Asymmetries in Italian Matrix Wh-questions: Word Order and Information Structure

Jessica Zipf  
University of Konstanz

Stefano Quaglia  
University of Konstanz

Proceedings of the LFG’17 Conference
University of Konstanz
Miriam Butt, Tracy Holloway King (Editors)
2017
CSLI Publications
pages 387–405

Keywords: Italian, wh-questions, information-structure, scope

Abstract

We develop an analysis that accounts for word order asymmetries found in Italian wh-questions. In particular, the different orderings of subject and verb in perché-questions as opposed to wh-questions involving other wh-items.

In line with Butt (2014) and Mycock (2006, 2013), we see a strong interaction of information structural content and the structure of questions. We argue that the word order differences mirror the semantic divergence of perché in contrast to other wh-items such as chi, quando or dove, among others. We suggest that perché differs from the other wh-items with regard to the partition it introduces at i-structure. While all wh-items introduce a partition between focus and background at i-structure, we argue for the background to be articulated with perché. We show that this correctly captures the fact that perché can scope over a focus operator, while all other wh-items are within the scope of a focus operator, as well as the word order asymmetries found in Italian matrix wh-clauses.

1 Introduction

In this paper, we propose an analysis of matrix wh-questions in Italian from a Lexical-Funntional Grammar perspective. The structure of Italian wh-questions has been investigated in derivational approaches for over two decades. However, this issue has been rarely tackled from a perspective that does not draw on transformational notions.

We focus on matrix wh-questions and address asymmetries between two classes of wh-items, which we call “Class 1” and “Class 2” (cf., among others, Rizzi, 2006; Bocci & Pozzan, 2014):

<table>
<thead>
<tr>
<th>CLASS 1</th>
<th>CLASS 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italian</td>
<td>English</td>
</tr>
<tr>
<td>chi</td>
<td>‘who’</td>
</tr>
<tr>
<td>(che) cosa</td>
<td>‘what’</td>
</tr>
<tr>
<td>dove</td>
<td>‘where’</td>
</tr>
<tr>
<td>quando</td>
<td>‘when’</td>
</tr>
<tr>
<td>come</td>
<td>‘how’</td>
</tr>
</tbody>
</table>

Table 1: Class 1 and Class 2 wh-items

The data we discuss in this paper demonstrate a strong interaction between c-structure and i-structure. The paper is structured as follows: Section 2 illustrates three systematic asymmetries between Class 1 and Class 2 wh-items in Italian matrix wh-questions; Section 3 reviews existing analyses; Section 4 presents our LFG analysis; Section 5 concludes this paper. We are currently working on including our findings and analysis into a grammar fragment of Italian, using the Xerox Linguistic Environment (XLE) (Crouch et al., 2011).
2 Asymmetries in matrix \(wh\)-questions: word order and information structure

In this section we outline the data, illustrating the different surface structures of Italian matrix \(wh\)-questions. Three important asymmetries can be observed between \(wh\)-questions displaying Class 1 vs. Class 2 \(wh\)-items. These asymmetries result from c-structural constraints, but are also connected to the informational status of the subject (and of other GFs). The distribution of subjects in \(wh\)-questions is a long standing issue in the literature on the syntax of Italian, and Romance languages in general (Rizzi, 1996; Poletto, 2000; Barbosa, 2001; Cardinaletti, 2007).

Observation 1:

While a \(wh\)-item is fronted in both (1) and (2), the placement possibilities for the SUBJ-NP Anna differ substantially:

\[
\text{(1) Quando \*Anna balla, Anna?}
\text{\hspace{1cm} when \*Anna dance.PRS.3SG Anna}
\text{\hspace{1cm} ‘When does Anna dance?’}
\]

\[
\text{(2) Perché Anna balla, Anna?}
\text{\hspace{1cm} why Anna dance.PRS.3SG Anna?}
\text{\hspace{1cm} ‘Why does Anna dance?’}
\]

While the subject cannot be realised between the main verb and a Class 1 item like quando in (1), this preverbal position is acceptable and in fact unmarked with the Class 2 element perché (‘why’) in (2).

Observation 2:

Bocci & Pozzan (2014) point out that questions introduced by perché, and with the subject in postverbal position, allow the subject to receive a contrastive focus interpretation without additional prosodic marking (3b). On the other hand, a \(wh\)-question with this type of interpretation for the subject is not felicitous if the subject is realised preverbally, as in (3c). In this case, the choice between a preverbal and a postverbal subject thus depends on information structure. Questions introduced by a Class 1 \(wh\)-element do not allow a focus interpretation of the subject.

\[
\text{(3) a. (Speaker A)}
\text{Gianni mi ha appena portato il libro}
\text{John me AUX.3SG just bring.PST.PTCP the book}
\text{‘John has just brought me the book.’}
\]

\footnote{The contrast between (3b) and (3c) is neutralised if the subject-NP is realised with focal stress.}
b. (Speaker B)

Perché te l’ha portato Gianni (e non why to-you CL.3SG.M-AUX.3SG bring.PST.PTCP John and not Leo)?
Leo

‘Why did John bring it to you and not Leo?’

c. (Speaker B)

#Perché Gianni te l’ha portato (e non why John to-you CL.3SG.M-AUX.3SG bring.PST.PTCP and not Leo)?
Leo

‘Why did John bring it to you and not Leo?’

Yes-No questions exhibit the same behaviour:

(4)

a. (Speaker A)

Gianni mi ha appena portato il libro
John to-me AUX.3SG just bring.PST.PTCP the book

‘John has just brought me the book.’

b. (Speaker B)

Te l’ha portato Gianni? Credevo
to-you CL.3SG.M-AUX.3SG bring.PST.PTCP John? think.PST.1SG
che te lo dovessi portare Leo.
that to-you it should bring Leo.

‘Did John bring it to you? I thought that Leo should have brought it.’

c. (Speaker B)

#Gianni te l’ha portato? Credevo
John to-you CL.3SG.M-AUX.3SG bring.PST.PTCP? think.PST.1SG
che te lo dovessi portare Leo.
that to-you it should bring Leo.

‘Did John bring it to you? I thought that Leo should have brought it.’
Regarding this property, yes-no questions and *perché*-questions behave like declarative sentences:

(5) a. Gianni ha telefonato.
    John AUX.3SG call.PST.PTCP
    ‘John called.’

b. Ha telefonato Gianni.
    AUX.3SG call.PST.PTCP John
    ‘John called.’

In (5b), the subject can receive a narrow (information, or merely contrastive) focus interpretation by occupying a postverbal position. The same focus reading is, however, unavailable for preverbal subjects (5a).

In contrast, in questions with a Class 1 *wh*-item, as in example (1), the postverbal subject is “marginalized” (Antinucci & Cinque, 1991; Cardinaletti, 2001), meaning that it is realised to the right of the verb (complex), destressed and interpreted as presupposed information. Hence, a postverbal subject can only get a narrow focus interpretation if the introducing *wh*-element is *perché*.

Alternatively, postverbal subjects in Class 2 *wh*-questions can get an interpretation similar to the one of “marginalized” subjects in Class 1 *wh*-questions:

(6) *Perché* te l’ha portato, Gianni?
    why to-you CL.3SG.M-AUX bring.PST.PTCP John
    ‘Why did John bring it to you?’

**Observation 3:**

Additionally, *perché* is compatible with a contrastive/corrective focus in the left periphery of the sentence, differing once more from Class 1 *wh*-expressions:

(7) *Perché* QUESTO avremmo dovuto dirgli e non qualcos’altro?
    why this AUX.3PL to-have.to.PST.PTCP say.him and not something-else
    ‘Why THIS we should have said to him, not something else?’ (Rizzi, 2001)

(8) *Quando* QUESTO avremmo dovuto dirgli e non qualcos’altro?
    when this AUX.3PL to-have.to.PST.PTCP say.him and not something-else
    ‘When THIS we should have said to him, not something else?’ (Rizzi, 2001)

\(^{2}\)We adopt Bocci & Pozzan’s (2014) definition of “merely contrastive focus” as a focus creating “a contrast internal to the sentence that does not imply the correction of a previously asserted proposition” (cf. Bocci & Pozzan 2014:33).
Note that in structures like (7), the word order is fixed: the contrastively focused XP can never precede perché:

(9) *QUESTO perché avremmo dovuto dirgli e non this why AUX.3PL to-have-to.PST.PTCP say.him and not qualcos’altro? something-else

‘Why THIS we should have said to him, not something else?’ (Rizzi, 2001)

3 Previous Accounts

As mentioned in the introduction, there are several analyses which seek to account for these asymmetries, and especially for the ban on preverbal subjects in questions introduced by a Class 1 wh-item (1). Since these are all situated within derivational theories, we will only outline the most prominent ones.

According to Rizzi (1996, 2001, 2006), Class 1 and Class 2 wh-items fulfill the so-called WH-Criterion in different ways. Put simply, the criterion requires the following: A wh-operator must be in the Specifier of a head X₀ that carries the interrogative feature [+wh] (Rizzi, 1996).

In his analysis, all wh-items are hosted within the CP system, but not in the same position. Class 1 items only meet the criterion by moving to the left periphery of the clause, thereby crossing the subject and triggering “subject-verb inversion”: the inflected verb carrying the wh-feature must move to the complementiser position to create the required Spec-head configuration with the wh-operator, depicted below:

![Figure 1: Spec-head configuration with the wh-operator](image)

Assuming an articulated left periphery, Rizzi (2006) argues that a Class 1 item moves to the specifier position of the Focus head, and the TP is subsequently moved to the left periphery as well, thereby crossing the subject:

(i) [CP whereᵣ [C: hasⱼ [IP John tⱼ gone tⱼ]]]
On the other hand, the Class 2 item perché is directly merged in a higher position that does not attract movement, namely the specifier of Int(errogative). In the latter case, no movement is needed and the subject can appear between the interrogative and inflected verb. However, the movement of a Class 1 item, with subsequent movement of the TP, prevents the subject from intervening between the interrogative and the main verb. Rizzi’s analysis leads to the following structure of the complementiser position:\(^4\)

\[
\begin{array}{c}
\text{Force} \\
\text{Spec(ForceP)} \\
\text{Int} \\
\text{Spec(IntP)} \\
\text{Foc} \\
\text{Spec(FocP)} \\
\text{“bare”wh} \\
\text{Spec(FinP)} \\
\text{Fin} \\
\text{Spec(IP)} \\
\end{array}
\]

Figure 2: Structure of the complementiser position (Rizzi, 2006)

In contrast to Rizzi, Cardinaletti (2007) argues that movement from I to C does not take place in Romance questions at all. While supporting the idea that \(wh\)-items are hosted in CP, and preverbal subjects in a lower SubjP position within the IP, she does not assume that the verb (or a projection of the verb) moves to the C-domain. According to her work, the ban against preverbal subjects as in (1) is due to a selective intervention effect which rules out subjects in the specifier of SubjP in questions. Adapting her previous hypothesis, she postulates more than one preverbal subject position, each one specialised to host different subjects. Full DPs and strong pronouns are hosted in the specifier of SubjP, pronominal subjects, e.g. the empty \(pro\) or weak subject pronouns, are hosted in the specifier of TP, leading to the following structures:

\(^4\)For the sake of convenience we left out optional topic projections that can be inserted below each functional projection.
The concluding generalisation is that “only strong subjects are to be excluded from occurring between the wh-phrase and the verb in wh-questions, whereas weak subjects (either null or overt) are permitted” (Cardinaletti 2007:66).

The proposal by Shlonsky & Soare (2011) differs in that it does not assume why to be directly merged in the specifier of IntP, but moved there from the specifier of the functional projection ReasonP.

Stepanov & Tsai (2008) suggest that why must be inserted in the left periphery within CP. As such, why takes the entire IP as its argument, allowing for the correct scope relations.

4 Our Analysis

In this paper, we focus on the differences between perché and Class 1 wh-items and do not consider the other members of Class 2. Given that this paper aims at an analysis within the framework of LFG, we integrate c- and f-structural constraints combined with i-structural information. We thereby take into account work on wh-questions in other languages, such as Urdu/Hindi, and the general relation between discourse functions and questions (Butt & King, 1996; Butt, 2014; Mycock, 2006, 2013).

In their work, Butt & King (1996) define discourse functions by using binary values for the two features [±NEW] and [±PROM(INENT)]:

<table>
<thead>
<tr>
<th>+PROM</th>
<th>+NEW</th>
<th>−NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FOCUS</td>
<td>TOPIC</td>
</tr>
<tr>
<td>−PROM</td>
<td>COMPLETIVE</td>
<td>BACKGROUND</td>
</tr>
<tr>
<td></td>
<td>INFORMATION</td>
<td>INFORMATION</td>
</tr>
</tbody>
</table>

Table 2: Discourse Functions (Butt & King, 1996)

Adopting Mycock (2013) and her account for the discourse functions that wh-words may have, we consider the feature [±NEW] to have a relational definition. Referring to Lambrecht (1997), Mycock (2013) argues that new information can be conveyed regardless of whether the designata of different parts of the sentence has been mentioned before or not: “Given an appropriate context a sentence consisting solely of anaphoric expressions such as She did it serves to establish relations between the various elements of the proposition, thus conveying new information and changing the addressee’s representation of the world.” (Mycock, 2013, p. 2). With regard to the feature [±PROM], we adopt the definition given by Butt & King (1996) in that it depicts whether or not information is of high importance to the information structure of the discourse at hand.
We first assume that the presuppositions, which we represent as BACKGROUND, associated to both Class 1 and Class 2 wh-items are “open” propositions where an operator binds a variable. However, we argue that these elements differ with regard to the partition they introduce at i-structure. In order to account for the differences between Type 1 wh-items and perché, we need to take a closer look at their respective semantic properties.

It has generally been observed that reason/cause wh-items like why interact differently with focused elements compared to other wh-items (irrespective of the argument vs. adjunct distinction). Consider the following sentence pairs (Stepanov & Tsai, 2008):

(10)  a. Why did Adam eat the apple?
    b. Why did ADAM eat the apple?
    c. Why did Adam eat the APPLE?
    d. Why did Adam EAT the apple?

(11)  a. When did Adam eat the apple?
    b. When did ADAM eat the apple?
    c. When did Adam eat the APPLE?
    d. When did Adam EAT the apple?

While each question in (10) has different expectations as to what counts as an acceptable answer ((b.) Because he (Adam) was the one that Eve worked on; (c.) Because it (the apple) was the only food around; etc.), all questions in (11) can be answered with one and the same answer (e.g. At 9 am on December 14). Thus, while the truth value of an answer depends on the focused element in a why-question, the truth value of the answer does not relate to the focused constituent with when (Stepanov & Tsai, 2008).

However, the contrast illustrated in (10) and (11) cannot straightforwardly be transferred to Italian. As we have shown in Section 2, Class 1 wh-items like quando ‘when’ cannot appear with a focused XP in matrix questions. Matrix constructions mirroring the structures in (11) are thus ungrammatical. Clefted structures must instead be used, as in (12), leading to following two paradigms:

(12)  a. Perché ADAM ha mangiato la mela?
    b. Perché Adam ha mangiato LA MELA?
    c. Perché Adam ha MANGIATO la mela?
    d. Perché Adam ha mangiato la mela?
    Why Adam AUX.3SG eat.PST.PTCP the apple
    ‘Why did Adam eat the apple?’

(13)  a. Quand’è che ADAM ha mangiato la mela?
    b. Quand’è che Adam ha mangiato la MELA?
c. Quand’è che Adam ha MANGIATO la mela?

Parallel to the English sentences, the presence of another focused XP influences the expectations of a felicitous answer in (12) but not in (13). Like the question in (10b), the question in (12a) implicates that someone else could have eaten the apple, but did not.

However, this reading is unavailable for the question in (13a) (which parallels (11b)). Such a question can be uttered in a situation in which other people apart from Adam ate the apple too, but the speaker is only interested in Adam. In order to explain this asymmetry, Stepanov & Tsai (2008:602) argue that while when is in the scope of the focus operator, the focus operator itself is in the scope of why:

\[
(14) \quad (\exists x = Adam) (\text{When } t) \ (x \text{ ate the apple at } t)
\]

\[
(15) \quad (\text{Why}) (\exists x = Adam) \ (x \text{ ate the apple})
\]

In other words, INTERROGATIVE FOCUS can have scope over CONTRASTIVE FOCUS with why, but not with when. Dretske (1972:412) describes what we have shown for why in (10) and (12) as:

If \(C(U)\) is a linguistic expression in which \(U\) is embedded, and \(U\) can be given different contrastive foci say \(U_1\) and \(U_2\), then it often makes a difference to the meaning of \(C(U)\) whether we embed \(U_1\) or \(U_2\).

We argue that the different scope possibilities just shown can be explained if we assume that Class 1 wh-items and perché introduce different partitions at the level of i-structure.

To account for the presented data, we suggest a slight modification of the notion of BACKGROUND: While Class 1 items introduce a partition in the i-structure between FOCUS and BACKGROUND, we assume the same partition for perché, but with an articulated BACKGROUND INFORMATION. Specifically, why introduces a partition between FOCUS and BACKGROUND, but the BACKGROUND can itself be articulated and contain other i-structure categories like TOPIC, FOCUS and BACKGROUND 5:

\footnote{For ease of exposition, we follow Rooth (1997) and use (abstract) trees to represent semantic composition. Nonetheless, we do not assume that semantic composition follows c-structure representation.}
Adapting Rooth (1997), $C$ is a semantic variable, denoting an alternative set and $\sim$ represents the focus interpretation operator that contributes a constraint on the variable $C$. According to Hamblin (1973), the value of a question is represented by a set of possible answers. Such a set is represented by $C$. The focus interpretation operator $\sim$ illustrates where the i-structure information is accessed (see Rooth 1997). In the semantic composition, the lowest BACKGROUND and FOCUS form one unit, which itself forms the BACKGROUND to the FOCUS why. If we reconsider questions with perché and a focused XP as in (12a), repeated in (16):

\begin{equation}
(16) \text{Perché Adam ha mangiato la mela? Why Adam AUX.3SG eat.PST.PTCP the apple 'Why did Adam eat the apple?'}
\end{equation}

Decomposing from top-to-bottom: perché corresponds to the highest FOCUS, while the focused XP Adam, together with IP ha mangiato una mela, form the BACKGROUND to the focal perché. In turn, that BACKGROUND consists of the FOCUS Adam, with BACKGROUND ha mangiato una mela. In contrast, the partition into FOCUS and BACKGROUND introduced by Class 1 items can be depicted as:

\begin{equation}
\text{Figure 4: Proposed i-structure configuration for Class 1 wh-items}
\end{equation}

In this case, the introduced partition places the wh-item in FOCUS and necessarily forces the subject to be backgrounded information, because the category BACKGROUND is not further articulated with Class 1 wh-items. Apart from the differing i-structure partitions, we argue that the two classes of wh-
items also differ with respect to the access of the i-structural information within semantic composition:

![Diagram of semantic composition of perché]

Figure 5: Semantic composition of perché

Here, the i-structural information is added to the semantic composition at XP, that is as a complement to, and embedded under, perché. As such, it is in the scope of perché.

However, with Class 1 wh-items, we assume the i-structure information is added to semantic composition at YP, in which the wh-item is included and embedded itself:

![Diagram of semantic composition of Class 1 wh-items]

Figure 6: Semantic composition of Class 1 wh-items

This accounts for the fact that only why but not Class 1 items can have scope over the focus operator.

With regard to c- and f-structural constraints we propose the following: Both, Class 1 wh-items, and contrastive foci are inserted in a c-structural position associated with a functional uncertainty path. Accordingly, in a structure like (1) quando ‘when’ corresponds to FOCUS, whereas the subject must be realised in a position associated with BACKGROUND INFORMATION because, as argued in the previous section, Class 1 wh-items introduce BACKGROUND that cannot be further split into additional i-structure categories. Importantly, the preverbal position, i.e. SpecIP, is not such a position, since it is associated with the discourse function TOPIC.

In contrast, perché allows for an articulated BACKGROUND. Consequently, the introduced partition at i-structure differs considerably: While subjects are forced into the BACKGROUND with Class 1 items, they are free to be positioned in c-structure nodes corresponding to TOPIC or BACKGROUND (the preverbal and postverbal positions in 17 respectively) or contrastive FOCUS (18) with perché:
(17) **Perché** {Anna} balla, {Anna}? why {Anna} dance.PRS.3SG, {Anna}? ‘Why does Anna dance?’

(18) **Perché** te l’ha portato GIANNI (e non why to-you CL.3SG.M-AUX.3SG bring.PST.PTCP John and not Leo)? Leo ‘Why did John bring it to you and not Leo?’

The resulting c-structure template that we propose is shown below. To allow for a more comprehensive presentation, we also illustrate each of the three important subtrees separately.

**Resulting C-Structure Template**

![C-Structure Template Diagram](image-url)

The main reason for assuming two CPs is the rich and highly articulated structure of the left periphery in Italian (Rizzi, 1997; Cinque, 1999; Rizzi & Bocci, 2017).

The c-structure thus reflects the ordering of complementisers and topics in Romance clitic left-dislocated constructions: while the **TOPIC** is preceded by the finite complementiser *che* (‘that’), it must be followed by the infinitival complementiser *di* (‘of’):

---

6The present c-structure is not meant to represent the case of CP recursion but rather two different C-related projections. As suggested by a reviewer, CP₁ can be thought of as the FinP projection and CP₂ as the ForceP projection by Rizzi (2006). However, we prefer remaining neutral to the specific implications of these projections.
(19) Ho deciso che, [TOPIC la macchina], la comprerò.
AUX.1SG decide.PST.PTCP that, [TOPIC the car], it buy.FUT.1SG
quest’anno
this’year
‘I decided that, the car, I will buy it this year.’
‘I decided to buy the car this year.’ (Rizzi & Bocci, 2017:3)

(20) Ho deciso, [TOPIC la macchina], di comprarla.
AUX.1SG decide.PST.PTCP, [TOPIC the car], of buy.INF.CL
quest’anno
this’year
‘I decided, the car, of to buy it this year.’
‘I decided to buy the car this year.’ (Rizzi & Bocci, 2017:3)

In sentence (2), repeated in (21), the preverbal subject corresponds to such a left-dislocated TOPIC situated in the specifier of CP₁. Alternatively, the preverbal subject can be situated in the specifier of IP. Both positions are available and probably lead to slight differences at i-structure. However, we do not go into further detail on this matter. On the other hand, the postverbal subject is an instantiation of BACKGROUND, placed in IP (illustrated in the configuration of IP on the following page).

(21) Perché {Anna} balla, {Anna}?  
why {Anna} dance.PRS.3SG, {Anna}?

‘Why does Anna dance?’

Subsequently, we illustrate each of the three important subtrees of our proposed c-structure template.

**Configuration of CP₂ - Position of perché**

[Diagram]

Figure 8: Position of perché
Configuration of CP₁ - Position of Class 1 wh-items

Figure 9: Position of Class 1 wh-items

Configuration of IP - further Subject Positions

Figure 10: Subject Positions
5 Conclusion

In this paper, we proposed an analysis for word order asymmetries in Italian matrix wh-questions. In particular, we investigated the differences between Class 1 wh-items and perché. Our proposal integrates and combines their different behaviours and properties at the levels of c- and f-structure, as well as i-structure. We suggest that perché introduces a partition at i-structure between FOCUS and BACKGROUND, like Class 1 wh-items, but unlike those, the BACKGROUND is articulated and can itself be further split into additional i-structure categories. This allows perché to co-occur with another focused XP, influencing the truth conditions for felicitous answers. However, the BACKGROUND associated with Class 1 wh-items is not articulated, and cannot contain additional i-structure categories. We propose that this difference is reflected at c-structure: Perché is inserted in a higher c-structure position hosting sentential operators corresponding to clause-level adjuncts. As such, subjects are free to be positioned in c-structure nodes corresponding to TOPICS or BACKGROUND or contrastive FOCUS. As for Class 1 wh-items, we take these to be in the same position as contrastive foci. Specifically, we suggest that Class 1 wh-items are inserted in a c-structural position associated with a functional uncertainty path, forcing the subject to be realised in a position associated with BACKGROUND information. Our current and future work involves an implementation of these results in a small grammar fragment of Italian.
References


