associated with wh-in situ. Specifically, wh-in situ is thought to be a hallmark of more colloquial French, which is why many researchers have confined corpora studies to oral ones, or have discussed non-standard dialects of French (see Quillard 2001; Myers 2007; Hamlaoui 2009; among others). Looking at the variation in rates of in-versus ex-situ form through a number of corpora studies, however, evidence for this register differentiation is not as straightforwardly obvious as believed; rates in adult spontaneous production, while higher than in written settings, still range from about a third of all wh-questions (Myers 2007) to more than half of all wh-questions produced (Hamlaoui 2009). The only clear pattern that emerges is that spontaneous production elicits

more wh-in situ. The exact role of colloquial vs formal discourse registers hinted at in the literature has not been systematically investigated in a controlled fashion.

Earlier corpus studies investigated the question of the relative frequency of a variety of question forms. In an early meta-analysis of a variety of corpora, Coveney (1996) showed that the frequency of ex-situ vs in-situ questions was highly variable depending on the type of corpus examined. He also found the in-situ form to be overall less frequent than the ex-situ ones.

Table 1. Meta-Analysis from Coveney (1996)

Type de public (chercheur)	%	Nb d'exemples
Paris - classe moyenne (Ashby)	38,8	85
Classe moyenne, registre informel (Behnstedt)	33	446
Classe moyenne, registre formel (Behnstedt)	25	4,367
Québec - adolescents (Fox)	24,5	805
Somme classes moyenne et ouvrière (Coveney)	15,6	122
Journalistes télévision belges (Lafontaine et Lardinois)	14,9	230
Montréal - enfants/adolescents (Lefebvre)	13	433
Enfants de neuf ans (Söll)	12,9	364
Classe ouvrière (Behnstedt)	12	587
Classe moyenne (Pohl)	10,3	155
Classe ouvrière (Pohl)	8,7	69
Couple âgé belge (Pohl)	2	184

However, in a meta-analysis of more recent corpora that included more oral productions, we see that the relative frequency seems to be more uniform, resulting in essentially a 50/50 split between in-situ and ex-situ.

Table 2. In-Situ and Ex-Situ in Oral Corpora

Authors	Total wh-Q	In-Situ	Ex-Situ
Quillard (2001)	633	43.7%	56.3%
Myers (2007)	189	33.3%	66.6%
Hamlaoui (2009)	175	59.6%	40.3%
Boucher (2010)	388	25%	75%
Bonan & Tual (2016)	344	54.9%	45.1%
TOTAL	1734	42.2%	57.7%

Finally, the most recent corpus analysis comes from the dissertation work of Glasbergen (2021), who analyzed a corpus of productions from French reality shows. Glasbergen was interested in all the different kinds of questions that can be found in French, but if we collapse in-situ and ex-situ options together, the rates for both are equivalent. As a result, this corpus-analysis, like the others before, essentially shows an equal distribution for these two forms. If we attempt to answer which of these options should be considered the base form, these studies are uninformative. From the point of view of frequency, either of these forms could do just as well. However, recently Faure and Palasis (2021) have argued that one French grammar, an informal one, has in-situ as its core base interrogative form with the ex-situ form being derived, while another French grammar, a formal one, manifests the opposite pattern. If this proposal is correct, we expect to find that in different registers, the proportion of in-situ vs ex-situ question is significantly distinct, such that the ex-situ option is favored in the formal register, but in-situ is favored in the informal one. For this to be testable corpora of essentially the same size from the formal register would need to be compared to an informal one and relative in-situ and ex-situ frequency compared. This is what the following corpus study aimed to achieve.

3. Corpus Analysis

We consulted the Enquêtes Sociolinguistiques à Orléans (ESLO) database (Abouda & Baude 2005, 2006; Baude & Dugua 2011; Eshkol-Taravella et al. 2011), an online collection of corpora of recorded conversations. ESLO is further subdivided based on the time frame in which the recordings were made; ESLO1 corresponds to recordings from 1968-1971, and ESLO2 corresponds to recordings from 2008 or later. These corpora feature speakers in a variety of discourse settings and situations, which range from some that are relatively informal

(e.g. conversations at meals, the market, or the cinema) to some that are more formal in nature (e.g. interviews and conferences). The goal of our analysis is to determine if a difference in formality results in a difference in the rates of *wh*-in situ, and whether we find empirical support in corpus studies for a diglossia concerning the grammar of questions. We focus first on ESLO1 and ESLO2 separately before comparing rates of in-situ across both sets of corpora.

3.1 Procedure

We selected a formal and informal corpus from ESLO1 and ESLO2 for analysis: Repas and Entretien from ESLO1, and Repas and Diachronie from ESLO2. In order to create samples of equivalent length for the purpose of comparison, we analyzed approximately 1000 lines from each corpus. We made classifications with respect to formality primarily based on the description of the discourse setting, as provided in the database. We also compared *tu/vous* counts, as an additional measure of formality. In French, there are two second person singular subject pronominal clitics: *tu* is informal, while *vous* is formal. We compared the rates of informal pronouns within the samples taken from the corpora we had selected. Because the second person plural pronoun in French is also *vous*, it was the rate of the informal pronoun *tu* that allowed us to gauge formality. As seen in Tables 3-4 below, informal pronouns are found more often in what we had categorized as the informal corpora.³

Table 3. Categorization (ESLO1)

ESLO1	Repas 1 (Informal)	Entretien (Formal)
Discourse Setting	mealtime interactions	interviews
# of Informal Pronouns	42	0

Table 4. Categorization (ESLO2)

ESLO2	Repas 2 (Informal)	Diachronie (Formal)
Discourse Setting	mealtime interactions	interviews
# of Informal Pronouns	40	5

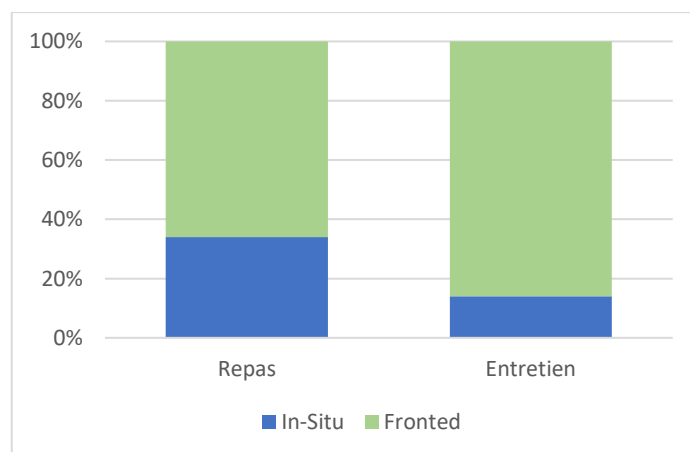
³ Tables 3-4 contain the counts of informal pronouns in the *wh*-questions we analyzed alone, and not for the entire sample. Overall, the informal pronoun is found more often in Repas1-2.

For the analysis, we started by searching for all questions within these corpora samples. We coded by type of question (e.g. yes/no questions, fragment questions, wh-questions). We excluded subject wh-questions for the purpose of this analysis, as there can be no distinction in root subject wh-questions between in-situ and ex-situ forms. We also exclude why-questions, which can only be fronted. All remaining wh-questions were categorized as being either fronted or in situ. This allowed us to then compare the proportion of wh-in situ between the formal and informal corpora, both separately within ESLO1 and ESLO2, and then overall.

3.2 Results

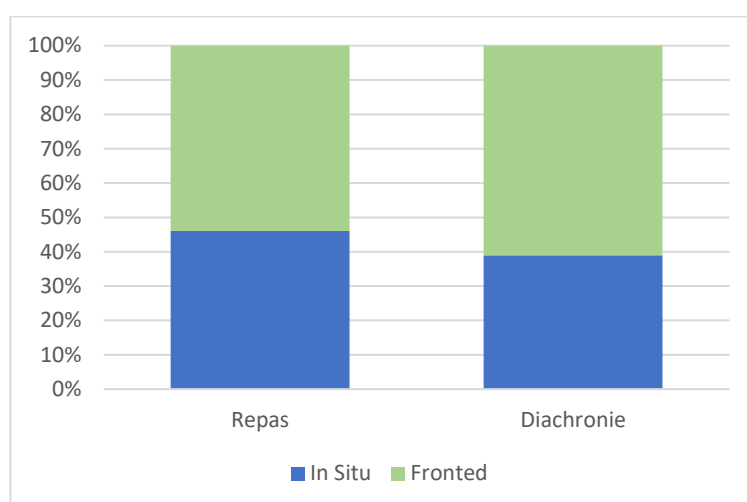
Overall, fronted wh-questions were produced more often in all corpora. We return to a discussion of this point after detailing further our results. Comparing the rates of in situ between Repas and Entretien in ESLO1, we confirmed that wh-in situ is much more frequent in the informal context; 34% of wh-questions in Repas (informal) are in situ, as compared to 14% in Entretien (formal). We analyzed these results with a mixed effects logistic regression model, with Formality as the dependent variable (with two levels: formal, informal) and wh-phrase as the random intercept. There is a significant main effect of Formality ($p < 0.01$), with speakers producing more in-situ in informal contexts.

Figure 2. ESLO1 Results



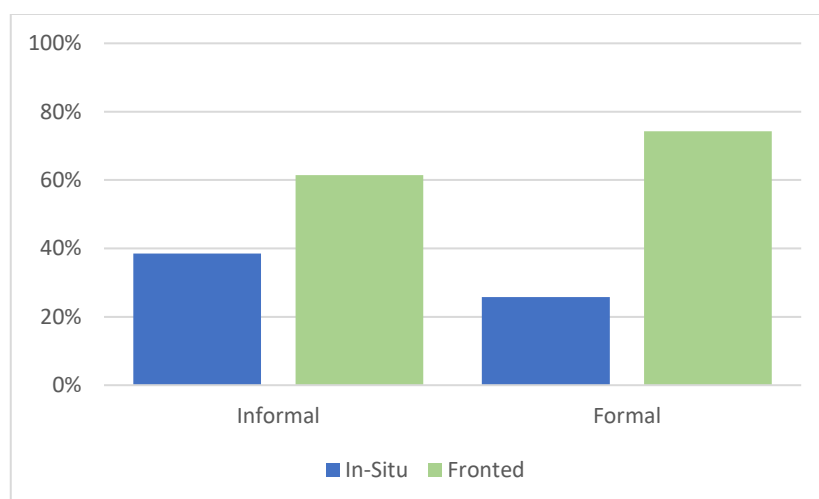
Comparing the rates of in situ between Repas and Diachronie in ESLO2, we find the same trend, with more in situ questions in the former corpus (see Figure 3), consistent with ESLO1. However, the difference (45% vs. 39%) is not as pronounced in ESLO2, and is not statistically significant as determined via the same logistic regression model.

Figure 3. ESLO2 Results



If we collapse ESLO1 and ESLO2 together, with Formality and Corpus as the dependent variables, and wh-phrase in the random effects structure, we find that there is a significant effect of formality ($p < 0.01$) and corpus ($p < 0.01$). There is also a significant interaction of Formality*Corpus ($p < 0.05$).

Figure 4. Overall Results (Formality)



When considering all of these results, we see that wh-in situ questions are produced more often in informal contexts, despite the fact that fronted wh-questions remain overall the preferred option within and across these corpora.

3.3 Discussion

In this analysis, we confirmed quantitatively that rates of *wh-in situ* increase in informal contexts, where speakers have casual discussions with one other. In all contexts, fronted *wh*-questions are preferred and relatedly neither form is either “required” or infelicitous in a given interactive situation. Lastly, rates of *wh-in situ* may be shifting over time. While it has been the case that *wh-in situ* has intuitively been seen as an informal choice, this research quantitatively confirms that formality is indeed a factor— but it simultaneously suggests that the picture is more nuanced. That is, although it is possible to predict that *wh-in situ* will be *more common* in informal contexts, the choice by a speaker in any discourse situation always remains, as *in situ* is never a stigmatized form⁴. The choice between the moved and *in-situ* options is thus not reducible to a parameter that would respectively characterize two different grammars for the two separate registers, *in-situ* for an informal grammar and *ex-situ* for a formal one.

In all corpora, formal and informal, old and new, fronted *wh*-questions are found more often. Fronted questions thus appear to be the default form. In effect, there is no evidence here for a register in which the *wh-in-situ* form would be either the most common one or the base form as hypothesized by Faure and Palasis (2021). Perhaps more importantly, the results we observe turn out to cast doubt on accounts that attempt to attribute production of *wh-in situ* to a single factor: controlling for discourse situations demonstrates that formality is a factor, but it alone does not account for when or why *in situ* can be favored. Conversely, it is also the case that *wh-in situ* is never prohibited based on discourse context alone, as it is found in both formal corpora that we consulted, as well as the various corpora that Coveney (1996) consulted and even in written texts (albeit less frequently (Boucher 2010)).

The results from ESLO2 are intriguing because, while they conform to the general pattern, such that *wh-in situ* is more common in informal contexts, the difference in rates of *in situ* production between the corpora is not as pronounced. There are a few reasons for why this might be. It could be that this sample size is just not large enough to highlight the difference between these corpora. Another possibility, however, is that the French language is evolving; *wh-in situ* may be used more now than it was between 1969 and 1974, when ESLO1 was recorded. In other words, it could be that *wh-in situ* was originally only produced in more informal contexts, but that the language has shifted causing it to become felicitous in formal

⁴ It did not turn out to be statistically feasible to cross *wh in-situ* usage with the use of the familiar pronouns and the discourse situation, as the counts of the informal *tu* were too low in the formal corpus to allow for a balanced statistical comparison.

contexts, and increasingly favored overall. We are unable to answer this with these results from relatively recent corpora alone, but a diachronic perspective aligns well with findings by Larrivée (2019) who notes an increase in the frequency of wh-in-situ forms of *comment* questions from the 18th to the 20th century in the written French Frantext corpus.

The influence of discourse register is confirmed with these results, but there are clearly other factors motivating wh-in situ that are independent of the formality of the discourse setting. To this end, we also conducted an experiment to explore the speaker-addressee relation, which has also been argued to influence the choice of question form (Engdahl 2006; Myers 2007; see also Section 2). Such factors cannot however be captured in corpus data where the knowledge state of the interlocutor is not accessible or manipulable. For this reason, we turned to conducting an experiment to observe such effects through active context manipulations.

4. Experiment

This experiment was designed to explore the importance of speaker-addressee pragmatics on the choice between in-situ and ex-situ. We sought to control for the familiarity between speakers, while manipulating the context of a question so as to emphasize (or not) the expertise of the addressee—we define expertise here as perceived capacity of the addressee to answer the question, based on Myers (2007) and Engdahl (2006). The corpus data we considered above cannot easily disentangle the effect of an informal context from that of speaker familiarity with one another, because it is not always clear what kind of relation the discourse participants entertain. Hence it could turn out that the discourse context alone does not suffice to trigger wh-in situ if both speakers are relatively unknown to one another. Here we attempted to test the answerability criterion of Engdahl (2006) and Myers (2007). On this view speakers are expected to produce more in situ questions when they believe the question to be highly “answerable”—that is, when speaker have good reasons to believe that the answer is part of the addressee’s knowledge base.

4.1 Stimuli and Design

We designed a forced-choice task that had a 2x2 factorial design, in which we manipulated familiarity (familiar vs. non-familiar) and addressee knowledge state or expertise (expert vs. non-expert). Since manipulation of the knowledge base of the participants has to be indirect, we provided participants with specific discourse contexts. In these contexts, interlocutors were labelled as friends or family members (familiar contexts), and as strangers or professionals,

such as a train conductor (unfamiliar contexts). Addressee knowledge was emphasized in the brief scenarios provided before participants are tasked with making a selection; the addressee was designated as competent professionals with respect to the question asked, or as being no more knowledgeable than the speaker. After reading the brief scenario, participants saw both a fronted wh-question or a wh-in situ question, and had to determine which form they thought was most natural for that context. The storyboard below illustrates the four different conditions that were designed to take place in the same discourse setting. The upper left corner illustrates the Unfamiliar-Expert condition (here the addressee is a professional), the bottom left corner corresponds to the Familiar-Expert condition (colleagues), the upper right corner corresponds to the Unfamiliar-Non-Expert condition (strangers), and the Lower right corner the Familiar-Non-Expert condition (a couple). The scenarios provided made these relations clear. Samples are illustrated in (11) to (14) below, each one highlighting a different scenario and condition.

Figure 5. Representative Experimental Images



(11) **Unfamiliar-Expert**

S'adressant à un chef de gare sur le quai d'un
 REFL-address to a boss of station on the platform of-a
 train qui arrive en gare :
 train who arrives in station

‘Addressing a station master on the platform for a train that is arriving at the station:’

- a. Pardon, J’ai besoin d’aide. Il va où, ce train ?
 Pardon, I-have need of-help it goes where, this train
 (lit.) ‘Excuse me, I need help. This train is going where?’

- b. Pardon, j'ai besoin d'aide. Où il va, ce train ?
 Pardon I-have need of-help where it goes this train
 'Excuse me, I need help. Where is this train going?'

(12) **Unfamiliar-Non-Expert**

Dans un abribus à un inconnu qui attend aussi :
 in a bus-shelter to an unknown who waits also
 'In a bus shelter, to a stranger who is also waiting:'

- a. Bonjour, il part quand le prochain bus 23?
 Hello it leaves when the next bus 23
 (lit.) 'Hello, the next 23 bus leaves when?'
- b. Bonjour, quand part le prochain bus 23?
 Hello when leave the next bus 23
 'Hello, when does the next 23 bus 23 leave?'

(13) **Familiar-Expert**

À un ami décorateur qui vous conseille pour
 to a friend decorator who you consult for
 rénover votre appartement :
 to-renovate your apartment
 'To a designer friend who you are consulting to renovate your apartment:'

- a. On peut mettre quelle couleur pour éclaircir une pièce?
 one can put which color for to-brighten a place
 (lit.) 'One should paint with which color to brighten a place?'
- b. Quelle couleur on peut mettre pour éclaircir une pièce?
 which color one can put for to-brighten a place
 'Which color should one paint with to brighten a place?'

(14) **Familiar- Non-Expert**

Une fillette à sa mère qui vient de rentrer à
 a daughter to her mother who come of to-return to
 la maison :
 the house
 'A daughter to her mother who has just returned home:'

- "Maman, la femme de ménage a encore rangé
 mom the woman of cleaning has again put-away
 mon nounours."
 my teddy-bear
 'Mom, the housekeeper put my teddy bear away again.'
- a. Elle l'a mis où cette fois?
 she it-have put where this time
 (lit.) 'She put it where this time?'
- b. Où elle l'a mis, cette fois?
 where she it-have put this time
 'Where did she put it this time?'

There were 4 target questions for each of these conditions, for a total of 16 target questions. The experiment also contained 16 fillers utilizing declarative sentences with either floated vs. non-floated quantifiers (*Tous les enfants aiment les brocolis.* vs *Les enfants aiment tous les brocolis.* (Eng. *All the children like broccoli.* vs. *The children all like broccoli.*)) or with negative quantifiers vs. negated indefinites (*Ils n'ont rien dit.* vs *Ils n'ont pas dit quoique ce soit.* (Eng. *They said nothing* vs *They did not say anything.*)). The goal was to provide distractors that compared essentially synonymous constructions with different forms, as a parallel to the in-situ and ex-situ wh-questions. Each participant saw a total of 32 questions overall. We designed the study entirely online via Ibex Farm, and we recruited adult participants (N= 82) through the Facebook page of the Institute for Cognitive Sciences-Marc Jeannerod UMR 5304 in Lyon, France. The participants received a direct link to the study, which had been distributed by the ISC to ensure that only native French speakers participated.

4.2 Results

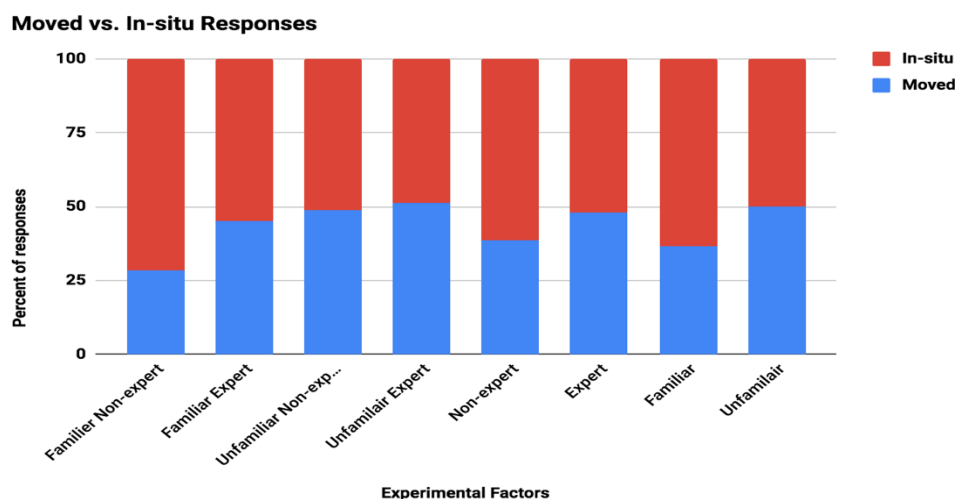
We found that throughout the experiment the participants selected wh-in situ questions slightly *more* often than fronted ones, with the former being selected 57% of the time, collapsing across conditions.

Table 5. Overall Response Pattern (Fronted vs. In Situ)

	Fronted	Wh-In Situ
All Conditions	43% (569)	57% (743)

The results of the four different conditions are reported below in Figure 6. We report choice as a function of each of the factors manipulated (i.e. rates of in situ for (non)-expertise, collapsing across familiarity, and rates of in situ for (un)familiarity, collapsing across expertise). For three of the four conditions (Familiar-Expert, Unfamiliar-Expert, Unfamiliar-Non-Expert), the percentage of fronted and in situ choices is around 50%, suggesting that both question forms are equally natural for participants in these scenarios. However, in the Familiar-Non-Expert condition, there is a strong preference for in situ, which is selected 72% of the time by participants.

Figure 6. Percentages of In-Situ vs. Fronted (Ex-Situ) Selections in Relation to Factors Examined



Overall, there was a significant effect of familiarity, $\chi^2=23.5$ (1), $p < 0.01$, as determined via a chi-square test. Participants selected more wh-in situ questions whenever the speaker was familiar with the addressee in some capacity (see also Table 6). Expertise was also a significant factor, $\chi^2=12.3$ (1), $p < 0.01$. Participants selected the wh-in situ option more often when their addressee was a non-expert. However, as participants were just as likely to select an in situ question when the addressee was perceived as an expert, regardless of the relationship between the speaker and the addressee in the scenario.

Table 6. Percentage of Wh-In Situ Responses Per Condition

	Familiar	Non-Familiar	TOTAL
Expert	55% (180)	49% (160)	52% (340)
Non-Expert	72% (235)	51% (168)	61% (403)
TOTAL	63% (415)	50% (328)	

4.3 Discussion

In this experiment, we see the reverse pattern from the corpus-based analysis, in that participants selected more in situ questions overall, when we collapse across conditions. This is initially surprising, considering the corpus findings. However, closer inspection revealed that participants only preferred wh-in situ in one particular context—when the speaker and addressee know each other, and the addressee is not perceived to be an expert. Otherwise, the choice between the two types of questions remains about equal. Although French preference for wh-in situ may shift depending on the discourse, fronted wh-questions always remain an option.

These results also suggest that a *combination* of factors is responsible for in situ licensing: both familiarity and (non-)expertise prompt wh-in situ. The findings thus stress the importance of the nature of the interaction between interlocutors on wh-in situ, and point to a subtle and complex interaction of pragmatic factors governing the preference for in situ in a given scenario. Taken together with the results from the corpus analysis, we confirmed that French speakers are more likely to produce wh-in situ questions in informal contexts, and less likely do so when in formal contexts with people they do not know. In other words, the formality of a given register should be understood as the convergence of a specific social discourse situation and a privileged interlocutor relation. These findings provide support for approaches to wh-in situ that have underlined the importance of discourse factors, while pointing to the need for a finer grained approach to the understanding of question choice that looks not just for the role of particular factors but rather for their combined effects.

The role of addressee expertise in question answerability seems a bit more difficult to understand. As mentioned, while Engdahl (2006) and Myers (2007) have both suggested that the speaker's perception of the addressee's knowledge should matter, and that in situ should be more permissible when the addressee is expected to be more knowledgeable. Our results do not align with this proposal. We find an increase in the choice of fronted wh-questions when the addressee is portrayed as an expert. There are several possibilities for why this might be the

case. First, it could be that the participants overlooked the expert status of the addressee which in our experiment was built in the contexts provided before the question. That is, our experiment manipulated answerability by foregrounding in the context the addressee's expertise. It could be argued that this manipulation was too subtle; as expertise is manipulated in the context and not directly in the question, it may not be obvious to our participants that the friend or stranger in these target scenarios is or is not an expert. Experts were also generally depicted as qualified professionals, which may have interfered with familiarity. Our design did not include a control to verify that this factor was taken into account by the participants. Yet, although we have no grounds to reject this possibility, ignorance of a factor usually results in the lack of its statistical significance, and seems rather unlikely to lead to the statistically significant reverse pattern that we observed. Another possible explanation then, is that our results here effectively argue against the claims that in-situ is favored when the speaker believes that the addressee knows the answer, that is, against the answerability factor of Myers (2007) and Engdahl (2006). This raises new questions as to why expertise should favor ex-situ. Gunlogson (2001) argues that in interrogatives, the authority to update the context is assumed to reside with the addressee, who moves the conversation along. This is consistent with the addressee being the seat of knowledge. In declaratives, this authority resides with the speaker. Moreover, Beyssade & Marandin (2006) have suggested that French *wh-in situ* questions resemble declarative polar questions in associating illocutionary force with a declarative syntax, while the illocutionary force aligns with an interrogative syntax for ex-situ forms. A possible way to understand our results then, is to speculate that while they contradict the Myers-Engdahl hypothesis, they could support the view that a perceived increase in the addressee knowledge in fact increases the “interrogativity” of a speech act, as expected under Gunlogson (2001). This potential connection to Gunlogson and its consequences for French in situ is speculative, and will be left for further research (see Loder & Déprez forthcoming).

5. Conclusions

In this research, we have investigated the force of several discourse factors that have been proposed to influence the choice of the *wh-in situ* question form in French in two empirical ways, through corpus and experimental research. Our comparative corpus research did not provide evidence to support a diglossia perspective for partial questions in French in which one register would feature the moved *wh*-form as a base form and the other the in-situ form. For both of the registers we compared, the ex-situ form was the dominant one, even if there appears

to be an evolution towards an increase of the in-situ form between the earlier ESLO1 corpus and the later ESLO2. Our corpus research provides quantitative evidence that the formality of a register is a factor in favoring wh-in-situ. Additionally, we argue that the social situation of the interaction and the relation between Speaker/Addressee both influence the likelihood of in-situ. These results confirm that the choice of a question form is not driven by a single factor but by a complex interaction of multiple ones. Nevertheless, the delicate components of this interaction remain to be better understood.

There are also still important open questions. We were not able to address how all of the factors that we explored interact with one another; the prediction is that the convergence of informal discourse contexts, in which speakers also know each other, would elicit more wh-in situ than informal contexts alone, for example. Unfortunately, we are unable to control for this with the corpora consulted, but we are attempting to do so in future research. As discussed in Section 3.3, there is also a question of diachronic shift in terms of overall frequency of wh-in situ over time. While it is possible that variation in rates is simply a result of the samples selected, the fact that there is an increase in wh-in situ between ESLO1 and ESLO2 suggests that there may be change in progress. Lastly, this research underscores the fact that fronted wh-questions are always an option in French, regardless of the discourse context. It does not seem to be the case that ex-situ and in-situ questions should be categorized in terms of a complementary distribution; rather, the preference for wh-in situ seems to be tied to a complex set of interacting, convergent factors. It is likewise possible that manipulating each individual factor alone is never sufficient for the fronted option to become infelicitous. This suggests that there may in fact be a degree of optionality concerning wh-in situ questions in French— but that it could still be possible to predict when wh-in situ is more likely to be selected than not.

References

- Abouda, L. and Baude, O. (2005). Du français fondamental aux ESLO. Grand corpus de français parlé, Bilan historique et perspective de recherche. *Cahiers de linguistique* 33 (2), 131-146.
- Abouda, L and Baude, O. (2006). Constituer et exploiter un grand corpus oral: Choix et enjeux théoriques. Le cas des ESLO. *Corpus en Lettres et Sciences Sociales, Des documents numériques à l'interprétation*. [halshs-01162506](https://halshs.archives-ouvertes.fr/halshs-01162506)
- Barra-Jover, M. (2004). Interrogatives, negatives, and the evolution of verbal formal features in spoken French. *Langue Française*, 141(1), 110-125.

- Baude, O. and Dugua, C. (2011). (Re)faire le corpus d'Orléans quarante ans après: quoi de neuf, linguiste? *Corpus* 10, 9-118.
- Beyssade, C., & Marandin, J. M. (2006). The speech act assignment problem revisited: Disentangling speaker's commitment from speaker's call on addressee. *Empirical Issues in Syntax and Semantics*, 6(37-68).
- Boeckx, C. (2000). Decomposing French questions. In *University of Pennsylvania Working Papers in Linguistics* 6.1.
- Bonan, C. & Tual, L. (2016). Interrogative sentences in French syntax and prosody. SynCart Workshop.
- Boskovic, Z. (2000). Sometimes in SpecCP, sometimes in-situ. In R. Martin, D. Michaels, & J. Uriagereka (Eds), *Step by step: Essays on minimalism in honor of Howard Lasnik*, MIT Press: Cambridge, MA, 53-87.
- Boucher, P. (2010). L'interrogation partielle en français : l'interface syntaxe / sémantique. *Syntaxe et Sémantique*. 55-82. DOI 10.3917/ss.011.0055.
- Cheng, L. L., and Rooryck, J. (2000). Licensing Wh-in situ. *Syntax* 3.
- Coveney, A. (1996). *Variability in spoken French : A socio-linguistic study of interrogation and negation*. Elm Bank Publications.
- Déprez, V., Syrett, K, and Kawahara, S. (2012). Interfacing information and prosody: French wh-in situ questions. In *Proceedings of Going Romance 2010*. Amsterdam: John Benjamins.
- Déprez, V., Syrett, K., and Kawahara, S. (2013). The interaction of syntax, prosody, and discourse in licensing French wh-in situ questions. *Lingua* 124.
- Déprez, V. (2018). On en est où? Intonation, information structure, and the embedding problem. Talk at GEN-wh2018. Geneva.
- Engdahl, E. (2006). Information packaging in questions. *Empirical issues in syntax and semantics* 6.1.
- Eshkol-Taravella, I., Baude, O., Maurel, D., Hriba, L., Dugua, C. and Tellier, I. (2011). A large oral corpus "available": The corpus of Orleans 1968-2012. *Ressources linguistiques libres* 52:17-46.
- Frantext. Base textuelle Frantext. <http://www.frantext.fr/>
- Faure, R., & Palasis, K. (2021). Exclusivity! Wh-fronting is not optional wh-movement in Colloquial French. *Natural Language & Linguistic Theory*, 39(1), 57-95.
- Glasbergen, A. (2021). Questions in context: The case of French wh-in situ. LOT dissertation series. LOT, Amsterdam.

- Gunlogson, C. (2001) . True to Form : Rising and Falling Declaratives as Questions in English. [Doctoral dissertation, University of California, Santa Cruz].
- Hamlaoui, F. (2009). *La focalisation à l'interface de la syntaxe et de la phonologie: le cas du français dans une perspective typologique*. [Doctoral dissertation].
- Kučerová, I. (2007). *The Syntax of Givenness*. [Doctoral dissertation, MIT].
- Larrivée, P. (2019) Historical pragmatics, explicit activation and wh in situ in French. In *Romance Languages and Linguistic Theory 15: Selected papers from 'Going Romance'* 30. John Benjamins. DOI: [10.1075/rllt.15.06lar](https://doi.org/10.1075/rllt.15.06lar)
- Mathieu, E. (2004). The mapping of form and interpretation: The case of optional wh-movement in French. *Lingua* 114 (9-10).
- Myers, L. L. (2007). WH-interrogatives in spoken French: A corpus-based analysis of their form and function. [Doctoral dissertation].
- Pires, A., and Taylor, H. (2007). The syntax of wh-in-situ and common ground. In *the Proceedings from the Annual Meeting of the Chicago Linguistic Society* (Vol. 43, No. 2).
- Poletto, C., and Pollock, J. (2004). On wh-clitics and wh-doubling in French and some North Eastern Italian dialects. *Probus* 16 (2), 241-272.
- Shlonsky, U. (2012). Notes on wh in situ in French. In L. Brugè, A. Cardinaletti, G. Giusti, N. Munaro, & C. Poletto (Eds.), *Functional heads. The cartography of syntactic structures* (pp. 242-252). New York: Oxford University Press.
- Tieu, L. (2012). Semantic-pragmatic conditions on wh-in-situ in English. Presentation at *the 86th Annual Meeting of the Linguistic Society of America, Portland*.
- Quillard, V. (2001). La diversité des formes interrogatives: comment l'interpréter? *Langage et société*, (1) 57-72.