QUESTIONS WITH DECLARATIVE SYNTAX TELL US WHAT ABOUT SELECTION?¹

JONATHAN DAVID BOBALJIK
SUSI WURMBRAND
University of Connecticut

1 Introduction

One of the many enduring themes of Chomsky's (1965) Aspects is the question of selection (broadly construed) and the distinction among syntactic and semantic properties (features) of linguistic expressions. In this brief contribution, we aim to reaffirm the role that syntactic selection plays in the domain of clausal embedding; that is, where verbs select for a complement of a particular syntactic type and a semantically (or pragmatically) equivalent utterance is sharply ungrammatical. Our specific focus is to synthesize a body of literature on the phenomenon of ‘optional’ (non-echo) wh-in-situ in wh-movement languages, arguing ultimately that syntactically, the phenomenon as such may not exist. What appears to be wh-in-situ in these languages may carry interrogative force as a speech act, but from a syntactic perspective is a declarative clause with a wh-expression in focus—a question with declarative syntax (DSQ). The key evidence for this claim comes from selection/subcategorization. The relevant facts have been noted for individual languages, including English, but we offer here a meta-study, of sorts, contending that the generalization in (1) holds systematically across all languages we have been able to examine, despite a wealth of variation along other dimensions.

(1) DSQ/wh-in-situ generalization:
   If a language has wh-movement (to Spec,CP), then wh-movement is obligatory in indirect questions. Equivalently: If a wh-movement language allows ‘optional’ wh-in-situ, the in-situ construction is blocked in selected questions.

¹ For examples from other languages we thank Meltem Kelepir (Turkish), Gísli Rúnar Harðarson (Icelandic), and the ARD and ZDF Mediatheken. The paper also greatly profited from feedback and discussions with Benjamin Girard-Bond, Željko Bošković, Jon Gajewski, Magdalena Kaufmann, Jairo Nunes, Lara Reglero, Koji Shimamura, Sandra Wood, and especially Cynthia Levart-Zocca.
The systematicity with which this holds, and the equally systematic absence of a restriction on embedding in *wh-in-situ* languages, provides a compelling argument, in our view, against analyses such as Cheng and Rooryck (2000) and Pires and Taylor (2007) which propose to assimilate the optional *wh-in-situ* facts of English, French or other languages to the *wh-in-situ* constructions of, say, Chinese, Japanese or Turkish. We suggest instead (partly in line with Ginzburg and Sag’s [G&S] 2000 analysis of English) that DSQs are syntactically, if not semantically, akin to echo questions. In contrast to true *wh-in-situ*, DSQs involve no interrogative syntax (no question operator or interrogative C) and find an interpretation as a question in the pragmatics (for accounts of echo questions along these lines, see Artstein 2002, Poschmann 2010). True *wh-in-situ* by contrast (contra G&S) involves an interrogative complementizer, C*WH*, and thus is syntactically typed as a question (Cheng 1991; whether this triggers covert movement in one form or another, or merely binds the *wh*-word is immaterial here). This C*WH* allows (true) *wh-in-situ* questions to be selected by a higher predicate, forming indirect questions. Since DSQs lack C*WH*, not only is there no *wh*-movement (overt or covert), but a DSQ clause cannot be selected as an interrogative, explaining the observed correlation. At its core, our explanation of (1) is thus that selectional compatibility reveals syntactic properties that are partially obscured in simple clauses by alternative (“pragmatic”) strategies. In their syntax languages are either *wh*-movement or *wh-in-situ* (more accurately, *wh-in-FOCUS*) but no language (that we know of) truly mixes both constructions. As a syntactic corollary, we note that the generalization presented here supports an account where *wh*-movement, when it occurs, is the consequence of a property of C, and not triggered by a property inherent to *wh*-expressions forcing them to move (or to be syntactically licensed in any way). We sketch a formal account within the Reverse Agree framework of Wurmbrand (2012a, b, 2014) which supports the various assumptions we are led to from the broader, cross-linguistic considerations just sketched.

2 DSQs

English is routinely described as lacking *wh-in-situ*, and thus requiring movement of a *wh*-expression in contexts other than echo questions. This is an over-simplification. In English, it is possible to ask an information-seeking, non-echo *wh*-question without fronting the *wh*-word, as noted, for example, in standard descriptions (Huddleston and Pullum 2002: 973; hereafter *CGEL*) and occasionally in the theoretical literature (Postal 1972, G&S 2000, Pires and Taylor 2007, Zocca DeRoma 2011). Some examples are given in (2); none of these fit the defining criteria of echo questions in the contexts in which they occur.²

(2) a. So, your boy’s name is what? [McNulty, *The Wire*, season 1, episode 1]

b. Major, you want this stuff where? [unnamed officer, *The Wire*, season 2, episode 1]

c. A: All the creative people — our R&D, marketing, in-house ad staff — that's all done here in Jersey.

   B: But the sneaks are made where? China? Malaysia? [Stabler, *Law and Order*]

d. Briscoe: What do you suspect?

   Doctor: She was poisoned, by her own hand or someone else’s.

   Briscoe: And you know this how?

² Examples (2c,e) are from the collection of examples in Zocca DeRoma (2011).
Doctor: Fresh needle mark on her left buttock. [Law & Order, season 10, episode 23]
e. Mrs. Valentine: Good news, darling. Your father's spirit guide has allowed him to
leave his crystals and meditation mat long enough to come to New York on business,
and he wants to see you tonight.
Drue: And this is good news on what planet? [Dawson's Creek, Season 4, Ep. 11]
f. “Now,” said Umbridge, looking up at Trelawney, “you've been in this post how long,
exactly?” [Harry Potter and the Order of the Phoenix, Ch. 15]

Some German examples illustrating the same point are given in (3):

(3) a. Und diese Teilhaber erreichen wir wie? ‘And we can reach these partners how?’ [Stark, Schweinegeld / Tatort ep. 746]
b. Köhler: Ich hab nicht viel Geld. Herr Klarmann hat mir geholfen…
Wilsberg: Und diese “Hilfe” dauert jetzt wieviel Jahre? ‘And this “help” has been lasting for how many years?’ [Wilsberg, Bullenball]

Though somewhat restricted to a colloquial register, examples of DSQs are not hard to find. They are distinct from echo questions, in that, for example, they occur in contexts that are not asking for repetition or clarification of an expression in a preceding utterance (see CGEL: 886-891). The examples in (2a) and especially (2b,d), as is clear from the contexts they occur in, are genuine requests for information. The early literature identifies, as species of unmoved questions in English, at least legalistic questions (Postal 1972) which further a string of questioning such as (2a,c), and quiz-show questions, such as (4).

(4) a. And now, for $5,000, London is the capital of which country? (Cooper 1983:148)
b. During the Korean War, the United Nations forces made largely of the troops of the
United Nations and South Korea fought against the troops of North Korea and what
[country]? [COCA]
c. Louis XIV was how old when he became King? studystack.com

Another context supporting such questions in English, noted by G&S 2000, is one in which further information is requested to flesh out a salient context, where there is no strict linguistic antecedent to count as an echo. Pires and Taylor (2007) illustrate with an example like (5a), while Poschmann (2010) and Kaufmann and Poschmann (2011) give (German) examples like (5b):

(5) a. [Seeing somebody reading]: You’re reading what?

3 Echo questions repeat a prior utterance, substituting a wh-expression such as what for a part of that utterance, requesting repetition or clarification (or rhetorically expressing surprise). In response to A’s utterance in (i), speaker B’s responses in either (ii) or (iii) are echo questions. Note that the string replaced by what in an echo question may be smaller than a word (Cooper 1983:150, CGEL)—this is not generally possible in the (other) DSQs we consider.

(i) A: I saw a sign in the market advertising “Grillhändchen.”
(ii) B: You saw a sign advertising WHAT?
(iii) B: You saw a sign advertising Grill-WHAT(-chen)?
b. [Discussing pot-luck plans]: Diane’s baking a cake, Magda’s buying bagels, and Harry’s bringing what?

DSQs in English often have a sarcastic or disdainful edge to them and can be used as rhetorical questions. Examples (2c-f) illustrate this flavor, but this is neither obligatory nor unique to wh-in-situ and would be available to the corresponding fronted questions as well.\(^4\)

In sum, despite their declarative (focus) syntax, DSQs are (or can be) genuine interrogative speech acts. In terms of their syntactic distribution, they are quite free. The apparently in-situ wh-word may be in an embedded clause, as in (6a), and even in an island (where the corresponding wh-movement would be impossible), as in the coordinate structure island in (6b) or the relative clause island in German (6c):

(6)  
\(a.\) And the defendant claimed that he was standing where?  
\(b.\) During the Korean War, the UN forces made largely of the troops of the UN and South Korea fought against the troops of [North Korea and what country]?  
\(c.\) Er hat den Mann, der aus welchem Grund nach Patmos gefahren ist, angerufen?  
He has the man who out which reason to Patmos gone is called ‘He called the man [who went to Patmos for what reason]?’

A prominent line of analysis (we discuss others below) treats DSQ in English, German, and other wh-movement languages as instances of the wh-in-situ constructions familiar from languages such as Chinese and Turkish:

(7)  
\(a.\) Hufei mai-le shenme?  
Hufei buy-ASP what  
‘What did Hufei buy?’  
\([\text{Cheng 2003: 103}]\)  
\(b.\) Ozan ne oku-du?  
Ozan what read-PAST  
‘What did Ozan read?’  
\([\text{M. Kelepir, p.c.}]\)

For example, Cheng and Rooryck (2000), analyze French ‘optional wh-in-situ’ as involving a null interrogative complementizer that licenses the in-situ wh-expressions in French in a manner directly analogous to wh-in-situ licensing in Chinese. Pires and Taylor (2007) offer such a proposal for English (and analogous examples in Brazilian Portuguese). Noting the island-insensitivity, they adapt a familiar analysis of wh-in-situ, positing a null C which binds the wh-words in its domain, requiring no movement (neither overt nor covert). In these analyses, languages that allow ‘optional’ wh-in-situ have two interrogative complementizers, one that is associated with (overt) wh-movement, the other licenses the wh-word without overt movement, (either with covert movement of one sort or another, or via unselective binding or equivalently, Agree). G&S (2000) propose a very different syntactic treatment of English DSQs, yet share with the analyses above the general suggestion that DSQs in a language like English invoke the same grammatical pieces as standard interrogatives in a wh-in-situ language. In the next section,

\(^4\) The text after (2f) draws this out: “Professor Trelawney scowled at her, arms crossed and shoulders hunched as though wishing to protect herself as much as possible from the indignity of the inspection. After a slight pause in which she seemed to decide that the question was not so offensive that she could reasonably ignore it, she said in a deeply resentful tone, ‘Nearly sixteen years.’”
we show that despite their syntactic freedom, DSQs obey one strikingly robust restriction— they are infelicitous as indirect questions, which sets them apart from true wh-in-situ.

3 Wh-in-situ vs. wh-in-situ

A systematic point of difference between ‘optional’ and ‘true’ wh-in-situ, as far as we have been able to determine, lies in the possibility of occurring as the complement to a predicate that syntactically selects for an interrogative: DSQs cannot occur as indirect questions. In optional wh-in-situ languages, if a wh-phrase occurs in the complement of an interrogative-selecting predicate, wh-movement is obligatory in the embedded clause and unmoved variants are sharply ungrammatical or parsed as direct quotes. This is shown in (8) for English (see also G&S 2000; CGEL: 973), and (9) for German.5

(8) a. *He asked me your boy’s name is what.
   b. *I wonder I should put this stuff where.
   c. *Umbridge asked Trelawney she’s been in the post how long.

(9) a. *Stark hat gefragt diese Teilhaber erreichen wir wie?
   Stark has asked these partners reach we how
   ‘Stark asked we can reach these partners how?’
   b. Stark hat gefragt wir diese Teilhaber erreichen.
   Stark has asked we these partners reach
   ‘Stark asked how we can reach these partners.’

In true wh-in-situ languages, on the other hand, this restriction is not found, and indirect questions show the wh-in-situ configuration in the embedded clause:

(10) a. Botong xiang-zhidao [Hufei mai-le shenme]? Mandarin
   Botong want-know [Hufei buy-ASP what] 
   ‘Botong wants to know what Hufei bought.’ [Cheng 2003: 103]
b. [Ozan’in ne oku-dug-un-u] merak ed-iyor-um Turkish
   [Ozan what read-NOM-POS-ACC] wonder do-IMPERF-1SG
   ‘I wonder what Ozan read.’ [M. Kelepir, p.c.]
c. [Masao-ga [CP Hanako-ga dono hon-o tosyikan-kara Japanese
   [Masao-NOM [CP Hanako-NOM which book-ACC library-from
   ‘the fact that Masao wants to know which book Hanako checked out from the library’ [Saito 1992: 84, (33a)]

Facts such as these demonstrate that there is nothing in principle preventing a wh-in-situ clause from occurring as the complement to an interrogative-selecting predicate. If (true) wh-in-situ in-

---

5 Since verb-second distinguishes main from embedded clauses, the direct quote parse of these examples, marginally available in English, is not an interfering factor in German (9a).
volves a null CWH, then that CWH (or the projection it heads) satisfies the selectional requirements of an embedding predicate. The question, then, is why the lack of movement in English and German correlates with an inability to host an indirect question construal. To the extent this question has been addressed, the general proposal (with the notable exception of Bošković 2000 on French, see below) is that the correlation is spurious – an accidental coincidence in English (G&S 2000) and French (Cheng and Rooryck 2000). A cross-linguistic survey (summarized below) shows instead that the correlation appears to be systematic, and thus in need of a principled explanation. A number of languages are syntactically like English and German, in that DSQs look, in their gross syntax, like declarative sentences. In all these languages, DSQs cannot serve as the complements to interrogative-selecting predicates.

Perfunctory inquiries suggest Dutch and Icelandic are like English and German in the relevant respects. For space reasons, we only include Icelandic examples here.6

(11) a. (a classroom situation)
...og Jón Sigurðsson fæddist hvænar?
...and Jón Sigurðsson be.born when
‘...and Jón Sigurðsson was born when?’

b. (a police questioning situation)
... og þú varst hvar þegar Jóna keyrði út af?
...and you were where when Jóna drove out.of
‘...and you were where when Jóna drove off the road?’

As in English, DSQs are possible in islands, such as coordination (12a) where overt wh-movement is robustly impossible (12b).

(12) a. Sigur Rós hélt 14 tónleika í Bandaríkjunum og hvaða landi?
Sigur Rós held 14 concerts in USA and what country?
‘Sigur Rós held 14 concerts in the US and what country?’

b. *Hvaða landi hélt Sigur Rós tónleika í Bandaríkjunum og hvaða landi?
what country held Sigur Rós concerts in USA and what country
*‘What country did Sigur Rós hold concerts in the US and what country?’

And as in English, DSQs are possible as matrix questions, even in embedded clauses, as in (13a), but in indirect questions, fronting is obligatory (cf. (13b–c)).

(13) a. Obama lét í ljós að hann hefði verið fæddur hvar?
Obama let in light that he had been born where
‘Obama revealed that he was born where?’

b. *Þingmaðurinn spurni Obama væri fæddur hvar?
senator the asked Obama was born where
‘The senator asked Obama was born where?’

c. Þingmaðurinn spurni hvar Obama væri fæddur?
senator the asked where Obama was born
‘The senator asked where Obama was born?’

6 Our thanks to Gísli Rúnar Harðarson for these examples. DSQs are perhaps not as freely available in Icelandic as they are in English, but the relevant contrast is evidently sharp, just as it is in English.
Wood (2009) reports the same effects for American Sign Language (ASL). ASL is especially interesting since there is both leftwards and rightwards movement of *wh*-expressions with some debate as to the analysis (see Petronio and Lillo-Martin 1997, Neidle et al. 2000). Nevertheless, ASL is like English in allowing apparent optional *wh*-in-*situ* as non-echo interrogatives (14b), including in embedded clauses (15b), but this option is prohibited in selected questions (16b), where movement in the embedded clause is obligatory.

\[\begin{align*}
\text{(14) a. } & \text{**WHO** \hspace{1em} JOHN \hspace{1em} SEE \hspace{1em} YESTERDAY?} \\
& \text{b. JOHN \hspace{1em} SEE \hspace{1em} **WHO** \hspace{1em} YESTERDAY?} \\
& \text{‘Who did John see yesterday?’}
\end{align*}\]

\[\begin{align*}
\text{(15) a. } & \text{**WHO** \hspace{1em} JOHN \hspace{1em} FEEL \hspace{1em} MARY \hspace{1em} KISS?} \\
& \text{b. JOHN \hspace{1em} FEEL \hspace{1em} MARY \hspace{1em} KISS \hspace{1em} **WHO**?} \\
& \text{‘Who does John think (that) Mary kissed?’}
\end{align*}\]

\[\begin{align*}
\text{(16) a. } & \text{JOHN \hspace{1em} ASK \hspace{1em} (MARY) \hspace{1em} **WHO** \hspace{1em} SEE?} \\
& \text{b. *JOHN \hspace{1em} ASK \hspace{1em} (MARY) \hspace{1em} SEE \hspace{1em} **WHO**?} \\
& \text{‘John asked Mary who she saw.’}
\end{align*}\]

Beyond Germanic and ASL, the Romance languages are well known for having apparently optional *wh*-in-*situ* to one degree or another, and there is a significant literature on this topic. Despite a wealth of interesting variation to which we cannot do justice, the core main/embedded asymmetry in (1) appears to hold wherever it is testable. Pires and Taylor (2007) and Zocca DeRoma (2011) discuss (Brazilian) Portuguese in comparison to English. They note that English is regularly described as having ‘obligatory’ *wh*-movement outside of echo contexts, while Portuguese is described as having optional *wh*-in-*situ*, as in (17):

\[\begin{align*}
\text{(17) a. } & \text{Quem você viu?} \\
& \text{who you saw} \\
& \text{b. Você viu quem?} \\
& \text{you saw who} \\
& \text{‘Who did you see?’}
\end{align*}\]

Yet closer inspection shows that the languages are essentially the same in their syntax, differing primarily in that the range of pragmatic contexts that permit DSQs. Important for our concerns is the observation that both languages disallow DSQs as indirect questions (18).\(^7\)

\[\begin{align*}
\text{(18) a. } & \text{O Pedro perguntou quem você viu?} \\
& \text{The Pedro asked who you saw} \\
& \text{‘Pedro asked who you saw.’} \\
& \text{b. *O Pedro perguntou você viu quem?} \\
& \text{The Pedro asked you saw who} \\
& \text{‘Pedro asked who you saw.’}
\end{align*}\]

\(^7\) The one syntactic difference is that BP allows apparent matrix questions with an overt complementizer. DSQs are excluded in such cases. An explanation may be that these constructions involved a concealed embedding predicate.
Like English, there is no general prohibition on *wh*-in-*situ* in embedded clauses (19a), or even in islands—(19b) permits a reading of the adjunct *wh*-word as a modifier of the predicate in the adjunct clause (how did she fix it?), a reading that is impossible for an overtly moved *wh*-word (19c).

(19) a. O João pensa que a Maria viu quem?
The João thinks that the Maria saw who
‘John thinks Maria saw who?’

b. O Pedro saiu depois que a Maria consertou o carro como?
The Pedro left after that the Maria fixed the car how
‘Pedro left after Maria fixed the car how?’ (ambiguous)

c. Como o Pedro saiu depois que a Maria consertou o carro?
How The Pedro left after that the Maria fixed the car
‘How did Pedro leave after Maria fixed the car?’ (matrix ‘how’ only)

Pragmatically, Brazilian Portuguese is freer in the range of contexts in which DSQs are permitted. Zocca DeRoma collected examples from Brazilian TV shows (and other sources) including their contexts of occurrence, and it is clear that examples in BP are felicitous in contexts where they are not in colloquial English. Despite this pragmatic variation, there is no appreciable syntactic variation in the key properties of DSQs.

Given the overwhelming prohibition of *wh*-in-*situ* in embedded interrogative clauses, the DSQ/*wh*-in-*situ* generalization in (1) thus appears to be correct and we offer an account of this generalization in the next section.

4 DSQ: An account

4.1 The system

In this section, we provide a syntactic account within an explicit feature system which we show naturally derives the properties of DSQs as described in the previous section. In short, the following are the main claims we argue for:

i. In all languages, a syntactic interrogative clause involves an interrogative $C_{WH}$. This element enters a dependency with (one or more) *wh*-expressions in a *wh*-question.

ii. The difference between *wh*-movement and true *wh*-in-*situ* [TWhiT] languages involves the features of $C_{WH}$: in a *wh*-movement language, $C_{WH}$ has features that require an overtly filled specifier whereas in a TWhiT language, the dependency between $C_{WH}$ and a *wh*-expression may be satisfied without overt movement, either by covert movement (however that is understood) or via a binding dependency.

iii. DSQs are (syntactically) declarative TPs or CPs which lack $C_{WH}$ and in which the *wh*-XP is in focus (position); an interrogative interpretation is derived via pragmatics.

The general architecture of our proposal is given in (20). In all languages, true syntactic questions involve a dependency (to be specified below) between $C_{WH}$ and a *wh*-XP. No such dependency exists in DSQs. Since there is no (movement) dependency in DSQs, the fact that DSQs are
Questions with Declarative Syntax

possible in islands is expected. Furthermore, since DSQs are formally declaratives they cannot occur in contexts which syntactically select for an interrogative (i.e., an attempted DSQ such as *He asked me your boy’s name is what* is excluded in exactly the same way the declarative *He asked me your boy’s name is Leo* is).

(20) a. Syntactic interrogative

\[ \text{CP}_{\text{WH}} \]

\[ \begin{array}{c}
\text{C}_{\text{WH}} \\
\text{TP} \\
\text{wh-XP}
\end{array} \]

b. DSQ

\[ \text{(CP)} \]  \[ \rightarrow \text{pragmatics: interrogative} \]

\( \text{(C)} \)

\[ \begin{array}{c}
\text{TP} \\
\text{[declar]} \\
\text{[wh-XP]_{FOC}}
\end{array} \]

The approach in (20) has some immediate consequences for the syntax of *wh*-phrases. Since there is no morphological difference between moved and unmoved *wh*-expressions, the null hypothesis is that *wh*-phrases in regular syntactic questions and DSQs are the same elements, that is, they are equipped with the same lexical features. We are not aware of any language, for example, in which moved and unmoved (DSQ) *wh*-expressions (which correspond to interrogatives) are morphologically distinct. On an account that would assign one class a feature that the other class lacked, the lack of an overt signal of this difference would be a strange coincidence. Since there is no *wh*-operator in the CP in (20b), *wh*-XPs must be syntactically independent of a C head; in other words, they cannot involve a feature which needs to be licensed by an interrogative C head or a feature that triggers movement of a *wh*-XP. We therefore propose that *wh*-phrases involve an *interpretable, valued* *wh*-feature: \( iQ: \text{wh} \). This has the advantage that no other Q feature must be in the structure to license a *wh*-XP, and DSQs are thus in principle possible, even in formally declarative contexts. Furthermore, being equipped with an interpretable interrogative feature, *wh*-XPs have the power to contribute an interrogative interpretation by themselves. This is what we suggest is the case in DSQs. Once set in an appropriate pragmatic and semantic context, *wh*-XPs can trigger an interrogative interpretation and yield a DSQ.

Regarding syntactic dependencies and feature licensing, we follow the view that the interpretability of features is independent of the notion of valuation (both interpretable and uninterpretable features can come as valued \([i/uF: \text{val}]\) or unvalued \([i/uF: \text{___}]\)) and that licensing is established under Agree which is valuation driven (see Pesetsky and Torrego 2007, Bošković 2009). The definition of Agree we adopt is the Reverse Agree definition given in (21) from Wurmbrand (2014).

(21) A feature F: ___ on \( \alpha \) is valued by a feature F: val on \( \beta \), iff

i. \( \beta \) c-commands \( \alpha \) AND

ii. \( \alpha \) is accessible to \( \beta \). [accessible: not spelled-out]

iii. \( \alpha \) does not value {a feature of \( \beta \)}/ {a feature F of \( \beta \)}.

Let us see how this system derives syntactic interrogatives. As shown in (22), interrogative C is first merged with an unvalued, but crucially *interpretable* Q-feature. Given Agree as in (21), this feature needs to be valued by a c-commanding, valued Q-element, such as a *wh*-phrase. This re-

---

8 We do, however, posit a featural difference between the interrogative and indefinite uses of *wh*-expressions in languages like German: interrogative *wh*-expressions involve an *interpretable* *wh*-feature, whereas the indefinite ones involve an *uninterpretable* *wh*-feature (see also Pesetsky and Torrego 2007).
quirement thus triggers movement of the \(wh\)-XP, which we assume is inherently valued as a \(wh\)-phrase (\(iQ: wh\)), to Spec,CP, a position from which it may value the unvalued \(iQ:\_\_\) of C. The result of this derivation (see (22a)) is an interrogative clause, that is, a CP typed as \(iQ: wh\). Such a clause may subsequently be merged with a verb that selects an interrogative, such as wonder or ask, as in (22b). Merging an interrogative selecting verb such as wonder with a CP lacking \(iQ: wh\) is filtered out by however selection is implemented (see, among others, Adger 2003, Wurmbrand 2014 for feature based proposals). It is important to note that \(C_{\text{WH}}\), and thus the CP it heads, is unvalued until such time as the \(wh\)-XP moves to Spec,CP and values the \(iQ:\_\_\) of \(C_{\text{WH}}\). This property forces \(wh\)-movement to be overt in embedded interrogatives—prior to movement of the \(wh\)-XP, the CP is \(iQ:\_\_\) and thus not selectable by an interrogative-embedding predicate.

(22) a. Matrix interrogative  
\[
\begin{array}{c}
\text{XP} \\
\text{\(iQ:wh\)}
\end{array}  
\begin{array}{c}
\text{C'}
\end{array}
\begin{array}{c}
\text{\(C\)} \\
\text{\(iQ:wh\)}
\end{array}
\begin{array}{c}
\text{TP}
\end{array}
\begin{array}{c}
\text{\(\text{CP} \ [iQ:wh]\)}
\end{array}
\]

b. Embedded interrogative  
\[
\begin{array}{c}
\text{VP} \\
\text{\(V_{\text{wh}}\)}
\end{array}  
\begin{array}{c}
\text{\(\text{CP} \ [iQ:wh]\)}
\end{array}
\begin{array}{c}
\text{\(\text{XP} \ [iQ:wh]\)}
\end{array}
\begin{array}{c}
\text{\(\text{C'}\)}
\end{array}
\begin{array}{c}
\text{\(\text{TP}\)}
\end{array}
\begin{array}{c}
\text{\(\ldots\text{\(iXP\)}\ldots\)}
\end{array}
\]

As noted above, we assume that TWhiS languages also involve a C–XP-\(wh\) dependency. Without delving into the many intricacies of this topic, it seems there are at least two broad families of analysis of TWhiS, plausibly reflecting two classes of phenomena or different language types (see, e.g., Cheng and Rooryck 2000, Bayer 1996, Cole and Hermon 1998). One type involves covert movement—a derivation essentially like (22), with the only difference that the \(wh\)-elements move covertly. There are a variety of ways of implementing this, including covert phrasal movement (Huang 1982, Pesetsky 2000), movement of an invisible feature or other subword element (Watanabe 1992, Pesetsky 2000, Cable 2007), or overt phrasal movement followed by pronunciation of the lower copy. For our purposes, all of these may be lumped together under the rubric ‘covert’ movement. In such a language, TWhiS should be subject to the same locality conditions as overt movement. A second strategy is (un)selective binding (see, e.g., Pesetsky 1987). In such languages/constructions, we propose that C can be inserted with a valued \(Q\) feature. If that feature binds another \(Q\) feature, the structure is interpreted as a \(wh\)-question. Like variable binding, unselective binding requires c-command, but is otherwise not sensitive to islands, which correctly characterizes certain TWhiS languages. Thus, while island-sensitivity is

\[9\] We leave open the possibility that one species of apparent \(wh\)-in-situ is derived by overt movement of the \(wh\)-phrase, but with the lower copy of that phrase pronounced (Bobaljik 1995). This cannot, on our view, be the derivation of DSQs, though, contra, e.g., Reglero (2007). Such a system does not discriminate between matrix and embedded interrogatives, and would incorrectly allow indirect questions without movement.
indicative of TWHiS, lack of island-sensitivity does not entail that the structure involves a DSQ—both DSQs and TWHiS established via unselective binding are possible in islands.\(^\text{10}\)

Crucially however under our proposal, a DSQ does not involve a C–XP-\textit{wh} dependency. For this reason—and this is the main difference between TWHiS and DSQs—DSQs may be interpreted as interrogatives, but lacking a CP marked as \textit{iQ: wh} cannot be embedded by a predicate selecting a syntactically interrogative complement. Embedability is thus indicative of TWHiS. Both covert movement and unselective binding produce CPs that are embeddable, thus we need not enter into debates about the merits of individual approaches to \textit{wh-in-situ}; the crucial point for us is that we have provided a formal characterization of the central difference between DSQs and TWHiS.

### 4.2 DSQs with \textit{wh-in focus}

The languages considered in section 3 share the property that DSQs resemble simple declarative clauses in their gross word order. These are all languages in which elements in focus may remain in situ. Since we propose that DSQs are focus constructions (cf. (20b)), in languages where focus constituents occupy dedicated positions, DSQs will be characterized by a \textit{wh}-expression in focus, rather than necessarily in situ. Exactly such distribution is found in Spanish, among others.

DSQs in Spanish are similar to Brazilian Portuguese, but with the additional requirement, discussed by Uribe-Etxebarria (2002), and Reglero (2007) that the non-fronted \textit{wh}-expression must be final in its intonational phrase. This order appears to be distinct from \textit{wh-in-situ}, in that it does not always coincide with the unmarked order of post-verbal constituents, but as Reglero discusses in detail, this requirement holds independently of post-verbal focus constituents in Spanish. For example, while the unmarked order in Spanish is DO\textsuperscript{»}IO, when the DO is an unmoved \textit{wh}-expression, it follows the IO, reflecting the broader generalization that focused constituents come last in their intonational phrase:

\begin{equation}
[y] \text{ tú le diste a María (el) } \textit{qué}?
\end{equation}

*‘And you gave Maria what?’*

The fact that Spanish \textit{in situ} \textit{wh}-expressions obey a word order generalization not evident in \textit{wh}-movement languages is thus, as Reglero originally argued, not indicative of any syntactic difference about \textit{wh-in-situ} among these languages, but is rather a consequence of the facts that DSQs involve \textit{wh-in-focus}, and that the syntax (or perhaps prosody) of focus differs between Spanish and the languages described in section 3.

Cross-linguistic variation in the syntax of focus thus conspires to sometimes obscure the fundamental cross-linguistic similarity in the syntax of DSQs. This phenomenon is particularly important when considering DSQs in Slavic. It is widely recognized in the literature on multiple \textit{wh}-fronting that not all fronted \textit{wh}-expressions in Slavic undergo \textit{wh}-movement—some occupy a lower, focus position (Bošković 1998). Focus movement (i.e., to a non-initial position) is possible for question formation even with a single \textit{wh}-word. If these are instances of DSQs (and not binding by an interrogative C), then we expect that the non-initial fronting will, like other DSQs

\(^{10}\) There is much more to be said here. As regards putative island-insensitivity in \textit{wh-in-situ} languages, there is a substantial literature debating this point (see Cheng 2003, Bayer 2006 for overviews).
conform to (1). This is correct for Polish (Lasnik and Saito 1984) and Russian (Dyakonova 2009), at least: as shown in (24) for Russian, wh-expressions in matrix questions may either undergo wh-movement or focus movement, but in an indirect question, a wh-expression must front to the initial position.\footnote{It appears there is some variation in Russian as to whether the wh-expression (and focused expressions generally) may occur in post-verbal position in (24a) (see, e.g., Dyakonova 2009).}

(24) a. \{kogda\} Ivan \{kogda\} budet \{kogda\} vystupat’?
   ‘When is Ivan going to present?’ [N. Radkevich, p.c.]

b. Boris sprašivaet kogda Ivan \{*kogda\} budet \{*kogda\} vystupat’.
   ‘Boris is asking when Ivan is going to present.’ [Dyakonova 2009: 197]

In other words, exactly as in all the other languages examined above, it is possible to ask a question with the syntax of a declarative clause, where a wh-expression occurs in a focus position, but a clause with this syntax cannot be syntactically selected by a predicate that selects an interrogative complement. Such a predicate requires a complement that is not only interrogative in force, but is syntactically marked as such, for example, by (true) wh-movement.

Similarly for Polish, Lasnik and Saito (1984) contend that wh-expressions must occupy an A’-position at S-structure, but need not front overtly to Spec,CP (on their account, they may do so at LF). Thus, in (25a,b) the wh-word is in non-initial position in the embedded clause, but still yields a matrix question interpretation. Yet even in Polish, if an interrogative complement is selected (as by wonder), then a wh-word must be fronted to clause initial position (embedded Spec,CP), as in (25c).

(25) a. Maria powiedziała, że co Piotr kupił?
   Maria said that what Piotr bought
   ‘What did Maria say that Peter bought?’

b. Spotkałeś mężczyznę, który jak rozwiązał zadanie?
   you.met man who how solved problem
   ‘How did you meet the man who solved the problem?’

c. Zastanawiam się \{co\} Maria przyjmię \{*co\}.
   I.wonder REFL \{what\} Maria will.bring \{*what\}
   ‘I wonder what Maria will bring.’ [Lasnik and Saito 1984]

As our system now stands, we must reject the characterization of languages like Polish, English or French, as having optional wh-in-situ, or optional fronting to non-initial A’-positions. The DSQs are crucially not simply limited instances of the wh-in-situ derivation that languages like Chinese use freely. Note again that TWhiS languages such as Mandarin happily allow cases exactly like (25c) (cf. (10)), thus there is no general ban on embedded TWhiS.

Instead we assume that Polish and Russian, like English, are wh-movement languages and that cases of unmoved wh-phrases (or rather cases where the wh-elements are moved to a focus position) are instances of DSQs and not TWhiS. This approach derives the distribution in (25). That these are DSQs is further suggested by the fact that (25a,b) are environments out of which overt wh-movement would be prohibited (Polish disallows wh-extraction out of finite comple-
ments), and provides a principled account for why apparently optional \textit{wh}-movement languages like Polish always lose the option of \textit{wh-in-situ} in embedded contexts: in syntactic questions, movement is always obligatory; the apparent optionality arises as a result of the option of DSQs. As we have surveyed briefly in section 3, the option of DSQs to form matrix interrogatives shows varying degrees of pragmatic freedom cross-linguistically, but is syntactically always restricted to matrix interrogative interpretations.

4.3 Alternatives?

In this section, we compare our feature valuation system with two other important proposals. Our assumption that \textit{wh}-XPs are inserted with a valued Q-feature and C is inserted with an unvalued Q feature is identical to the feature assignment proposed in Pesetsky and Torrego [Pe&To] (2007). However, in contrast to our system, Pe&To follow an upward valuation approach to Agree. In derivations like (26), C probes downward until it finds a valued Q-feature, the \textit{uQ: wh} of the \textit{wh}-XP, and is then valued (upwards) by that XP. To derive movement of the \textit{wh}-XP, C is equipped with an EPP feature in movement languages. The difference between movement and TWhiS languages could thus be encoded as the presence vs. absence of an EPP feature. To derive the properties of languages like Polish or BP, with (descriptively) optional \textit{wh}-movement in matrix interrogatives, it would be necessary to have an optional EPP feature on matrix C, but an obligatory EPP feature on embedded Cs. While this technically yields the correct results, it raises the question of why embedded C must always involve an EPP feature in languages with an optional EPP feature on matrix C.

(26) a. Matrix interrogative

\begin{center}
\begin{tikzpicture}
  \node (cp) {CP \([iQ: wh]\)};
  \node (xp) [below left of=cp] {XP \([uQ: wh]\)};
  \node (c) [below of=cp] {C \([iQ: \underline{=}EPP]\)};
  \node (t) [below of=c] {TP \([\ldots tXP \ldots]\)};
  \draw (cp) -- (xp);
  \draw (xp) -- (c);
  \draw (c) -- (t);
\end{tikzpicture}
\end{center}

b. Embedded interrogative

\begin{center}
\begin{tikzpicture}
  \node (vp) {VP \([wonder]\)};
  \node (cp) [below left of=vp] {CP \([iQ: wh]\)};
  \node (xp) [below of=cp] {XP \([uQ: wh]\)};
  \node (c) [below of=cp] {C \([iQ: \underline{=}EPP]\)};
  \node (t) [below of=c] {TP \([\ldots tXP \ldots]\)};
  \draw (vp) -- (cp);
  \draw (cp) -- (xp);
  \draw (xp) -- (c);
  \draw (c) -- (t);
\end{tikzpicture}
\end{center}

Furthermore, this account does not explain why of the four possible EPP combinations only three seem to be attested. If both matrix and embedded Cs lack EPP features, a TWhiS languages is derived (e.g., Mandarin); if both matrix and embedded Cs involve EPP features, an obligatory movement language is derived (e.g., English, though DSQs are then not accounted for; see also below); if matrix C involves no or an optional EPP feature, but embedded C involves an obligatory EPP feature, a language like Polish is derived. What is not attested is a language with an obligatory EPP feature on matrix C, but no or an optional EPP feature on embedded C. This would yield a language which involves obligatory movement in matrix clauses but no or optional movement in embedded clauses. Under a Pe&To feature system, there does not seem to be a
principled reason for why such languages are missing. In our account, on the other hand, this is predicted: A language is either a TWhiS language or an obligatory movement language (matrix and embedded clauses always behave alike); all languages allow DSQs, however, DSQs are only possible in matrix questions. Hence, the only ‘mismatch’ between matrix and embedded clauses that is possible is cases in which a matrix clause involves no movement (due to DSQ), but the embedded clause requires movement (due to the unavailability of embedded DSQs).

Lastly, a Pe&To feature system does not readily extend to DSQs in a language like English. Since wh-XPs are specified as $uQ$: wh, they do not carry interrogative force and thus may not be able to license an interrogative interpretation by themselves. A possible solution may be to assume a special DSQ C which licenses the uninterpretable Q-feature of wh-XPs in DSQ contexts but by a means different from Agree. In this regard, Pe&To’s approach would become very similar to the proposal of Pires and Taylor (2007), where it is explicitly argued that DSQs must be licensed by a special Q operator. The following discussion thus applies to both proposals and similar ones involving a syntactic licensing requirement of wh-XPs in DSQ contexts (see for instance Cheng and Rooryck 2000, Zocca DeRoma 2011, among others). Pires and Taylor (2007) suggest that DSQs in English are cases of wh-in-situ, specifically, wh-in-situ involving the strategy of unselective binding. As mentioned above, the main hurdle for such proposals is to block DSQs in embedded questions. Since TWhiS languages readily allow indirect questions with wh-in-situ, it is hard to see how the special DSQ C could be blocked from being selected, in a principled way. Why, for instance, can’t this special C be selected (it can in TWhiS languages)? Why is it that all languages, even ones that most persistently have been described as obligatory wh-movement languages (such as English) allow DSQs? If the special DSQ C is a lexical item, we would expect it to be absent in some languages. However, once we control for pragmatic factors, DSQs seem to be available cross-linguistically, exactly as expected under our approach where wh-XPs involve an interpretable Q-feature which licenses an interrogative interpretation even in the absence of an interrogative syntax. In sum, in approaches involving (or requiring) licensing of wh-XPs, the unembedability of DSQs becomes an accidental, idiosyncratic restriction arising only in (obligatory and optional) wh-movement languages, whereas it follows on principled grounds from our approach.

4.4 A final note—French

Up to this point, we have put aside discussion of French, although it is one of the most widely discussed cases of ‘optional’ wh-in-situ in a wh-movement language. French is somewhat of a two-edged sword, in part as there are competing descriptions of the facts, and evidently variation among speakers. All analysts agree that French respects the basic contrast, consistent with the generalization in (1). Wh-in-situ is freely available to form non-echo questions in simple clauses, as in (27a), but such questions are strongly ungrammatical as an indirect question; wh-movement in the embedded clause is obligatory as in (27b,c).

---

12 Compare verb movement (to $C^0$), which across (most of) Germanic has exactly the property of being obligatory in matrix questions but impossible in indirect questions.

13 Pires & Taylor also propose that the pragmatic restrictions on DSQs are attributable to (in some way or another) the presence of this special complementizer. But see Tieu (2011) for critical remarks.
Questions with Declarative Syntax

(27) a.  Il parle de quoi?
     He talks of what
     ‘What is he talking about?’
b.  *Je me demande tu parles de quoi.
     I me ask you talk of what
     ‘I wonder what you are talking about.’
c.  Je me demande de quoi tu parles.
     I me ask of what you talk
     ‘I wonder what you are talking about.’ [Obenauer 1994: 291-2; P. Schlenker, p.c.]

Beyond these poles, the French are disagreeable. Broadly speaking, there are at least three different characterizations of the data. An important starting point is Obenauer (1994) who offers an extended comparison of French and English. Under Obenauer’s characterization (see also Pollock 1998, Starke 2001), French is like English in allowing *wh-in-situ* in embedded (non-interrogative) clauses as in (28), and even in (weak) islands as in (29) and (30) (see also Starke 2001, Shlonsky 2012).

(28) a.  Tu veux que je le fasse quand?
     you want that I it do when
     ‘You want me to do it when?’ [Obenauer 1994: 319]
b.  Tu crois que Jean a acheté quel livre?
     you believe that Jean has bought which book
     ‘You believe that Jean bought which book?’ [Pollock 1998: 189]

(29) a.  Il s’est défendu [ en accusant qui ]?
     he REF-l-is defended [ by accusing who ]
     ‘He defended himself by accusing who?’
b.  *Qui s’est-il défendu [ en accusant t ]?
     who REF-l-is-he defended [ by accusing t ] [Obenauer 1994: 296]

(30) a.  Vous connaissez des gens qui pourraient héberger combien de personnes?
     you know of people who could host how many people
     ‘You know people who could host how many people?’ [Obenauer 1994: 297]
b.  *Combien de personnes connaissez-vous des gens qui pourraient héberger?
     how many people know-you of people who could host

However, unlike English, French on this characterization does obey various types of apparently syntactic locality conditions. On Obenauer’s characterization, embedding the islands in (29)-(30) still further yields unacceptability; later authors (Starke 2001, Shlonsky 2012) characterize this as a weak-strong island asymmetry. Many authors also contend that various quantificational elements, including negation, occurring above the *in-situ wh*-expression lead to degradation.

Chang (1997) (and following her Bošković 2000 and others) presents a much more restrictive variety than Obenauer (1994) and Starke (2001). In this variety, while the contrast in (27) obtains, *wh-in-situ* is significantly degraded in embedded clauses generally—these authors mark sentences like (28) as unacceptable. Thus while the generalization in (1) holds, it does for rather trivial reasons, as a special case of a broader generalization.
In the other direction, Starke (2001) presents a more permissive, colloquial variety of French than Obenauer. Starke’s characterization is essentially similar to Obenauer’s, noting the strong vs. weak island asymmetry, and also the intervention effect of elements such as negation, but Starke contends that these do not yield unacceptability per se. Rather, these sentences are limited in their contexts and intonation patterns; but when these variables are controlled for, *wh-in-situ* is seen to be acceptable even in strong islands and under negation and other putative interveners (Starke notes carefully that the relevant contexts are nevertheless distinct from echo questions).

We have neither the space nor the resources to sort out the French situation in the remaining pages of this short paper. We note on the one hand that all varieties of French thus far reported are consistent with the generalization in (1)—French thus supports our main contention that the lack of movement and unembeddability are intimately entwined. On the other hand, we note that most authors, cutting across the three varieties reported, argue for LF-movement of *wh-in-situ* and thus a covert version of the dependency with C that we reject on the basis of languages like English.

Of the competing accounts, Bošković (2000) analysis stands out in offering a proposal that restricts the construction to matrix questions. In brief, under Bošković’s account (like those of its competitors), the French interrogative complementizer $C_{WH}$ always triggers movement, but $C_{WH}$ may be inserted into the derivation early, in which case movement is overt, or late (at LF), in which case the movement is covert (yielding apparent *wh-in-situ*). Given other assumptions of the framework in which Bošković (2000) is couched, the covert movement option is restricted to matrix clauses: because of the cyclicity of derivations, in order for $C_{WH}$ to be inserted in an embedded clause it would need to be inserted overtly, prior to the merge of the embedded clause as complement to the matrix predicate. In this way, it is only in matrix questions that insertion of $C_{WH}$ and thus movement of the *wh*-expression, may be delayed until LF.

While Bošković’s account thus meets the general desideratum we have set forth, relating the *in-situ* property to the inability to be selected, Bošković treats the numerous restrictions of the narrow variety of French as key support for his proposal: the inability of the *wh-in-situ* to be beneath negation or quantifiers, or to occur in embedded clauses at all, are taken as key evidence for the specific account he sets forth. Thus, as other authors have noted (Reglero 2007, Wood 2009, Zocca DeRoma 2011), the properties that support Bošković’s account of the variety he describes keep that account from generalizing to languages that fail to show a general ban on embedding, that allow *wh-in-situ* under ‘interveners’ such as negation, and which allow *wh-in-situ* in islands. Thus, while Bošković’s proposal provides an account of the language particular properties of one variety of French, it does not generalize. Conversely, an account such as ours, which focuses on the general property of DSQs, does so at the expense of the language-particular properties of various languages.14

8 Conclusion

Although we have left several loose ends, the landscape of DSQs can be characterized as follows. In all languages that we have investigated with classical *wh*-movement (i.e. to the leftmost periphery of the clause), an ‘optional’ interrogative strategy exists—even in non-echo contexts—

---

14 Dyakonova (2009) claims that Russian patterns with the restrictive variety of French, but no other language we have investigated does. There are additional differences among individual languages which are not captured by our account. For example, Bayer (2006) notes that in German *warum* ‘why’ resists forming DSQ contexts, contrasting with the roughly synonymous *aus welchem Grund* ‘for what reason’.
in which the clause has declarative syntax, with a *wh*-expression in focus position (i.e., *in-situ* where there is no special focus syntax). There is variation (both cross-linguistic and intra-speaker) in the pragmatic contexts in which this strategy is available, and, most strikingly in French, also in the syntactic contexts that tolerate DSQs. Despite this variation, there is one point of absolute stability: the *in-situ* construction is blocked in selected questions. The very systematicity of the last point constitutes an argument that ‘optional’ *wh-in-situ* in *wh*-movement languages should not be assimilated to the *wh-in-situ* strategy of TWhiS languages, which lack such a restriction. This property is theoretically important in that it speaks to the nature of what is selected; but it is also of interest in that it constitutes the one island of stability in a sea of other cross-linguistic variation in both the clause-internal syntax and overall distribution of DSQs. We have sketched a feature-based account that weds the unmoved nature of the *wh*-expression to its unembeddability, relating this centrally to the notion of syntactic selection set out in *Aspects*.

Finally, our approach leads us to expect that DSQs should be possible in principle in all languages with *wh*-expressions. In a TWhiS language, these will be very hard to distinguish from *wh-in-situ*. But not impossible. Under our approach, DSQs and TWhiS should both be possible as matrix questions, but DSQs should be impossible as indirect questions. To the extent there are TWhiS languages that show island effects, we thus predict a class of languages in which, as a matter of observation, island effects with *in-situ* *wh*-expressions are avoided in matrix question interpretations (which may be DSQs) but create violations when the intended interpretation is an indirect question (where DSQs are impossible). We do not know at this time whether such a class of languages exist, but leave this as an open conjecture.

References


Cable, Seth. 2007. The Grammar of Q: Q-Particles and the Nature of Wh-Fronting, as Revealed by the Wh-Questions of Tlingit. Doctoral dissertation, MIT, Cambridge, MA.


