Some notes on the Acquisition of Relative Clauses: New Data and Open Questions*

FLAVIA ADANI
Universität Potsdam, Department Linguistik
flavia.adani@uni-potsdam.de

This paper presents new experimental evidence on the relevance of feature mismatch between the head noun and the embedded noun in the comprehension of relative clauses. Exp. 1 shows that number mismatch determines an amelioration of both subject- and object-extracted relatives in English children. Exp. 2 shows that animacy mismatch per se does not play a role in German but 4-year-olds show a frequency effect, which disappears in older children and adults. The implications of these findings for recent approaches to the acquisition of relative clauses are discussed.

For Luigi,
I am still grateful for your inspiring introductory course to linguistics in 1999 - my first opportunity to jump into the realm of language acquisition- and for your guidance and advice ever since.

1. Introduction

The correct understanding/processing of relative clauses (henceforth RC) requires the computation of grammatical relations between constituents that are pronounced in a position different from the one where they are interpreted. This is shown in (1) and (2):

(1) The man that is scratching the boy …
(2) The man that the boy is scratching …

For instance, the DP *the man* is always pronounced at the beginning of the sentence in (1) and (2), but it is interpreted as the subject of the verb *scratch* in (1) and as its object in (2). Specifically, (1) is an example of a subject-extracted RC (henceforth SRC) and (2) is an object-extracted RC (henceforth ORC).

Within the generative framework (Chomsky, 2000; Rizzi, 2004) and within linguistically-oriented psycholinguistic approaches (Carminati, 2005; Franck, Lassi, Fraunfelder & Rizzi, 2006, a.o.), the necessity to satisfy feature checking

* I thank all the children who took part in the studies presented in this paper, their parents and their teachers; colleagues and students who helped me in conducting them: Maria Teresa Guasti, Matteo Forgiarini, Heather van der Lely for the experiment in English; Talea Glaw for the one in German and Jule Bergt for drawing the pictures. I am also grateful to Marinella Carminati for discussing these data in several occasions. All remaining errors are, of course, my own.

Enjoy Linguistics! Papers offered to Luigi Rizzi on the occasion of his 60th birthday

requirements is the trigger of a movement operation, by means of which the original merging position (where the constituent is interpreted) and the final landing position (where the constituent is pronounced) are connected.

Over the years, several accounts have been proposed of how children fine-tune their abilities to compute these complex dependencies. Although diverging in several respects, the earlier hypotheses (see Guasti, 2002 for a summary) shared the common idea that, up to 5 years of age, children are unable to use adult-like linguistic processes, and they have resort to a number of linear heuristics to process RCs. Since then, a lot of experimental work has been conducted and a number of theoretical approaches have been proposed to capture old and new data. In what follows, I will provide a summary of two of these recent approaches, namely the frequency/usage-based approach and the grammatical/intervention approach (section 2). Then, I will focus on some recent data which, in my opinion, raise some new questions for both approaches. In particular, I will discuss the following results:

a) The comprehension of center-embedded RCs is enhanced in English-speaking children when number dissimilarities on subject and object DPs are manipulated. This means that children display a more accurate performance when one of the nominal constituents is singular and the other one is plural (e.g. *The hippo that the rhinos are washing has climbed onto the stool*), rather than when these constituents have the same number (both singular or both plural, e.g. *The hippo that the rhino is washing has climbed onto the stool*). This facilitation effect is attested in both SRC and ORC and it is of comparable amplitude in the two structures (section 3).
b) When tested in comprehension, German-speaking 4-year-olds are significantly more accurate on ORC with inanimate heads (e.g. *The pullover that the man is scratching*) than on ORC with animate heads (e.g. *The boy that the man is scratching*). However, this effect disappears in 5-year-olds, who perform equally accurately on both conditions (section 4).

2. Two perspectives on the acquisition of RCs

According to the frequency/usage-based approach (cf. references cited in Brandt, Kidd, Lieven, & Tomasello, 2009), children's mastery of RCs is limited to structures that occur in their own spontaneous speech and which have a simple communicative function. Evidence supporting this claim is based on the analysis of early production corpora and elicited imitation data. More recently, controlled comprehension studies have been conducted (e.g. Brandt et al., 2009; Arnon, 2009). These experiments show that children are sensitive to input frequencies, i.e. they are more accurate in producing and comprehending RCs which conform to regularly occurring patterns (e.g. ORC with an inanimate head noun and/or with an embedded pronominal subject).

Moving to the grammatical/intervention approach, recent work by Grillo (2009) and Friedmann, Belletti & Rizzi (2009) proposed that the correct interpretation and production of ORC is hindered by the presence of the embedded subject DP (*the boy* in (2)). This constituent plays a role as competitor of the object DP (*the man*) in the resolution of the relevant grammatical dependency. What makes the subject DP a potential competitor is its structural similarity with the object DP: they are both full DPs (or lexically-restricted DPs, using Friedmann's et al. terminology). More recently, Belletti, Friedmann, Brunato & Rizzi (2011/under revision) refined the definition of structural similarity between the moved object and the intervening subject. The authors present cross-linguistic evidence that only Hebrew- speaking children, but not Italian-speaking children, show amelioration in the comprehension of ORC when the subject and object DPs have a different gender value. They relate

this effect to the specific morpho-syntactic properties of gender in Hebrew (i.e. it is a trigger of movement and it is overtly marked on inflected verbs).

To recap, whereas the frequency/usage-based approach mainly concentrate on the communicative function and the discourse/pragmatic factors that influence children's early linguistic knowledge and its usage, the grammatical/intervention approach focuses on the structural and linguistically-motivated processes that hinder or enhance the emergence of this knowledge.

In the next sections, I will present some findings from two experiments and a proposal on how these results could be interpreted. I hope that the discussion of these data may raise some novel questions and, possibly, suggest some point of communication between the approaches presented above.

3. Number effects in comprehension of SRC and ORC by English-speaking children

Looking at center-embedded ORC only, Adani, van der Lely, Forgiarini & Guasti (2010) found that those conditions where the subject and object DPs presented different number marking (singular vs. plural) were significantly more accurate than conditions where the two DPs differed in terms of gender (masculine vs. feminine). These results are consistent with the proposal that only features that trigger syntactic movement (and are therefore overtly spelled out on verbs) modulate intervention, as recently argued by Belletti, Friedmann, Brunato & Rizzi (2011/under review).

Using the same procedure, Adani (2008) investigated center-embedded SRC (3 A & B) and ORC (4 A & B) in a group of English-speaking typically-developing children (N=28; age range: 6;0-9;11). What distinguishes conditions A and B is the number marking on the subject and object DPs: in conditions A, both nouns are singular (i.e. matching) whereas in condition B, one is singular and the other is plural (i.e. mismatching):

(3) A. The cat that is washing the goat has climbed onto the stool.
B. The cat that is washing the goats has climbed onto the stool.
(4) A. The hippo that the rhino is washing has climbed onto the stool.
B. The hippo that the rhinos are washing has climbed onto the stool.

Mean response accuracy percentages across children in each condition are reported in the table below:

Table 1: Mean response accuracy percentages.

Sentence type	DP match	% correct
SRC	**Match**	**78.6**
	Mismatch	**88**
ORC	**Match**	**56.6**
	Mismatch	**67**

All children were significantly more accurate on SRC (3 A & B) than on ORC (4 A & B) and they were also significantly more accurate in mismatch conditions (3 & 4 B) than in match conditions (3 & 4 A).

Besides confirming the well-established subject-object asymmetry (i.e. SRC are usually easier than ORC, at least in head initial languages) and the result previously

found in Italian (i.e.–number mismatch makes ORC easier to comprehend), the English experiment revealed a facilitation effect in the SRC mismatch condition as well.

Such effect in SRC is hardly described in the literature, since SRC are usually taken as a baseline to which different types of ORC are compared (Gordon, Hendrick & Johnson, 2001, a.o.). This observation also holds for more recent studies, where the number feature was specifically manipulated in a way similar to the experiment described above (e.g. Volpato, 2009; Contemori & Marinis, 2011). Therefore, it is rarely investigated under which conditions SRC become easier to process, especially when number is manipulated. Frequency/usage-oriented approaches do not predict a SRC facilitation either (Gennari & MacDonald, 2008; Arnon, 2009; Brandt, Kidd, Lieven, & Tomasello, 2009).

Moreover, another question that arises is whether such effect may have a common underlying origin in both sentence types. In this respect, the English results just described are in apparent contrast with what was found using a similar gender manipulation (one DP masculine, the other feminine) in Hebrew (Belletti, Friedmann, Brunato & Rizzi, 2011/under revision). The effect of gender mismatch in Hebrew was significantly stronger in ORC and the authors conclude that it is the amplitude of this effect (smaller in SRC, bigger in ORC) which distinguishes between the contexts where intervention is at play (i.e. ORC) and contexts where it is not (i.e. SRC). This difference between the results of these studies raises the question whether English number marking should be treated differently from Hebrew gender marking[1] and whether the intervention approach could indeed explain the English results.

To this aim, I would like to propose an alternative explanation, which takes into account the effect of intervening constituents at the final stage of the derivation, as well as the intervening effect of traces (indicated within "< >" in 5 & 6) during the derivation of a sentence (cf. Franck, Lassi, Frauenfelder & Rizzi, 2006; Franck, Soare, Frauenfelder & Rizzi, 2010, for a similar approach to other linguistic contexts). Let us consider the derivation of a SRC, such as (5):

(5) $[_{CP}\ DP_S$ that $[_{AgrS} <DP_S> [_T$ AUX $[_{AgrO}\ DP_O [_{VP} <DP_S>$ V $<DP_O>$]]]]]

Under the assumption that an object DP (DP_O) always raises to Spec AgrOP to check its case and agreement features and that the subject DP (DP_S) originates within the VP (where thematic roles are assigned), we noticed that the critical interfering configuration holds not only for ORCs but also for SRCs, at an earlier step of the derivation. What becomes crucial here is the configuration established between the three elements indicated between the square brackets in (6):

(6) DP_S $<DP_S>$ .... [DP_O .... $<DP_S>$ ... < $DP_O>$]

However, what makes SRCs overall easier to process is the fact that only one constituent (the $<DP_S>$ instance between DP_O and its copy) intervenes during the derivation process. In contrast, in ORCs, the copy of the embedded subject intervenes within the VP (in the same way as it does in SRCs) in the initial step of the derivation, but in addition, the embedded subject DP also intervenes at the final stage of the sentence derivation. One could therefore argue that these two movement steps add up and create the well-known ORC extra computational complexity.

[1] Both features are overtly marked on nouns and verbs, with the difference that gender in Hebrew appears systematically on lexical verbs whereas the paradigm of number in English is less productive on lexical verbs but rather systematic on auxiliary verbs, thus making the test trials between the two experiments quite comparable.

More cross-linguistic research is needed to shed light on the role of grammatical features in sentence comprehension and how these effects vary depending on the properties of the individual language. In the next experiment, I examine the role of a semantic feature in RC comprehension, namely animacy.

4. Animacy effects in comprehension of ORC by German-speaking children

ORCs where the subject- and object-constituents are both animate referents (e.g. (2)) are notoriously harder to parse/understand than SRCs with two animate referents (e.g. (1)) (Gibson, 1998, a.o.). However, it has recently been shown that both children and adults find ORCs with an inanimate object and an animate embedded subject (e.g. (7)) easier to understand than ORC with two animate DPs (e.g. (1)) (Brandt, Kidd, Lieven, & Tomasello, 2009, a.o.):

(7) The pullover that the man is scratching …

This facilitation in ORC with inanimate head nouns (as well as other ORC types) has been explained in terms of input frequencies, ORCs with an inanimate head noun being more frequent in the spoken language than ORC with two animate DPs (Roland, Dick & Elman, 2007; Mak, Vonk, & Schriefers, 2002). However, it could be argued that, in (7), distinct semantic properties (+/- animate) on the subject and object make these constituents dissimilar and, thus, less prone to interference in working memory (Gordon, Hendrick, & Johnson, 2001).

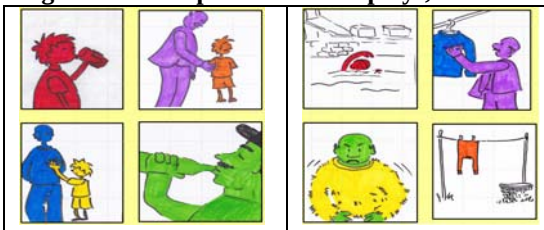
The current experiment was designed to compare these accounts by testing another type of ORCs, those with an animate object/head and an inanimate embedded subject, (e.g. 8). These are rarer than (7) in German (Mak, Vonk, & Schriefers, 2002) but are similar to (7) in that the nominal constituents differ in animacy:

(8) The man that the pullover is scratching ...

In summary, both accounts predict SRC with two animate DPs to be more accurate than ORC with two animate DPs; however, the frequency account predicts ORCs with an inanimate head (7) to be more accurate than ORCs with an animate head (8), which, in turn, would be as accurate as ORCs with 2 animate DPs (1). Under the working memory account, ORCs with an inanimate head (7) and ORCs with an animate head (8) are predicted to be equally accurate and, in turn, more accurate than ORCs with two animate DPs.

Three groups of German monolingual speakers participated in the experiment: 4-year-olds (N=19), 5-year-olds (N=19), and adults (N=13) as a control group. Participants were presented with displays on a computer screen containing 4 pictures (cf. Fig. 1a/b from Glaw, 2010; procedure adapted from Arnon, 2009):

Figure 1: A sample of visual displays, .



(a) Test pictures for (1) & (2)	(b) Test pictures for (7), (8) & (9)

Participants were asked questions about the color of some characters depicted in these pictures. To this end, each test condition (1, 2, 7, 8) and fillers (9) was embedded into a question starting with 'Which color is ... ?':

(9) The man that is swimming …

In Table 2, a summary of the experimental stimuli in German and expected answer is provided:

Table 2: Experimental conditions, English translation, expected answer.

German Stimuli	English (Which color is …?) cf. e.g.s in the text	Expected answer, cf. Fig. 1
Welche Farbe hat der Mann, der schwimmt?	**(9)**	Red
Welche Farbe hat der Mann, der den Jungen kratzt?	**(1)**	Purple
Welche Farbe hat der Pulli, den der Mann kratzt?	**(7)**	Blue
Welche Farbe hat der Mann, den der Pulli kratzt?	**(8)**	Green
Welche Farbe hat der Mann, den der Junge kratzt?	**(2)**	Blue

A trial was scored as correct if the participant uttered the correct color. Proportion of correct responses is reported in Table 3:

Table 3: Mean correct answer percentages, in each condition for each age group.

	4-year-olds	5-year-olds	Adults
Fillers	**100**	**99.3**	**99.5**
SRC with 2 animate DPs	**81.5**	**89.4**	**100**
ORC with inanimate head	**61.8**	**43.4**	**96.1**
ORC with animate head	**48.6**	**47.3**	**98**
ORC with 2 animate DPs	**40.7**	**40.7**	**100**

Adults performed at ceiling in all conditions, and so did children on the fillers. We fitted a linear mixed model with Participants and Items as random factors and with Age Group and Condition as ordered fixed factors. This analysis revealed that SRC are overall more accurate than all ORC types, with adults > 5 y.o. = 4 y.o. Although there was no significant difference between the three ORC types, there was a significant interaction of Age by Condition, showing that, only for 4-year-olds, ORC with inanimate heads are significantly more accurate than ORC with animate heads.

By collecting cross-sectional data in different age groups, we were able to identify age-specific performance patterns in the comprehension of ORCs. These results do not fall straightforwardly within any of the two accounts that were compared in this experiment. Specifically, the memory account does not seem to capture the findings from either the children or the adults. However, the data of the 4 year-olds are in line with the predictions of the frequency account, suggesting that children at this age are relying on input frequencies to parse ORCs. However, the facilitation in ORCs with an inanimate head (wrt ORCs with an animate head and ORCs with two animate DPs) disappears in 5 year-olds, where all ORC types are found to be equally hard.

These data seem to indicate that animacy dissimilarities *per se* do not enhance the comprehension of ORCs either in children or in adults, at least as far as off-line comprehension processes are concerned. It is an open question whether an animacy dissimilarity effect would be detectable during the on-line processing of the same structures. Furthermore, these data revealed that 4-year-olds' ability to comprehend

ORCs correctly is enhanced when the unfolding structure conforms to distributional patterns that occur frequently in the input. Whether this effect also holds in younger children and whether it persists in on-line processing are intriguing questions which remain to be addressed. However, between 4- and 5-years of age, children seem to start processing ORCs by relying on qualitatively different processes and this shift is already detectable using an off-line measure. Five-year-olds do not rely on input frequencies as heavily as 4-year-olds do but, arguably, they fully use adult-like structural-based linguistic processes. Again, it is possible that the effects that were detected off-line might emerge at an earlier age using on-line methods.

Importantly, in this preliminary study the frequency account and the grammatical account were not directly compared, but this experimental design can be used to compare the effect of frequency and the effect of grammatical features systematically, within the same experiment. In my opinion, this is the kind of study which is needed to disentangle the specific contribution and/or potential interactions of frequency and grammar during language acquisition and its development.

References

Adani, F. (2008). *The role of Features in Relative Clause Comprehension: a study of typical and atypical development.* Unpublished PhD dissertation University Milano-Bicocca.

Adani, F., van der Lely, H. K. J., Forgiarini, M., & Guasti, M. T. (2010). Grammatical Feature Dissimilarities Make Relative Clauses Easier: a Comprehension Study with Italian Children. *Lingua, 120*(9), 2148-2166.

Arnon, I. (2009). Rethinking child difficulty: The effect of NP type on children's processing of relative clauses in Hebrew. *Journal of Child Language, Forthcoming*(-1), 1-31.

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2011). *Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian*.Unpublished manuscript.

Brandt, S., Kidd, E., Lieven, E., & Tomasello, M. (2009). The discourse bases of relativization: An investigation of young German and English-speaking children's comprehension of relative clauses. *Cognitive Linguistics, 20*(3), 539-570.

Carminati, M. N. (2005). Processing reflexes of the Feature Hierarchy (Person>Number>Gender) and implications for the linguistic theory. *Lingua*(115), 259-285.

Chomsky, N. (2000). Minimalist Inquiries. In R. Martin, D. Michaels & J. Uriagereka (Eds.), *Step by step: Minimalist syntax in honor of Howard Lasnik* (Vol. 3, pp. 89-155). Cambridge: MIT Press.

Contemori, C., & Marinis, T. (2012). Real-time processing of number mismatch in relative clauses by English-speaking children and adults. In S. Stavrakaki, X. Konstantinopoulou & M. Lalioti (Eds.), *Advances in Language Acquisition*: Cambridge Scholars Publishing.

Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition 101*(1), 173-216.

Franck, J., Soare, G., Frauenfelder, U. H., & Rizzi, L. (2010). Object interference in subject–verb agreement: The role of intermediate traces of movement. *Journal of Memory and Language, 62*(2), 166-182.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized Relatives. Types of intervention in the acquisition of A-bar dependencies. *Lingua, 119*, 67-88.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminancy in object relative clauses. *Journal of Memory and Language, 58*, 161-187.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*, 1-76.

Glaw, T. (2010) *What factors are responsible for the facilitating effects found in the processing of object relative clauses?* Unpublished BSc Thesis, Univeristy of Potsdam.

Gordon, P., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*(27), 1411-1423.

Grillo, N. (2009). Generalized Minimality: Feature impoverishment and comprehension deficits in agrammatism. *Lingua, 119*, 1426-1443.

Guasti, M. T. (2002). *Language Acquisition: the growth of grammar*. Cambridge: MA: MIT Press.
Mak, W. M., Vonk, W., & Schriefers, H. (2002). The Influence of Animacy on Relative Clause Processing. *Journal of Memory and Language, 47*(1), 50.
Rizzi, L. (2004). Locality and left periphery. In A. Belletti (Ed.), *Structure and Beyond* (pp. 223-251). New York: Oxford University Press.
Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language, 57*, 348-379.
Volpato, F. (2010). *The acquisition of relative clauses and phi-features : evidence from hearing and hearing-impaired populations.* Università Ca' Foscari di Venezia.