

A contrastive perspective on French and Italian *wh-in situ* questions.

A discourse-pragmatic approach

Davide Garassino

University of Zurich & ZHAW, Winterthur

Abstract

This paper offers a qualitative and quantitative analysis of French and Italian *wh-in situ* questions based on spontaneous spoken data. A pragmatic analysis relying on two parameters, propositional activation and pragmatic function, reveals that the licensing conditions and the use of this structure largely differ in the two languages. While French *wh-in situ* do not require an activated proposition and can introduce a discourse-new topic, Italian *wh-in situ* mostly require an activated proposition and, at least in the analyzed corpus data, do not introduce discourse-new topics. An examination of the contexts also reveals that the different licensing conditions influence the interactional uses of these questions. Finally, we claim that both French and Italian *wh-in situ* require a pragmatic condition, which is their ‘anchoring’ to given, or at least inferable, information in the linguistic context (as is typical of Italian) or to predictable situations in the extralinguistic context (such as expected discourse moves in social interactions, as is the case for French).

Keywords: Romance linguistics, pragmatics, syntax and discourse, Question Under Discussion.

1. Introduction

From a typological perspective, some languages rely on a double strategy for formally encoding partial interrogatives (i.e., non-polar questions). They can either place the interrogative phrase in the same position as the corresponding non-interrogative phrase (*wh-in situ* questions; henceforth *in situ*) or they can move it to the left periphery of the sentence (*wh-ex situ* questions; henceforth *ex situ*). Such languages, however, differ in how they use the two strategies. For instance, English, which is a predominantly *ex situ*-language, as shown in (1), allows *in situ* only under special circumstances, as in (2):

- (1) When are you leaving?
- (2) a. You are leaving when?!?
b. Napoleon died in what year? (Dryer 2013: 378)

In (2a), an echo question, the speaker can either signal that she was unable to hear part of the interlocutor’s previous turn or she can express surprise or disbelief with regard to what she

has just heard. In (2b), the *in situ* can be used by a teacher to ask her students a question, to which she obviously knows the answer. Similarly, French and Italian are deemed to be *ex situ*-languages that allow *in situ* in specific contexts.

In this paper, we intend to compare French and Italian *in situ* questions from a discourse-pragmatic perspective (section 2), based on a corpus of spoken language. Following some recent studies on the use of this interrogative form in Brazilian Portuguese (Rosemeyer 2019), French (Larrivée 2019) and Spanish (Rosemeyer 2018), we mainly focus on two pragmatic parameters: activation status and pragmatic functions (section 3). This type of analysis is particularly compelling for Italian since the use of (non-echoic) *in situ* questions is understudied in this language.

The corpus analysis (section 4) reveals significant differences between *in situ* questions in the two languages as regards their frequency, licensing conditions and pragmatic functions. In French, *in situ* interrogatives occur much more frequently and show a greater functional flexibility (i.e. they do not require activation in discourse). In Italian, *in situ* questions are rarer and their propositional content is required to be activated in the previous context. The different licensing conditions have consequences on the pragmatic functions of the structure: in French, *in situ* can introduce new discourse topics, whereas in Italian they mostly elaborate on a previously introduced topic. Other functions (such as challenge and echo questions) are much less frequent in both languages. In the last part of section 4, we also offer a preliminary comparison between *ex situ* and *in situ* questions, in order to better appreciate the functional specificities of the latter.

Furthermore, we examine how the licensing conditions of information-seeking *in situ* questions influence their use in interaction (section 5). While *in situ* with activated propositional material tend to react to specific communication problems (i.e., to questions left unanswered in discourse), *in situ* with non-activated propositional content can be used to formulate ‘routine’ questions or to anticipate predictable discourse moves.

In conclusion (section 6), we claim that both French and Italian information-seeking *in situ*—despite the cross-linguistic differences—require a pragmatic condition. While in Italian, it

is the presence of a discourse-given or inferable propositional antecedent, in French it can also be information that is made available by the extralinguistic context. In numerous cases, new discourse topics introduced by French *in situ* are highly expected because of the situational context or world knowledge.

2. French and Italian *in situ*

French *in situ* can be used both as echo-question, as in (3), and information-seeking interrogatives, as in (4):

- (3) A: Mon fils, il lit [inaudible]
‘My son, he is reading [inaudible]’
B: *Il lit quoi ?*
he read.PRS.3SG what
‘Is he reading what?’ (Engdahl 2006: 104)¹
- (4) A: *Ton fils, il lit quoi ?*
your son he read.PRS.3SG what
‘What is your son reading?’
B: Des bandes dessinées.
‘Comic books.’ (Engdahl 2006: 104)

According to the literature, information-seeking *in situ* in French are subject to sociolinguistic, syntactic and pragmatic constraints.

Regarding their sociolinguistic distribution, Coveney (1996), relying on a quantitative survey of different corpora gathered in the 1960s and the 1970s, claims that *in situ* are mostly preferred in informal spontaneous speech (*français familier*; cf. also the discussion in Adli 2006:

¹ The abbreviations used in the glosses are based on the Leipzig Glossing Rules (Comrie *et al.* 2015).

169–171), but they cannot be considered typical of lower registers, such as *français populaire*. The data discussed in Boucher (2010) hint at a more complex situation, since the *in situ* in his data also occur in the written language as well as in numerous text genres. However, their distribution is biased overall towards orality and informality: for instance, while *in situ* represent 47% of all interrogative structures in informal conversations between friends (Boucher 2010: 106),² their frequency falls dramatically in formal speech.

At the syntactic level, some researchers have highlighted the incompatibility of *in situ* with contexts involving, among others, the presence of negation and modals (Cheng & Rooryck 2000: 11). However, counterexamples to these observations are provided and discussed in Adli (2006: 176–179). Recently, the corpus-based inquiry by Zimmermann & Kaiser (2019) has shown that *in situ* structures in contemporary (Hexagonal) French are not subject to many of the alleged formal restrictions discussed in the literature. Another strand of research, represented by the corpus-based studies of Coveney (1996) and Myers (2007), has suggested that there are several formal factors favoring the use of *in situ*, such as the argumental status of the *wh*-pronoun or -phrase (cf. section 4.2.1) or its weight and complexity: longer and more complex forms, such as a *wh*-word occurring within a prepositional phrase, seem more likely to occur in the sentence-final position than to be fronted.

Furthermore, many scholars have analyzed the different pragmatic conditions behind the licensing of French *ex situ* and *in situ* questions. Mathieu (2009: 28), for instance, has claimed that the two strategies receive different interpretations: *in situ* are preferred when the propositional content of the question belongs to the informational background. According to Cheng & Rooryck (2000: 4), “a strongly presupposed context” is needed for the licensing of *in situ*. Other researchers, such as Myers (2007) and Boucher (2010), have acknowledged the importance of “highly answerable, active and highly expected contexts” (Myers 2007: 108). More recently, Larrivée (2019) has argued that in order to understand the use of *in situ* in discourse, it is necessary

² For recent analyses of interrogatives in French, see Béguelin *et al.* (2018); Coveney (2020); Lefevre & Rossi-Gensane (2015). See also Kellert (2017) for a survey on interrogatives in Romance.

to rely on the notion of activation (Dryer 1996). Based on a diachronic analysis, he shows that in past stages of the language (from Middle French to the first half of the nineteenth century) the rare *in situ* occurrences mostly required the explicit activation of their propositional content (henceforth activation of *p*) in the previous context. In contemporary French, this does not seem to be the case anymore. Examples such as (5) show that these syntactic structures can be uttered out of the blue, without any previous activation of *p*:

- (5) *Pardon, il est quelle heure ?*
 sorry it be.PRS.3SG which hour
 ‘Sorry, what time is it?’ (Adli 2006: 184)

In regard to Italian, the scarce literature available has been mostly devoted to echoic *in situ* (recent examples are Badan *et al.* 2017; Badan & Crocco 2019), as in (6):³

- (6) A: Ho visto [inaudible]
 ‘I saw [inaudible]’
 B: *Hai visto cosa?*
 have.PRS.2SG see.PST.PTCP what
 ‘You saw what?’

Except for Badan & Crocco (2021), non-echoic *in situ* have not yet received much attention in the literature. In Badan & Crocco (2021), a syntactic and prosodic study of Italian *in situ* is offered, in which the authors show that these structures are attested in dialogical interactions and are felicitous “when they express a request for clarification relying on the

³ Unlike in Italian, *in situ* questions are widely employed and often used as a non-marked interrogative strategy in several Northern Italo-Romance varieties, such as the Lombard and Venetan dialects. Among others, cf. Donzelli & Pescarini (2019) and Bonan (2021).

preceding discourse context and when the request for information is emphatic, as, for instance, in a teaching context” (Badan & Crocco 2021: 762).

As reported in this study, information-seeking *in situ* in Italian are very rare. The authors only found 19 occurrences in a corpus of dialogical and monological speech.⁴ Examples (7) and (8) represent two typical occurrences:⁵

(7) A: comunque tu parti

lit.: ‘Anyway you leave.’

B: e io parto da dove?

and I leave.PRS.1SG from where

lit.: ‘And I leave from where?’ (Badan & Crocco 2021: 771)

(8) A: volevo sapere se venivi al cinema?

‘I wanted to know if you were coming to the cinema?’

B: eh no no dove

‘Eh no no where?’

A: all’Astra

‘To the Astra.’

B: a vedere cosa?

⁴ Of these 19 occurrences, 4 were drawn from the *Corpora Linguistici per l’Italiano Parlato e Scritto* (CLIPS), ‘Linguistic Corpora for Spoken and Written Italian’ (www.clips.unina.it; last access: 05.06.2021), containing task-oriented interactions gathered by map tasks and spot-the-difference tasks and whose size is estimated to be about 72,400 graphic words. The other 15 examples were extracted from the *Lessico di frequenza dell’Italiano Parlato* (LIP), ‘Frequency Lexicon of Spoken Italian’ (badip.uni-graz.at/it/corpus-lip/descrizione; last access: 05.06.2021), including dialogical and monological speech recorded in several contexts and containing about 490,000 words.

⁵ For the sake of simplicity, the transcription conventions used in Badan & Crocco (2021) were omitted.

to see.INF what

‘To see what?’ (Badan & Crocco 2021: 772)

In these occurrences the propositional material is already activated, either because it is discourse-given, as in (7), or inferable, as in (8).

Kaiser & Quaglia (2015) offer a similar analysis, claiming that Italian *in situ*, as the made-up example in (9), are possible when “the question is associated with a presupposition that contains a variable ranging over a restricted set already provided by the discourse” (Kaiser & Quaglia 2015: 96):

(9) A: Per l’esame di Latino devo leggere Virgilio, Orazio, Seneca e Cicerone.

‘For the Latin exam I have to read Virgil, Horace, Seneca and Cicero.’

B: *E adesso stai leggendo (che) cosa?*

and now stay.PRS.2SG read.GER what

‘And what are you reading now?’ (Kaiser & Quaglia 2015: 96)

Based on a corpus of Brazilian Portuguese, Italian and Spanish translations of a French crime novel, the authors determine that *in situ* in Italian and Spanish represent less than 1% of all partial interrogatives in the respective subcorpora. Moreover, the only Italian example involving an *in situ* is not even information-seeking, but rather involves a rhetorical question:

(10) Sei rimasto per questo? Avrei potuto...

‘Did you stay because of this? I could have...’

Avresti potuto cosa?

have.COND.PRS.2SG can.PST.PTCP what

Piazzarti qui ad aspettare lo squillo del telefono? Come ora.

‘What could you have done? Lock yourself in here, waiting for the telephone to ring? Like now.’ (Kaiser & Quaglia 2015: 99)

This brief survey of the literature permits us to conclude that Italian *in situ* are very rare. It is perhaps for this reason that systematic studies of their use and functions have thus far been neglected. In the rest of this paper, we offer a contribution to a pragmatic analysis of Italian *in situ* questions as well as a comparison between French and Italian, which, to the best of our knowledge, has not yet been pursued.

3. Research questions and discourse parameters

3.1 Research questions

Capitalizing on recent empirical work on Brazilian Portuguese (Oushiro 2011; Pires & Taylor 2009; Rosemeyer 2019), French (Larrivée 2019) and Spanish (Rosemeyer 2018), we have chosen as relevant pragmatic parameters for the analysis of French and Italian *in situ* the activation of *p* (see section 3.2) and the pragmatic functions of the structure (section 3.3). These parameters are intertwined, since, as we shall see below, certain activation statuses of *p* are required for some pragmatic functions (such as Elaboration or New Topic). In other words, the activation of *p* plays a key role in licensing the structure and is to some extent responsible for its pragmatic functions. In the rest of the paper, we shall tackle the following research questions:

- a. What are the licensing conditions for French and Italian *in situ*?
- b. What are the preferred discourse-pragmatic functions of French and Italian *in situ*? How are these structures used in discourse?

Regarding the first question, it is not easy to have clear expectations since we do not know much about non-echoic Italian *in situ*. In light of current research (Larrivée 2016, 2019), however, we know that the frequency of a certain construction and its pragmatic activation in discourse are interrelated. Larrivée (2019: 125) has observed that the rise in frequency of French *in situ*

interrogatives from the 1960s onwards was accompanied by a loosening of their pragmatic licensing conditions. Nowadays, these structures can also be licensed by a non-discourse given *p* (i.e., inferable or new). Larrivée also suggested that a ‘threshold effect’ may be at work. In his own words: “There is a categorical association between *in situ* and an explicit activation pragmatic value when the *in situ* represents less than 1% of occurrences of *wh*-questions” (Larrivée 2019: 126). If this principle can be extended to other languages, such as Italian, in which *in situ* seem to have an extremely low frequency (Badan & Crocco 2021; Kaiser & Quaglia 2015), we may expect that Italian *in situ* have stricter licensing conditions compared to French.

Concerning the second question, in light of the partial dependence of the pragmatic functions of *in situ* on the activation statuses of *p* (cf. section 3.3), a broader spectrum of functions is expected to be available for French, as compared to Italian.

3.2 Activation of the proposition

As mentioned in section 2, Larrivée (2019) convincingly showed that in order to understand the pragmatic motivation behind French *in situ* questions, one must look at their relation to the preceding discourse, “in terms of whether the proposition is activated—that is, already present to the attention of the discourse participants—or not” (Wallage 2015: 200).

Based on Dryer (1996), we distinguish three activation statuses (cf. also Chafe 1994: 71–81), which are illustrated in the following examples drawn from the French and the Italian corpora (for a presentation of the corpus data, see section 4.1). The first status is *activated*, which means that the propositional content of the question has already been mentioned in discourse, as in the Italian example (11), drawn from the Perugia Corpus (henceforth PEC; cf. section 4.1). The propositional content of the *in situ* uttered by B can be considered ‘primed’ by the proposition *Camilla has changed* in the antecedent context:

- (11) A: Camilla è cambiata tantissimo in quest’ultimo periodo.
‘Camilla has changed a lot in this last period.’

B: *Ma cambiata come?*
 but change.PST.PTCP how
 ‘But how did she change?’ (PEC, movies)

In other examples, the propositional content of the *in situ* is *inferable*, as in (12). The placement of the table and the chairs, which is inferable in JOS’s question, was not explicitly mentioned in the conversation, but is easily accessible thanks to our world knowledge:⁶

- (12) NAT: Et qu’est-ce que tu as acheté d’autre alors ?
 ‘What else did you buy?’
 MAI: Et ben on a acheté &euh la table avec les quatre chaises /# sept-cent balles//#
 ‘Well, we bought a table with four chairs, 700 euros.’
 JOS: *Pour mettre où ? #*
 for put.INF where
 ‘Where are you going to put them?’ (C-ORAL-ROM, ffamcv05)

Finally, the propositional content of the question can be *non-activated*, i.e., neither explicitly mentioned nor inferable based on the preceding context. In example (13), BET asks a question that is not related to the previous discourse context:

- (13) BET: D’accord // O.K. / O.K. // et alors
tu commences la journée en faisant quoi ?

⁶ In this paper, we maintain the prosodic annotations used in the C-ORAL-ROM (Cresti & Moneglia 2005: 14–40), to which we refer for more details on the annotation procedure. In particular, the symbols used in the examples are: ‘//’ for a terminal break; ‘/’ for a non-conclusive break, ‘[]’ for a non-conclusive break followed by a false start. Moreover, ‘<’ and ‘>’ mark the beginning and the end of the overlapped text of a given speaker, ‘&’ indicates an incomplete word or phonetic fragment, and, finally, ‘#’ stands for a pause in the speech flow. The Perugia Corpus data, by contrast, do not have any annotations at the prosodic level.

you start.PRS.2SG the day while do.PRS.PTCP what

‘All right, what do you do to get your day started?’ (C-ORAL-ROM, fpubd106)

3.3 Pragmatic functions of *in situ* questions

In order to describe the pragmatic functions of *in situ* in French and Italian, we have mainly relied on Rosemeyer (2018, 2019) for Spanish and Brazilian Portuguese, respectively. Similar to the previous paragraph, we shall briefly illustrate each of them with the aid of several French and Italian corpus examples.

An initial broad distinction can be made between information-seeking and non-information-seeking questions. Information-seeking *in situ* can be further subdivided into interrogative structures having an Elaboration or a New Topic function. Elaboration *in situ* are used to clarify or to add further details to a discourse topic raised in the previous context. They thus require an activated, as in (14), or an inferable *p*:

(14) ALE: E ieri / mentre ero lì / arrivano quei due che si son sposati / ti dissi / la
cugina di Marco / e questo Massimiliano [...]

‘Yesterday while I was there, the two who got married arrived, I told you
about them: Marco’s cousin and Massimiliano [...]

FRA: *E si son sposati quando?*

and RECP be.PRS.3PL marry.PST.PTCP when

‘When did they get married?’ (C-ORAL-ROM, ifamcv15)

The New Topic function establishes instead a new discourse topic and associates with a non-activated proposition. In (15), for instance, the new topic concerns information about the birthplace of the interviewee:

- (15) EUL: *Alors vous /vous êtes né où ? #*
 then you you be.PRS.2PL be.born.PST.PTCP where
 ici à Bayonne ? #
 ‘So where were you born? Here in Bayonne?’ (C-ORAL-ROM, ffamd105)

Non-information-seeking functions include Challenge, Rhetorical and Repeat, which are characterized by specific interactional goals.

Challenge, as the name itself suggests, has the function of critically rebutting a previous conversation move made by the interlocutor. By using a challenge *wh*-question, the speaker “refers anaphorically back to a prior utterance” with the aim of challenging “the grounds for that utterance” (Koshik 2003: 59). Challenge *wh*-questions convey “the epistemic stance of the questioner, a negative assertion” (Koshik 2003: 59) and are functionally considered akin to exclamatives (Rosemeyer 2018: 307, based on Chernova 2015: 166). In the literature, they have often been conflated with rhetorical questions, although it would be preferable to keep them separated, since challenge questions convey an interactional challenge which is usually not typical of ‘pure’ rhetorical interrogatives. In (16), speaker B’s *in situ* can be paraphrased as *you didn’t understand anything*. In the same turn, after uttering the question, B also provides details to support his rebuttal of A’s utterance (*we understood who the woman you were with was*):

- (16) A: *Io ero con Marianna e abbiamo capito chi era la donna che stava con te [...]*
 ‘I was there with Marianna and we understood who the woman you were with was [...].’
 B: *Avete capito cosa?*
 have.PRS.2PL understand.PST.PTCP what
 Era la mia quasi fidanzata, che tu non conosci [...]
 ‘What did you understand? She was my ‘almost girlfriend’ whom you don’t know.’ (PEC, conversations)

The Rhetorical function is used to convey the speaker's train of thought. As is well known, such questions do not require an answer from interlocutors. In spoken language, they are often used in monologic turns in conversations or in lectures, as in the following example, drawn from a conference on psychoanalysis:

- (17) MAR: On va en parler toute la journée / hein ? c' est un temps de la vie // mais ce qui m' intéresse // bien sûr / hein vous savez que je parle # en me référant à Freud &l [/]
- la vieillesse c' est quoi ?*
- the old.age it be.PRS.3SG what
- pour le sujet vivant / # ...
- 'We're going to talk about it all day long, right? It's a phase of life, but what I'm interested in, of course, you know, I'm referring to Freud, what is old age? For the living subject... ' (C-ORAL-ROM, fnatco02)

Strictly speaking, the question in (17) can be considered semi-rhetorical (Oushiro 2011: 151), since an answer is provided by the speaker herself. The use of a question-answer pair uttered by the same speaker is part of a discourse strategy that aims to emphasize a proposition or discourse referent (in this case, *old age*) and, at the same time, to create a kind of suspense so as to captivate the audience's attention.

Finally, the Repeat function corresponds to the echo questions seen in section 2. It characterizes those occurrences in which the interlocutor is asked to repeat a word or phrase:

- (18) ELE: Quenotte c' est une souris // #
- 'Quenotte is a mouse.'

FRA: *C' est qui ? #*
 it be.PRS.3SG who
 'Who is it?'

ELE: *Une souris // #*
 'A mouse.' (C-ORAL-ROM, fpubcv07)

In the next section, after a brief introduction to our working corpus, we offer a quantitative analysis of French and Italian *in situ*.

4. Corpus analysis

4.1 The corpus

For the empirical investigation, we relied on the French and Italian sections of the C-ORAL-ROM corpus (Cresti & Moneglia 2005). The C-ORAL-ROM is a corpus of the spoken language, which has the advantage of providing comparable data belonging to different types of interactions (face-to-face conversations, phone calls, as well as television speech)⁷ and with different levels of formality (formal vs. informal). The French data are mostly representative of the regional variety of French spoken in Aix-en-Provence and surroundings. The Italian data mostly reflect the regional variety of Florence and its neighboring area. The recordings also involved speakers of other diatopic varieties of French and Italian (for more information about the corpus data and design, see Cresti & Moneglia 2005: 1–70).

Since the number of Italian occurrences is very low (see Table 1), we also queried a larger corpus, the Perugia Corpus (PEC, Spina 2014), consisting of 26 million words (tokens) and freely available online.⁸ For this inquiry, only a subset of the corpus was considered, representing

⁷ The context of the occurrences in the French media files was unfortunately not accessible. For this reason, we did not consider the *in situ* retrieved in the French television speech section.

⁸ <https://www.unistrapg.it/cqpwebnew/pec/>; last access: 05.06.2021.

spontaneous dialogical and monological speech.⁹ Dialogical speech contains face-to-face interactions and phone calls as well as ‘institutional’ speech (such as interrogatories and court trials). Monological speech is mainly made up of conferences and lectures. Finally, the TV section was also considered, but only with respect to interviews, talk shows and newscasts (scripted speech from fictional TV broadcasts was excluded).

Even if the data from the C-ORAL-ROM and the PEC might not be perfectly comparable, they still belong to similar genres and, in our opinion, allow for a first cross-linguistic comparison.

Table 1. The corpus.

French	Italian	Source	Size
C-ORAL-ROM FR	C-ORAL-ROM IT	Conversational and non-conversational speech	Ca. 600,000 words (ca. 300,000 words for the French section; ca. 300,000 words for the Italian section)
	PEC	Conversational and non-conversational speech	Ca. 3,060,000 words

In the corpus, we searched for the following French and Italian *wh*-words, as shown in Table 2:¹⁰

Table 2. French and Italian *wh*-words.¹¹

French	<i>combien</i>	<i>comment</i>	<i>pourquoi</i>	<i>quand</i>	<i>lequel/s</i>	<i>qui</i>	<i>quoi</i>	<i>où</i>
	(<i>de</i>)				<i>laquelle/s;</i>			

⁹ The spoken section of the PEC mainly consists of material already gathered in other corpora, such as parts of the LIP and CLIPS corpora (cf. section 2). We refer to Spina (2014) for detailed information about the design and content of the PEC.

¹⁰ Due to some limitations of the free version of the concordance software used for the corpus search (*Contextes* 1.1.0; cf. Cresti & Moneglia 2005: 2–3), in the C-ORAL-ROM we searched for every occurrence of each *wh*-word listed in Table 2. In the case of the PEC, we used the CQP interface available on the corpus website to search for all strings containing a *wh*-word and a question mark, separated from zero up to five tokens.

¹¹ In the corpus search, we did not consider either bare *wh*-forms (i.e., *wh*-elements occurring in one-word utterances, such as *Quoi?*, ‘what?’) or *wh*-forms occurring in nominal sentences (e.g., *un’assicurazione di chi?*, ‘an insurance of whom?’).

					<i>quel/s</i> <i>quelle/s</i> N(oun)			
Italian	<i>quanto</i>	<i>come</i>	<i>perché</i>	<i>quando</i>	<i>quale/i;</i> <i>quale/i</i> N(oun)	<i>chi</i>	<i>(che)</i> <i>cosa</i> ¹²	<i>dove</i>
	‘how much’	‘how’	‘why’	‘when’	‘which’	‘who’	‘what’	‘where’

The data extracted from the corpus were manually labeled according to the relevant parameters (activation of *p* and pragmatic function) in an Excel spreadsheet. The statistical analysis was carried out in R (R Core Team 2020).

4.2 Quantitative analysis

4.2.1 *The frequency of in situ questions in French and Italian*

A grand total of 196 *in situ* (French: 139; Italian: 57) were extracted from the working corpus. Their distribution is shown in Table 3.

Table 3. Absolute frequency of French and Italian *in situ*

French	Italian
C-ORAL-ROM FR n = 139	C-ORAL-ROM IT n = 11
	PEC n = 46

Comparing the C-ORAL-ROM data, we can see a striking asymmetry between the two languages. Even in the absence of data concerning other types of partial interrogatives in the corpus, one can safely claim that Italian *in situ* are very rare. Their frequency differs dramatically

¹² In the *wh*-form *che cosa*, *che* is the determiner of the noun *cosa*, ‘thing’. *Che cosa* usually occurs in the reduced form *cosa* (De Santis 2010).

from French *in situ* and resembles more closely the distribution observed in the Spanish section of the C-ORAL-ROM (Rosemeyer 2018: 295), in which, according to Rosemeyer’s search, only 17 occurrences were gathered (which represent just over 1% of all *wh*-questions in the Spanish corpus).

One possible structural explanation for the higher frequency of *in situ* in French compared to Italian involves the syntax-phonology interface. As suggested by Mathieu (2009) and Hamlaoui (2011), the occurrence of *in situ* in spoken French can be favored by its prosodic specificities, i.e., the fixed position of the accent, which occurs in the final position of the word, the intonational unit or the sentence. In this respect, *in situ* questions, by placing the *wh*-form at the end of the sentence, seem particularly suitable in contexts in which the *wh*-word needs to be stressed.

Regarding the form of French and Italian *in situ*, the absolute frequency of the different types of *wh*-words is displayed in Figure 1:

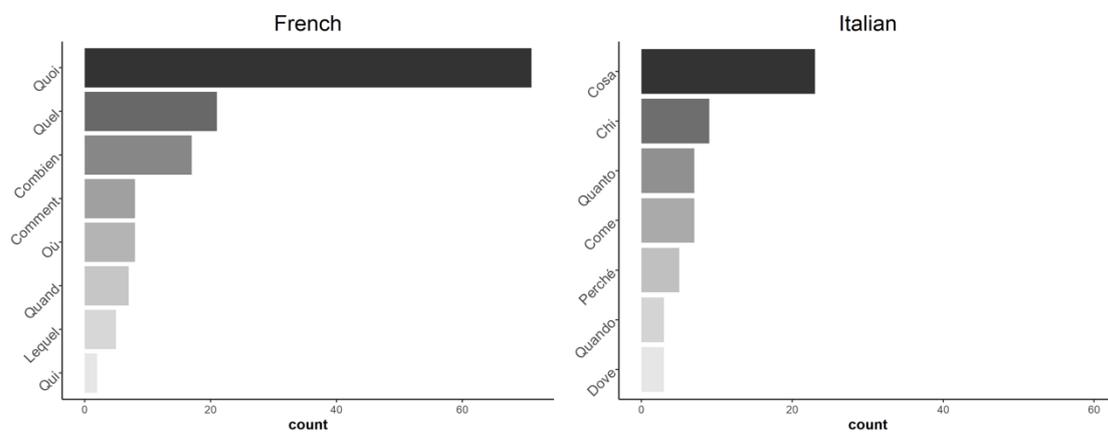


Figure 1. The distribution of the *wh*-words in the corpus (absolute frequencies).

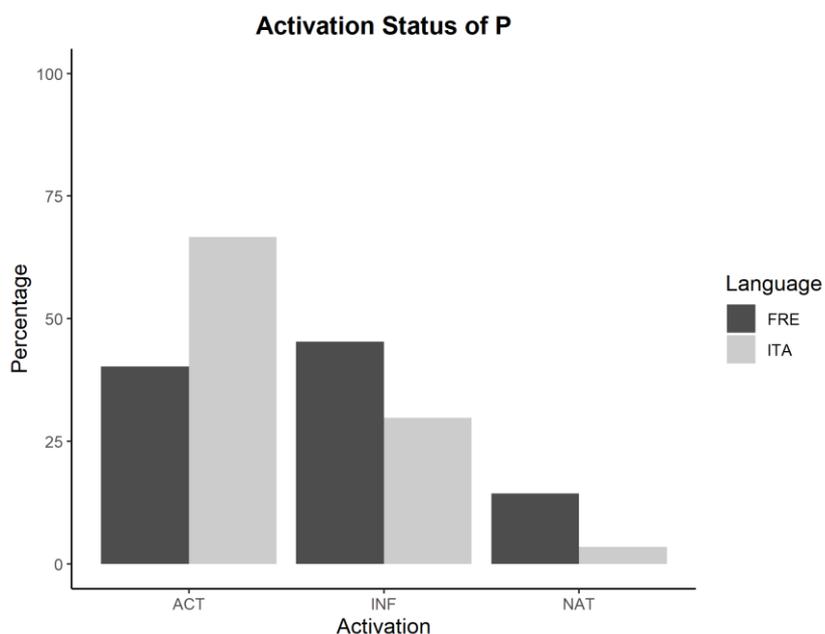
As observed in the literature, the distribution of *wh*-words within *in situ* seems to reflect a ‘scale of likelihood’ for the pronoun to be a verb argument (as already observed by Bolinger 1978; Myers 2007: 133). The stronger the semantic connection between the *wh*-word and the main verb, the higher the chances for an *in situ* to occur. This prediction seems partly confirmed in our data. In both languages, French *quel* and *quoi* and Italian *(che) cosa* and *chi*, expressing

either a direct or an indirect object, are the most frequent pronouns.¹³ Interestingly, *qui* and *lequel* in French have a much lower frequency than expected according to this prediction.¹⁴

Finally, our French data also confirm the results of other corpus-based inquiries: according to Lefevre & Rossi-Gensane (2015), *combien*, ‘how much’, occurs quite frequently within *in situ* questions. On the contrary, *pourquoi*, ‘why’, is considered to be rare within *in situ* (Lefevre & Rossi-Gensane 2015) or altogether avoided, although its appearance in this interrogative structure—according to Coveney (2020)—seems to be on the rise in contemporary informal spoken French.

4.2.2 The activation of p

Figure 2 shows the distribution of *in situ* questions in French and Italian according to their activation statuses (activated, inferable and non-activated):



¹³ *In situ* questions in which the *wh*-element is a subject are not considered to be grammatical (Kellert 2017: 579). An exception is represented by echo (or Repeat) questions such as *c'est qui ?*, *è chi?*, ‘who is it?’.

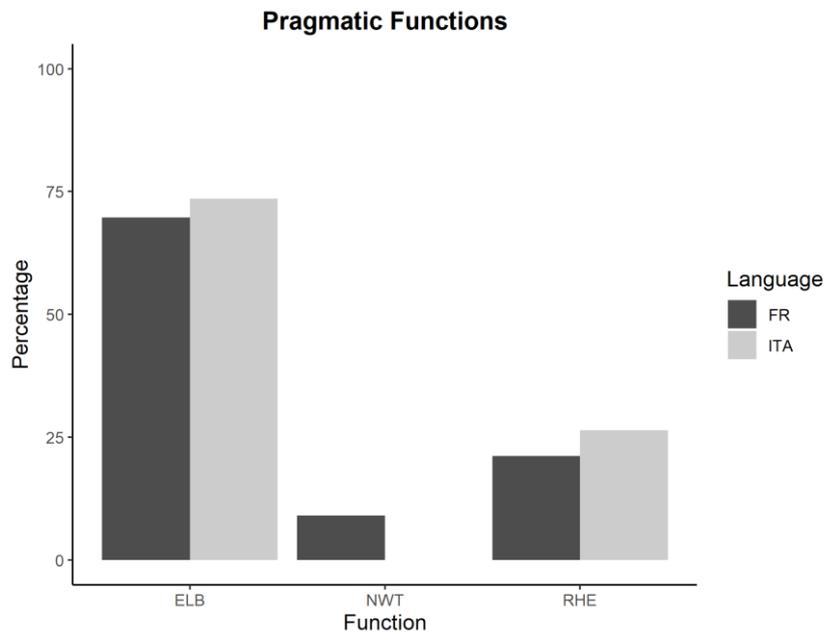
¹⁴ The very low frequency of *lequel* may well be related to its overall scarce frequency in the spoken (informal) language (Coveney 2020). As for *qui*, we do not have any convincing explanation at the moment.

Figure 2. Activation statuses of *p*: ACT = activated, INF = inferable, NAT = non-activated (French, n = 139; Italian, n = 57).

The two languages differ significantly (*Chi-squared test*, $\chi^2 = 12.51$, $df = 2$, $p < .01$).¹⁵ Compared to French, explicit activation in Italian is more frequent. By contrast, French occurrences show a higher frequency of inferable and non-activated propositional content.

4.2.3 *The pragmatic functions of in situ questions*

In light of the results presented in the previous paragraph, we expect that the two languages also differ in terms of their pragmatic functions in discourse. As shown by Figure 3, this seems to be the case (*Fisher's exact test*, $p < .05$).



¹⁵ In this article, we rely on the Pearson's Chi-Squared test for significance testing. When the assumptions behind the Chi-Square test do not apply, the Fisher's exact test is used (for a technical discussion, cf. Levshina 2015: 210–215).

Figure 3. Pragmatic functions. Elb = Elaboration, Nwt = New topic, Rhe = Rhetorical (French, n = 132; Italian, n = 53).¹⁶

Cross-linguistic differences become more evident if we focus on the information-seeking questions (Elaboration and New Topic) and, more specifically, on the New Topic function, which is altogether absent in the Italian dataset. By contrast, Elaboration represents the most common function in the whole corpus (70% of all occurrences), regardless of the language.

Exploring the association between the variables ‘activation of *p*’ and ‘pragmatic function’, more fine-grained differences between French and Italian emerge, as shown in Figure 4:

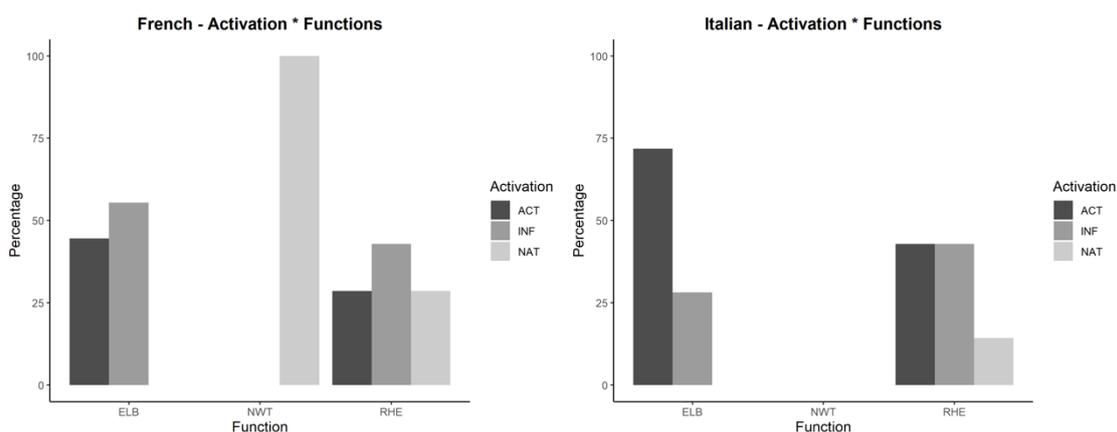


Figure 4. Activation statuses of *p* and pragmatic functions in the corpus.

It is worth observing that the internal distribution of *in situ* reveals significantly distinct patterns within the French and Italian data (*Fisher's exact Test*, French, $p < .0001$; Italian, $p < .05$). From a contrastive perspective, Elaboration preferably associates with an activated *p* in Italian and an activated as well as an inferable *p* in French. Moreover, *in situ* questions with a non-activated *p* in French mostly convey a New Topic (although the frequency of New Topic *in*

¹⁶ Due to their low frequency, echoic or Repeat *in situ* (six French and three Italian examples) as well as Challenge *in situ* (two occurrences, one for each language) were not considered.

situ in the whole corpus is low overall, cf. Figure 3), whereas in Italian they are restricted to the Rhetorical function only.

Finally, our inquiry revealed that other functions besides Elaboration, New Topic and Rhetorical are very rare. The Challenge function is, in particular, marginal in both languages. However, a first look at the scripted speech in the PEC hints at a different picture for Italian. Considering the *wh*-word (*che*) *cosa*, Challenge *in situ* questions account for 26% of the whole *in situ* occurrences (10 occurrences; the remaining 29 *in situ* questions are represented by Elaboration questions). Interestingly, 7 out of 10 examples are to be found in the movie section, as in (19):

- (19) A: Va bene, vedo che ha risolto
‘Ok, it seems you’ve fixed that.’
B: *Risolto* *cosa?*
fix.PST.PTCP what
Vieni fuori.
‘Fixed what? Come outside.’ (PEC, movies)

The association between the Challenge function and scripted dialogues in movies is noteworthy because it may tell us something about the ‘naive’ perception of *in situ* questions in Italian, which seem to be considered by native speakers (or at least by Italian screenwriters) as an ideal means for challenging an interlocutor’s statement, although this use is usually avoided in the spontaneous data of the corpus (see section 4.2.3).¹⁷

¹⁷ In Rosemeyer (2019), a similar situation is observed regarding Brazilian Portuguese *in situ* occurrences drawn from theater plays. The high percentage of *in situ* in contexts in which the proposition is discourse-new does not reflect a generally valid property of the structure, but is due instead to a genre-specific overrepresentation of rhetorical questions, which tend to associate with new information in Rosemeyer’s data.

4.3 A brief look at *ex situ* questions in French and Italian

Although a detailed comparison of *ex situ* and *in situ* in French and Italian is well beyond the scope of this paper, we can assess in broad strokes the functional motivations behind *ex situ* questions, in examples (20) and (21), to better understand the specificities of *in situ*.

(20) SAN: *et qu'est-ce que tu as acheté*
and what you have.PRS.2SG buy.PST.PTCP
d'autre alors ?
else then
'And what else did you buy then?' (C-ORAL-ROM, famcv05)

(21) MAX: *in che cosa ti vuoi laureare / <ora>?*
in what REFL want.PRS.2SG graduate.INF now
'In what subject do you want to major?' (C-ORAL-ROM, itelpv08)

In the next paragraphs, we offer an analysis of the licensing conditions and the pragmatic functions of *ex situ* carried out on a small sample drawn from the C-ORAL-ROM corpus. The *wh*-words that we searched for were *qu'est-ce que*,¹⁸ *che cosa*, 'what', as well as Italian *dove*, 'where', and French *comment*, 'how'. The analysis was limited to the first 120 occurrences retrieved: 60 for French (*qu'est-ce que* = 32; *comment* = 28) and 60 for Italian (*che cosa* = 36; *dove* = 24).

4.3.1 The activation of *p* in *ex situ* questions

Figure 5 shows two important facts. First, the distribution of the activation statuses of *p* in *ex situ* does not significantly differ between the two languages (*Chi-squared test*, $\chi^2 = 1.06$, $df = 2$, $p =$

¹⁸ The form *qu'est-ce que*, which cannot occur *ex situ*, is considered an allomorph of *quoi* (Meyers 2007: 54–55).

.06), which suggests that *ex situ* in French and Italian have analogous licensing conditions: they are mostly used when their propositional content is discourse-new.

Secondly, Figure 5 reveals a striking asymmetry with respect to Figure 2: while *in situ* questions preferentially require an activated (in Italian) or an inferable (in French) proposition and disprefer non-activated propositions, *ex situ* questions show the reverse pattern. In both French and Italian, these structures mostly occur with non-activated propositions and associate much less frequently with activated ones:

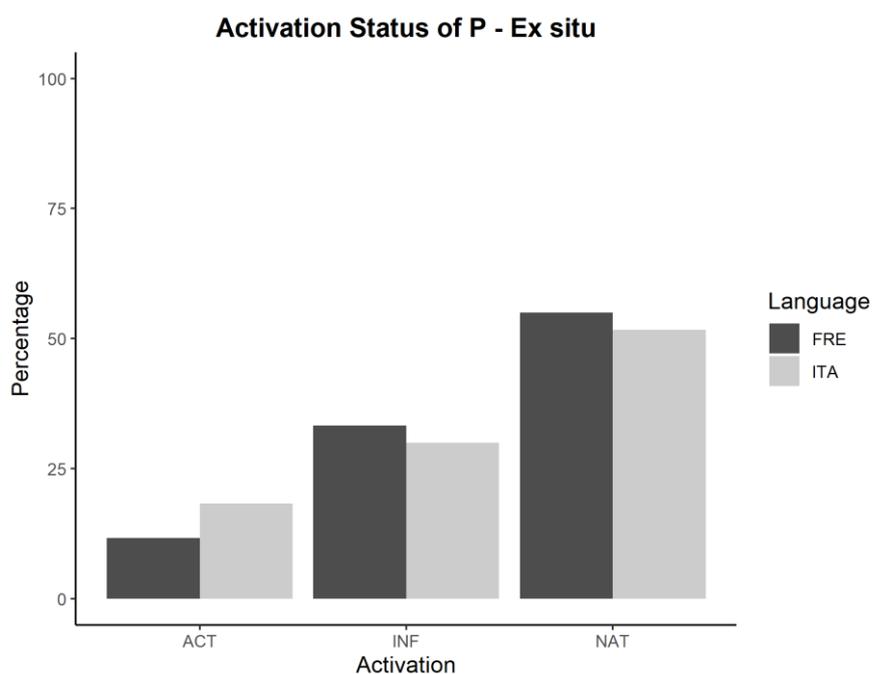


Figure 5. Activation statuses of *p*: ACT = activated, INF = inferable, NAT = non-activated (French, n = 60; Italian, n = 60).

4.3.2 The pragmatic functions of *ex situ* questions

According to the data depicted in Figure 6, *ex situ* interrogatives are preferentially used to introduce a new discourse topic or to elaborate on a previously introduced topic.

Regarding the difference between the two languages, which are significant (*Chi-squared test*, $\chi^2 = 6.6$, $df = 2$, $p < .05$), *ex situ* questions tend to occur with a New Topic function more

frequently in the Italian than in the French data, whereas the frequency of Rhetorical *ex situ* questions is higher in French than in Italian. The comparison with the data in Figure 3 reveals an almost complementary pattern: in both languages, the proportion of the New Topic function is much higher for *ex situ* than *in situ*, whereas the proportion of the Elaboration function in *ex situ* is lower than in *in situ*:

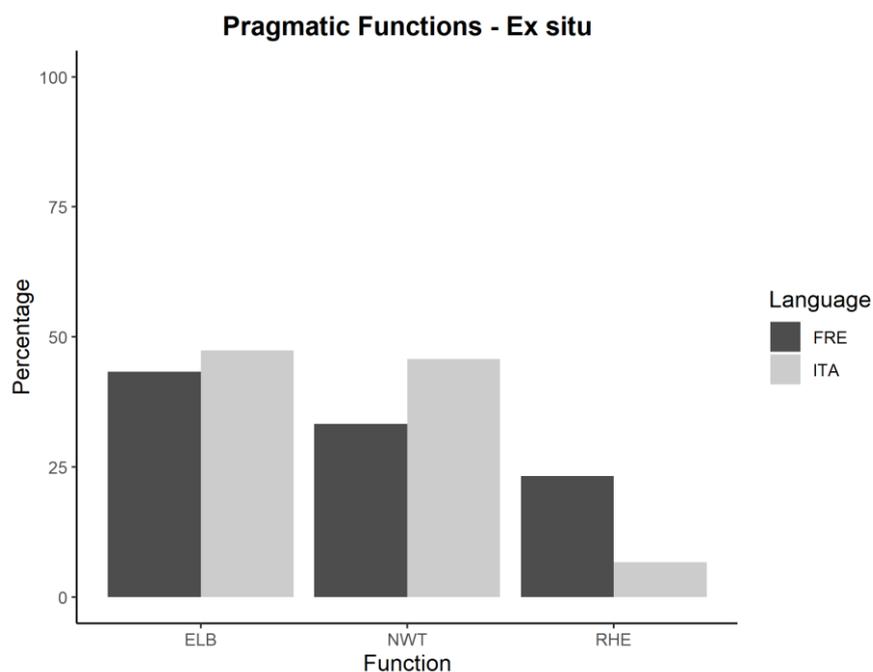


Figure 6. Pragmatic functions. Elb = Elaboration, Nwt = New topic, Rhe = Rhetorical (French, n = 60; Italian, n = 59).¹⁹

Based on the data analyzed in the present and previous paragraph (and keeping in mind the limited amount of data considered), we can conclude that *ex situ* and *in situ* questions considerably differ: they appear to have divergent licensing conditions and pragmatic functions, in both French and Italian. This comparison also suggests that the activation of *p* is not an important parameter for *ex situ* (cf. Larrivée 2019: 126, footnote 12). On the contrary, the

¹⁹ One Italian *ex situ* occurrence with echoic function was discarded.

propositional material of *in situ* interrogatives tends to be discourse-given, or at least inferable, while new propositional content is disfavored (even if French is less constrained than Italian; see also section 5).

5. Information-seeking *in situ* and discourse structure

As observed in section 4, French and Italian *in situ* mostly have an information-seeking function in our corpus: they are used either to elaborate on a discourse-given (or inferable) proposition or to introduce discourse-new material (as it seems possible in French). In this section, we intend to take a step forward in understanding the relation between *in situ* and discourse by focusing on their interactional properties. In order to do so, we rely on a Question Under Discussion model (henceforth QUD; see, among others, Riester *et al.* 2018; Roberts 2012 [1996]; Velleman & Beaver 2016) which has the advantage of providing an explicit discourse representation.²⁰

In QUD models, discourse is conceived of as a hierarchically organized structure (Riester *et al.* 2018: 405). Its most basic units are questions and assertions: each assertion is interpreted as an answer to a relevant question (in more formal terms, assertion nodes are dominated by question nodes in a discourse tree), which in most cases remains implicit (or covert). To illustrate the reconstruction of QUDs in a specific context, let us briefly consider a simple case, such as the one provided by the German example in (22):

- (22) Wir haben ja nun alle von Konflikten gehört, aber es gibt immer Konflikte
'Now, we have all heard of conflicts but there are always conflicts.' (Riester *et al.* 2018: 413).

The utterance *aber es gibt immer Konflikte*, 'but there are always conflicts', can be thought of as answering different implicit questions, such as the more generic *{What is going on?}*

²⁰ In this paper, we rely on the model outlined in Riester *et al.* (2018).

or the more specific *{What about conflicts?}*. As observed by Riester *et al.* (2018), the reconstruction of implicit questions needs to be constrained by explicit principles in order to be objective and reproducible.²¹ In particular, QUDs “should make reference to the immediately preceding context” (Riester *et al.* 2018: 412) by containing as much given material as possible (the ‘Maximize-Q-Anaphoricity’ principle; cf. also Rosemeyer this SI). For this reason, the question *{What about conflicts?}*, presenting more discourse-given information than the alternative *{What is going on?}*, can be selected as the best candidate in (22). Following Riester *et al.* (2018), a QUD analysis of this example is provided in (23), in which the indentation symbol (>) signals a hierarchical relation (Q and A stand for Question and Assertion):

(23) Wir haben ja nun alle von Konflikten gehört,

Q₁: {What about conflicts?}

> A₁: aber es gibt immer Konflikte

We are now equipped to analyze several French and Italian *in situ* occurrences, (24)–(26), drawn from the corpus:

(24) VAL: Non// # non // c’est pas ça //# c’est que pour l’instant / pour me faire à l’idée / c’est bien qu’il soit pas trop prêt // #

‘No, no, it’s not that, it’s that for now for me to get used to the idea, it’s good that it’s not too soon.’

CYR: *Pour te faire <à l’ idée de quoi> ?*

for REFL do.INF to the idea of what

‘To get used to the idea of what? (C-ORAL-ROM, ffamcv01)

²¹ For reasons of space, we cannot delve into the different principles put forward by Riester and colleagues. The interested reader is referred to Riester *et al.* (2018: 411–416).

(25) A: Allora questa sarebbe una famosa lettera spedita all'agenzia interinale di Genova [...] e appunto il contenuto della lettera invitava in un certo senso, il responsabile dell'agenzia, a chiudere in un periodo... in quella settimana in cui si verificava, si svolgeva il vertice lì a Genova...

'So this would be one of the famous letters that were sent to the temporary agency in Genoa [...] and the content of that letter invited, let's say so, the manager to close in a period... during the week in which the summit was held in Genoa...'

B: *A chiudere cosa?*

to close.INF what

'To close what?'

A: A chiudere le agenzie interinali [...]

'To close the temporary agencies [...]' (PEC, interrogatories)

(26) A: Sì questa è la nuova Regione, sarà pronta...

'Yes, this is the new regional headquarters, it will be finished...'

B: *Sarà pronta quando?*

be.FUT.3SG ready when

'When will it be finished?' (PEC, interviews)

It is worth observing that in these examples, *in situ* questions are used in contexts in which a QUD was left unanswered (see Biezma 2018: 29–34 on Spanish *in situ*), as we can see by closely examining a portion of (25):

(27) Q₁: {What did the content of the letter suggest closing?}

> A₁: appunto il contenuto della lettera invitava in un certo senso, i responsabili dell'agenzia, a chiudere in un periodo...

>> Q₂: {When was the suggestion contained in the letter supposed to apply?}

>>> A₂: in quella settimana in cui si verificava, si svolgeva il vertice lì a Genova...

In this excerpt, two covert questions can be reconstructed, Q₁ and Q₂. Q₁ is left unanswered: the focal part of the question, the *wh*-word, is not assigned any value (i.e., speaker A is not revealing *what* was supposed to be closed). Instead of providing this crucial piece of information, the speaker answers another QUD (Q₂), which elaborates on Q₁ (i.e., *the period* in which the closure was supposed to happen). The interlocutor reacts to this discourse ‘derailment’ by uttering an *in situ* interrogative, which makes explicit the content of Q₁. A similar analysis can be applied to examples (24) and (26), in which the unanswered QUDs concern the content of the idea mentioned by VAL and the time by which the new regional headquarters should be ready, respectively.

In her analysis of Spanish *in situ*, Biezma (2018) discusses another relevant context for the licensing of this question type, represented by (28):

- (28) A: Ana y Susana fueron ayer de compras. Ana se compró una falda preciosa!
‘Ana and Susana went shopping yesterday. Ana got herself an amazing skirt!’
- B: ¿Y Susana se compró qué?
And Susana REFL buy.PST.3SG what
‘And what did Susana buy?’ (Biezma 2018: 27)

This example can be described as a parallel structure or, from a QUD perspective, as a question-subquestions strategy (a Contrastive Topic in Büring 2016; see also Krifka & Musan 2012: 30–31), in which an implicit QUD containing two variables can be reconstructed, *{What did x buy?}*, i.e., one variable representing the subject and the other standing for the direct object. From this QUD two subquestions follow: *{What did Ana buy?}* and *{What did Susana buy?}*.

(29) Ana and Susana went shopping yesterday

Q₁: {What did x buy?}

> Q_{1.1}: {What did Ana buy?}

>> A_{1.1}: Ana got herself an amazing skirt!

> Q_{1.2}: {What did Susana buy?}

Answering Q_{1.1} and Q_{1.2} is a strategy for resolving the hierarchically higher Q₁. Only a few similar examples are to be found in the French section of the C-ORAL-ROM (none in Italian), as the one in (30), involving a conversation between friends (PER, STE, JEA):

(30) PER: Et là non plus // il y avait pas de préface // <il y avait même pas la vie de l'auteur> //

'There was no preface, there wasn't even the author's biographical notes.'

STE: <Mais l'auteur moi / j' ai mis juste ce que je> t'ai dit là //

'But as for the author, I wrote what I told you.'

PER: Ouais ben moi aussi // # moi aussi //

'Yes, well, me too, me too.'

STE: *Et toi t' as mis quoi sur*

and you you have.PRS.2SG put.PST.PTCP what on

l' auteur ? #

the author

'And you, what did you write about the author?' (C-ORAL-ROM, ffamcv09)

In (30), we can acknowledge a question-subquestions discourse strategy, which can be represented in detail in (31):

- (31) Q₁: {What did x write?}
- > Q_{1.1}: {What did STE write?}
- >> A_{1.1}: [...] moi j'ai mis juste ce que je t'ai dit là
- > Q_{1.2}: {What did PER write?}
- >> A_{1.2}: ouais ben moi aussi moi aussi
- > Q_{1.3}: {What did JEA write?}

The question posed by STE to its friend JEA is expected, since *{What did JEA write?}* is entailed by Q₁, as are the subquestions Q_{1.1} and Q_{1.2}. As in the previous examples, the speaker STE reacts to a question not yet answered (Q_{1.3}) by uttering an *in situ*, in order to acquire the last piece of information that would resolve the superordinate question at issue.²²

Crucially, in all these dialogical contexts, *in situ* can be conceived of as a repair move by which the speaker reacts to a communication problem (see Rosemeyer this SI).²³ This problem can be real (if the interlocutor refuses to provide relevant information) or perceived (if the interlocutor has no reason to conceal information, but is just temporizing or hesitating). In some cases, the repair move can also overlap or interrupt the interlocutor's turn (as we can infer from (26)). By doing so, this discourse move may contribute to characterizing the speaker as 'impatient' or even 'aggressive'. Not by chance, there is a communicative asymmetry between the participants in (25) and (26), since the *in situ* is produced by a prosecutor, in (25), and a journalist, in (26). In these contexts, both the prosecutor and the journalist want to obtain all the relevant bits of information and, for different reasons, try to minimize digressions or hesitations on the part of the interlocutors.

²² As observed by Rosemeyer (2018: 312, footnote 9), copulative conjunctions are indicative of 'QUD-advancing discourse functions', as we can see in the case of Spanish *y* in (28) and French *et* in (30).

²³ We do not intend to claim that *in situ* are always used as repair moves in conversation. Interestingly, though, this pattern seems very consistent in the dialogical contexts in our corpus.

Let us now examine some information-seeking occurrences whose propositional content is discourse-new (in our dataset, these are only French). As discussed in section 4.2, these *in situ* questions can be used for introducing a new discourse topic, such as in example (32):

- (32) CHR: *Et vous faites quoi / alors pour les vacances ?*
 and you do.PRS.2PL what then for the holidays
vous avez prévu ? # toi tu pars la semaine prochaine non ? # tu vas [/] tu
retournes dans ta famille // à Toulon / non ?
 ‘What are you doing for the holidays? Have you made any plans? You’re
 leaving next week, aren’t you? You’re going back to your family in
 Toulon, aren’t you?’ (C-ORAL-ROM, ffamcv11)

This kind of *in situ* tends to appear either at the very beginning of a conversation (Adli 2006: 184), as in (33), or serves to express a change of conversational topic, as in (34):

- (33) *Pardon, il est quelle heure ?*
 sorry it be.PRS.3SG which hour
 ‘Sorry, what time is it?’ (Adli 2006: 184)
- (34) C. *Peut-être il faut un divan oui*
 ‘Maybe we need a sofa.’
- M. *Vous allez faire quoi ce dimanche*
 you be.going.PRS.2PL do.INF what this Sunday
 ‘What are you going to do this Sunday?’ (Myers 2007: 154)

Unlike (24)–(26) and (30), the propositional content of the *in situ* is not activated in the linguistic context; thus, there cannot possibly be an interplay with implicit questions in discourse. However, the *in situ* in (32)–(34) occur in certain situations in which the proposition is made

readily available by the extralinguistic context (cf. also Boucher 2010: 108; Myers 2007: 154). More precisely, *in situ* occurrences such as the ones in (32)–(34) can be considered as ‘routine’ questions in certain situations, such as a conversation between friends or work colleagues before the summer break.

According to Clark (1996: 109), these are pieces of information contained in the common ground as part of larger social ‘scripts’, which, for instance, can “specify the expected course of the joint activities that take place in restaurants, doctor’s offices, supermarkets”. As reported by Rossi-Gensane & Ursi (2020), French *in situ* are frequently used in similar scripts, such as in the following example involving a transactional context in which a client anticipates a predictable discourse move (i.e., the seller disclosing the price):

- (35) *ça fait combien s’ il vous plaît*
 that do.PRS.3SG how much if it you like.PRS.3SG
 ‘How much is this please?’ (Rossi-Gensane & Ursi 2020: 154)

The final type of *in situ* introducing a new discourse topic is represented in our corpus by example (36):²⁴

- (36) [phone call between a mother and her daughter]
 MAM: Voilà // # et toi ? # qu’est-ce que tu deviens ? #
 ‘And you, what have you been up to?’
 EST: Oh bien rien je travaille / il pleut / alors bon
 ‘Nothing really, I’m working, it’s raining, well’
 MAM: Tu travailles / il pleut / tu restes à la maison // <ben oui hein >
 ‘You’re working, it’s raining, you’re staying at home, yes.’

²⁴ Example (36) is the only occurrence of this kind in the corpus.

EST: <xxx> je [/] je monte un peu à cheval et puis voilà // # <qu'est-ce qu' il dit
papa>

'I'll go riding a little bit, that's it. What is dad saying?'

MAM: *Ton discours à Grenoble il est pour quand ?*

your talk in Grenoble it be.PRS.3SG for when

'When are you giving your talk in Grenoble?'(C-ORAL-ROM, ftelpv03)

In (36), the *in situ* content is discourse-new (it has not been mentioned in the preceding utterances) and is not amenable to a general social script. Instead, this occurrence relies on some specific piece of shared knowledge between the interlocutors (the 'personal common ground', as described in Clark 1996: 112–116). The fact that the *in situ* content, in spite of being discourse-new, is already known by the participants is shown by the left dislocated constituent *ton discours à Grenoble*, 'your talk in Grenoble'. This left dislocation signals both the topical status of the NP and a topic shift (which are both typical functions of left dislocation, see Frascarelli 2017: 475–476). Even if one can claim that the *in situ* is still dependent on known information (the discourse-new *in situ* reactivates and elaborates on information that is already part of the common ground), its use seems to be functionally more similar to *ex situ* questions, which do not need to be activated and mostly introduce discourse-new material, as can be seen in section 4.3.

In conclusion, it is also important to mention that the choice of an *in situ* question in (36) can be favored by other functional factors. As suggested by Coveney (1996: 219–223) and Meyers (2007: 130–132), a *wh*-word occurring within a PP (*pour quand*, 'for when') tends to appear in sentence-final position, obeying a functional 'end-weight' principle. In any case, the mere fact that French *in situ* can be used in similar contexts suggests that this structure, unlike the Italian counterpart, is situated in a more advanced conventionalization phase.

6. Conclusion

In this paper, we have shown that French and Italian (information-seeking) *in situ* questions differ from a pragmatic point of view: in the corpus, they have different licensing conditions and different discourse-pragmatic functions. In the Italian data, *in situ* mostly require an explicit activation of *p* and express an Elaboration function, whereas the French occurrences are preferably licensed by an inferable or a non-activated *p* and can also be used to introduce new discourse topics.

The activation (or not) of *p* also has an impact on the interactional functions of the structure. When *p* is activated, *in situ* can be used by both French and Italian speakers to react to a QUD that was left unanswered in previous discourse. When *p* is not activated, *in situ* can be used instead to convey ‘routine’ questions or to anticipate a predictable discourse move in ‘scripted’ contexts, such as client-seller interactions, as we see in the French data.

In conclusion, the use of *in situ* in both French and Italian is pragmatically motivated, since it requires either propositional activation in the linguistic context, as in Italian, or predictable situations occurring in the extralinguistic context (e.g. social scripts), as in French. A preliminary comparison with *ex situ* questions also revealed some general functional differences in the use of these two types of interrogatives: unlike *in situ*, *ex situ* questions do not require contextual activation.

It is important to keep in mind that these observations only concern information-seeking *in situ*. Further empirical analyses are necessary in order to assess the discourse properties of less frequently occurring functions, such as Challenge or Rhetorical, as well their use in specific text genres (as observed in section 4.2.3, Challenge, which is marginal in the present corpus overall, appears more frequently in data drawn from Italian movies).

Acknowledgments. I would like to thank the participants at the workshop *Discourse-pragmatic perspectives on interrogatives in Romance and elsewhere* (Freiburg im Breisgau, July 16, 2019) as well as the audience at the *Innsbrucker Linguistischer Arbeitskreis* (Innsbruck, December 1, 2020). Discussing my work on *in situ* questions on both occasions has been of great help in the

elaboration of the present paper. I am also very grateful to Pierre Larrivée, Malte Rosemeyer, María Sol Sansiñena Pascual and an anonymous reviewer for their insightful remarks on a first draft of this article. All remaining mistakes are, of course, my own responsibility.

Appendix. Data tables (absolute frequency)

Table 1. *Wh*-words in *in situ* questions (in decreasing order of frequency).

French <i>wh</i> -forms	Freq.	Italian <i>wh</i> -form	Freq.
<i>Quoi</i>	71	<i>Cosa</i>	23
<i>Quel</i>	21	<i>Chi</i>	9
<i>Combien</i>	17	<i>Quanto</i>	7
<i>Comment</i>	8	<i>Come</i>	7
<i>Où</i>	8	<i>Perché</i>	5
<i>Quand</i>	7	<i>Quando</i>	3
<i>Lequel/s, laquelle/s;</i>	5	<i>Dove</i>	3
<i>Quel/s, quelle/s</i> N(oun)			
<i>Qui</i>	2		

Table 2. Activation statuses of *p* in *in situ* questions.

	French	Italian
Activated	56	38
Inferable	63	17
Non-activated	20	2

Table 3. Pragmatic functions of *in situ* questions.

	French	Italian
Elaboration	92	39
New Topic	12	0
Rhetorical	28	14

Table 4. French *in situ*: activation statuses and pragmatic functions.

	Activated	Inferable	Non-activated
Elaboration	41	51	0
New Topic	0	0	12
Rhetorical	8	12	8

Table 5. Italian *in situ*: activation statuses and pragmatic functions.

	Activated	Inferable	Non-activated
Elaboration	28	11	0
New Topic	0	0	0
Rhetorical	6	6	2

Table 6. Activation statuses of *p* in *ex situ* questions.

	French	Italian
Activated	7	11
Inferable	20	18
Non-activated	33	31

Table 7. Pragmatic functions of *ex situ* questions.

	French	Italian
Elaboration	26	28
New Topic	20	27
Rhetorical	14	4

References

- Adli, Aria. 2006. French *wh*-in-situ questions and syntactic optionality: Evidence from three data types. *Zeitschrift für Sprachwissenschaft* 25. 163–203.
- Badan, Linda & Claudia Crocco. 2019. Focus in Italian echo *wh*-questions: An analysis at syntax-prosody interface. *Probus* 31(1). 29–73.

- Badan, Linda & Claudia Crocco. 2021. Italian *wh*- questions and the low periphery. *Linguistics* 59(3): 757–784.
- Badan, Linda, Stella Gryllia & Gaetano Fiorin. 2017. Italian echo-questions at the interface. *Studia Linguistica* 71(3). 207–240.
- Béguelin, Marie-José, Aidan Coveney & Alexander Guryev (eds.). 2018. *L'interrogative en français*. Bern: Peter Lang.
- Biezma, Maria. 2018. *Givenness* and the difference between *wh*-fronted and *wh*-in-situ questions in Spanish. In Janine Berns, Heike Jacobs & Dominique Nouveau (eds.), *Romance languages and linguistic theory 13. Selected papers from 'Going Romance' 29, Nijmegen*, 21–39. Amsterdam: Benjamins.
- Bolinger, Dwight. 1978. Asking more than one thing at a time. In Heryn Hiz (ed.), *Questions*, 107–150. Dordrecht: D. Reidel.
- Bonan, Caterina. 2021. *Romance interrogative syntax. Formal and typological dimensions of variation*. Amsterdam: Benjamins.
- Boucher, Paul. 2010. *Wh*-questions in French and English. Mapping syntax to information structure. In Carsten Breul & Edward Göbbel (eds.), *Comparative and contrastive studies of information structure*, 101–137. Amsterdam: Benjamins.
- Büring, Daniel. 2016. (Contrastive) Topics. In Caroline Féry & Shinichiro Ishihara (eds.), 64–85.
- Chafe, Wallace L. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: UCP.
- Cheng, Lisa Lai-Shen & Johan Rooryck. 2000. Licensing *wh*-in-situ. *Syntax* 3. 1–19.
- Chernova, Ekaterina. 2015. *The Syntax of wh-movement in multiple (true and echo) questions: A Q-particle approach*. Girona: Universitat de Girona PhD Thesis.
- Clark, Herbert H. 1996. *Using Language*. Cambridge: CUP.
- Comrie, Bernard, Martin Haspelmath & Balthasar Bickel. 2015. *Leipzig glossing rules. Conventions for interlinear morpheme-by-morpheme glosses*. Leipzig: Max Planck

- Institute for Evolutionary Anthropology. <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf> (Last access: 05.06.2021).
- Coveney, Aidan. 1996. *Variability in spoken French. A sociolinguistic study of interrogation and negation*. Exeter: Elm Bank Publications.
- Coveney, Aidan. 2020. L'interrogation directe. In *Encyclopédie grammaticale du français*. http://www.encyclogram.fr/notx/002/002_Notice.php (Last access: 05.06.2021).
- Cresti, Emanuela & Massimo Moneglia. 2005. *C-ORAL-ROM. Integrated reference corpora for spoken Romance languages*. Amsterdam-Philadelphia, John Benjamins.
- De Santis, Cristiana. 2010. Aggettivi e pronomi interrogativi. In Raffaele Simone (ed.), *Enciclopedia dell'italiano*, vol. I (A-L), 684–686. Roma: Istituto dell'Enciclopedia Italiana G. Treccani.
- Donzelli, Giulia & Diego Pescarini. 2019. Tre tipi di *wh in situ* nei dialetti lombardi. In Annamaria Chilà & Alessandro De Angelis (eds.), *Bollettino del Centro di studi filologici e linguistici siciliani* 21. 183–197.
- Dryer, Matthew S. 1996. Focus, pragmatic presupposition, and activated propositions. *Journal of Pragmatics* 26(4). 475–523.
- Dryer, Matthew S. 2013. Position of interrogative phrases in content questions. In Matthew S. Dryer & Martin Haspelmath, *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/93> (Last access: 05.06.2021).
- Dufter, Andreas & Elisabeth Stark (eds.). 2017. *Manual of Romance morphosyntax and syntax*. Berlin: Mouton.
- Engdahl, Elisabeth. 2006. Information packaging in questions. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics*, 93–111. Paris: CSSP.
- Féry, Caroline & Shinichiro Ishihara (eds.). 2016. *The Oxford handbook of information structure*. Oxford: OUP.

- Frascarelli, Mara. 2017. Dislocations and framings. In Andreas Dufter & Elisabeth Stark (eds.), 472–501.
- Kaiser, Georg & Stefano Quaglia. 2015. In search of *wh*-in-situ in Romance: An investigation in detective stories. In Ellen Brandner, Anna Czypionka, Constantin Freitag & Andreas Trotzke (eds.), *Charting the landscape of linguistics*, 92–103. Konstanz: Konstanzer Online-Publikations-System. <https://kops.uni-konstanz.de/handle/123456789/34161> (Last access: 05.06.2021).
- Kellert, Olga. 2017. Interrogatives. In Andreas Dufter & Elisabeth Stark (eds.), 569–602.
- Koshik, Irene. 2003. Wh-questions used as challenges. *Discourse Studies* 5(1). 51–77.
- Krifka, Manfred & Renate Musan. 2012. Information structure: Overview and linguistic issues. In Manfred Krifka & Renate Musan (eds.), *The expression of information structure*, 1–43. Berlin: de Gruyter.
- Hamlaoui, Fatima. 2011. On the role of phonology and discourse in Francilian French *wh*-questions. *Journal of Linguistics* 47(1). 129–162.
- Larrivée, Pierre. 2016. The pragmatics of marked configurations: Negative doubling in French. *Journal of Pragmatics* 95. 34–49.
- Larrivée, Pierre. 2019. Historical pragmatics, explicit activation and *wh in situ* in French. In Ingo Feldhausen, Martin Elsig, Imme Kuchenbrandt & Mareike Neuhaus (eds.), *Romance languages and linguistic theory 15. Selected papers from 'Going Romance' 30, Frankfurt*, 114–132. Amsterdam: Benjamins.
- Lefeuve, Florence & Nathalie Rossi-Gensane. 2015. Interrogation. In Pierre Larrivée & Florence Lefeuve (eds.), *Projet Fracov*. <http://www.univ-paris3.fr/index-des-fiches-227311.kjsp?RH=1373703153287> (Last access: 05.06.2021).
- Levshina, Natalia. 2015. *How to do linguistics with R. Data exploration and statistical analysis*. Amsterdam: Benjamins.

- Mathieu, Éric. 2009. Les questions en français: Micro et macro-variation. In France Martineau, Raymond Mougeon, Terry Nadasdi & Mireille Tremblay (eds.), *Le français d'ici: Études linguistiques et sociolinguistiques de la variation*, 37–66. Toronto: GREF.
- Myers, Lindsay Lee. 2007. *WH-interrogatives in spoken French: A corpus-based analysis of their form and function*. Austin: University of Texas Austin PhD Thesis.
- Oushiro, Livia. 2011. Wh-interrogatives in Brazilian Portuguese: The influence of common ground. *University of Pennsylvania Working Papers in Linguistics* 17 (2). 145–154.
- Pires, Acrisio & Heather Taylor. 2009. The syntax of wh-in-situ and common ground. *CLS* 43 (2). 201–215.
- R Core Team 2020. *A language and environment for statistical computing*, v 3.6.2. Wien: R Foundation for Statistical Computing. <https://www.r-project.org/> (Last access: 05.06.2021).
- Riester, Arndt, Lisa Brunetti & Kordula De Kuthy. 2018. Annotation guidelines for Questions Under Discussion and information structure. In Evangelia Adamou, Katharina Haude & Martine Vanhove (eds.), *Information structure in lesser-described languages. Studies in prosody and syntax*, 403–444. Amsterdam: Benjamins.
- Roberts, Craige. 2012 [1996]. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 5. 1–69.
- Rosemeyer, Malte. 2018. The pragmatics of Spanish postposed wh-interrogatives. *Folia Linguistica*, 52(2). 283–317.
- Rosemeyer, Malte. 2019. Brazilian Portuguese *in-situ wh*-interrogatives between rhetoric and change. *Glossa* 4(1). 1–29.
- Rosemeyer, Malte. This issue. Modeling the discourse pragmatics of interrogatives. *Functions of Language*.
- Rossi-Gensane, Nathalie & Biagio Ursi. 2020. Syntaxe et genre interactionnel : le cas des interrogatives directes partielles dans les interactions commerciales. *Langages* 219(3). 147–160.

- Spina, Stefania. 2014. Perugia Corpus: una risorsa di riferimento per l'italiano. Composizione, annotazione e valutazione. In Roberto Basili, Alessandro Lenci & Bernardo Magnini (eds.), *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it*, vol. 1, 354–359. Pisa: Pisa University Press.
- Velleman, Leah & David Beaver. 2016. Question-based models of information structure. In Caroline Féry & Shinichiro Ishihara (eds.), 86–107.
- Wallage, Phillips. 2015. Identifying the role of pragmatic activation in changes to the expression of English negation. In Pierre Larrivé & Chungmin Lee (eds.), *Negation and polarity: Experimental perspectives*, 199–227. New York: Springer.
- 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.