In search of wh-in-situ in Romance:
An investigation in detective stories *

Georg Kaiser & Stefano Quaglia

1 Introduction or a personal note

It is well known that languages differ with respect to the position of the wh-element in constituent questions. While in many languages these questions are generally formed by the fronting of the wh-element to a sentence-initial position, other languages require the wh-element to remain in what seems to be its canonical position, i.e. in situ. One generally distinguishes a third type of languages forming constituent questions either by fronting the wh-element or by leaving it in situ.

It is also well known that Josef Bayer feels a strong affection for question formation with wh-phrases in situ (cf. Bayer, 1996; Bayer, 2006; Bayer & Cheng, forthcoming). This is certainly one of the reasons—among others—why he likes Bangla, a language with (almost) obligatory wh-in-situ. Interestingly, there is another language for which Josef feels a strong affection, namely Italian. This is remarkable, since Italian is a language that generally does not allow wh-in-situ. What is also remarkable is that Josef maintains this affection by reading Italian detective stories by authors like Loriano Macchiavelli.

Given these affections and contradictions and given that the authors of this paper feel a strong affection for Romance languages—one of the author actually being a native speaker of a Romance language, namely Italian—, our purpose is to provide a search on wh-in-situ questions in some Romance detective stories. Since French is known as a language that optionally allows for wh-in-situ, our study is based on a series of French detective stories, written by Jean-Claude Izzo, which will be compared with the translations into other Romance languages, namely Italian, Spanish, and Portuguese. The goal is to look for differences between these languages with respect to the use of the wh-in-situ option and to compare these results with the observations and explanations which have been made in the literature on (optional) wh-in-situ questions in Romance so far.

2 wh-in-situ questions in Romance: A brief state of the art

Romance languages are generally described as belonging to the language type instantiating obligatory wh-fronting in (information-seeking) constituent questions. The in-situ option is

* We are grateful to Simon Dold and Janina Reinhardt for helpful comments on a previous version of this paper as well as to Barbara Krisl-Kaiser for her helpful support when building our corpus.
Georg Kaiser & Stefano Quaglia

normally considered to be restricted to multiple questions and echo questions. Yet, it has long been observed that Colloquial French stands out in optionally allowing the wh-element to remain in situ in information-seeking questions (Aoun et al., 1981; Cheng & Rooryck, 2000). Similar observations have been made with respect to (European and Brazilian) Portuguese (Ambar & Veloso, 2001).

(1)  a. Jean a acheté quoi?
    John has bought what
    ‘What did John buy?’
  b. Jean a vu qui?
    John has seen whom
    ‘Whom did John see?’

(2)  a. O João comprou o quê?
    DET John bought DET what
    ‘What did John buy?’
  b. O João viu quem?
    DET John saw whom
    ‘Whom did John see?’

Both languages are therefore classified as ‘optional wh-in-situ languages’ (e.g. Kato, 2013). Although it is still far from being clear what exactly triggers the wh-in-situ option in these languages (Bayer & Cheng, forthcoming), it has been acknowledged that “optional” wh-in-situ is constrained by restrictions which do not exist in “real” wh-in-situ languages (Cheng & Rooryck, 2000). There is, however, some disagreement with respect to the exact determination of the syntactic and discourse-pragmatic constraints on this kind of questions.

As for syntactic conditions, most authors agree that wh-in-situ (without echo-interpretation) is not possible in the scope of sentential negation (Shlonsky, 2012: 243; Mathieu, 2004: 1093):

(3)  a. Il ne voit pas qui?
    he NEG sees not who
    ‘Who doesn’t he see?’
  b. Qui est-ce qu’il ne voit pas?
    who is-this that he NEG sees not
    ‘Who doesn’t he see?’

In addition, most authors argue that wh-in-situ is excluded in embedded questions (Shlonsky, 2012: 245) as well:

(4)  a. Jean a vu qui?
    John has seen who
  b. “Tu te demandes Jean a vu qui.
    you refl.2sg wonder John has seen who
  c. Tu te demandes qui Jean a vu.
    you refl.2sg wonder who John has seen
    ‘Who doesn’t he see?’
However, the ban on wh-in-situ does not extend to all dependent clauses. In particular, some authors claim that the in-situ option is allowed in embedded contexts when the matrix verb is non-intensional, i.e. factive (compare (5a) vs. (5b)) (Boeckx et al., 2001: 59):

(5)  a. Jean sais que Marie a acheté quoi?  
    John knows that Mary has bought what  
    ‘What does John know that Mary bought?’

b. *Jean pense que Marie a acheté quoi?  
    John thinks that Mary has bought what  
    ‘What does John think that Mary bought?’

Other authors assume that the varying acceptability of ex-situ and in-situ questions in embedded contexts is not due to the matrix verb type, but rather to diatopic or diaphasic variation (Mathieu, 2004: 1092). Discussing a controversial example provided by Bošković (2000: 64) where the embedded sentence is introduced by an non-factive matrix verb, Bayer & Cheng (forthcoming) contend that this example is “quite natural in everyday conversation” at least for young French speakers:

(6)  ?*Jean et Pierre croient que Marie a vu qui?  
    John and Peter believe that Mary has seen whom  
    ‘Whom do John and Peter believe that Mary saw?’

A similar piece of data is also provided for Portuguese by Pires & Taylor (2009: 202) without further comments:

(7)  O Bill acha que a Sue comprou o quê?  
    Bill thinks that Sue bought what  
    ‘What does Bill think that Sue bought?’

As far as the interpretive dimension is concerned, some authors claim that wh-in-situ questions are only felicitous if certain discourse-pragmatic conditions are met. With respect to French, it has been claimed that wh-in-situ is associated to a stronger presupposition by the speaker than it is the case in ex-situ questions. This explains why the answer in (8a) is considered to be inappropriate (Coveney, 1989: 96; Chang, 1997: 42-46; see also Cheng & Rooryck, 2000):

(8)  a. Q: Marie a acheté quoi?  
    Marie has bought what  
    ‘What has Marie bought?’

    Nothing

c. Q: Qu’est-ce que Marie a acheté?  
    what Marie has bought  
    ‘What has Marie bought?’

d. A: Rien.  
    Nothing
Similar observations have been made for Portuguese by Ambar (2002) and Pires & Taylor (2009). Other authors, however, do not mention any such requirement (Shlonsky, 2012 for French; Kato, 2013 for French and Brazilian Portuguese). Mathieu (2004: 1100) presents some counterexamples taken from the internet, which according to him show that “not all dialects/registers of French contain in-situ wh[-]phrases that are presuppositional:”

(9) a. Q: Vous faites quoi exactement dans la vie?  
   your make exactly in the life  
   ‘What do you do exactly for a living?’

   b. A: En ce moment rien. J’avais un mi-temps chez MacDonald.[…]  
      in this moment nothing I had a half-time at MacDonald  
      ‘Nothing at the moment. I had a half-time job at MacDonald.’

Pires & Taylor (2009) also concentrate on discourse-pragmatic conditions, which they model in terms of a “Common Ground” requirement. Descriptively, they identify four types of questions, corresponding to different configurations, which allow (non-echo) wh-in-situ. Interestingly enough, they claim that this holds not only for optional wh-in-situ languages, but also for English, which is usually not considered to belong to the optional in-situ type. The typology sketched out by Pires & Taylor (2009) is reported in detail in what follows:

a. ‘[+specific] questions’, i.e. questions which “request more specific information about something mentioned immediately prior” (Pires & Taylor, 2009: 203):

(10) a.  A: I made desserts.
   b.  B: You make what ↑kind of desserts↓?

b. ‘Expect-questions’, i.e. questions “when further questioning for new information is expected, as in legal questioning” (Pires & Taylor, 2009: 203):

   b.  A (Defendant): I was driving along Andrews Avenue.
   c.  B: And you were driving in which ↑direction↓?

c. ‘Reference-questions’, i.e. questions asking for “a paraphrase or repetition of an immediately prior antecedent” (Pires & Taylor, 2009: 204)

(12) a.  A: I did not sell those strange pictures.
   b.  B: You didn’t sell what ↑↓ strange pictures↓?

d. Questions requiring a particular extra-linguistic context. Pires & Taylor (2009: 204) provide the example in (13), which can be used felicitously in a daily routine mother-daughter conversation where the daughter is asking for extra pocket money:

(13)  B (mother): So, you want how much today?
According to Pires & Taylor (2009: 204), these four question types share one crucial property, namely that “the set of possible answers [...] is part of the Common Ground,” which can be informally described as knowledge shared by both speaker and hearer (see also Bayer & Obenauer, 2011 for the relevance of common ground with respect to the use of German discourse particles like denn in questions).

A further discourse factor has been mentioned by Jiménez (1997) with respect to Spanish (see also Uribe-Etxebarria, 2002; Etxepare & Uribe-Etxebarria, 2012: 253). According to the author, the wh-element may remain in situ provided that the question is associated with a presupposition that contains a variable ranging over a restricted set already provided by the discourse. This is exemplified by (14), where the value for the variable is picked up from the set of referring expressions {huevos, leche, café}:

(14)  

a. A: Fuimos a la tienda a comprar huevos, leche y café. Mi madre compró los huevos. ‘We went to the store to buy eggs, milk and coffee. My mother bought the eggs.’

b. B: ¿Y tu padre compró qué? ‘What did your father buy?’

Incidentally, the same seems to hold for Italian:

(15)  

a. A: Per l’esame di Latino devo leggere Virgilio, Orazio, Seneca e Cicero. ‘For the exam I have to read Virgil, Horace, Seneca, and Cicero.’

b. B: E adesso stai leggendo (che) cosa? ‘And what are you reading now?’

Even for German it has been observed that wh-in-situ is not only restricted to echo-questions stricto sensu and multiple wh-questions. Bayer (2006: 378) argues that a question like (16a) “also seem[s] to function as normal information-seeking question [...] under certain circumstances.” On the other hand, Reis (2013: 107) notes that what Pires & Taylor (2009) call ‘expect-questions’ is actually possible in German, too, as shown in (16b).

(16)  

a. Der Hans hat wen gesehen?  
   DET John has whom seen  
   ‘Whom did John see?’

b. Sie haben den Angeklagten gesehen? Und dann sahen Sie, dass er wo
   you have the defendant seen and then saw you that he where
   stopped
   ‘Did you see the defendant? And then, did you see where he stopped?’
Interestingly, Reis (2013) provides arguments for classifying questions like (16b) as a subkind of echo-questions. But she acknowledges that the exact status of these questions is still controversial.

In sum, our brief overview allowed us to identify two major research threads in the current literature on wh-in-situ in Romance languages. On the one hand, a number of contributions are devoted to Romance ‘optional wh-in-situ languages’, i.e. Colloquial French and Portuguese. Both languages are considered on a par in allowing both ex-situ and in-situ strategies in order to form constituent questions. It is systematically pointed out that the latter strategy is only licit in certain syntactic environments. Nonetheless, there is some disagreement about what the exact environments are. On the other hand, the phenomenon of wh-in-situ has been investigated with respect to other Romance languages, too. Although these languages are generally assumed to allow wh-in-situ only in echo and multiple questions, some authors have shown that wh-in-situ is indeed allowed if certain discourse-pragmatic conditions are met. These results, if empirically confirmed, may well lead to the hypothesis that wh-in-situ is an option actually displayed by all wh-fronting languages, albeit under more severe non-syntactic requirements.

We believe that both research threads would benefit from extensive empirical research. For example, scrutiny of large corpora may be a highly profitable tool for investigating variation both within a single language and across languages. In this paper, we carry out a little corpus study based on parallel texts in order to prove that such a method would indeed help clarifying some highly debated or unclear issues found in the literature.

In particular, given limitations in space, we would like to explore the following research questions:

i. Can any (qualitative and quantitative) differences be noticed between French and Portuguese, i.e. the Romance languages considered to be ‘optional wh-in-situ’ languages?

ii. Are the syntactic restrictions on wh-in-situ in French confirmed?

iii. Do we find wh-in-situ in Spanish and Italian? If yes: under which conditions and with which kind of interpretation?

In what follows, we illustrate the method adopted for our empirical study and present its most interesting results.

3 In search of wh-in-situ in Romance detective stories

Our investigation is inspired by Hans-Georg Obenauer’s talk on the occasion of Josef Bayer’s 60th birthday, where he investigated the uses of ‘wh-the-hell’-like constructions in English and French by comparing some detective stories of John Le Carré in English and French.

As a basis for our corpus, we took the three detective novels Total Khéops, Chourmo, and Solea written by Jean-Claude Izzo (1945–2000), a French novelist from Marseille. The novels have been published in the mid-1990s and are known as the ‘Marseille Trilogy’. The main character is a former policeman who is faced with several criminal cases happening
in Marseille and its surroundings. The novels contain a high number of dialogues written in Colloquial French. Other parts, in particular those where the narrator tells the story or reports interior monologues, are written in Standard French.

We built our corpus as follows. We first extracted all constituent questions from the three original French novels. We then turned to the book translations into Italian and Spanish and extracted again all constituent questions. We did the same for Brazilian Portuguese, but since a Portuguese translation only exists for the first novel, we could integrate only Total Khéops in our corpus. We classified the data according to the distinction between finite and non-finite matrix and dependent clauses. We subsequently analyzed the questions with respect to the position of the wh-element, distinguishing between ex-situ and in-situ. Multiple questions in which one wh-element appears in situ were classified as in-situ. We finally marked the questions according to the type of wh-phrase: subject, direct object, indirect object, prepositional or adverbial phrase.

The results of the analysis are listed in table 1 and table 2.

<table>
<thead>
<tr>
<th></th>
<th>wh-in-situ</th>
<th>wh-ex-situ</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>142</td>
<td>341</td>
<td>483</td>
</tr>
<tr>
<td></td>
<td>29.4%</td>
<td>70.6%</td>
<td></td>
</tr>
<tr>
<td>Italian</td>
<td>3</td>
<td>450</td>
<td>453</td>
</tr>
<tr>
<td></td>
<td>0.6%</td>
<td>99.4%</td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>2</td>
<td>456</td>
<td>458</td>
</tr>
<tr>
<td></td>
<td>0.4%</td>
<td>99.6%</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Matrix finite questions featuring a wh-element in all three novels

<table>
<thead>
<tr>
<th></th>
<th>wh-in-situ</th>
<th>wh-ex-situ</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>51</td>
<td>95</td>
<td>146</td>
</tr>
<tr>
<td></td>
<td>34.9%</td>
<td>65.1%</td>
<td></td>
</tr>
<tr>
<td>(Brazilian) Portuguese</td>
<td>3</td>
<td>131</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td>2.2%</td>
<td>97.8%</td>
<td></td>
</tr>
<tr>
<td>Italian</td>
<td>1</td>
<td>135</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>0.7%</td>
<td>99.3%</td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>0</td>
<td>132</td>
<td>132</td>
</tr>
<tr>
<td></td>
<td>0.0%</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Matrix finite questions featuring a wh-element in Total Khéops

What catches the eyes first is the huge quantitative asymmetry between French on the one hand and Italian, Portuguese, and Spanish on the other. This finding is quite striking if,
as claimed in the literature, Portuguese can be truly classified as an ‘optional wh-in-situ language’ on a par with French. The immediate question, then, is what factors are responsible for the observed difference in frequency. Two hypotheses can be formulated. Either French and Portuguese differ as far as the stylistic conditions for the use of the in-situ option are concerned, or Portuguese actually instantiates a language type that is closer to Italian and Spanish.

As far as the latter languages are concerned, the number of wh-in-situ does not even reach 1% of all constituent questions found in the respective sub-corpora. Given that the number of both examples in Italian and only examples in Spanish are instances of multiple wh-questions, we are left with only one wh-in-situ question for both languages. Note that both examples are the translation of the very same sentence.

(17) Sei rimasto per questo? Avrei potuto…
Be.2sg stayed for this have.cond.1sg could
Avresti potuto cosa? Piazziarti qui ad aspettare lo squillo del
have.cond.2sg could what place.cl.refl.2sg here to wait the ringing of-the
telefono? Come ora.
telephone like now

(18) – No te has quedado sólo por eso, ¿no? Yo podría haber...
Neg cl.refl.2sg have.2sg stayed only for this no I can.cond.1sg have
– ¿Podrías haber qué? ¿Haberte encerrado aquí, mientras que esperabas
Can.cond.2sg have what have.cl.refl.2sg in-locked here while that waited
a que sonara el teléfono? Como ahora.
to that ringed.subj.past.3sg the telephone like now
‘Did you stay because of this? I could have...
What could you have done? Lock yourself in here, waiting for the telephone to ring?
Like now.’

As for the interpretation of these questions, our intuition is that they are not real information-seeking questions but rhetorical questions. As a matter of fact, in the following utterance the speaker provides himself the (in his opinion) only possible value for the wh-phrase cosa/què. Thus, in both the Italian and the Spanish sub-corpus we did not find a single wh-in-situ question with an authentic information-seeking illocutionary import.

Now turning to Portuguese, three instances of wh-in-situ questions have been found in the sub-corpus, which only consists of the novel Total Khéops. Since one of these is a multiple constituent question, we will discuss only the remaining two instances:

(19) Olhei meu vizinho nos olhos. Os outros pararam de bater nas paredes
looked.1sg my neighbor in-the eyes the others stopped to beat in-the walls
do vagão. Estava claro que a coisa se complicava. Eles me
of-the coach was clear that the thing refl.2sg complicated they cl.acc.1sg.
cercaram, cada vez mais perto.
surrounded every time more close
I looked my neighbor in the eye. The others stopped beating on the walls. This was serious now. They pressed around me.

“What are you talking about, man? What is it you don’t like? Rap? Our faces?”

Yasmine came to join us. She slipped her arm into Kader’s, and snuggled up to him. Tenderly. Kader smiled at her. A loving smile.

“How much longer are you staying?” I asked Kader.

As regards (19), this question could be considered as an instance of question type exemplified in (14), where the value for the variable is picked up from a given restricted set. The only difference, we contend, is that in this case the set is provided by the speaker himself after uttering the question. The question (20), on the other hand, can be interpreted as corresponding to type d. in Pires & Taylor’s (2009) typology. In this particular case, Kader and Yasmine, respectively the brother and the best friend of Leila, victim of an assassination, meet up and hold a brief conversation after Leila’s funeral. Importantly, Kader lives in Paris and not in Marseille, where the funeral takes place. Asking a question such as the one in (20) is indeed part of routine small-talk in the extra-linguistic context of meeting a person who pays a visit.

In sum, all the instances of (non-multiple) wh-in-situ questions we found in the Italian, Portuguese, and Spanish sub-corpora seem to be licensed by special discourse-pragmatic configurations. As far as French is concerned, by contrast, our results clearly suggest that this language is far more liberal in this respect. This is illustrated by the following dialogue, where wh-in-situ questions, marked in boldface, are formulated out of the blue.

(21) – Ça va ?
   Je fis oui de la tête, fermai les yeux. Malgré la faible lumière, j’avais du mal à les garder ouverts. Elle enleva le gant de mon front. Puis elle le reposa. Il était de nouveau froid.
   C’était bon.
   – Il est quelle heure? je dis.
   – Trois heures vingt.
   – T’as une cigarette?
   Elle en alluma une et me la mit entre les lèvres. J’aspirai, puis amenai ma main gauche pour l’ôter de mes lèvres. Ce seul mouvement me déchira le ventre. J’ouvris les yeux.
   – Tu fais quoi là?
   – T’as eu mon adresse où?
   – Le Minitel.
“Are you alright?”
I nodded, and closed my eyes. Despite the dim light, I found it hard to keep my eyes open. She took the glove off my forehead. Then put it back. It was cold again. It felt good.

“What time is it?” I asked.
“Twenty after three.”
“Got any cigarettes?”
She lit one for me and put it between my lips. I sucked on it, then lifted my left hand to take it out of my mouth. It was a small movement, but it gave me an excruciating pain in my stomach. I opened my eyes.

“What are you doing here?”
“I had to see you. I mean, I had to see someone, and I thought of you.”

“How did you get my address?”
“Minitel.”

Our little investigation also enabled us to further study the distribution of in-situ \textit{wh}-phrases in French. In particular, two results merit mention. We found three cases of \textit{wh}-in-situ in the scope of negation. But upon closer examination, though, all these questions turned out to have either an echo or a rhetoric interpretation. On the other hand, we found two cases of \textit{wh}-in-situ in an embedded clause taken as a complement by a non-factive verb. One of them corresponds to Pires & Taylor’s (2009) d. type. The other one, though, does not appear to be bound to any particular pragmatic restrictions:

\begin{equation}
\end{equation}

– \textit{Tu penses qu’il est allé où?}
– Ici. À Marseille. (Chourmo 43)

‘Guitou, the youngest of her three boys, had run away. Friday morning. Without leaving anything written. He had just taken a thousand francs from the shop’s cash register. Afterwards, the silence. She had hoped that he would call her up, just like when he used to leave for holiday to go to his cousins in Naples. She had hoped that he would come back on saturday. She had waited for him the whole day. Then the whole sunday. That night, she had broken down.

“Where do you think he went?”
“Here. In Marseille.”’

This result provides positive evidence for the acceptability of a highly debated construction, namely \textit{wh}-in-situ appearing in a CP subcategorized for by a non-factive verb.
4 Conclusions

In this little contribution, we have shown that wh-in-situ is a quite natural strategy in Colloquial French, while it seems to be quite unusual and more heavily pragmatically restricted in Romance languages like Italian and Spanish. Quite strikingly, Portuguese seems to be oriented to the second language type, and not to French, contrary to what is usually claimed. wh-in-situ in French, on the other hand, seems to require less special discourse-pragmatic conditions.

We hope that both the spirit and the results of our investigation will prove useful for future research on wh-in-situ in Romance in general and in French in particular. But of course, we more heartily hope that this little investigation will have pleased Josef Bayer as much as investigations in detective stories usually please him.

References


