



**HAL**  
open science

## French Interrogatives in an OT-LFG Analysis

Anna Gazdik

► **To cite this version:**

Anna Gazdik. French Interrogatives in an OT-LFG Analysis. LFG08 Conference, 2008, Stanford, United States. halshs-00751116

**HAL Id: halshs-00751116**

**<https://shs.hal.science/halshs-00751116>**

Submitted on 18 Nov 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **FRENCH INTERROGATIVES IN AN OT-LFG ANALYSIS**

Anna Gazdik

Eötvös Loránd University, Université Paris 7

Proceedings of the LFG08 Conference

Miriam Butt and Tracy Holloway King (editors)

2008

CSLI Publications

<http://csli-publications.stanford.edu/>

## *Abstract*

The present paper approaches the French interrogative system from a non-derivational perspective, aiming to provide an analysis that accounts not only for its own complexity (optional simplex and complex variants of the same question in formal French and in situ questions in informal French), but for its differences with other languages (English) as well. As will be shown, the choice of the theoretical framework (OT-LFG) is motivated by important lexical differences between the two languages, which cannot be captured solely by OT-type constraints.

## **1 Introduction**

The syntax of interrogatives (be it constituent or polarity, single or multiple) has vast literature in transformational grammar, dealing mainly with English, and is considered as an important field of analysis in non-derivational frameworks as well (Grimshaw 1995, Ackema & Neeleman 1998, Ginzburg & Sag 2000).<sup>1</sup> The present analysis focuses on French main clause interrogatives and shows how the non-derivational framework of OT-LFG (Optimality Theory and Lexical Functional Grammar) (Kuhn 2001, Sells 2001, Abeillé 2007) can account for this system in its whole complexity. As some significant lexical differences between languages (in this case, English and French) show, the OT-type constraint ranking needs to be completed by a lexicalist theory encoding these lexical differences.

In Optimality Theory (Prince & Smolensky 1993, Grimshaw 1995, Kager 1999) Universal Grammar consists of a set of soft (violable) constraints on well-formedness and individual grammars are constructed by the reranking of these constraints. The expression that satisfies the higher ranked constraints is the optimal, thus, the grammatical one; the others are then considered ungrammatical.

The OT architecture contains an Input that consists of elements from the lexicon. A general structure generator (Gen) constructs candidate expressions from the Input that constitute the candidate set. In the evaluation part, the optimal candidate is selected, based on the language-specific constraint-hierarchy.

The paper is structured as follows. After a short overview of previous OT analyses, the system of French interrogatives will be presented with examples, followed by some important conclusions that have to be taken into account by the analysis. Then comes the analysis itself, starting with the presentation of the framework and concluding with its major claims and their realizations within this architecture. Finally, some examples illustrate the proposed analysis.

---

<sup>1</sup>I would like to thank Anne Abeillé and Mark Newson for their useful comments about earlier versions of this paper, as well as the anonymous reviewer for the equally important remarks.

## 2 Previous analyses

Let us first consider analyses in the framework of Optimality Theory treating interrogatives and French phenomena. For this, a short introduction of OT is offered. In Optimality Theory (Prince & Smolensky 1993, Grimshaw 1995, Kager 1999) Universal Grammar consists of a set of soft (violable) constraints on well-formedness and individual grammars are constructed by the reranking of these constraints. The expression that satisfies the higher ranked constraints is the optimal, thus, the grammatical one; the others are then considered ungrammatical.

The OT architecture contains an Input that consists of elements from the lexicon. A general structure generator (Gen) constructs candidate expressions from the Input that constitute the candidate set. In the evaluation part, the optimal candidate is selected, based on the language-specific constraint-hierarchy. In Pesetsky's account (1997) there is a preference in French for the complementizer to be pronounced instead of the adjacent relative pronoun, if this latter is not modified (2–5):<sup>2</sup>

In non-subject relatives the complementizer *que* is pronounced instead of the relative pronoun (*qui*) preceding it:

- (1) *la fille que j'ai vue*  
the girl that I have seen

Pronouncing the relative pronoun only would lead to ungrammaticality:

- (2) \**la fille qui j'ai vue*  
the girl who I have seen

Pronouncing both is ungrammatical, but is possible in some dialects:

- (3) \**la fille qui que j'ai vue*  
the girl who that I have seen

The relative pronoun must be pronounced when modified:

- (4) *la fille avec qui il est parti*  
the girl with whom he has left

In the OT framework, analyses of interrogatives have been proposed, concentrating on English. According to Grimshaw (1995, 1996), subject-auxiliary inversion and do-insertion

---

<sup>2</sup>In subject relatives the complementizer appears in another form: *qui*

- (i) *la fille qui est partie*  
the girl that has left

in English interrogatives can be accounted for by the following constraints: interrogative operators must appear in specifier positions (CP), no headless projections (auxiliary movement and do-insertion), Stay (economy constraint: no movement, lexical verbs cannot move to C). Ackema & Neeleman (1998) give an OT analysis of multiple questions in English, Bulgarian, Czech and Chinese/Japanese, with different constraint-rankings: QMark (a question must be overtly question-marked), QScope (Q+ elements must c-command the VP at surface structure), Stay. However, neither of these analyses can be applied to French without some necessary adaptations. The present analysis is based on the constraints defined by Ackema & Neeleman (1998), with some modifications, elaborated later on. In what follows the French main clause interrogative system will be presented, to point out some crucial differences with English, and problems the analysis can encounter.

### 3 The French main clause interrogative system

In French main clause interrogatives, depending partly on the register, there are three ways of asking the same question. The first type, at first sight, contains subject-auxiliary inversion. In the second, the question is biclausal, the main clause containing inversion of a dummy verb and a pronoun, whereas the content question follows it with indicative word order in the embedded clause. The third version is present in colloquial French, which is considered as a different grammar. In this latter case, the monoclausal question is used without inversion and interrogativity is indicated only by rising intonation at the end of the sentence. Consider the following examples:

#### 3.1 Polar Interrogatives

All 3 types are present.

- (5) *Parlez-vous français ?*  
 speak<sub>pl2pres</sub> you French  
 Do you speak French?
- (6) *Est-ce que vous parlez français ?*  
 is it which/that you speak<sub>pl2pres</sub> French  
 Do you speak French?
- (7) *Vous parlez français ? (informal)*  
 you speak<sub>pl2pres</sub> French  
 Do you speak French?

## 3.2 *Wh*-Interrogatives

In *Wh* subject interrogatives, on the other hand, this is not always the case. In animate questions there are only two possibilities, since the colloquial question falls together with the monoclausal formal one. However, in inanimate subject interrogatives, the would-be monoclausal question is not grammatical. In object interrogatives the tripartite difference as described above is again present.

- Subject:

- Inanimate:

(8) *\*Que vous dérange ?*  
what cl<sub>pl2acc</sub> disturb<sub>sg3pres</sub>

What disturbs you?

(9) *Qu'est-ce qui vous dérange ?*  
what is it that cl<sub>pl2acc</sub> disturb<sub>sg3pres</sub>

What is it that disturbs you?

(10) *\*Quoi vous dérange ?*

- Animate

(11) *Qui va à Paris ?*  
who go<sub>sg3pres</sub> to Paris

Who is going to Paris?

(12) *Qui est-ce qui va à Paris ?*  
who is it who go<sub>sg3pres</sub> to Paris

Who is going to Paris?

- Object

- Inanimate

(13) *Que faites-vous ?*  
what do<sub>pl2pres</sub> you

What are you doing?

(14) *Qu'est-ce que vous faites ?*  
what is it that you do<sub>pl2pres</sub>

What are you doing?

(15) *Vous faites quoi ? (informal)*  
you do<sub>pl2pres</sub> what

what are you doing?

(16) *\*Quoi faites-vous ? (inversion)*

(17) \*Vous faites que ? (informal)

– Animate:

(18) *Qui cherchez-vous ?*  
who look for<sub>pl2pres</sub> you  
Who are you looking for?

(19) *Qui est-ce que vous cherchez ?*  
who is it that you look for<sub>pl2pres</sub>  
Who are you looking for?

(20) *Vous cherchez qui ? (informal)*  
you look for<sub>pl2pres</sub> who  
Who are you looking for?

(21) \*Cherchez-vous qui ?

(22) \*Est-ce que vous cherchez qui ?

It has to be pointed out that unlike English, French interrogatives do not contain subject-auxiliary inversion at all.<sup>3</sup> Instead, French contains interrogative verb forms in its lexicon. According to Huot (1987) and Miller & Sag (1997) French bound pronominal-clitics (complement or enclitic) are best analyzed as lexical pronominal affixes forming one single lexical unit (word) with the verb<sup>4</sup> (Huot (1987), Miller and Sag (1997)). Similarly, a class of suffixes qualify verbs as interrogative in French, incorporating the subject as well (the hyphenated verb forms in the examples (9-23) all illustrate this phenomenon). The clitic status of these elements can be justified by the fact that since they cannot appear as subjects, the subject has to be present in the clause as well:

(23) *Paul part-il ? (clitic doubling)*  
Paul leave<sub>pl2pres</sub> he  
Paul, is he leaving?

In addition, they cannot take scope over coordinated verbs, unlike ordinary pronouns:

(24) *Il vient et repart aussitôt.*  
He come<sub>sg3pres</sub> and leave<sub>sg3pres</sub> soon.  
He comes and leaves soon.

*Il* as a personal pronoun can take scope over coordinated verbs, unlike the homonymous lexical affix, which must be present on each verb:

---

<sup>3</sup>I thank Anne Abeillé for pointing this out to me.

<sup>4</sup>This is a basic lexical difference between English and French, which shows that the different ranking of the same constraints is not enough when determining differences between particular languages.

(25) \**Vient et repart-il* ?  
 come<sub>sg3pres</sub> and leave<sub>sg3pres</sub> he?

(26) *Vient-il et repart-il* ?  
 come<sub>sg3pres</sub> he and leave<sub>sg3pres</sub> he?  
 Does he come and leave?

From the above data we can draw the following conclusions. French simple questions contain interrogative verbs with enclitics, the bi-clausal complex alternative is constructed with clitic-inversion in the matrix clause (and indicative word order in the embedded clause), and both the *wh*-word and the verb are *in situ* in colloquial French. The presence of interrogative verb forms with enclitics is, therefore, a lexical difference to be taken into consideration by any analysis.

As shown by example (23), the *wh*-word must precede *est-ce que*, and cannot be *in situ* in the case of interrogative verb forms; see (22).<sup>5</sup> On the whole, a conflict between two tendencies can be observed: question-marking with interrogative verb forms (simple questions), or the avoidance of these (indicative word order) in the matrix clause (colloquial French), and the biclausal alternative, which is the combination of the two, satisfying one in the matrix and the other in the embedded clause.

There are three main problems to be solved by the analysis. First of all, in inanimate subject interrogatives only the complex, bi-clausal alternative is possible, neither the simplex one with the interrogative verb form nor the *in situ* question is allowed. Secondly, in the other cases, however, optionality can be observed between the simplex and complex versions that cannot be accounted for by a register difference. Thirdly, the role of *qui/que* in the *qu'est-ce qui/qu'est-ce que* constructions has to be made clear.

Let us deal with the third problem first and compare the French and English systems of *wh*-words and complementizers:

WH operator	Complementizer
qui <sub>(anim)</sub> que <sub>(inanim)</sub>	que, qui
who, what, etc.	that what (non-standard) <sup>6</sup>

As seen in the tableau, both *qui* and *que* can function as a *wh*-operator or as a complementizer. In the present analysis, they both appear in the *que/qui est-ce qui/que* constructions, as *wh* operators in initial positions and as complementizers in the other (both *qu'est-ce qui* and *qui est-ce qui* refer to subjects; in all other cases, however, these constructions end in

<sup>5</sup>Adjunct interrogatives follow the pattern of examples (14-23), and the analysis works identically in their case as well.

<sup>6</sup>I thank Mark Newson for pointing out the non-standard use of *what* to me.



*que*, the non-subject complementizer). From this it follows that there are two lexical entries belonging to *qui* and *que*. This is a second issue in which LFG, a lexicalist theory, facilitates the representation.

As for the second problem, optionality, we can anticipate that with the help of the OT-LFG architecture, the analysis can capture the free choice between the simplex and complex alternatives, as shown below.

Concerning the first issue, it is important to emphasize that this can also be stated in lexicalist terms. More precisely, it is a lexical characteristic of *que* that it cannot appear in subject positions, or, from another point of view, cannot be in nominative case. This accounts for the fact that the sentence \**que vous dérange*, see example (9), is ruled out as ungrammatical and there is no optionality in inanimate subject interrogatives.

Before going on to the proposed analysis, the *que/quoi* alternance needs some clarification. As can be seen from the above examples, *quoi* is the *wh*-word in *in situ* questions, but it is ungrammatical with interrogative verb forms. A possible explanation goes as follows. *Que*, as a weak form, is always cliticized on to a verb, whereas *quoi*, as a strong form does not have to be. *Quoi*, on the other hand, cannot be fronted, whereas *que* cannot appear *in situ* (*qu'est-ce qui* is used as the only possibility). In an OT-LFG analysis, this is also treated in the lexical entries: *que* cannot be a subject (as already pointed out) and *quoi* cannot be focalised (unless made heavy by an adjunct<sup>7</sup>).

---

<sup>7</sup>However, when modified, *quoi* can be fronted:

(i)      *Quoi*    *d'autre*    *as-tu*      *vu*    ?  
             what    else           have you    seen

This can be integrated into the analysis by postulating that *quoi* cannot be focused unless it is syntactically heavy (Abeillé & Godard 1999b). This is made explicit by the following equation:

(ii)                      *quoi*    *d'autre*  
                               what    else  
     NP →    NP        AP  
                               ↑=↓      ↓=(↑ADJUNCT)  
   (↑WEIGHT)=c heavy

The extraction rule is:

(iii)    S →    NP    S  
                   (↑FOC)=↓    ↑=↓  
                   (↑Q)=<sub>c</sub> +     (↑Q)= +  
                   (↑FOC)= GF(↑SUBJ/OBJ/...)  
                   (↑WEIGHT)= heavy

## 4 The proposed OT-LFG analysis

The OT principles are used in the present analysis<sup>8</sup> in order to account for the optionality between the simplex and complex versions of the interrogatives. However, as has already been shown, certain factors (interrogative verb forms in French, *qui/que* as *wh*-word and complementizer and the *que/quoi* difference) necessitate lexical specifications that would be problematic to include in an OT framework.

In OT-LFG (Sells 2001, Kuhn 2001) the OT framework is completed by elements of Lexical-Functional Grammar. The Input is built from the lexical entries of the particular items and is presented as an LFG-type feature matrix, containing the common information content of the candidates. From this,  $G_{inviol}$ , comprising a set of universal rewrite rules, constructs various *c*-structures with a corresponding, more detailed *f*-structure. In this version of the theory the Generator component is conceived as a function between the set of *f*-structures and that of the candidates (the power set of the analyses in  $G_{inviol}$ ). The evaluation of these candidates happens in an OT manner. Let us have a look at the lexical entries of the already treated *wh* words:

*qui*<sub>1</sub> OP (↑PRED)=’pro’  
(↑Q)=+  
(↑ANIM)=+

*qui*<sub>2</sub> COMP (↑Q)=-

*que*<sub>1</sub> OP (↑PRED)=’pro’  
(↑Q)=+  
(↑ANIM)=-  
(↑CASE)≠nom  
(FOC↑)= ↓  
(SUBJ↑)¬= ↓

*que*<sub>2</sub> COMP (↑Q)=-

*quoi* OP (↑PRED)=’pro’  
(↑Q)=+  
(↑ANIM)=-  
(↑CASE)≠nom

As shown above, different entries belong to *qui/que* as *wh*-words and complementizers: Q+ and Q-, respectively, referring to the interrogative or non-interrogative nature of the given element. Moreover, *que* is specified as non-subject, and the lexical entry of *quoi* is neutral with respect to focus status.

---

<sup>8</sup>Based on Gazdik 2006

In what follows we turn to the core of the analysis. It is proposed that the simplex and complex versions of the questions belong to different, but very similar inputs. This accounts for the fact that although their meaning is practically the same (the dummy interrogative verb in the main clause of the complex question does not contribute to it semantically), they are structurally different. The Input belonging to the complex version does not contain the dummy interrogative verb, which can be added to it in some of the candidates. Adding non-input elements to certain candidates is perfectly possible in OT-terms. Such candidates are penalized by the Faithfulness constraint which requires that the Output contain only Input elements and only those, but can, at the same time, satisfy higher-ranking constraints and thus be selected as the optimal candidate.

The input of the complex question is, therefore, an embedded clause (as indicated by the complementizer *que-*) provided with a *wh*-word, or a Q+ feature in polar interrogatives, embeddedness being the only difference between the two inputs. Due to the similarity of these inputs, Gen, in both cases, generates candidate sets that intersect with each other, i.e. the candidate set belonging to the simplex question contains the bi-clausal alternative and *vice versa*; however, in both cases, the evaluations rule out the candidates containing more faithfulness violations than the one closer to the input. The evaluation belonging to the simplex one yields the simplex question as an optimal candidate outruling the complex that is less faithful to the input and *vice versa*.

Before looking at some examples, consider the constraints used in the present analysis:

1. QMark: a question must be overtly Q (question) marked. This can be fulfilled by the presence of an interrogative (Q+) operator followed by an interrogative verb (with enclitic) in French or by subject-auxiliary inversion in English, or, in matrix yes-no questions, by clitic inversion in French or subject-auxiliary inversion in English.
2. QScope: Q+ elements must have scope over the clause they Q mark
3. Lexical verb: lexical verbs cannot Q-mark a question
4. Faithfulness: the output contains all elements that are also included in the input and only those

Proposed rankings:

Language/Order				
French	1	2	3	4
English, colloquial French	3	1	4	2

As the different rankings already reveal, formal and informal French are treated as different grammars. Interestingly, it seems that English and informal French follow the same pattern in constraint-ranking. In what follows, we will examine some examples of formal French and then compare them to informal French and English.

## 5 Examples

### 5.1 Inanimate object interrogatives

In inanimate object interrogatives, optionality can be observed between the simplex and complex versions:

(27) *Que faites-vous ? / Qu'est-ce que vous faites ?*

What are you doing?

The Inputs, differing only in the EMB+ feature are as follows:

1. *Que faites-vous ?*  $\left[ \begin{array}{l} \text{PRED} \quad \text{'do} \langle (\uparrow \text{SUBJ}), (\uparrow \text{OBJ}) \rangle \\ \text{TNS} \quad \text{pres} \\ \text{MOOD} \quad \text{ind} \\ \text{Q} \quad + \\ \text{SUBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'pro'} \\ \text{NUM} \quad \text{pl} \\ \text{PERS} \quad 2 \end{array} \right] \\ \text{OBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'pro'} \\ \text{ANIM} \quad - \\ \text{FOC} \quad + \\ \text{CASE} \quad \text{acc} \\ \text{Q} \quad + \end{array} \right] \end{array} \right]$

2. Qu'est-ce que vous faites ?

$$\left[ \begin{array}{l} \text{PRED} \quad 'do \langle (\uparrow \text{SUBJ}), (\uparrow \text{OBJ}) \rangle' \\ \text{TNS} \quad \text{pres} \\ \text{MOOD} \quad \text{ind} \\ \text{Q} \quad - \\ \text{SUBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad 'pro' \\ \text{NUM} \quad \text{pl} \\ \text{PERS} \quad 2 \end{array} \right] \\ \text{OBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad 'pro' \\ \text{ANIM} \quad - \\ \text{CASE} \quad \text{acc} \\ \text{Q} \quad + \end{array} \right] \end{array} \right]$$

The candidate sets are built from these inputs. Nothing prevents the structure generator ( $G_{invio}$ ) from adding or suppressing elements in the inputs. The second input is, in fact an embedded clause with a non-interrogative verb form (indicated by the complemetizer Q-), in which a Q+ *wh*-word is also present. In order for the question to get Q-marked, a dummy interrogative verb must be inserted in the optimal candidate. Given the similarity of the inputs, the two candidate sets intersect. In some candidates belonging to the simplex question, the dummy interrogative verb might appear, whereas others belonging to the complex question might be constructed with an interrogative verb, a type of faithfulness violation. A possible example of the intersection of the candidate sets is shown in the following:

The intersection of the candidate sets (a subset):

1. que<sub>1</sub> faites-vous
2. faites vous que<sub>1</sub>
3. que<sub>2</sub> vous faites
4. que<sub>1</sub> est-ce que<sub>1</sub> vous faites

All candidates are then submitted to the evaluation, based on the constraint hierarchy.

Evaluations:<sup>9</sup>

1. 

Cand/Constr	QMark	QScope	Faith	LexV
☞ candidate 1				
candidate 2		*!		
candidate 3	*!		**	
candidate 4			**!	

<sup>9</sup>Legend: ☞ optimal candidate, \* constraint violation; !\* fatal violation

2.

Cand/Constr	QMark	QScope	Faith	LexV
candidate 1			**!	
candidate 2		*!		
candidate 3	*!			
☞ candidate 4			*	

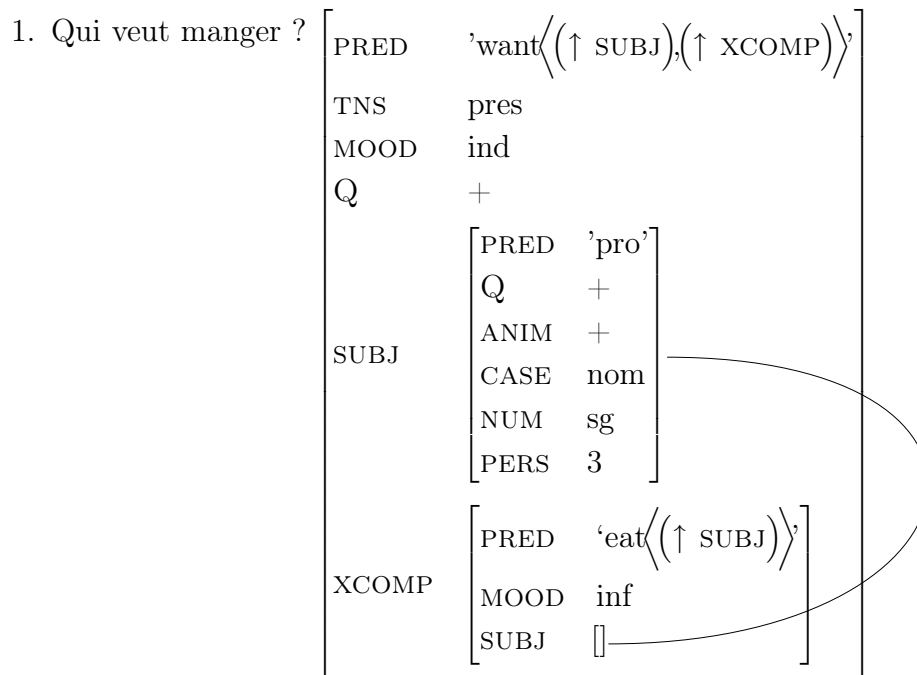
The first evaluation, belonging to the first input, selects the simplex question as the optimal candidate, whereas the complex is chosen in the second. The main reason behind this is the different number of faithfulness violations they commit with respect to the input the evaluation corresponds to. In the first case, the dummy verb is introduced in the complex candidate, which becomes Q- whereas in the second, the simplex candidate is not embedded and changes the verb interrogativity. All the other candidates violate higher-ranking constraints, such as the scope of the *wh*-word or question-marking.

## 5.2 Animate subject interrogatives

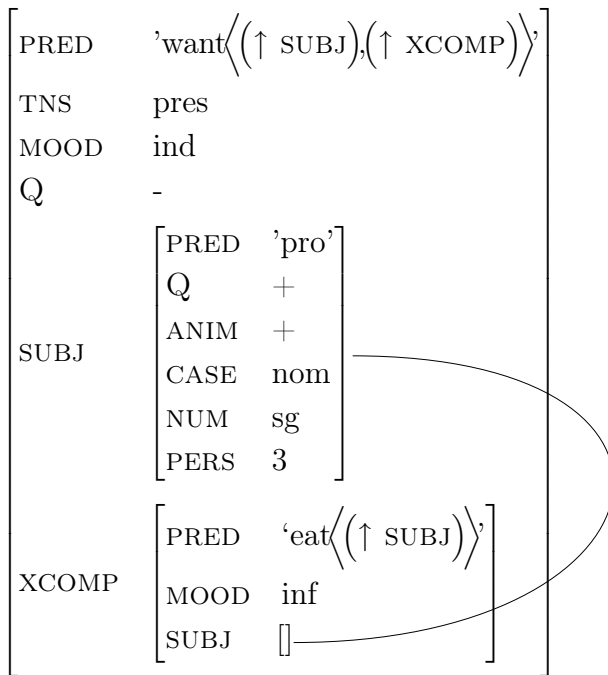
In animate subject interrogatives the same optionality can be observed between the simplex and complex questions.

- (28) *Qui<sub>i1</sub> veut manger ?* / *Qui<sub>i1</sub> est-ce qui<sub>i2</sub> veut manger ?*  
 who want<sub>sg3pres</sub> eat  
 Who wants to eat?

The inputs are the following:



2. Qui est-ce qui veut manger ?



The difference is the same as for object questions, the complex candidate appears as an embedded clause, whereas the simplex is a main clause. Logically, in the case of subject questions we cannot suppose interrogative verb forms, since there is no subject to be incorporated by the complex verb form. Therefore, it is only the *wh*-word, *qui*, that question-marks the interrogative. Some examples from the intersection of the candidate sets:

1.  $qui_2$  veut manger
2.  $qui_1$  veut manger
3.  $qui_1$   $qui_2$  veut manger
4.  $qui_1$  est-ce  $qui_2$  veut manger

The evaluations will bring the same results, the simplex candidate in the first case and the complex in the second, due to the different number of faithfulness violations with respect to the inputs.

1.

Cand/Constr	QMark	QScope	Faith	LexV
candidate 1	*!		**	
☞ candidate 2				
candidate 3	*!		*	
candidate 4			**!	

2.

Cand/Constr	QMark	QScope	Faith	LexV
candidate 1	*!		**	
candidate 2			**!	
candidate 3	*!		*	
☞ candidate 4			*	

The faithfulness violations committed are the following. The complex question contains the complementizer and the dummy verb (not present in the simplex input), whereas the simplex question is not embedded and does not contain the complementizer (as required by the input belonging to the complex question).

After the formal French examples let us turn to informal French and English, which operate, with the same constraint hierarchy.

### 5.3 Informal French

In informal French, the *wh*-word is always *in situ*. The input is similar to that of the simplex candidate in formal French, but due to the different constraint hierarchy, another candidate is selected as the optimal one.

(29) Vous cherchez qui ?

Input:

PRED	'look for'⟨⟨(↑ SUBJ),(↑ OBJ)⟩⟩								
TNS	pres								
MOOD	ind								
Q	+								
SUBJ	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding-right: 10px;">PRED</td> <td style="padding-left: 10px;">'pro'</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 10px;">NUM</td> <td style="padding-left: 10px;">pl</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 10px;">PERS</td> <td style="padding-left: 10px;">2</td> </tr> </table>	PRED	'pro'	NUM	pl	PERS	2		
PRED	'pro'								
NUM	pl								
PERS	2								
OBJ	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding-right: 10px;">PRED</td> <td style="padding-left: 10px;">'pro'</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 10px;">ANIM</td> <td style="padding-left: 10px;">+</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 10px;">CASE</td> <td style="padding-left: 10px;">acc</td> </tr> <tr> <td style="border-right: 1px solid black; padding-right: 10px;">Q</td> <td style="padding-left: 10px;">+</td> </tr> </table>	PRED	'pro'	ANIM	+	CASE	acc	Q	+
PRED	'pro'								
ANIM	+								
CASE	acc								
Q	+								



A subset of the candidate set:

1. tu vois qui
2. qui<sub>2</sub> tu vois
3. vois-tu qui
4. est-ce que tu vois qui
5. qui vois-tu

The Evaluation:

Cand/Constr	LexV	QMark	Faith	QSc
☞ candidate 1		*		*
candidate 2		*	*!	
candidate 3	*!			
candidate 4	*!			
candidate 5	*!			

As can be seen from the tableau, interrogative verb forms are penalized by the first constraint in the hierarchy, thus the formal French-type question is out. The winner is the candidate in which the verb form is non-interrogative, but it is question-marked and more faithful to the input.<sup>10</sup> Let us now turn to English, which we propose has the same constraint hierarchy in the evaluation as informal French.

## 5.4 English

In English, we will examine a polar interrogative with do-insertion. The analysis offered here differs from that of Grimshaw (1995) in that it does not imply movement and from that of Grimshaw (2001) as well. In this latter, structure itself is submitted to economy (although Head and Specifier are obligatory, their absence violate the corresponding constraints) and word order is accounted for by alignment constraints (ex. Head Left or Specifier Left). The same goal is attainable in the present approach by the flexible c-structures, where no constituent is obligatory, they are present in order to satisfy more abstract requirements, such as Q-marking (that can be done in several ways).

(30) Do you speak French?

---

<sup>10</sup>A challenge to the analysis is the question *qui tu vois ?*, which is an amalgamation of formal and informal French (the verb form is not interrogative, and the *wh* word is not in situ). This type is present in informal French, which points towards the reconsideration of the role of the QScope constraint in the grammar of informal French.

The input looks as follows:

Input:

$$\left[ \begin{array}{l} \text{PRED} \quad \text{'speak} \langle (\uparrow \text{SUBJ}), (\uparrow \text{OBJ}) \rangle \\ \text{TNS} \quad \text{pres} \\ \text{MOOD} \quad \text{ind} \\ \text{Q} \quad + \\ \text{SUBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'pro'} \\ \text{NUM} \quad \text{sg} \\ \text{PERS} \quad 2 \end{array} \right] \\ \text{OBJ} \quad \left[ \begin{array}{l} \text{PRED} \quad \text{'French'} \\ \text{PERS} \quad 3 \\ \text{NUM} \quad \text{sg} \end{array} \right] \end{array} \right]$$

A possible subset of the candidate set:

1. do you speak French
2. you speak French
3. speak you French
4. you do speak French

Evaluation:

	<b>Cand/Constr</b>	LexV	QMark	Faith	QScope
☞	candidate 1			*	*
	candidate 2		*!		
	candidate 3	*!			
	candidate 4		*!	*	

Since candidate c violates the LexV constraint, ranked high in English, and candidates b and c are not question-marked, the question with *do*-insertion is selected as the optimal candidate.

## 6 Conclusion

Going back to the main problems such an analysis has to account for, we can draw finally the following conclusions. Optionality in formal French between the simplex and complex

questions is accounted for in an OT manner, by the different number of faithfulness violations with respect to the different but very similar inputs. On the other hand, problems, such as that of *que* and *quoi* and thus that of inanimate subject interrogatives are treated in the lexical entries, specifying *que* as a non-subject and *quoi* as a non-focus, justifying the role of a lexicalist theory in the analysis. Finally, we have also shown that the differences between languages cannot only come from the different constraint rankings, but from lexical differences as well, as English does not possess interrogative verb forms, unlike French.

## References

- [Abeillé 2007] Abeillé, Anne (2007). *Les grammaires d'unification*. Paris: Lavoisier.
- [Abeillé & Godard 1999a] Abeillé, Anne & Godard, Danièle (1999a). La place de l'adjectif épithète en français : le poids des mots. *Recherches Linguistiques*, 28, pp. 9-31.
- [Abeillé & Godard 1999b] Abeillé, Anne & Godard, Danièle (1999b). *Some aspects of interrogative constructions in French*. International HPSG conference, Edinburgh.
- [Ackema & Neeleman 1998] Ackema, Peter & Neeleman, Ad (1998). WHOT?. In Barbosa, Pilar, Fox, Danny, Hagstrom, Paul, Mc Ginnis, Martha and Pesetsky, David, (Eds.) *Is the Best Good Enough?: Optimality and Competition in Syntax*. (pp. 15-34.). Cambridge Massachusetts: MIT Press.
- [Bresnan 2000] Bresnan, Joan (2000). Optimal Syntax. In Legendre, G eral-dine, Grimshaw, Jane, Vikner, Sten, (Eds.), *Optimality-Theoretic Syntax*. (pp. 334-385.). Cambridge, Massachusetts, London, England: MIT Press.
- [Bresnan 2001] Bresnan, Joan (2001). *Lexical-Functional Syntax*. Oxford: Blackwell Publishers.
- [Bresnan & Kaplan 1982] Bresnan, Joan & Kaplan, Ronald (1982). Lexical-Functional Grammar: a formal system for grammatical representation. In: *The Mental Representation of Grammatical Relations*. MIT Press: Cambridge
- [Gazdik 2006] Gazdik, Anna (2006). *Interrogatives in English and French. An OT-LFG Analysis*. MA Thesis, ELTE, Budapest.

- [Ginzburg & Sag 2000] Ginzburg, Jonathan & Sag, Ivan (2000). *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. Stanford: CSLI.
- [Grimshaw 1995] Grimshaw, Jane (1995). *Projection, Heads, and Optimality*. <http://roa.rutgers.edu/files/68-0000/roa-68-grimshaw-3.pdf>
- [Grimshaw 2001] Grimshaw, Jane (2001). *Economy of Structure*. MS. Rutgers University. <http://rucss.rutgers.edu/roa.html>
- [Huot 1987] Huot, H el ene (1987). Morpho-syntaxe du verbe fran ais et inversion du clitique-sujet. In: *Travaux de linguistique* 14-15, 159-176.
- [Kager 1999] Kager, Ren e (1999). *Optimality Theory*. Cambridge: Cambridge University Press.
- [Koml osy 2001] Koml osy, Andr as (2001). *A lexikai-funkcion alis grammatika mondattan nak alapfogalmai*. Budapest: TINTA K onyvkiad o.
- [Kuhn 2001] Kuhn, Jonas (2001). Generation and Parsing in Optimality Theoretic Syntax: Issues in the Formalization of OT-LFG. In Sells, Peter, (Ed.), *Formal and Empirical Issues in Optimality Theoretic Syntax*. (pp. 313-366.). Stanford, California: CSLI Publications.
- [Pesetsky 1997] Pesetsky, David (1997). Optimality Theory and Syntax: Movement and Pronunciation. In Archangeli, Diana, Langendoen, D., Terence, (Eds.), *Optimality Theory: An Overview*. (pp. 134-170.) Oxford: Blackwell.
- [Pesetsky 1998] Pesetsky, David (1998). Some Optimality Principles of Sentence Pronunciation. In Barbosa, P., Fox, D., Hagstrom, P., Mc Ginnis, M. and Pesetsky, David, (Eds.) *Is the Best Good Enough?: Optimality and Competition in Syntax*. (pp. 337-384.). Cambridge Massachusetts: MIT Press.
- [Prince & Smolensky 1993] Prince, Alan, & Smolensky, Paul (1993): *Optimality Theory: Constraint Interaction in Generative Grammar*. <http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF>
- [Sells 2001] Sells, Peter (2001). Introduction. In Sells, P., (Ed.), *Formal and Empirical Issues in Optimality Theoretic Syntax*. (pp. 1-16.). Stanford California: CSLI Publications.