SUBEXTRACTION IN ROMANCE INTERROGATIVES

JAUME SOLÀ & ANNA GAVARRÓ
Universitat Autònoma de Barcelona

In Gavarró & Solà (2004), we present evidence for non-target wh-subextraction in child Catalan and provided an account in terms of Case theory. Our account crucially relies on a specific analysis of the Case licensing strategies available in adult Catalan, according to which Catalan does in fact permit wh-subextraction, although of a more restricted type than that attested in child Catalan. In this paper we aim to substantiate the analysis of adult grammar for wh-interrogatives, drawing from Kayne (2002) and arguing that wh-subextraction in Catalan is an instance of remnant movement.

The paper proceeds as follows. First we consider the data on wh-subextraction in the literature, from Romance languages and others, and summarise the original data on child Catalan from Gavarró & Solà (2004). Second, we review the analyses proposed for subextraction in the literature and point out some of their empirical and/or conceptual shortcomings. Third, we provide an analysis of the adult grammar of Catalan which can encompass the facts of child Catalan in a natural way and is also consistent with the facts from other languages.[*]

1. Wh-subextraction: the data

Wh-subextraction is a kind of wh-movement where the wh-determiner moves without pied-piping its noun phrase. It is attested in some adult languages with rich Case morphology (Corver 1990), such as Slavic Languages or Latin, where it is freely allowed as an alternative to pied-piping:

(1)   a. *Skolko   Tania   prochitala knig?*
        how-many Tania   she-read   books-GEN
        "How many books did Tania read?"
                     (Russian, Gavruseva & Thornton 2001)
   b. *Jaki   wykrecilés numer?*
        which you-dialed number
        "Which number did you dial?"
                     (Polish, van Kampen 1994)

In other languages, it is attested only with certain determiners or DP specifiers:

(2) a. *Combien as-tu   lu   de livres?*
         how-many have-you read of books
         "How many books have you read?"
                                      (French, Obenauer 1984)
     b. *Wat heb je   voor auto's gekocht?*
         what have you for cars   bought
         "What kind of cars have you bought?"
                                      (Dutch, Corver 1990)
     c. *Com és de   gruixut?*
         how is   of   thick
         "How thick is it?"
                                      (Catalan)
     d. *Kínek   veszett el   kalapja?*
         Whose   got   lost   hat
         "Whose hat got lost?"
                                      (Hungarian, Gavruseva & Thornton 2001)

For Romance languages, the presence of the preposition *de* between the wh-determiner and the rest of the phrase (as in *combien de livres*) seems to be a necessary (but not a sufficient) condition for subextraction.

 Regarding child language, subextraction has been attested in Dutch and English. Van Kampen (1994, 1997) found that Dutch children produced questions like those in (3), which are not found in the input the child is exposed to.

 (3) a. *Welk$_i$ wil   jij [t$_i$ boekje]?*   (age 3;7)
         which   want you   book
         "Which book do you want?"
     b. *Hoe$_i$ is het [t$_i$ laat]?*   (age 6;5)
         how is it   late
         "How late is it?"
                                      (van Kampen 1994)

For English, subextraction cases have been reported in Hoekstra, Koster & Roeper (1992), Thornton & Gavruseva (1996), and Chen, Yamane & Snyder (1998). Gavruseva and Thornton (2001) investigate, through an elicited production experiment, wh-possessive questions and find that, contrary to adult English, children allow wh-extraction of *whose*, without pied-piping of the entire *whose*-DP.

(4) a. Who do you think's flower fell off?
     b. Whose do you think ball went in the cage?

These are all cases of (spontaneous or elicited) production. Gavarró & Solà (2004) present new experimental evidence, which we summarise here, that in child Catalan non-target wh-subextraction can arise in comprehension tasks, in cases like (5):

(5)   *Qui   necessita sabates?*
       who needs   shoes
       "Who needs shoes?"

These questions are unambiguous in adult Catalan, involving wh-movement of a bare wh-word (*qui* 'who'). In two experiments, however, Gavarró and Solà (2004) found that Catalan children can understand them in two ways, as illustrated for question (5):

(6)   a. *Adult interpretation:*
          Question interpretation: Qui$_{subject}$ necessita t$_{subject}$ sabates?
                                      who   needs   shoes
          Corresponding answer: La germana petita
                                      the sister   young "The young sister"
       b. *Non-adult wh-subextraction interpretation*
          Question interpretation: Qui$_i$ necessita [t$_i$ sabates]?
                                      which needs   shoes
          Corresponding answer: Les blaves.
                                      the blue   "The blue ones"

The children's non-adult interpretation involves misinterpreting *qui* as meaning 'which', that in adult Catalan has the similar form *quin* (with gender and number agreement: *quina, quins, quines)*, but does not allow subextraction (*Quines$_i$ necessita* [t$_i$ *sabates*]).

        The experiment in Gavarró and Solà (2004) had as subjects twelve 2-year-olds and nine 3-year-olds (age range 2;5,27–3;8,27). These children produced target answers (of the type in (6a)) in 72.9% of cases for 2-year-olds and 86% of cases for 3-year-olds; wh-subextraction interpretations (of the type in (6b)) in 20.8% of cases for 2-year-olds and 13.9% of cases for 3-year-olds. Only 6.2% of non-adult answers were something other than wh-subextraction errors; the results were statistically significant by Fisher's exact test. Controls performed 100% as expected. To our knowledge, non-target subextraction was not previously attested in the literature on Romance. The case reported in Gavarró & Solà (2004) bears a close resemblance to that in Dutch: the subextracted quantifier is of the 'which' type (*welk* in Dutch/ *qui(n)* in Catalan) and is extracted from object position. It is worth noting that the percentage of subextractions in child Dutch is over 50%; this represents a quantitative

difference with child Catalan. Also, as mentioned, the Dutch results come from spontaneous production, while those in Catalan belong to interpretation; we would expect production to pair with interpretation, but studies of production in Catalan and comprehension in Dutch remain a topic for future research.

2.   *Possible analyses for wh-subextraction*

Let us now proceed to consider the theoretical options available to account for the cases of subextraction as reported in the literature. Leaving aside the performance approaches put forward for child productions in English (see Chen, Yamane & Snyder 1998 and Yamane, Chen and Snyder 1999), all the analyses of subextraction in adult grammar are grammatical approaches.[1]

Ross (1967) was the first to characterise subextraction in Slavic languages as a violation of the Left Branch Constraint. He also noted the correlation between subextraction and rich nominal morphology in Slavic languages and Latin, where determiners and nouns fully agree in phi- and Case features.

Capitalising on the Slavic data, Uriagereka (1988) and Corver (1990) developed an account for adult language, later adopted for child language by Jordens & Hoekstra (1991), Hoekstra, Koster & Roeper (1992) and Hoekstra & Jordens (1994). Corver's analysis is based on the idea that nominal arguments may differ in their categorial structure: either they are DPs in the standard sense (with the NP in complement position) or they are NPs and the DP phrase is in fact an adjectival phrase left adjoined to the NP. The latter situation is claimed to arise in Slavic languages, where the lack of definite articles would be consistent with the adjunct status of determiners (the correlation between Left Branch Extractions and the absence of common determiners in Slavic languages had been previously noted by Uriagereka 1988).

According to Corver's proposal, subextraction out of DP is not possible, either because the DP is a barrier or because subextraction is improper movement of some sort (head movement, non-constituent movement). However, in languages in which nominal arguments are NPs, and where the DP material (demonstratives, wh-words or genitive specifiers) simply constitutes an adjunct, extraction of this adjunct becomes possible. Adjuncts to a maximal projection can move away, since this projection does not constitute a barrier for them:

---

[1] As argued in Gavarró and Solà (2004), performance approaches based on some processing disturbance (such as overload) are unlikely for our case, since the miscomprehended sentences are short (three words) and they are set in straightforward context where the adult interpretation is perfectly available.

(7)   a. * Wh$_i$ ... [$_{DP}$ t$_i$ NP]
        b. √ Wh$_i$ ... [$_{NP}$ t$_i$ NP]

This proposal must ensure that movement from the adjoined position respects the Condition on Extraction Domains, since subextraction is possible from objects but not from subjects or adjuncts.

There is a possible problem with this analysis, from a minimalist point of view. In minimalist terms, movement targets a feature F. Then, to move a feature F, F must "carry along just enough material for convergence" (Chomsky 1995:262). Moving the whole word containing the feature is already a minimal case of pied-piping. Moving larger constituents is a stronger case of pied-piping that must be due to further requirements for convergence at the interfaces. Then, if the NP structure in Slavic languages allows subextraction of the determiner, this may predict that subextraction is obligatory, since no extra material is predicted to be needed for convergence. This prediction is incorrect: subextraction is only optional in Slavic languages.

In addition, on the theoretical side, it is questionable that UG should allow for such an essential part of syntactic structure as is DP to be reshuffled as something as different as an adjunction structure. Anything departing from universal functional structure poses the question as to how to define the limits of syntactic variation in a non-stipulative way. We adhere to the view, defended in Cinque (1999), that there is no variation in the hierarchical structure of functional categories.

Let us, then, consider the hypothesis that subextraction involves no such structural variation with respect to pied-piping. In particular, we consider the minimalist contention that it is not syntax itself, but the phonological component, that imposes pied-piping. In line with this idea, pied-piping or the lack thereof should be sensitive to formal properties of the DP whose satisfaction at the interface would force pied-piping. Let us explore the intuition that those formal properties have to do with morphology and, more specifically, Case.

Recent approaches that capitalise on Case licensing within the DP as the key to its pied-piping possibilities include van Kampen (1997, 2000). Van Kampen proposes that there are two strategies for the Case licensing of DPs and NPs. In languages not allowing subextraction, such as (adult) Germanic languages, D^o is the head that gets Case. For the N(P) to get Case, it has to be string adjacent to D^o, a PF condition. This requirement makes subextraction of D^o impossible. In contrast, in languages generally allowing subextraction, such as the Slavic languages, N^o can be Case-licensed independently of D^o, due to the morphological strength of N^o Case morphology. This independence is also consistent with the absence of obligatory determiners (articles) in these languages, in contrast with the Germanic languages, where N^o needs to be adjacent to D^o to be Case-licensed.

In order to extend this analysis to child Dutch, where subextraction is possible, van Kampen assumes that the PF condition that requires D^o-N^o adjacency does not hold in child language. In addition she argues that subextraction fits LF requirements better than pied-piping. First, it dispenses with the need for reconstruction of the NP in a wh-phrase (as in well known cases like: *Which picture of himself$_i$ did John$_i$ choose*). Second: in child Dutch, when the subextracted determiner or degree quantifier undergoes scrambling, it receives narrow focus, leaving the stranded part as unstressed, so that the mapping of the informational structure accurately corresponds to the representation provided by subextraction (see van Kampen 2000). Therefore, the minimalist problem concerning optionality between subextraction and pied-piping may be addressed in terms of a tension between optimising LF representations (subextraction) and PF representations (pied-piping).

Van Kampen's analysis does not involve, like Corver's, a dual analysis of nominal arguments (as DPs or bare NPs). However, the unitary structural analysis forces van Kampen to analyse subextraction as head-movement (D^o can strand NP only as a head), and this forces the author to some special assumptions on wh-movement, in order to deal with the Head Movement Constraint and the landing site of wh-movement.

If we (uncontroversially) assume that wh-movement is phrasal movement, then either left branch subextractions can only involve a specifier (or an adjunct), or they involve more complex derivations (remnant movement). We will adopt the latter option in section 3 in our account of child subextraction.

With regards to phrasal subextraction, consider the proposal in Gavruseva & Thornton's (2001). They provide a Case-based account of the subextraction of possessor phrases. In order to account for possessive extraction in Hungarian and its impossibility in adult English, they claim that Genitive case blocks *whose*-extraction in English-like languages. In Hungarian, in contrast, possessed phrases are not licensed as Genitive, but get an independent Case (Dative), as extensively discussed by Szabolcsi (1983/84,1994):

(8)   Kinek$_i$ veszett   el   [$_{DP}$ t$_i$ a kalap-ja]?
       Who-DAT lose-PAST away   the hat-3sg.NOM
       "Whose hat got lost?"   (Gavruseva & Thornton 2001)

This approach contrasts with Corver's and van Kampen's in two respects. First, it does not involve any special assumptions about phrase structure (Corver's duality of nominal arguments) or wh-movement (van Kampen's head wh-movement). Second, it is not a theory about a general subextraction parameter: it deals with the specific Case licensing requirements of possessor phrases, as they apply to subextraction.

This suggests that there may be no macro-parameter of the type [+/– subextraction language]: strategies for Case licensing, responsible for subextraction possibilities, may vary both between and within languages, as they may depend on lexical specifications in functional heads (such a Genitive Case in D^o). (9) illustrates the attested variation that should be predicted:

(9)

	Subextraction of		
Language	'which'	'how-many'	'whose'
Slavic	+	+	+
Hungarian	-	(?)+	+
French	-	+	?+
English	-	-	-

Therefore deviations in child language should occur to the extent that the Case licensing intricacies of the adult language are hard to acquire. The options for subextraction in adult languages appear to be those in (10).

(10) a) some general Case licensing strategy for DPs/NPs (Slavic, Latin).
      b) some specific Case licensing strategy for DP possessors (Hungarian).
      c) some specific Case licencing strategy for NPs preceded by a preposition (French *combien de* NP, Dutch *wat voor* NP)

As we pointed out, from a minimalist perspective, there is an important issue to be addressed by any approach to subextraction: there is no attested case, either in child language or in adult language, of obligatory subextraction. The same language, and the same individuals, allow for both pied-piping and subextraction. In other words, there are instances of obligatory pied-piping, instances of optionality, but no instance of obligatory subextraction. If, in minimalist terms, we characterise pied-piping as a last resort strategy, to be applied only when subextraction is not available, the prediction should be that, when pied-piping is not required, it is not possible, contrary to fact.

If both subextraction and pied-piping are possible, this must mean, in minimalist terms, that they are equally economical. Most plausibly, it must mean that they involve different numerations, that are equally available. With these considerations in mind, let us now turn to the child Catalan subextraction problem. In what follows, we advance a proposal for child Catalan subextraction which closely relates it to French *combien* subextraction. We do not make any proposal for the Slavic type of subextraction; we just speculate that its availability across all DPs should derive from general Case properties of DPs in these languages.

3.   Wh-subextraction in Catalan

Adult Catalan does not seem to allow subextraction generally (except for cases like (2c) above). For child Catalan, the questions to be addressed are (a) which kind of misanalysis leads the Catalan child to allow subextraction in the attested cases; and (b) what features of the input adult Catalan (if any) might favor or make possible this misanalysis.

Regarding the first question, we propose that subextraction in child Catalan can be essentially analysed as French *combien* subextraction, for which we will adopt Kayne's (2002) proposal. As for the second question, we argue that, despite all appearances, adult Catalan already features this kind of subextraction, although in a disguised fashion.

Kayne (2002) argues that apparent subextraction in French is a case of remnant movement. Essentially, and simplifying the details of his account, for sentences like (11), first studied in Obenauer (1984), we have a derivation like that in (12) (expressed as successive merge/move).

(11) *Combien   a-t-il   acheté de livres?*
     how-many has-he bought of books
     "How many books has he bought?"

(12) a.   acheté [livres combien]
     b.   livres$_i$ ...[acheté [t$_i$ combien   ]]
     c.   de [livres$_i$ ...[acheté [t$_i$ combien   ]]]
     d.   [acheté [t$_i$ combien]]$_k$ [de [livres$_i$ t$_k$]]
     e. [t$_i$ combien]$_j$ ...[acheté   t$_j$]$_k$ [de [livres$_i$ t$_k$]]

The derivation in (12) can be summarised in the following way:
  • We start by merging the verb with its object DP [*livres combien*], where the NP *livres* is a specifier.[2]
  • In (12b) we extract the NP [*livres*] to the specifier of a FC responsible for Case licensing of NPs (K-de, in Kayne's terms)
  • Step (12c) merges the functional head *de*, which does not form a constituent with *livres*.

---

[2] An anonymous reviewer claims that Kayne's (2002) derivation (12) involves a special argument structure, where the "partitive" (*de*) *livres* is a "separate argument" from the object (*combien*). However. Kayne explicitly argues (p. 76, in the discussion of his +N Case Filter (25)) that one single argument containing various nominal heads (D and N) may be Case-licensed for each of those heads by separate strategies. In (12), there is one internal argument [*livres combien*] and, while *livres* is Case-licenced by an oblique functional head (visible as *de*), *combien* will check Accustive Case in the usual way.

• Step (12d) is VP movement to Spec, *de*P.

• Finally, and essentially to our proposal, in step (12e) remnant movement of [t *combien*] takes place.

Essential to Kayne's proposal is that steps (12b,c) are determined by Case requirements: *livres* moves to the specifier of a functional category that licences Case. While step (12e) (wh-movement) is uniform across many languages, the possibility of steps (12b,c,d) may vary from language to language and, apparently, also within a language. In languages like French, in a QP like *combien de livres*, the NP [*livres*], would be Case-licensed via movement to some specific FC, independently of the Case licensing of the Quantifier Phrase [t *combien*]; in others, like Italian, the NP would be Case licensed together with the quantifier, probably by agreement. Let us call these possibilities split Case licensing and agreement Case licensing. (Apparent) subextraction is predicted to occur only with split Case licensing.

Admittedly, this proposal looks like a step back from the minimalist ideal of a "uniform syntax except for PF requirements". The non-uniform analysis, however, can be defended if it can be shown that the two derivations find independent motivation.

Let us see then whether there are proofs of the availability of split Case licensing in Catalan. In adult Catalan, subextraction is only available in one case, (13c), but not in cases like (13a-b), where there is obligatory pied piping as in (14).

(13) a. *Quants   has   comprat t (de) llibres?*
          how-many   have-you bought   (of) books
          "How many books have you bought?"
      b. *Quins has   comprat t (de) llibres?*
          which   have-you bought   (of) books
          "Which books have you bought?"
      c. *Com és   de llarg?*
          how is-it   of long
          "How long is it?"
(14) a. *Quants   llibres   has   comprat t?*
          how-many books   have-you   bought
          "How many books have you bought?"
      b. *Quins llibres   has   comprat t?*
          which books   have-you bought
          "Which books have you bought?"

Specifically, (13b) is not allowed in contrast with our reported cases in child language. What evidence could lead the child to allow subextraction in this

case? We propose that independent evidence in the adult language provides a
clue. Consider (15):

(15) a. *Quants   n'has   comprat, de llibres?*
          how-many NE-have-you bought   (of) books
          "How many books have you bought?"
      b. *Quins   has   comprat,   de llibres?*
          which have-you bought   (of) books
          "Which books have you bought?"
      c. *Com n'és,   de llarg!*
          how NE-is-it   of long
          "How long it is!"

Examples (13) and (15) differ minimally in that in the latter the inner NP *(de)
llibres* appears as right dislocated (it can also appear as left dislocated). The
point is: why are the dislocated NPs marked with the preposition *de*? A
possible answer to this question is that this an idiosyncratic requirement on
dislocated NPs (and APs, as in (15c)). A more interesting answer is that these
dislocated constituents have been Case-licensed through split Case licensing.
(15b) would have the derivation in (16):

(16) a.   comprat [*llibres quins*]
      b.   *llibres*$_i$ ...[comprat [t_i *quins*]]
      c.   *de* [*llibres*$_i$ ...[comprat [t_i *quins*]]]]
      d.   [comprat [t_i *quins*]] [*de* [*llibres*$_i$ t_k]]
      e. [t_i *quins*]$_j$ ...[comprat t_j]] [*de* [*llibres*$_i$ t_k]]

        The derivation in (16) differs from Kayne's (12) in two respects. First,
it involves a 'which' wh-phrase, while French only allows subextraction of
'how-many' wh-phrases. For this, we assume that the availability of the
structure [*llibres quins*] in (16a) in Catalan, where the NP *llibres* occupies the
highest specifier, depends on the lexical properties of the head *quins*, not
shared by its French counterpart *quels*.
        Second, the stranded nominal is (right) dislocated. Now, suppose that
the constituent [*de llibres* t_k] in (16e) is obligatorily assigned a Topic feature
in Catalan, and that this feature forces it to move to the appropriate specifier,
the specifier of a TopicP that licences dislocated elements, following Villalba
(2000). It is well known that languages differ in the obligatoriness of overt
syntactic focus/topic marking. Now let us try to put things together into the
following proposal:

(17) a. Languages like French or Catalan share the possibility of split Case licensing, as in derivations (12) and (16).
b. All languages share the possibility of agreement Case licensing (which gives pied-piping for wh-phrases).[3]
c. Catalan differs from French in allowing split Case licensing not only for 'how-many' phrases, but also for 'which' phrases.
d. Catalan differs from French in the obligatoriness of marking the stranded part of step (16d) as Topic, which implies it must end up as (right or left) dislocated.[4]
e. We assume that (left/right) dislocation is not base-generated but consists in moving a Topic-marked constituent to a certain (higher/lower) TopP specifier (Villalba 2000).
f. For Catalan, examples like those in (15), which are frequent in adult speech, constitute robust evidence for split Case licensing.

We argue, then, that adult Catalan provides the child with evidence for split Case licensing, which occurs only together with dislocation of the remnant [*de* NP...] constituent, as in (15b), and the child generalises it to cases without dislocation. In Gavarró & Solà (2004) we assume that, for the child, there is not robust enough evidence for setting the obligatoriness of Topic marking (dislocation), so that the child is led to admit the possibility of split Case licensing without Topic marking. In the experiment, the child's interpretation involves two adjustments with respect to the adult grammar: the wh-determiner *qui* 'who' is interpreted as *quin* 'which'; and the absence of the Case-related preposition *de* is disregarded in the comprehension cases. This leads the child to assume that Catalan has split Case licensing without dislocation.

This analysis does not presuppose that children are insensitive to the presence of *de* or to the prosodic properties of dislocation; rather, they are crucially sensitive to the presence of *de* in the adult input with dislocation, which constitutes robust evidence for split Case licensing. And also children may perfectly recognise the phonological pattern for dislocation. But as children overgeneralise split Case licensing to cases without dislocation, they may accidentally parse as such sentences where the preposition (after all, an

---

[3] As an anonymous reviewer points out, we do not provide an answer to the question why pied-piping is always an option for all (child and adult) languages. We just assume that Kayne's Case licensing by agreement is always available. Perhaps this is so because of the more local character of this option.

[4] In French, dislocation of the stranded part is not obligatory (i), but it is certainly possible (ii):
(i)   Combien   a-t-il acheté   de livres?
        How-many   has he bought   of books
(ii)   Combien   en   a-t-il acheté,   de livres?
        How-many   PART has he bought,   of books

unstressed monosyllabic word) is not in the input. As a consequence, the analysis here is consistent with results in the literature indicating that children are aware of prosodic properties such as sentence and word level stress (de Cat 2000, Baauw et al. 2003, de Cat 2004).

A remaining question is: how do children abandon the grammar which allows for the deviant interpretation of (6a)? For comprehension, the prediction is that miscomprehesion will decrease as errors in morphological parsing (such as *qui* 'who' being interpreted as *quin* 'which') decrease; the child's sensitivity to the phonological contrast between *qui* and *quin* increases through his 2s and 3s and so the child may at least partly grow out of the early grammar thanks to that awareness. Obviously, this reasoning does not extend to the expected (but yet unattested) cases of production, a matter for further research.

To conclude, we argue that the deviations in the interpretation of wh-questions found in child Catalan are grammatical in nature and stem from the fact that UG makes (apparent) subextraction available if there is split Case assignment to internal arguments. We crucially assume that cases like (15b) involve split Case licensing plus left/right dislocation of the non-focused [*de* NP ...] remnant constituent. How plausible is this assumption for adult Catalan? Put another way: isn't the *de* NP constituent in (15b) just a loose topic, or an aboutness phrase, not really linked to the wh-determiner? We will argue that it is not, and propose it is a case of Clitic Left/Right Dislocation.

As extensively argued in Villalba (2000), Romance Clitic Left/Right Dislocation is an instance of movement to a Top specifier, a higher one for left Clitic Left Dislocation and a lower one for Clitic Right Dislocation. In this way, split Case licensing is expected to be able to feed Clitic Dislocation. Let us then check if the dislocated constituent in (15b) is likely to be a case of Clitic Right Dislocation.

Obviously, (15b) does not feature a clitic, unlike similar cases with other determiners. Compare the following:

(18) a. *Quants*   en *tens,   de llibres?*
         how-many NE have-you, of books
         "How many books do you have?"
      b. En *tinc   molts,   de llibres.*
         NE   have-I   many   of books
         "I have many."

(19) a. *Quin(s) tens,   de llibre(s)?*
         which   have-you of book(s)
         "Which books do you have?"
      b. *Tinc   aquest(s), de llibre(s).*
         have-I this/these of book(s)
         "I have this/these."

In (18), quantifier determiners like *quants* 'how many' or *molts* 'many' allow for Clitic Right (and Left) dislocation of *de llibres* (of books). Let us call this case *quants-de*-NP dislocation. With determiners like *quin* 'which' and *aquest* 'this', dislocation of *de llibres* does not involve a clitic. Let us call this case *quin(s)-de*-NP dislocation.

We propose that *quin(s)-de*-NP dislocation, like *quants-de*-NP dislocation, is essentially an instance of Clitic Left/Right Dislocation, but that with *quins* the *en* clitic is not available for the dislocated *de llibres*.[5] We therefore assume that cases of *quin(s)-de*-NP dislocation involve Clitic Right/Left Dislocation without a clitic, where Clitic Left/Right Dislocation stands for a well defined configuration of Topic movement to certain functional specifiers and the appearance of a clitic depends on its morphosyntactic availability.

Let us now check whether *quin(s)-de*-NP dislocation behaves like other well known cases of Clitic Left/Right dislocation. Various kinds of evidence give support to this hypothesis.

First of all, *quin(s)-de*-NP dislocates can occupy exactly the same positions as other clitic dislocates, and stand in free order with respect to each other:

(20) a. *Quins   li   donaràs,   a en Joan,   de llibres?*
       which   him   you'll-give to the Joan   of books
       "Which books will you give Joan?"
   b. *Quins   li   donaràs,   de llibres, a en Joan?*
       Which   him   you'll- give of books   to the Joan
       "Which books will you give to Joan?"

Second, like clitic left dislocates, they can move to a superordinate clause:

(21) *De llibre, no sé   quin compraré.*
    of book   not know-I which I-will-buy
    "I don't know which book I'll buy."

In this case, like in other cases of Clitic Left Dislocation, they show sensitivity to strong islands (Villalba 2000):

---

[5] Presumably this is so because the NP *llibres* that appears with both *quants* and *quins* is in fact embedded in different nominal functional categories in each case. We thank an anonymous reviewer for pointing out that our analysis of *quins*-NP is not a trivial extension of Kayne's analysis, which involves only partitive cases like *combien de livres*.

(22) a. *De llibre, he   marxat [sense saber quin llegir —]
         of book   have-I left   without knowing which to read]
         "I have gone without knowing which book to read."
     b. *De llibre, conec [l'autor   que ha escrit aquest —]
         of book   know-I the-author that wrote   this
         "I know the author who has written this book."

    Also like the other cases of Clitic Left/Right Dislocation, and unlike other cases of dislocation, they show connectivity effects: the NP must agree in number with the determiner *quin(s)*.

(23) a. *Quin   vols,   de llibre?*
         which-SG   want-you   of book-SG
         "Which book do you want?"
     b. *Quins   vols,   de llibres?*
         which-PL   want-you   of book-PL
         "Which books do you want?"
     c. **Quin   vols,   de llibres?*
         which-SG want-you of book-PL
     d. **Quins   vols,   de llibre?*
         Which-PL want-you   of book-SG

    A further piece of evidence for connectedness relates to the split Case licensing hypothesis we have adopted. Split Case licensing is an option for internal arguments. And indeed, *quin(s)-de*-NP dislocation is only possible with internal arguments:

(24) a. *Quin   vols,   de llibre?*   (Transitive object)
         which   want-you of book
     b. *Quin ha arribat, de llibre?*   (Unaccusative subject)
         which has arrived   of book
     c. *[*]Quin   vol   llibres, de professor?* (Transitive subject)
          which wants   books   of profesor
     d. *^{?*}Quin ha protestat, d'alumne?*   (Unergative subject)
          which has protested, of student

    We conclude, then, that *quin(s)-de*-NP dislocation, like Clitic Left/Right Dislocation, is a syntactically well defined phenomenon involving movement to a Top Specifier, and is fed by split Case licensing.
    As we pointed out, Catalan differs from French in forcing the dislocation of the stranded part of subextraction cases, due to its topic character. This assumption might seem arbitrary, but turns out to be quite

natural: in a wh-sentence, only the wh-material is focused. Now we can assume that subextraction involves focusing of the subextracted part only, so that the stranded part belongs to the background, and can be plausibly promoted to topic.

This view receives support from Mathieu's (2004) analysis of the informational structure of *combien* subextraction in French. He argues that the stranded nominal in *combien* subextraction contexts is a *new* topic, unlike a pied-piped nominal, which is part of focus. Also, van Kampen (2000) shows that in subextraction in child Dutch the stranded material is always unstressed, while the subextracted part is the one being focused. This makes dislocation of the stranded part a quite natural option in a language like Catalan where, as shown by Vallduví (1992), topics cannot remain in the undifferentiated background and must be dislocated.[6]

This seems to indicate that, even if the pied-piping/subextraction alternative may be formulated in terms of the formal properties of syntax (availability of split Case licensing), it has interpretative consequences: different derivations give different interpretations at the interface.

4. *Summary and some further speculations*

Split Case licensing, as defined in Kayne (2002), creates configurations in which (apparent) subextraction is possible, by way of extracting the NP part of a wh-phrase to a Case position. We have argued that adult Catalan features split Case licensing, although in a disguised way, since the remnant [*de* NP ...] constituent is forced to move to a Top specifier, in the same way as dislocates do in Clitic Left/Right Dislocation. We have adduced evidence that split Case licensing feeds *de*-NP dislocation. This result is furthermore consistent with the facts of child Catalan in Gavarró & Solà (2004): according to our analysis, in child Catalan subextraction is also dependent on Case licensing strategies.

If presence of *de* in nominal dislocates is indeed evidence of previous split Case licensing, the child will be led to assume its presence in the target language and may possibly misgeneralise it to other cases. Now, Catalan and French are languages where *de* preceeding NPs in those cases constitutes robust evidence. Other Romance languages are clearly different. Spanish does not feature *de*-NP in dislocation generally. Specifically, Catalan and Spanish contrast in crucial cases like:

(25) a. *Quins vols,   de llibres?*
      which want-you   of books
      "Which books do you want?"

---

[6] We thank an anonymous reviewer for pointing out the interpretative consequences of subextraction/pied-piping alternation.

    b. **Cuales quieres,   (de) libros?*
      which want-you of   books
      "Which books do you want?"

This strongly suggests that split Case licensing is not available, at least in this instance, for Spanish. This would predict that Spanish children, in the absence of such evidence, would not be led to produce wh-subextraction, assuming that split Case licensing is a strategy to be learned on positive evidence – a topic for future research. If this is on the right track, research on child non-target wh-subextraction could benefit from trying to trace it to available evidence for adult subextraction, perhaps of a more subtle or marginal nature.

References

Baauw, Sergio, Esther Ruigendijk & Fernando Cuetos. 2004. 'The interpretation of contrastive stress in Spanish-speaking children'. *Proceedings of GALA 2003*, ed. by J. van Kampen & S. Baauw. Utrecht: LOT Occasional Series, 103–114.

Chen, D., M. Yamane & W. Snyder. 1998. 'On the nature of Children's Left Branch Extractions'. *Proceedings of the 22nd Annual Boston University Conference on Language Development.*

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, Mass.: The MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective.* Oxford: Oxford University Press.

Corver, N. 1990. *The Syntax of Left-Branch Extractions.* PhD dissertation, University of Tilburg.

De Cat, Cécile (2000) 'Structure building and the acquisition of dislocations in child French'. *Proceedings of BUCLD 24.* Somerville: Cascadilla Press, 242–252.

De Cat, Cécile (2004) "Early 'pragmatic' competence and its implications regarding the null subject phenomenon". *Romance Languages and Linguistic Theory 2002*, ed. R. Bok-Bennema, B. Hollebrandse, B. Kampers-Manhe & P. Sleeman. Amsterda/Philadelphia: John Benjamins, 17–31.

Gavarró, Anna; Solà, Jaume. 2004. 'Wh-subextraction in Child Language'. Paper presented at the Second Lisbon Meeting on Language Acquisition, Universidade de Lisboa, June 1st 2004.

Gavruseva, Elena & Rosalind Thornton. 2001. 'Getting it right: acquisition of whose-questions in child English'. *Language Acquisition* 9:3.229–267.

Hoekstra, Teun & Peter Jordens. 1994. 'From adjunct to head'. *Language Acquisition Studies in Generative Grammar,* ed. by Teun Hoekstra & Bonnie Schwarz. Amsterdam/Philadelphia: John Benjamins, 119–149.

Hoekstra, Teun, Jan Koster & Thomas Roeper. 1992. 'Left Branch violations in acquisition'. Paper presented at the *Boston University Conference on Language Development.* October 24, Boston.

Jordens, Peter & Teun Hoekstra. 1991. 'The acquisition of negation'. Paper presented at the GLOW Workshop 'The development of bound variables and operators'. March 28, Leiden.

van Kampen, Jacqueline. 1994. 'The learnability of the left-branch condition'. *Linguistics in the Netherlands 1994*, ed. by R. Bok-Bennema & C. Cremers, 83–94. Amsterdam: John Benjamins.

van Kampen, Jacqueline. 1997. *First Steps in Wh-Movement.* Doctoral dissertation, Utrecht University.

van Kampen, Jacqueline. 2000. 'Left-branch extractions as operator movement: evidence from child Dutch'. *The Acquisition of Scrambling and Cliticization*, ed. by Susan Powers & Cornelia Hamann. Dordrecht: Kluwer.

van Kampen, J. 2004 'An acquisitional view on optionality'. *Lingua* 114.1133–1146.

Kayne, Richard. 2002. 'On some prepositions that look DP-internal: English of and French de'. *Catalan Working Papers in Linguistics* 2.71–115.

Mathieu, Eric (2004) 'The mapping of form and interpretation: the case of optional wh-movement in French'. *Lingua* 114.1090–1132.

Obenauer, Hans. 1984. 'On the identification of empty categories'. *The Linguistic Review* 4-2.153–202.

Ross, J. R. 1967. *Constraints on Variables in Syntax.* Doctoral dissertation, MIT.

Szabolcsi, Anna. 1983/84. 'The possessive that ran away from home'. *The Linguistic Review* 3. 89–102.

Szabolcsi, Anna. 1994. 'The Noun Phrase'. In: Kiefer, F.; Kiss, K. (eds.). *The Syntactic Structure of Hungarian*. San Diego, California: Academic Press, 179–274.

Thornton, Rosalind. 1990. *Adventures in long-distance moving: The acquisition of complex wh-questions.* Doctoral dissertation, University of Connecticut.

Thornton, Rosalind & Elena Gavruseva. 1996. 'Children's split "*Whose*-questions" and the structure of possessive NPs". Paper presented at the *21st Annual Boston University* Conference on Language Development.

Vallduví, Enric. 1992. *The informational component.* New York/London: Garland Publishers.

Villalba, Xavier. 2000. *The syntax of sentence periphery*. Doctoral dissertation, Universitat Autònoma de Barcelona.

Yamane, M., D. Chen & W. Snyder. 1999. 'Subject-object asymmetries and Children's Left Branch Extractions' *Proceedings of the 23nd Annual Boston University Conference on Language Development*, vol. 2: 732–740.