Catalan child relative contrasts as a processing effect

Anna Gavarró, Arnau Cunill, Míriam Muntané & Marc Reguant
(Universitat Autònoma de Barcelona)

1. Introduction

Research in various languages indicates that children interpret subject relatives in an adult-like manner substantially earlier than they interpret object relatives. This asymmetry, while grounded in a grammatical contrast, may be attributed to processing of the corresponding syntactic structures, as in Gibson (1998), Johnson (1998), Morrill (2000). One question that emerges is: if processing can be argued to be the source of poor performance in the interpretation of object relatives, does this carry over to production? Here we address this issue with the acquisition of Catalan; we present original results for relative clause elicitation and compare them with those of a relative clause interpretation experiment.

The paper proceeds as follows: first we provide the background to the present study, and we give an analysis for the subject/object relative asymmetry found in comprehension (Section 2). Then we describe an experiment for the elicitation of clitics carried out in Catalan and provide the results (Section 3); we compare production and comprehension and draw conclusions (Section 4).

2. Background: relative clause comprehension in Catalan

2.1. An experiment on relative clause comprehension

The acquisition literature reports on work in numerous languages showing a contrast between the interpretation of subject and object relative clauses. This is attested for English (Brown 1972, Sheldon 1974, de Villiers et al. 1979), French (Frauenfelder, Seguí & Mehler 1980), German (Schriefers, Friederici & Kühn 1995), Italian (Arosio, Adani & Guasti 2009), Greek (Guasti, Stavrakaki & Arosio 2008), and Hebrew (Friedmann et al. 2009), among others. This asymmetry is, however, not universal: languages with prenominal relative clauses such as Chinese and Basque are known to behave differently: see for example the work of Hsiao & Gibson (2003) on Chinese and that of Carreiras et al. (2010) on Basque. Here we centre our attention on the earlier language type, that of head-initial, postnominal relative clauses, and consider one particular language, Catalan, for which comprehension results are available, and we revisit prenominal relatives at the end of the paper.

Catalan relative clauses are postnominal, headed by a relative pronoun que which, unlike e.g. the Spanish relative pronoun, is not identical to a wh-word. Que introduces both subject and object relatives – only prepositional relatives present qui (La nena a qui he enviat el llibre ‘The girl to whom I sent the book’).

Gavarró, Adani, Ramon, Rusiñol & Sánchez (2009) carried out an experiment on the comprehension of Catalan relative clauses, modelled on the experiment of Arosio et al. (2009) and Adani (in press). The experiment was an agent identification task with picture support. Children had to identify a character in a picture when the experimenter requested it as in (1).

(1) a. Assenyala el camell que segueix els elefants!
point to the.SG camel that follows the.PL elephants
‘Point to the camel that follows the elephants!’
b. Assenyala el camell que els elefants segueixen!
   ‘Point to the camel that the elephants follow!’

c. Assenyala el camell que segueixen els elefants!
   ‘Point to the camel that the elephants follow!’

(1a) includes a subject relative, (1b) an object relative with a preverbal subject in the embedded clause, and (1c) an object relative with a postverbal subject in the embedded clause. Object relatives of this last type are potentially ambiguous as Catalan has no overt Case marking and therefore postverbal subjects may be interpreted as subjects or objects, unless subject–verb agreement disambiguates the sentence. (This ambiguity is in fact found in all the Romance null subject languages.)

All sentences in the experiment were unambiguous and reversible, so that the interpretation rested on the linguistic input only. The task was run with 33 children, 12 of which younger than 4;6, 11 of ages between 4;6 and 5;6, and 10 older than 5;6; the age range was 3;5,9–6;2,30 and the mean age 4;11,4. The results appear in (2).

(2) Results for relative clause comprehension: target comprehension for the relevant age range

<table>
<thead>
<tr>
<th>Age Range</th>
<th>SR</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;4;6</td>
<td>57/72</td>
<td>79%</td>
</tr>
<tr>
<td>4;6–5;6</td>
<td>64/66</td>
<td>97%</td>
</tr>
<tr>
<td>5;6&lt;</td>
<td>60/60</td>
<td>100%</td>
</tr>
<tr>
<td>adults</td>
<td>131/132</td>
<td>99%</td>
</tr>
</tbody>
</table>

These results indicate how subject relatives are interpreted in an adult-like manner from early on (79% of the time in the younger group), while object relatives develop at a slower pace, in fact object relatives with postverbal subjects are misinterpreted for the whole period investigated. Very similar results are found in a related language, Italian (see Arosio et al. 2009, Adani in press).

2.2. A categorial analysis

In Morrill & Gavarró (to appear) we analyse these comprehension contrasts as the result of the relative processing load of the structures involved. We propose to account for the findings by adopting an analysis based on Morrill’s (2000, 2010) metric of syntactic complexity, an implementation of Gibson’s (1998) and Johnson’s (1998) insight that processing difficulties increase as a function of the number of unresolved dependencies that the speaker must keep in memory. Morrill (2000, 2010) proposes a metric of processing cost that can account for the relative difficulty of Catalan object relatives as opposed to subject relatives (notice that both are low in processing cost if compared to centre embedding, e.g. the cheese that the rat that the cat saw ate stank, for which adults are found to have difficulties).

Categorial grammar (Morrill 2010) classifies words and expressions by means of fractional types built over basic types such as sentence (S) and nominal (which we parameterise here with number singular, N(sg), plural, N(pl), or unspecified, N(_)). An expression of type A\B is one which concatenates with any expression of type A to the left to form an expression of type B. An expression of type B/A is one which concatenates with any expression of type A to the right to form an expression of type B. Formally:

\[
A\B = \{s | \text{ for all } s' \in A, s+s' \in B\}
\]
\[
B/A = \{s | \text{ for all } s' \in A, s+s' \in B\}
\]

Morrill (2000, 2010) describes a complexity metric founded on incremental categorial processing in terms of proof nets. In this view of processing, types are marked with input polarity (•), meaning that a resource is given, or output polarity (◦), meaning that a resource is wanted. Polar types are unfolded upwards into polar type trees as follows:
We refer the reader to the references above for the details, which are quite involved, but we illustrate the basic idea here with the processing of this sentence:

(5)  John loves Mary.

Initially, a sentence is sought and after hearing the first word its type is given:

(6)  \[ S^0 \quad N(sg)^* \]

John

When the second word is heard, its unfolded type is connected by two dependencies: the subject sought is given by the first word *John* and the sentence projected is matched by the initial expectation of a sentence. We represent this as follows:

(7)

When the final word is heard, the parse is completed thus (the unspecified object number on the verb type becomes instantiated by unification with the type with which it is matched):

(8)
The derivations proposed for a subject relative and an object relative in Catalan appear in (9) and (10) respectively.

(9) Derivation of sentence (1a)

(10) Derivation of sentence (1b)

The relative pronoun *que* seeks to the right a category in turn seeking a DP, and this is satisfied earlier in the subject relative clause (in subject position) than in the object relative clauses. The processing load at each point in the sentence is mechanically worked out by the metric. The complexity profile of a sentence describes the incremental load at each word boundary by counting the syntactic dependencies that are unresolved at each point (where syntactic dependencies include major categories and feature values, both counting for 1). When the last word of the sentence has been heard, all dependencies are resolved and the profile reaches 0. We can observe the differences in the number of dependencies to be resolved in the subject and object relatives exemplified above. The complexity profile of (1b) in (12) is higher than that of (1a) in (11), thus predicting lower acceptability, which by hypothesis results in higher comprehension problems.
3. An experiment of relative clause production

3.1. Experimental design

In order to test the production of relative clauses in child Catalan, we ran an elicitation experiment, our version for Catalan of the elicitation task designed in the context of COST Action A33, in turn based on Novogrodski & Friedmann (2006) – results for the sixteen languages tested in the framework of COST Action A33 are in preparation (Friedmann et al. in prep.). The elicitation method is construed as a questionnaire in which we ask children about their preferences. The child is presented with situations such as: ‘There are two children. One child drinks milk, the other child drinks water’ and then is asked ‘Which child would you rather be?’. The target answer involves a relative clause: ‘I’d rather be the child who drinks milk’. Children were asked to start their answers with ‘I’d rather be the child…’; the experimenter could repeat the question at the request of the child, and suggest that the child starts with ‘I’d rather be the child’, but was not to mention the relative pronoun. This elicitation method is exemplified for subject relatives and object relatives for Catalan in (13). The prompts crucially induced the production of relative clauses, and the method proved to be very effective.

(13) a. Hi ha dos nens. Un nen beu llet i un altre nen beu aigua. Quin nen t’agradaria ser?
   CL have two children A child drinks milk and another child drinks water which child
   CL like to be
   ‘There are two children. One child drinks milk and the other child drinks water. Which child would you rather be?’

b. Hi ha dos nens. Un pare abraça un nen i un pare gronxa un nen. Quin nen t’agradaria ser?
   CL have two children A father hugs a child and a father swings a child. Which child
   CL like to be
   ‘There are two children. A father hugs a child and a father swings a child. Which child would you rather be?’

There were a total of twenty items, of which ten corresponded to subject relatives and ten to object relatives. The verbs in the embedded sentence were all transitive. Of each ten subject and ten object relatives, six were reversible (both referents involved could fulfill the Agent and Theme theta role, as in (13b)), while four were irreversible (like (13a) above).
The children were all native speakers of Catalan (in particular Central Catalan) recruited in the schools Lloriana in Sant Vicenç de Torelló and Maria Borés in La Pobla de Claramunt; adult controls also came from the same areas. Details of the children and of the controls appear in (14).

<table>
<thead>
<tr>
<th>Subjects</th>
<th>#</th>
<th>Age range</th>
<th>Mean age</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-year-olds</td>
<td>20</td>
<td>5;0,11–5;11,24</td>
<td>5;5,15</td>
</tr>
<tr>
<td>Adults</td>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Children were tested individually in a quiet room in their schools. The input was not recorded, but rather produced by the experimenters, who also transcribed the answers. The responses of all subjects were coded as established in the COST A33 working group.

### 3.2. Results

The response types relevant for Catalan were, for subject relatives (i) adult-like subject relatives with a gap (15a), headless relatives (15b) and a fragment without a relative clause (15c).

(15)

a. M’agradaria ser el nen que beu llet.
   CL would-like to-be the child that drinks milk
   ‘I would like to be the child who drinks milk.’

b. M’agradaria ser el que beu llet.
   CL would-like to-be the that drinks milk
   ‘I would like to be the one who drinks milk.’

c. M’agradaria ser el de la llet.
   CL would-like to-be the of the milk
   ‘I would like to be the one of the milk.’

In the case of object relatives, the strategies that Catalan speakers can adopt are more diverse, and likewise the errors found are also diverse. Answers included: (i) adult-like object relatives with a gap (16a), (ii) relatives with a postverbal argument, possibly the subject (object relative) or the object (subject relative) (16b), (iii) object relatives with a resumptive pronoun (16c), (iv) object relatives with a reflexive resumptive (16d), (v) object relatives with a resumptive full DP (16e), and fragments and inappropriate subject relatives.

(16)

a. M’agradaria ser el nen que el pare gronxa.
   CL would-like to-be the child that the father swings
   ‘I would like to be the child that the father swings.’

b. M’agradaria ser el nen que gronxa el pare.
   CL would-like to-be the boy that swings the father
   ‘I would like to be the boy who swings the father’ or ‘I would like to be the child who the father swings.’

c. M’agradaria ser el nen que el desperta la música.
   CL would-like to-be the boy that him wakes-up the music
   ‘I would like to be the child that the music wakes up.’

d. M’agradaria ser el nen que em desperten.
   CL would-like to-be the child that REFL wake up-3pl
   ‘I would like to be the child they wake up.’

e. M’agradaria ser el nen que el veï pentina el nen.
   CL would-like to-be the boy that the neighbour combs the boy
   ‘I would like to be the child that the neighbour combs.’

While all the options encountered amongst the productions of subject relatives were well-formed, not all the productions of object relatives were so; resumptive pronouns are grammatical in colloquial Catalan, but resumptive full DP (16d) are not well-formed.

The total number of answers by the children was 400, of which only 13 were not relative clauses. The results for 5-year-olds and adults for subject and object relatives appear in (17):
(17) a. Results for subject relative production

<table>
<thead>
<tr>
<th>Type</th>
<th>5-year-olds</th>
<th>adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>headless</td>
<td>173/200</td>
<td>86.5%</td>
</tr>
<tr>
<td>fragment</td>
<td>23/200</td>
<td>11.5%</td>
</tr>
<tr>
<td>other</td>
<td>3/200</td>
<td>1.5%</td>
</tr>
</tbody>
</table>

b. Results for object relative production

<table>
<thead>
<tr>
<th>Type</th>
<th>5-year-olds</th>
<th>adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>target unambiguous</td>
<td>15/200</td>
<td>7.5%</td>
</tr>
<tr>
<td>ambiguous</td>
<td>76/200</td>
<td>38%</td>
</tr>
<tr>
<td>resumptive pronoun</td>
<td>34/200</td>
<td>17%</td>
</tr>
<tr>
<td>reflexive resumpt</td>
<td>6/200</td>
<td>3%</td>
</tr>
<tr>
<td>null subject</td>
<td>9/200</td>
<td>4.5%</td>
</tr>
<tr>
<td>subject rel</td>
<td>26/200</td>
<td>13%</td>
</tr>
<tr>
<td>DP filled gap</td>
<td>25/200</td>
<td>12.5%</td>
</tr>
<tr>
<td>passive</td>
<td>9/200</td>
<td>4.5%</td>
</tr>
</tbody>
</table>

The results for the children are graphically represented as in the following two graphs. The first graph shows how children produce subject relatives straightforwardly, and although they produce more headless relatives than adults, the pattern of production is very adult-like.

(18) Results of the production of subject relative clauses, Catalan

While at age 5 children produced 98% of subject relatives as adults do (including headed and headless relatives), they produced fewer object relatives. Only 7.5% of relatives were object relatives with a gap, and 38% corresponded to ambiguous relatives (to be compared to 19% by adults).
Object relatives with resumptive pronouns occurred in 17% of cases in the children’s production, but 53% in the adults’ (although these resumptives are considered substandard, they are widely used and identified in the literature, see Solà 2002). Children, unlike adults, produced relatives with a full DP copy of the relativised element, and unlike adults never produced a subject relative with a passive verb (lack of passives by children is hardly surprising given that passives are a late development, as shown in the literature for many languages). Answers considered erroneous in the experimental setting are those where children produced a subject relative, or an object relative clause with a null subject (a completely uninformative answer).

As we have shown, resumptive pronouns in Catalan relatives are commonly used by children and adults. Child production of resumptives has also been noted for languages which disallow it in the adult grammar, such as English and French (data from Pérez-Leroux 1995 and Labelle 1990 respectively):

(20) the one that he lifted it
(21) sur la balle qu’il l’attrappe.

Pérez-Leroux (1990) ran a relative clause elicitation experiment with Spanish- and English-speaking children (26 Spanish-speaking children, aged 3;5 to 6;8, mean age 5;3; eleven English-speaking children aged 3;5 to 5;5; mean age 4;10); she argued that resorting to resumptive pronouns was to be considered on a part with DP filled gaps in relatives, as in (22)–(23), and was found across languages in a systematic way.

(22) the one that the cowboy is pulling the horse
(23) sur la balle qu’il lance la balle

In Catalan we found resumptives with object relatives only, the same as Pérez-Leroux for English and Spanish, and in a proportion of 17%; DP filled gaps were found in 12.5% of cases; together this represents 29.5% of answers. Pérez-Leroux found that the percentage of resumptives in a broad sense (including resumptive pronouns and DP filled gaps) was: 36.2% in Spanish, 25% in English and 40.9% in French (French data from Labelle’s 1990 study); she also found the difference in the production of resumptives in the different languages not to be significant. Our results are clearly in line with those of Pérez-Leroux, and this is particularly relevant given that Catalan adults produce more resumptives than English adults (other than with relatives with such that, English speakers appear not to commonly resort to resumptive pronouns). So we can conclude that our results are consistent with
the claim by Pérez-Leroux that, in acquisition, the proportion of resumptives does not differ substantially across languages.

As for the general outcome of the experiment reported, there is a sharp contrast between subject and object relatives in their production by 5-year-olds. It remains to be seen at what age the behaviour of children attains adult levels.

4. Comprehension and production compared

The research question which we address here is: Does the production of relative clauses parallel comprehension in acquisition? As a first step, let us compare the results of comprehension available to those of production for 5-year-olds. The age range of the two groups of children does not coincide in its span, since in the comprehension experiment the age groups were of children younger than 4;6, children 4;6 to 5;6, and children older than 5;6; still, given that the pace of development in relative clause comprehension is gradual, it seems legitimate to compare a large group of 5-year-olds with two groups of children spanning 6 months in the older and younger range.

(24) Subject and object relative clause comprehension and production

a. Comprehension

<table>
<thead>
<tr>
<th>Age Range</th>
<th>SR</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>4;6–5;6</td>
<td>64/66</td>
<td>53/121</td>
</tr>
<tr>
<td>5;6&lt;</td>
<td>60/60</td>
<td>63/110</td>
</tr>
<tr>
<td>Total</td>
<td>124/126</td>
<td>116/231</td>
</tr>
</tbody>
</table>

b. Production

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-y-o</td>
<td>98%</td>
</tr>
</tbody>
</table>

Target object relatives in production here include: object relatives with a gap, relatives with a resumptive pronoun and ambiguous relatives; inclusion of this last type may overestimate the percentage of correct answers (it may include some or many disguised subject relatives). With this caveat, comprehension and production are graphically represented in (25).

(25) Percentage of correct production and comprehension of relative clauses, Catalan

The similarity between comprehension and production is notable, even more so if we take into account that production of object relatives may have been overestimated. Here we will argue that, far from being accidental, this parallelism is to be expected under our approach.
There are some differences which the graph does not portray: the way in which miscomprehension of object relatives requires that children sometimes ignore the disambiguating morphology in the embedded verb (as found also in Arosio et al. 2009 in Italian), while children never produce errors in agreement in relative production (this contrast actually extends beyond relatives: subject-verb agreement is sometimes disregarded by children in comprehension, but not in production: see Johnson et al. 2005, Pérez-Leroux 2005 for main clauses).

Gibson’s and Morrill’s proposals are neutral with respect to whether linguistic knowledge is put to use in production or comprehension: here we claim that, in fact, for the empirical domain considered, production and comprehension are equally taxed. Whether production and comprehension follow the same path is, in our view, an empirical matter, and it is well known that in many domains in language acquisition the parallelism does not hold. Here, however, we have an empirical domain where they go in parallel. It makes sense to think that if comprehension of a sentence is more taxing to the speaker when there is a higher number of syntactic dependencies that s/he must keep in mind, building such a structure would also be more taxing.

The fact that we are analysing the children’s difficulties in comprehending and producing object relatives in Catalan as a processing effect implies that the grammar that we are attributing to them is fully adult-like: we do not claim that children have any problem with relativisation per se, with wh-movement, or any other basic syntactic operation. Indeed, they do not always fail with object relatives, which is what we would expect if their grammars were immature to handle them. Briefly, we claim that processing resources may be more limited in children than in adults, as shown by the results here, and such limitations have a gradual impact on performance. Equally, in adult populations the same gradual effect of processing load can be found: Catalan-speaking adults also produced some more errors with object relatives than with subject relatives, and under pressure we would expect that to become more visible.

Recently, Friedmann, Belletti & Rizzi (2009) have also taken the stand that comprehension and production run in parallel in acquisition. Friedmann et al. ran a series of comprehension and production experiments of Hebrew subject and object relative clauses: headed subject and object relatives with and without resumptive pronouns, free subject relatives, free object relatives, object relatives with an arbitrary pro subject. The experiments were ran with Hebrew-speaking children aged 3;7 to 5;0 and they found that not all object relatives were poorly understood. The structural similarity between the moved element and the intervening subject was argued to be the source of the interpretation/production problems; when the moved element and the subject of the embedded clause were sufficiently dissimilar, performance improved significantly. Friedmann et al. capitalised in the parallelism in (26) and claimed that the case of problematic object relatives could be subsumed by Relativised Minimality.

(26) a. the boy that the monkey hugs t
   b. *How do you wonder who behaved t?

Unlike theirs, the approach proposed here does not require that processing of relatives be analysed as an instance of a Relativised Minimality violation, thereby equating acceptability with well-formedness. Rather, there is a gradation from acceptable to unacceptable which is dependent on the grammatical structure, as opposed to binary well- or ill-formedness.

Furthermore, the processing load of subject relatives and object relatives is structure dependent, and therefore may vary across languages: as mentioned earlier, prenominal relatives have been shown to behave quite differently, with subject relatives being more costly that object relatives in terms of processing; Hsiao & Gibson (2003) show in a self-paced reading task that object relatives are processed faster than subject relatives.1 In Basque, a head-final language with prenominal relative clauses, Carreiras et al. (2010) ran a self-paced reading task and an ERPs experiment, and showed that Basque subject relative clauses are not easier to process than object relatives. The analysis here extends to those cases. We will illustrate it with the relatives of Chinese.

The sentences in (27), taken from Hsiao & Gibson (2003), exemplify subject and object relatives:

---

1 Guasti (2002) references the work by Lee (1992) on the acquisition of Chinese relatives and reports that both object and subject relatives were well comprehended by Chinese speaking children at age 4 under certain circumstances.
(27)  a.  Yaoching fuhao de guanyuan shinhuaibugui danshi shanyu yintsang.
   invite tycoon gen official have bad intentions but good at hiding
   ‘The official who invited the tycoon had bad intentions but is good at hiding them.’

b.  Fuhao yaoching de guanyuan shinhuaibugui danshi shanyu yintsang.
   tycoon invite gen official had bad intentions but good at hiding
   ‘The official who the tycoon invited had bad intentions but is good at hiding.’

The derivation of the two sentences appears in (28) on the next page; the complexity profiles of the two sentences are those in (29) and (30).

(29)  Complexity profile for (27a)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td></td>
<td>a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>a</td>
<td>a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Yaoching fuhao de guanyuan shinhuaibugui…

(30)  Complexity profile for (27b)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
<td>a</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>a</td>
<td></td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fuhao yaoching de guanyuan shinhuaibugui…

In terms of the complexity metric here, the subject relative in (27a) represents a higher processing load than the object relative in (27b). The results of the psycholinguistic experimental work by Hsiao & Gibson can be thus accounted for. To our knowledge, other than Lee 1992, neither for Chinese nor for Basque are there results on the acquisition of relatives against which our predictions may be checked, and so this remains for future research.
Derivations of (27a) and (27b)
To summarise, we have presented new results on the production of relative clauses by Catalan-speaking children, and shown that the asymmetry previously found in comprehension between subject and object relatives also holds in production. We have argued that a categorial metric of processing can account for this asymmetry in a precise, non-stipulative way, and that the analysis extends to other typologically different languages.

**References**


Friedmann, Naama et al. (in prep) ‘The child who prefers not to produce object relatives: The acquisition of relative clause production in 16 languages’.


