INVERSION IN WH-QUESTIONS IN CHILD ROMANCE AND CHILD ENGLISH

GRANT GOODALL
University of California, San Diego

0. Introduction

Many Romance languages show an obligatory inversion effect in some types of wh-questions, in that the verb is required to be adjacent to the wh-word, as in (1) ((1a-c) from Hulk and Pollock (2001)).

(1) a. Che cosa ha detto Maria? [Italian]
what has said
‘What did Maria say?’
b. Onde foi a Maria? [European Portuguese]
where went the
‘Where did Maria go?’
c. Què farà en Joan? [Catalan]
what do the
‘What will Joan do?’
d. Adónde fue María? [Spanish]
where went
‘Where did María go?’

This is of course reminiscent of the English phenomenon of subject-auxiliary inversion in non-subject wh-questions, as in (2).

(2) What will John say?

The similarity between the Romance and English cases makes it tempting to

---

1 Many thanks to Joana Rosselló for first pointing out to me that the analysis of inversion that I was then developing could have interesting implications for acquisition. Thanks also to the anonymous reviewers of this article and to audiences at UC Davis, University of Chicago, University of Maryland, Universitat de Barcelona, and LSRL for valuable comments on various aspects of the analysis.
claim that the same mechanism underlies both, and that has in fact often been suggested in the literature (e.g., Rizzi (1996), Pesetsky and Torrego (2001)).

Given this background, it is surprising that there are major differences between Romance and English in the pattern of acquisition of inversion in wh-questions in children. In English, children show an unusually wide range of individual variation in their use of inversion in this context. Stromswold (1990) reports an overall inversion rate of 93% among children in wh-questions, but individual rates range from 60.1% to 99.3%. In addition, some children show a long period of very unstable development, with inversion rates jumping between 40% and 100% at different times, and with the adult rate of 100% not becoming fixed until at least age 5.

In Romance, the picture is very different. There is no individual variation, as all children show a 100% adult-like inversion rate from their very first production of wh-questions. This pattern has been attested in Catalan (Grinstead (2001), Serrat & Capdevila (2001)), European Portuguese (Soares (2003)), Italian (Guasti (2000)), and Spanish (Serrat & Capdevila (2001)).

In this paper, I will propose an explanation for why there should be such a dramatic contrast between these Romance languages and English. I will claim that an explanation requires attention not just to the syntactic properties of the languages in question, but also to the way in which syntactic structures are processed. I will restrict my discussion here to Spanish and English, but the basic approach may be extendable to other Romance languages as well.

1. An analysis of the Spanish case

My basic claim will be that the facts of inversion in Spanish wh-questions could be no other way, and that children thus essentially have no choice but to do adult-like inversion from the very beginning. This is similar in spirit to the many attempts in the literature at deriving Spanish inversion from properties of UG (see Goodall (2001) and the references there), but a satisfactory solution along these lines has been elusive and in any event, the acquisition contrast between Spanish and English has not been addressed (though see Guasti (2000) for Italian and English).

In order to make such an analysis go through, we need to take into account both the syntactic (i.e., purely grammatical) properties of the language and the way natural language is processed. On the syntax side, we need to assume wh-movement, of course, as well as V-to-T movement and “optional”

---

2 Children are also adult-like in not requiring inversion in environments such as relative clauses and questions with why. See Goodall (2004, 2006) for an account of cases like these.
raising of the subject to SPEC/T (however such optionality is to be accounted for). The latter two assumptions jointly allow for subjects to appear in either preverbal or postverbal positions, as has been widely observed for Spanish:

(3) a. *Juan llegó.
    
    b. Llegó Juan.

   
    ‘Juan arrived.’

   

These assumptions about the syntax are standard in the literature on Spanish, and on their own, they would lead us to expect both (4a) and (b) to be acceptable, contrary to fact.

(4) a. *Qué Juan compró?
    
    b. Qué compró Juan?

   
    ‘What did Juan buy?’

Clearly, then, we still have some work to do in order to account for the obligatory inversion effect exemplified in (4).

On the processing side, we will adopt the standard view that wh-movement creates a filler-gap dependency (Fodor (1978)), which poses difficulties for the processor because of the need to process the filler (wh-phrase), to process the intervening material while holding the filler in working memory, and then to integrate the filler to the gap position once the subcategorizing verb is encountered. The more intervening material there is to process, the more difficult it is to process the filler-gap dependency.

It is notable that given these very minimal assumptions, we now predict that (4a) will be dispreferred, in that there is intervening material between the filler qué ‘what’ and the gap, indicated by the subcategorizing verb compró ‘bought’. Can this processing dispreference be strengthened to full unacceptability? To do this, we would need to say that the intervening material in (4a) (i.e., the preverbal subject Juan) is so difficult to process that it interferes greatly with the processing of the filler-gap dependency, resulting in a processing bottleneck that is perceived as unacceptability.

In fact, there are reasons to think that preverbal subjects in Spanish are especially hard to process. To begin with, subjects in general have been
claimed to pose a particularly high processing burden (Clark and Wasow (1998), Kluender (2004)), but Spanish subjects have additional properties that could increase the burden even more. First, subjects in Spanish may be either overt or null, with overt subjects taken to be (very roughly) contrastive or emphatic (e.g., Luján (1999)). Overt subjects should be more difficult to process not just because they are overt, but also because they are less accessible in discourse than null subjects and are more likely to have rich descriptive content, to be definite, or to consist of open-class lexical items. All of these properties have been claimed to contribute significantly to increasing the processing burden (Kluender (1998), Kluender (2004), Kutas, Besson, and Van Petten (1988), Warren and Gibson (2002)), and this in turn increases the difficulty of processing the filler-gap dependency established by wh-movement. Note that part of this difficulty arises because of the special discourse role that overt subjects have in Spanish: they are relatively low in discourse-accessibility (otherwise they would be null) and there are additional discourse properties (e.g., contrast or emphasis) that must be processed in addition to the lexical material and syntactic structure.

A second property of subjects in Spanish that seems likely to contribute to processing difficulty in sentences like (4a) is the fact that subjects may generally appear in either preverbal or postverbal position, as we saw in (3). Our concern here is the preverbal subject, since this is the type that can intervene in a filler-gap dependency. Preverbal subjects have been characterized in a number of ways, including as topics, subjects of individual-level predicates, or subjects of all-focus sentences (e.g., Casielles-Suárez (1999), Contreras (1976), Zubizarreta (1998)). Byrne (1998) argues that the preverbal/postverbal distinction is best characterized in terms of default judgment types, with preverbal subjects associated with a default interpretation of categorical judgment and postverbal subjects with thetic judgment. Interestingly, categorical judgment is the type in which the subject is perceived as highly individuated, and it is just this property of individuation which has been argued to be associated with a significantly increased processing burden (Kluender (1998)).

There are thus very plausible reasons for believing that overt preverbal subjects pose processing difficulties for wh-questions above and beyond the usual disruptive effect of any type of intervening material, and that this could be responsible for the unacceptability of (4a). Note that these difficulties arise largely because Spanish allows for subjects to be either overt or null, and either preverbal or postverbal. Each of these choices has interpretational and discourse consequences, and the consequences associated with the choice
“overt and preverbal” lead to a type of subject that is particularly difficult to process, as we have seen. Given that the nature of the intervening material is known to affect the processability of filler-gap dependencies (i.e., the harder the intervening material is to process, the less likely the filler-gap dependency is to be processed successfully), we thus arrive at a plausible explanation for why inversion is required in *wh*-questions as in (4).

I have reported elsewhere (Goodall (2004), Goodall (2006)) on the results of work suggesting that attributing the inversion effect in Spanish to processing in this way is on the right track. Assuming that this is so, let us now return to the question of how children acquire inversion in Spanish. Recall that children’s *wh*-questions are error-free in this regard from their earliest stages of production (both in Spanish and in similar Romance languages).

Given the logic of the account sketched above, once the child is aware of the overt/null and preverbal/postverbal contrasts, including the discourse consequences of these contrasts, interrupting a filler-gap dependency with an overt preverbal subject, as in (4a), should be at least as difficult for the child as it is for the adult.

The question to be faced now is whether children have acquired these contrasts by the time they begin producing *wh*-questions. We shall now explore this question in some detail, first with regard to null and overt subjects, and then with regard to preverbal and postverbal subjects.

2. **Null and overt subjects in child Spanish**

There is substantial evidence in the literature that children converge early on the idea that both null and overt subjects are allowed in Spanish, as well as in Catalan and Italian, which have very similar properties in this regard (Bel (2003), Valian (1991)). Bel (2003) discusses the case of María, for instance, a Spanish-speaking child whose null subject use is given in Figure 1. Figure 1 shows the percentage of subjects that are null out of María’s overall subject use. This may be compared with the adult range of use of null subjects in Spanish (indicated by dashed lines in Figure 1), which has been reported in the literature to be between 61% and 80% (Bentivoglio (1987), Enríquez (1984), Silva-Corvalán (1977)). María reaches this adult level at 1;10 and mostly remains within that range from then on.

These data tell us first of all that María is clearly making use of both null and overt subjects. Second, they suggest that at a relatively early point, María’s use of null subjects becomes at least similar to adults’ in terms of discourse properties, given that her proportions of null vs. overt subjects are so
much like adults’. This latter point is supported further by the fact that when looked at in the larger discourse context, children’s use of null subjects appears to be fully adult-like, as Bel points out. (5) and (6), for instance, show examples where a null subject is used to refer to an entity that is highly accessible in the discourse or environment.

(5) Mother: Que se ha caído, estaba aquí puesto y se ha caído. That CL has fallen was here placed and CL has fallen ‘It fell, it was here and it fell.’
Child: (Es)tá aquí. [Signaling the telephone] (María 1;8) is here ‘It is here.’

(6) O loto. [Showing a broken pair of glasses] (María 1;9) has broken ‘It has broken.’

In (7) and (8), in contrast, the entity referred to is less accessible and an overt subject is used.
Bel reports that no cases were found of non-adult-like use of null subjects.

Bel’s finding, that children converge early on the possibility of both null and overt subjects and that they differentiate between the two in an apparently adult-like way, is partially at odds with the results of Grinstead (2004), who finds that children go through a period of using only null subjects first and develop overt subjects later. Nonetheless, the Spanish-speaking children in Grinstead’s study still show relatively early convergence on the full range of possibilities for subjects, with preverbal overt subjects first appearing between 1;10 and 2;1 (and between 1;10 and 2;4 with the Catalan-speaking children).

3. **Preverbal and postverbal subjects in child Spanish**

Having seen that children seem to converge early both on the syntactic possibility of null and overt subjects and on the discourse use of each, we now turn to the distinction between preverbal and postverbal subjects. Here again, the available evidence indicates that there is early convergence on the possibility of these two subject positions. Bel (2003) shows that in the case of María, both preverbal and postverbal subjects exist from a very early age, as shown in Figure 2. Moreover, it appears that María has assigned adult-like discourse roles to each of these positions, choosing the preverbal position for subjects that are contrastive or otherwise focused:

(9) *Este se quita.* [%ca]

this CL removes
‘It falls.’

(10) *Guauguau no pasa.*

dog not pass
‘Dog doesn’t go by.’

(11) *¿Sabes? Se rompe, e(s)te patito se rompe.*

you know CL breaks this little-duck CL breaks
‘You know? It breaks, this little duck breaks.’
Bel reports that *este* ‘this’ in (9) and *e(s)te patito* ‘this little duck’ in (11) receive contrastive focus, and *guauguau* ‘dog’ in (10) is simply focused. Both of these roles are standardly associated with the preverbal position.

Similarly, María’s use of postverbal subjects appears to be completely adult-like:

(12)  *Se cae e(l) nene.*  
     CL fall the baby  
     ‘The baby falls.’

(13)  *Fal(l)ta el ot(r)o pendiente.*  
     lacks the other earring  
     ‘The other earring is missing.’

(14)  ¡*Ah!* *No (se) para (e)se patito.*  
     ah! not CL stop that little-duck  
     ‘That little duck doesn’t stop.’

All of these involve arguably unaccusative verbs, which are known to very strongly favor the postverbal position for subjects in the adult language, barring a focus or other special interpretation for the subject.

The difference in discourse roles for preverbal and postverbal positions for subject is subtle in Spanish, and it is admittedly difficult to be certain that
children’s and adults’ use of the two positions match up exactly, but at least in those cases where the adult preference for one position or another is reasonably clear, children seem to follow these adult preferences very closely. This suggests that children acquire early not just the syntactic possibility of preverbal and postverbal positions, but the discourse distinctions between the two as well.

4. Timing of the acquisition of subjects and wh-questions in child Spanish

We have now seen considerable evidence for early convergence on the properties of the subject in Spanish, i.e. for the idea that children show knowledge of the null/overt and preverbal/postverbal distinctions beginning with their earliest production of subjects, including sensitivity to the discourse conditions governing these distinctions. This leads us to expect that as with adults, overt preverbal subjects in Spanish will be particularly difficult for children to process (though clearly not impossible), and they will be fatally disruptive when they intervene in a filler-gap dependency. If this is correct, it provides us with a straightforward explanation for why Spanish-speaking children never make errors with regard to inversion in wh-questions.

This explanation rests on the assumption that the properties of the subject are mostly in place by the time wh-questions are first produced. If this were not the case, i.e. if wh-questions began to be produced when the child had not yet acquired the null/overt and preverbal/postverbal contrasts, we might expect that children would go through a non-inversion stage in their wh-questions until they attach enough additional properties to overt preverbal subjects that processing becomes difficult.

The only study where we have direct information on the relative timing of the acquisition of subject properties and the acquisition of wh-movement is Grinstead (2004). The Spanish-speaking children in his study begin using overt preverbal subjects between 1;10 and 2;1, but they apparently do not begin performing wh-movement until at least 2;3 – 2;5, since none of them have given evidence of wh-movement by the time of their final file (between 2;2 and 2;4). The Catalan-speaking children do begin producing wh-questions, but this occurs at the same time as, or more commonly, after the onset of overt preverbal subjects, as seen in Table 1. The available evidence thus suggests very strongly that by the time children begin producing wh-questions, their knowledge of the properties of subjects is already in place, and their processing of a filler-gap dependency should thus be at least as disturbed by an intervening subject as it is for adults.
This explanation rests on the further assumption that children are no better at processing subjects and filler-gap dependencies than adults, but this assumption seems completely uncontroversial. In fact, it has been argued that children are in general worse at processing subjects than adults (Bloom (1990), (1993), Kluender (2004)). If this is true, the error-free inversion rate shown by Spanish-speaking children in *wh*-questions would be even less surprising, since children would be even less able than adults to tolerate an intervening overt subject within a filler-gap dependency.

<table>
<thead>
<tr>
<th></th>
<th>Overt subjects</th>
<th>Wh-movement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laura</td>
<td>2;4.11</td>
<td>2;4.11</td>
</tr>
<tr>
<td>Pep</td>
<td>1;10</td>
<td>1;11.6</td>
</tr>
<tr>
<td>Guillem</td>
<td>1;11.13</td>
<td>2;2.28</td>
</tr>
<tr>
<td>Gisela</td>
<td>2;1.23</td>
<td>2;8.0</td>
</tr>
</tbody>
</table>

Table 1: *Age of initial use of overt preverbal subjects and wh-movement in Catalan*

5. **Inversion in English *wh*-questions**

We have now seen that the inversion effect in *wh*-questions in Spanish is forced by the fact that subjects in the language may be null or overt, and preverbal or postverbal. Each of these choices has discourse consequences, and the result is that a *wh*-question with an intervening overt preverbal subject will be ruled out by processing considerations. We correctly predict that children will be adult-like in their inversion in *wh*-questions, since they are subject to the same (or perhaps more stringent) constraints on processing as adults.

Let us now turn to the case of English. We saw at the outset that English also has an obligatory inversion effect in *wh*-questions, but that unlike in Romance, children show great individual variation in their rate of inversion, and many do not attain a consistently adult-like rate of inversion until around age 5. This picture of inversion in child English might seem disturbing for the analysis proposed here, in that we might expect that the same factors that force inversion in Spanish would also do so in English, whether with children or adults. In particular, we might expect that *wh*-questions without inversion, as in (15), would run afoul of processing considerations just as we saw for
Spanish.

(15) *What John will say?

Note that in (15), the subject John intervenes between the filler what and the gap, indicated by the subcategorizing verb say.

It is clear, however, that English differs from Spanish in some important ways, the most crucial of which for our purposes concerns the properties of the subject. Whereas in Spanish, an overt preverbal subject represents an explicit choice along two dimensions (null/overt and preverbal/postverbal), this is not the case in English, where an overt preverbal subject is simply the default, null and/or postverbal subjects not generally being allowed by the syntax. We then do not expect (15) to present the same level of processing difficulty as in Spanish. In fact, if (15) were ruled out by the processor, the version with inversion in (2) would have to be also, since it too has an intervening subject.

There thus appears to be little hope of attributing the obligatory inversion effects in English wh-questions to processing factors (see Goodall (2004), (2006) for a discussion of ways in which the behavior of English inversion supports this conclusion), and it is plausible that they are due to syntactic (i.e., purely grammatical) factors instead. The standard analysis invokes T-to-C movement, and I will adopt that here. From the standpoint of the child, then, English inversion involves something that must be learned. The child hears wh-questions like (16) and concludes that there must be some process such as T-to-C movement operating in order to produce this non-canonical order. In Spanish, on the other hand, there is nothing special to learn about wh-questions beyond the fact of wh-movement (note that no equivalent of T-to-C movement is needed). Overt preverbal subjects are impossible in this environment, but the child does not learn this; it follows from the limited capabilities of the processor.

The different courses of development between Spanish and English now follow straightforwardly. Spanish-speaking children never produce wh-questions without inversion because to do so would involve more processing capability than they have. English-speaking children, however, do produce wh-questions without inversion because nothing prevents them from doing so, and in any event, producing wh-questions with inversion requires that they have first learned T-to-C movement. It is not surprising that children acquire this rule at different rates and apply it with varying degrees of reliability.

Notice that this account of the acquisition of inversion assumes that T-
to-C movement is not forced by properties of UG. This runs counter to what has often been proposed (e.g., Rizzi (1996), Guasti (2000)), but it is supported by the fact that T-to-C movement does not seem to occur in Spanish, as seen in the analysis here, and as argued extensively in the literature.

Under this view, then, T-to-C movement is a language-particular property that the English-speaking child must simply learn. This learning task is perhaps made more difficult by the fact that non-subject wh-questions in English inevitably involve an intervening subject. As seen earlier, subjects may in general be difficult for children to process, and if so, this could mean that even in a language like English, where overt preverbal subjects are the norm, an intervening subject might make a filler-gap dependency relatively difficult to process. Producing a wh-dependency would then be a relatively complex task, involving the processing of a filler-gap dependency, a subject, and the dependency formed by T-to-C movement, and one could very plausibly expect to see wide variation among children in their ability to handle this complexity. 

6. Conclusion

The puzzling difference between Romance and English with regard to the acquisition of inversion in wh-questions now seems much more understandable. In the course of acquisition in English, we are in effect seeing children struggle with a language-particular process that is not mandated by UG, with the individual variation and extended period of acquisition that we would expect. In Spanish, we are seeing children do the best they can with the syntactic knowledge that they have acquired, but producing wh-questions with an overt preverbal subject goes beyond their (and adults’) processing abilities. I have not attempted here to extend this analysis to Romance languages such as European Portuguese, Catalan, and Italian, which behave similarly in many respects, but such an extension appears promising. It should be obvious that the analysis I have developed here has only been possible by taking seriously the roles of both the grammar and the processor in determining sentence acceptability, so to the extent that this analysis has been successful, it constitutes an argument in favor of pursuing such an approach.

References

3 My thanks to Julia Herschensohn for pointing this out to me.


Guasti, Maria Teresa (2000). “An excursion into interrogatives in early English and Italian.” In Marc-Ariel Friedemann and Luigi Rizzi (eds.), The Acquisition of Syntax: Studies in Comparative Developmental Linguistics,
105-128. Harlow: Longman.