The emergence of relative clauses in early child language

Holger Diessel
University of Jena

Introduction

This paper examines the development of relative clauses in early child language. It is argued that relative clauses constitute a network of related constructions that children acquire in a piecemeal bottom-up way, starting with relative clauses that are only little different from simple sentences which are gradually extended into more complex grammatical patterns. The acquisition process is driven by pragmatic and cognitive factors that are involved in the process of language use.

The analysis draws on previous research with Michael Tomasello supplemented by a new corpus investigation of children’s spontaneous relative clauses in English (cf. Diessel and Tomasello 2000, 2005; Diessel 2004, 2008; Brandt, Diessel, and Tomasello 2007). The paper reports the results of three studies. The first study is a corpus investigation of the external properties of children’s early relative clauses; the second study is an experimental study investigating the way children process the internal structure of English and German relative clauses; and the third study is another corpus study examining the meaning of children’s subject and object relative clauses.

Study 1

Relative clauses are subordinate clauses that are embedded in complex sentences. The first study investigates the structure and meaning of the sentence in which children’s early relative clauses are embedded. In the experimental literature on the acquisition of relative clauses, children are commonly confronted with complex sentences in which the relative clause modifies the subject or object of a transitive main clause including a prototypical agent and an activity verb as in examples (1) and (2) (adopted from Tavakolian 1977).

(1) The pig jumped over the horse that pumped into the lion.
(2) The horse that kicked the cow pushed the donkey.

The relative clauses of spontaneous child language are different. As shown in Diessel (2004) and Diessel and Tomasello (2000), the vast majority of the children’s spontaneous relative clauses are attached to the predicate nominal of a copular clause or an isolated noun phrase. Extending this analysis, the current study examines the external properties of children’s spontaneous relative clauses in the transcripts of four English-speaking children from the CHILDES database (MacWhinney 2000): Adam (Brown 1973), Sarah (Brown 1973), Nina (Suppes 1974), and Abe (Kuczaj 1976). The data include 460 files of one hour recordings that occurred at regular intervals between the ages of 2;0 and 5;0.¹ Using a similar coding schema

¹ Diessel and Tomasello (2000) investigated the relative clauses of five children: Adam, Sarah, Nina, Naomi, and Peter. Since Naomi’s and Peter’s transcripts include only few relative clauses, they were excluded from the current analysis, which was supplemented by new data from Abe.
as Diessel and Tomasello (2000), I divided the children’s relative clauses into four categories: (1) SUBJ-relatives, i.e. relative clauses that are attached to the main clause subject; (2) OBJ-relatives, i.e. relative clauses that are attached to an object (or adverbial) in the main clause; (3) PN-relatives, i.e. relative clauses that are attached to the predicate nominal of a copular clause; and (4) NP-relatives, i.e. relative clauses that are attached to an isolated noun (phrase).

Overall there are 583 finite relative clauses in the data, but only a minority of them are attached to the main clause subject or object. Figure 1 shows the mean proportions of the four types of relative clauses.

As can be seen, an average of only 2.1 percent of all relative clauses are attached to the main clause subject. Note that these are center-embedded relative clauses that interrupt the associated main clause. Relative clauses that are attached to an object (or adverbial) are more frequent; overall a mean proportion of 29.3 percent of all relative clauses are of this type. However, the vast majority of the children’s relative clauses do not occur with the main clause subject or object, but are attached either to the predicate nominal of a copular clause or to an isolated noun phrase: an average of 44.3 percent are PN-relatives, i.e. relative clauses that are embedded into a copular clause, and an average of 21.8 percent are NP-relatives, i.e. relative clauses that occur with an isolated noun phrase. Two of the four children, Nina and Sarah, began to use PN- and NP-relatives before they used SUBJ- and OBJ-relatives; the two other children, Adam and Abe, began to use all relative clauses except for SUBJ-relatives at around the same age.

How do we account for the early and frequent use of PN- and NP-relatives in child language? One of the reasons why children begin to use them so early is that these relative clauses are very frequent in the ambient language. Like children, adults make common use of PN- and NP-relatives when they talk to the young children (cf. Diessel 2004: 144-6). However, in addition to input frequency there are two other factors that are relevant for the early and frequent use of PN- and NP-relatives: First, these relative clauses suit the particular communicative needs of young children, and second, they are less complex than other types of relative clauses. I will discuss the two points in turn.
PN-relatives and NP-relatives are grammatical constructions with particular communicative functions. PN-relatives function to focus the hearer’s attention on a referent in the speech situation or in the universe of discourse, providing a reference point for the information expressed in the relative clause (examples 7-9), or they occur in questions, drawing the hearer’s attention onto a referent that is characterized by the information in the relative clause (cf. examples 10-11).

(7) *MOT: What's the baby patting?  
*CHI: A cat.  
*CHI: And here's a rabbit that I'm patting.
(8) *MOT: We'll have to go to the San Francisco Zoo then and see all the animals.  
*CHI: And there's the penguins that we saw.
(9) *MOT: That's gonna be very funny tea.  
*CHI: That's the kind of tea that I'm making for them.
(10) *MOT: To the fire house or to a house that's on fire?  
*CHI: To a firehouse.  
*CHI: Is that house that's on fire?
(11) *MOT: You don't mean razor blades, you mean a razor?  
*CHI: Yeah.  
*CHI: Mommy, what is dat thing dat shaves?

NP-relatives are commonly used to answer to a previous question. An average of almost 90 percent of the children’s NP-relatives are produced in response to a content question (cf. examples 10-12), but occasionally they also occur in other contexts resuming a referent from the previous discourse (cf. examples 16).

(12) *FAT: No what did you eat?  
*CHI: Some apples that were sweet.
(13) *MOT: What are those?  
*CHI: Animals that are chasing that.
(14) *FAT: What lion face?  
*CHI: The lion face you were gonna draw.
(15) *MOT: What are those?  
*CHI: Animals that are chasing that.
(16) *MOT: What do we make in our factory?  
*CHI: We don't make nothing.  
*CHI: I a cowboy maker.  
*CHI: A cowboy who shoot makers.

In accordance with Givon’s (2008) hypothesis that the development of relative clauses is determined by their communicative function, these data suggest that children learn the use of relative clauses in the communicative interaction with their parents. PN-relatives occur in copular constructions focusing the hearer’s attention on a referent that is defined or characterized by the relative clause, and NP-relatives occur in constructions answering a content question. In both constructions, the relative clause serves to establish or to retrieve a referent in the interactive discourse between parent and child.

The early and frequent use of PN- and NP-relatives is facilitated by the fact that these constructions are less complex than other types of relative clauses. SUBJ- and OBJ-relatives are embedded in bi-clausal constructions that express a relationship between two propositions, but PN- and NP-relatives occur in complex sentences that denote only a single state of affairs.
NP-relatives occur in topicalization constructions consisting of a single clause, and PN-relatives occur in copular constructions that are ‘propositionally empty’ (cf. Lambrecht 1988); that is, copular clauses do not denote an independent situation but function to establish a referent in focus position, which is subsequently integrated into the relative clause. Thus, both constructions contain only a single proposition expressed by the relative clause.

What is more, many of the OBJ-relative clauses are embedded in complex sentences in which the matrix verb has little semantic content. Very often, children’s OBJ-relatives include a stative verb expressing possession (cf. example 17) or a (mental) state (cf. examples 18), or they consist of a perception verb in the imperative drawing the hearer’s attention to a referent in the surrounding situation (cf. examples 19). While OBJ-relatives are semantically more complex than PN- and NP-relatives, only 25 percent of the children’s OBJ-relatives occur in prototypical transitive constructions including a goal-directed activity verb and an object functioning as patient in the main clause (cf. examples 20).

(17) You have tow things that turn around. Adam 3;8
(18) I like everything you fix for me. Abe 3;5
(19) Look at this dog wags his tail. Nina 3;2
(20) I punched someone that had white hair like me. Abe 3;6

In general, children’s early relative clauses occur in constructions that are low on the transitivity scale (cf. Hopper and Thompson 1980; Thompson and Hopper 2001). This is reflected in the semantic role of the noun modified by the relative clause. Distinguishing the following semantic roles—agent, patient, experiencer, location, recipient, instrument, and theme—I found that an average of 83.2 percent of all relative clauses are attached to a theme; all other thematic roles are infrequent (see Figure 2), supporting the hypothesis that the main clauses of children’s early relative clauses are low in transitivity.

Figure 2. Thematic role of head

---

2 A referent was classified as a theme if it is expressed by an isolated NP, by the subject or predicate nominal of a copular clause, by the object of a mental state verb or perception verb, or by the object of a verb of saying or verb of possession.
Note that the children’s relative clauses are semantically more substantial than the main clauses. As can be seen in Figure 3, an average of 50 percent of the verbs in the relative clause denote a physical activity and only 29.1 percent denote a state (expressed by a copular or some other stative verb). This is in sharp contrast to the main clause in which stative verbs are dominant (mean of 52.7 percent) and activity verbs are relatively rare (mean of 10.6 percent).

In sum, the development of relative clauses originates from particular constructions that are similar to simple sentences in that they denote a single state of affairs. Two types of constructions are dominant: copular constructions that focus the hearer’s attention on a particular referent, and topicalization constructions that are commonly used to answer to a previous question. There are only very few relative clauses that modify an agent or patient of a transitive activity. The vast majority of the children’s relative clauses occur in constructions in which the main clause is either propositionally empty or low in transitivity.

Interestingly, similar types of relative clauses have been found in other languages. For instance, Dasinger and Toupin (1994) noticed the predominance of presentational relative constructions in the speech of Spanish- and Hebrew-speaking children, which they collected in a picture book task, and Hudelot (1980) reports that the vast majority of children’s relative clauses in French are attached to the predicate nominal of a copular clause. Moreover, Hermon (2004) argued that there are some striking parallels in the development of relative clauses in English and Indonesian: like English-speaking children, Indonesian-speaking children begin to produce relative clauses in structures that denote only a single state of affairs. Finally, Brandt, Diessel, and Tomasello (2007) investigated a large corpus of relative clauses in the speech of a German-speaking boy who began to use relative clauses in topicalization constructions consisting of the relative clause and an isolated head noun.³

³ Ozeki and Shirai (2005) have shown that relative clauses in Japanese occur in different types of constructions; they are more often attached to the main clause subject and main clause object than children’s relative clauses in English. Interestingly, Ozeki and Shirai note that early relative clauses in Japanese are only little different from adjectives: they usually include a stative verb and involve the same morphology as adjectives (Kim 1987 found similar types of relative-clause constructions in the speech of Korean-speaking children). Since adjectives express properties rather than full propositions, Diessel (2007) suggests that children’s early relative clauses in Japanese (and other East Asian languages) are similar to children’s relative clauses in English (and other European languages) in that they denote only a single state of affairs, although the source constructions are rather different. In English, relative clauses originate from structures in which the main clause is propositionally empty, whereas in Japanese, relative clauses originate from attributive constructions in which the relative clause specifies a semantic feature of the head noun. In both types of languages children begin to produce relative clauses in constructions that contain only a single proposition.
Study 2

The second study is concerned with the internal syntactic properties of relative clauses that influence the acquisition process. The internal structure of relative clauses is defined by the syntactic function of the relativized element, which can be expressed by a pronoun or gap in the argument structure. The developmental literature has concentrated on two basic types of relative clauses: subject relatives, i.e. relative clauses in which the subject is gaped or relativized, and object relatives, i.e. relative clauses in which the object is gapped or relativized (e.g. Sheldon 1974; Tavakolian 1977; Hamburger and Crain 1982; Corrêa 1995; Kidd and Bavin 2002). However, subject and object are not the only syntactic roles that can be relativized. As can be seen in examples (21) to (25), the relativized syntactic role can be the subject, the direct or indirect object, an adverbial, or a genitive attribute.

(21) The man who met the woman. Subject
(22) The man who the woman met. Direct object
(23) The man who the woman gave the book to. Indirect object
(24) The man who the woman went to. Adjunct
(25) The man whose dog bit the woman. Genitive attribute

The earliest relative clauses that English-speaking children produce are subject relatives, but direct object relatives are also quite early. In fact, two of the four children examined in Study 1, Adam and Abe, began to use subject and direct object relatives at around the same age; only Nina and Sarah produced subject relatives before object relatives. Apart from subject and direct object relatives, the children produced adverbial relatives, which are often used to modify a location, but indirect object relatives and genitive relatives did not occur in the data.

Brandt, Diessel, and Tomasello (2007) observed a similar developmental pattern in German. Examining a corpus of 783 finite relative clauses produced by a German-speaking boy aged 2;0 to 5;0, they found that subject relatives are dominant among the earliest relative clauses; but with age the proportion of direct object relatives and adverbial relatives increased. Indirect object relatives and genitive relatives did not occur in the data.

In what follows I present the result of an experimental study that sheds some light on the acquisition of the internal properties of relative clauses. The study compares the development of relative clauses in English and German, in which the formation of relative clauses involve two different strategies (cf. Diessel and Tomasello 2005). Disregarding who-relatives, English uses the gap strategy in which the relativized syntactic role is indicated by a missing element in the argument structure, whereas German uses the relative-pronoun strategy in which the relativized syntactic role is indicated by a case-marked relative pronoun at the beginning of the relative clause. Since the relativization strategies involve different processing procedures (see Diessel and Tomasello 2005 for a detailed discussion), it is a plausible hypothesis that the development of relative clauses proceeds differently in English and German.

In order to test this hypothesis, Diessel and Tomasello (2005) conducted a sentence repetition task (cf. Slobin and Welsh 1973) in which 21 English-speaking children and 24 German-speaking children repeated six different types of relative clauses: (1) transitive subject relatives, (2) intransitive subject relatives, (3) direct object relatives, (4) indirect object relatives, (5) adverbial relatives, and (6) genitive relatives. We distinguished between transitive and intransitive subject relatives because previous studies hypothesized that transitivity plays an important role in the formation and processing of relative clauses (cf. Fox 1987; see also Fox and Thompson 1990). Table 1 provides an example of each of the six test items that were used in the English and German study.
Table 1. Experimental stimuli (Diessel and Tomasello 2005)

<table>
<thead>
<tr>
<th>English</th>
<th>German</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is the girl who played in the garden yesterday.</td>
<td>Das ist der Mann, der gestern hier gearbeitet hat.</td>
</tr>
<tr>
<td>This is the girl who saw Peter on the bus this morning.</td>
<td>Das ist der Mann, der mich gestern gesehen hat.</td>
</tr>
<tr>
<td>This is the girl who the boy teased at school yesterday.</td>
<td>Das ist der Mann, den ich gestern gesehen habe.</td>
</tr>
<tr>
<td>This is the girl who Peter borrowed a football from.</td>
<td>Das ist der Mann, dem ich das Buch gegeben habe.</td>
</tr>
<tr>
<td>This is the girl who Peter played with in the garden.</td>
<td>Das ist der Mann, mit dem ich gesprochen habe.</td>
</tr>
<tr>
<td>This is the girl whose horse Peter heard on the farm.</td>
<td>Das ist der Mann, dessen Hund mich gebissen hat.</td>
</tr>
</tbody>
</table>

Intransitive subject (=S)  
Transitive subject (=A)  
Direct object (=P)  
Indirect object (=IO)  
Adverbial (=ADV)  
Genitive (=GEN)

As can be seen, the relative clauses were attached to the predicate nominal of a copular clause. We also used test sentences with transitive main clauses, but since we were especially interested in relative clauses that children commonly use in spontaneous speech, the focus was on PN-relatives. All test items were of the same length and were controlled for various semantic and pragmatic factors. Figure 4 shows the percentages of the children’s correct responses to the six types of relative clauses.

Figure 4. Correct responses (Diessel and Tomasello 2005)

<table>
<thead>
<tr>
<th>English</th>
<th>German</th>
</tr>
</thead>
<tbody>
<tr>
<td>82,70</td>
<td>82,30</td>
</tr>
<tr>
<td>59,50</td>
<td>54,70</td>
</tr>
<tr>
<td>40,50</td>
<td>32,80</td>
</tr>
<tr>
<td>31,00</td>
<td>31,50</td>
</tr>
<tr>
<td>31,50</td>
<td>21,40</td>
</tr>
<tr>
<td>2,40</td>
<td>12,00</td>
</tr>
<tr>
<td>0,50</td>
<td>0,50</td>
</tr>
</tbody>
</table>

A/S vs. P  p < 0.001  
P vs. IO  p > 0.173  
P vs. ADV  p > 0.169  

A/S vs. P  p < 0.001  
P vs. IO  p > 0.061  
P vs. ADV  p < 0.001
As can be seen, in both studies subject relatives (S- and A-relatives) caused fewer errors than direct object relatives (P-relatives), which in turn caused fewer errors than indirect object relatives (IO-relatives) and adverbial relatives (ADV-relatives); genitive relatives (GEN-relatives) were almost always incorrect. The overall results are similar for English and German. Where the two studies differ is in the domain of object and adverbial relatives. The English-speaking children basically produced the same number of errors in response to these relative clauses (i.e. the differences were not significant); but the German-speaking children had significantly fewer problems with direct object relatives than with indirect object relatives and adverbial relatives. In particular, the adverbial relatives caused many more problems in the German study than in the English study.

How do we interpret these data? Let me begin with the subject relative clauses. Why did subject relatives cause little problems? What makes them so easy? In order to answer this question, we have to look at the errors in the children’s responses.

One of the most striking outcomes of this study was that both English- and German-speaking children made one very common type of mistake: they often converted object and adverbial relatives to subject relatives. The English-speaking children converted them by changing the word order (cf. example 26), and the German-speaking children converted them by changing the case role of the relative pronoun (and other case markers in the relative clause) (cf. example 27).

(26) TEST ITEM: This is the girl who the boy teased at school this morning.
CHILD: This is the girl that teased … the boy … at school this morning.

(27) TEST ITEM: Da ist der Mann, den das Mädchen im Stall gesehen hat.
CHILD: Da ist der Mann, der das Mädchen im Stall gesehen hat.

But interestingly, children were not consiste nt in making this type of error. Sometimes they converted a given relative clause, and sometimes they repeated the clause correctly. What is more, the children often noticed that they had made a mistake and repaired the conversion error before the end of the sentence (cf. examples 28-29), suggesting that at least some of the children were able to produce object and adverbial relative clauses correctly despite the fact that they often changed them to subject relatives.

(28) This is the girl who bor/ Peter borrowed a football from.
(29) Da ist der Junge, der/ dem Paul … die Mütze weggenommen hat.

These data suggest that the bulk of the conversion errors did not result from a lack of grammatical knowledge. But how then do we account for the errors? I suggest that the conversion errors are primarily due to the fact that subject relatives are more easily activated than other types of relative clauses.

One of the factors determining the ease of activation is frequency of occurrence: the more frequently a grammatical construction occurs, the more deeply entrenched it is in mental grammar, and the easier it is to active in language use (cf. Bybee 2006; Bybee and Hopper 2001). Thus, one might hypothesize that subject relatives are more easily activated than other types of relative clauses because they are more frequent.

However, if we look at children’s spontaneous relative clauses we find that while some children begin to use subject relatives before object relatives, older children make common uses of both types of relative clauses. In fact, two of the English-speaking children examined in Study 1, Adam and Abe, used object and adverbial relatives eventually more frequently than subject relatives. What is more, there is no evidence that subject relatives are more frequent in

---

4 The difference between direct and indirect object relatives is only marginally significant (see Figure 4).
the ambient language than object relatives. Diessel (2004) examined the relative clauses of four English-speaking mothers from the CHILDES database. In his data, more than 50 percent of the mothers’ relative clauses are direct object relatives and only 35.6 percent are subject relatives (the rest are adverbial relatives), suggesting that input frequency alone does not explain why subject relatives are so easily activated. But what then accounts for the ease of activation?

I suggest that children tend to activate subject relatives more easily than other types of relative clauses because subject relatives denote the actor (or agent) prior to any other thematic role. In fact, when subject relatives are attached to the predicate nominal of a copular clause, they are only little different from simple sentences: they involve the same order of grammatical relations than simple main clauses, whereas all other relative clauses express the object before the subject (cf. Bever 1970). This does not explain why intransitive subject relatives caused fewer errors than transitive subject relatives, but it provides an explanation for children’s good performance on subject relatives, which has also been observed in many other studies (cf. de Villiers et al. 1979; Tavakolian 1977; Clancy et al. 1986; Hamburger and Crain 1982; Corrêa 1995; Kidd and Bavin 2002). I will come back to the difference between transitive and intransitive subject relatives in Study 3 and will now concentrate on the four other types of relative clauses.

I begin with the English study. Why did the English-speaking children basically produce the same amount of errors in response to direct object relatives, indirect object relatives, and adverbial relatives? Given that direct object relatives are more frequent than adverbial relatives and that indirect object relatives are basically absent from the ambient language, one would expect that direct object relatives cause fewer errors than indirect object relatives and adverbial relatives, but the differences between these three types of relative clauses is insignificant (see Figure 4). I suggest that these three types of relative clauses basically caused the same amount of errors in the English study because they involve the same word order, which is essential for the formation of relative clauses in English. As can be seen in (30), direct object relatives, indirect object relatives, and adverbial relatives include the same sequence of constituents (i.e. … NP NP V …), which contrasts with constituent order in subject and genitive relatives.

(30) NP [V …]_{REL} subject
    NP [NP V …]_{REL} direct object
    NP [NP V …]_{REL} indirect object
    NP [NP V …]_{REL} adverbial
    NP [[GEN N] V …]_{REL} genitive

Note that in German object and adverbial relatives do not form a natural class. Each relative clause is marked by a different case form of the relative pronoun, so that structural similarities between object and adverbial relatives cannot affect the children’s performance. Since direct object relatives are more frequent than indirect object relatives and adverbial relatives, the German-speaking children had significantly fewer problems with direct object relatives than with the two other types of relative clauses.

Note, however, that adverbial relatives caused more problems than indirect object relatives although the latter are basically absent from the ambient language. I suggest that the German-speaking children had particular difficulties with adverbial relatives because these relative clauses are structurally very different from all other types of relative clauses in German: they include a preposition before the relative pronoun whereas all other relative clauses, including indirect object relatives, begin with the relative pronoun.

(31) der Mann, der … subject
    der Mann, den … direct object
Finally, we have to ask why genitive relative clauses were almost always incorrect. One of the reasons why children had great difficulties with genitive relatives may be that genitive relatives do not occur in the ambient language; but input frequency alone cannot account for children’s poor performance on genitive relatives because indirect object relatives caused significantly fewer problems than genitive relatives despite the fact that both types of relative clauses are basically absent from the ambient language. Both genitive and indirect object relatives are extremely infrequent in the input, but children had fewer problems with indirect object relatives than with genitive relatives because genitive relatives are very different from all other types of relative clauses: they involve a different semantic link between the head and the relativized element and their constituent structure is completely different.

To summarize, there are various factors influencing the acquisition of the internal properties of relative clauses. One important factor is the frequency of the various types of relative clauses in the ambient language. As we have seen, certain types of relative clause, notably subject and object relatives, are very frequent, whereas other types such as genitive and indirect object relatives are extremely rare. This is part of the reason why children had fewer difficulties with subject and object relatives than with other types of relative clauses. However, input frequency alone does not account for the data. In addition to frequency, there is another general factor that plays a key role in this study, namely the similarity (or relationship) between the various types of constructions:

- Subject relatives caused the fewest problems because they are similar to simple sentences, which children learn before they begin to produce relative clauses.
- English object and adverbial relatives caused basically the same amount of problems because they have the same word order.
- Indirect object relatives caused fewer problems than genitive relatives despite the fact that both types of relative clauses are basically absent from the input to preschool children because indirect object relatives are similar to other types of relative clauses.
- And genitive relatives and German adverbial relatives caused tremendous problems because they are very different from all other types of relative clauses.

Why is similarity so important? It is important because relative clauses are grammatical constructions, i.e. form-function pairings, that are related to each other in an associative network like lexical expressions (cf. Goldberg 1995, 2006; see also Diessel 2004: chap 2). Children acquire this network in a piecemeal, bottom-up fashion by relating new relative clause construction to constructions they already know. The development begins with subject relatives, which are only little different from simple sentences—they contain a single proposition and involve the same word order as simple main clauses (if they are embedded in copular constructions)—and it ends with genitive relatives that are most distinct from all other types of relative clauses.

Inspired by this research, Fitz and Chang (to appear) conducted a connectionist study in which a recurrent localist network (cf. Elman 1990) had to learn the various types of relative clauses from a training sample of simple and complex sentences. Interestingly, the model learned the various types of relative clauses in an order that reflects the children’s difficulty in the above experiment; that is, S-relatives were mastered before A-relatives, which in turn were learned before P-; IO-, and OBL-relatives (GEN-relatives were not included in the study). One of the factors determining the network’s performance was input frequency; but in accordance with the Diessel and Tomasello study, the network’s performance was also affected by the similarity between constructions. Manipulating the constructions in the training sample, Fitz
and Chang observed that the network’s performance on relative clauses varied with the types of simple (and complex) sentences to which the model was exposed during training, suggesting that the emergence of a particular type of relative clause is determined by its similarity to simple sentences and other types of relative clauses. Specifically, Fitz and Chang argued that it is the frequent occurrence of the fragment ‘THAT VERB’ as opposed to ‘THAT ARTICLE NOUN’ that facilitated the emergence of subject relative clauses.

Study 3

In accordance with much previous research, Study 2 showed that children have fewer difficulties with subject relatives than with object relatives. The same asymmetry between subject and object relative clauses has been found in numerous experimental studies in adult psycholinguistic (e.g. Wanner and Maratsos 1978; Frauenfelder, Segue, Mehler 1980; Holmes and O’Regan 1981; Ford 1983; MacWhinney and Pleh 1988; King and Just 1991; Just and Carpenter 1992; Cohen and Mehler 1996; Waters and Caplan 1996). What all of these studies suggest is that adult speakers find object relative clauses more difficult to process than subject relative clauses. However, recent research has shown that the processing difficulty of object relatives is crucially affected by semantic and pragmatic factors that have been ignored in older studies (Trueswell et al. 1994; Traxler et al. 2002, 2005; Warren and Gibson 2002, 2005; Mak et al. 2002, 2006; Reali and Christiansen 2007; Gennari and MacDonald in press).

Second, several experimental studies have shown that the processing difficulty of an object relative clause is affected by the type of subject it includes. For instance, Warren and Gibson (2002) found that object relative clauses including a first or second person pronoun as subject (i.e. I, you, or we) have shorter reading times than object relative clauses including a proper name, which in turn have shorter reading times than object relatives including a lexical subject, especially when the subject is indefinite (see also Warren and Gibson 2005). Warren and Gibson argue that the NP type of the subject influences the processing of object relative clauses because it correlates with the accessibility of the referent (cf. Ariel 1990; see also Givón 1983). Other things being equal, the higher the subject on the accessibility scale, the lower the processing load of the relative clause (see Gordon et al. 2001, 2004 and Reali and Christiansen 2007 for alternative explanations).

Building on this research, Kidd et al. (2007) conducted a sentence repetition experiment with 4-year-old English- and German-speaking children in which they manipulated the

<table>
<thead>
<tr>
<th>TEST ITEMS</th>
<th>SUBJECT</th>
<th>OBJECT</th>
<th>TYPE OF RC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 The burglars who stole the computer …</td>
<td>animate</td>
<td>inanimate</td>
<td>subject</td>
</tr>
<tr>
<td>2 The computer that the burglars robbed …</td>
<td>animate</td>
<td>inanimate</td>
<td>object</td>
</tr>
<tr>
<td>3 The burglars who robbed the occupant…</td>
<td>animate</td>
<td>animate</td>
<td>subject</td>
</tr>
<tr>
<td>4 The occupant who the burglars robbed …</td>
<td>animate</td>
<td>animate</td>
<td>object ***</td>
</tr>
</tbody>
</table>

Second, several experimental studies have shown that the processing difficulty of an object relative clause is affected by the type of subject it includes. For instance, Warren and Gibson (2002) found that object relative clauses including a first or second person pronoun as subject (i.e. I, you, or we) have shorter reading times than object relative clauses including a proper name, which in turn have shorter reading times than object relatives including a lexical subject, especially when the subject is indefinite (see also Warren and Gibson 2005). Warren and Gibson argue that the NP type of the subject influences the processing of object relative clauses because it correlates with the accessibility of the referent (cf. Ariel 1990; see also Givón 1983). Other things being equal, the higher the subject on the accessibility scale, the lower the processing load of the relative clause (see Gordon et al. 2001, 2004 and Reali and Christiansen 2007 for alternative explanations).

Building on this research, Kidd et al. (2007) conducted a sentence repetition experiment with 4-year-old English- and German-speaking children in which they manipulated the
animacy of the head and the NP-type of the subject. In accordance with the literature in adult psycholinguistics, they found that an inanimate head and a pronominal subject reduce the children’s difficulties with object relative clauses.

In what follows I report the results of a corpus investigation examining the semantic features of subject and object relative clauses in the speech of two English-speaking children from the CHILDES database, Adam and Abe (see above). Adam’s corpus includes a total of 178 finite relative clauses, and Abe’s corpus consists of 305 finite relative clauses. The study is limited to these two children because the transcripts of the two other children, Nina and Sarah, did not include enough relative clauses to investigate the correlations between semantic and syntactic features in their data.

In a first step, I examined the relationship between the relativized syntactic role and the animacy of the noun that is modified by the relative clause. As can be seen in Table 3, both categories were coded as dichotomous variables. Specifically, I distinguished between subject and non-subject relatives and animate and inanimate nouns.

Table 3. Frequencies of relativized role and animacy of the head

<table>
<thead>
<tr>
<th></th>
<th>RELATIVIZED ROLE</th>
<th>ANIMACY OF THE HEAD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subject</td>
<td>Non-subject</td>
</tr>
<tr>
<td>Adam</td>
<td>62 (35.4%)</td>
<td>113 (64.5%)</td>
</tr>
<tr>
<td>Abe</td>
<td>130 (42.6%)</td>
<td>175 (57.4%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>39.1%</td>
<td>60.9%</td>
</tr>
</tbody>
</table>

Note that the vast majority of the children’s relative clauses are attached to an inanimate noun. Overall, an average of 78.1 percent of their relative clauses modify an inanimate noun and only 21.9 percent occur with an animate noun. Pre-examination of the data revealed no significant difference between the two children, suggesting that they basically produced the same types of relative clauses. Figure 5 shows the relationship between the animacy of the head and the relativized syntactic role.

As can be seen in this figure, subject relatives are common with both animate and inanimate nouns; there is only a small difference between them. But non-subject relatives (i.e. object and adverbial relatives) are much more frequent after inanimate nouns: overall more than 90
percent of the non-subject relatives modify an inanimate noun. A \( \chi^2 \)-test for independence revealed a significant association between the animacy of the head and the syntactic function of the relativized role, suggesting that the semantic feature of animacy is an important determinant of the children’s spontaneous relative clauses (\( \chi^2 = 75.15; df = 1; p < 0.001 \)).

Interestingly, the majority of the children’s non-subject relatives include a transitive verb; intransitive verbs occur only in some of the adverbial relative clauses (e.g. That’s the pumpkin that I was standing next to), whereas subject relatives are mostly intransitive: 66.7 percent of the non-subject relatives include an intransitive verb and only 33.3 percent include a transitive verb (cf. Diessel 2004).\(^5\) Interestingly, transitive and intransitive subject relatives are headed by different semantic types of nouns. As can be seen in Figure 6, while transitive subject relatives are slightly more frequent after animate nouns, intransitive subject relatives are much more frequent after inanimate nouns (\( \chi^2 = 16.29; df = 1; p < 0.001 \)).

![Figure 6. Animate and inanimate heads of transitive and intransitive subject relatives](image)

Note that about one third of the intransitive subject relatives are copular clauses including the copula be (Some apples that were sweet; Abe 3,6); but even if we disregard copular clauses, 67 percent of the intransitive subject relatives are attached to an inanimate noun. If we consider the intransitive subject relatives more closely we find a correlation between the meaning of the head and the meaning of verb: while unergative verbs occur with both animate and inanimate nouns (cf. examples 32 and 33), unaccusative verbs are exclusively used with inanimate nouns. Note that most of the unaccusative verbs are transitives in the passive (cf. example 34); true unaccusative verbs are rare (cf. example 35).

\[
\begin{align*}
(32) & \quad \text{The doggie that runs away.} \quad [\text{Adam 3,8}] \\
(33) & \quad \text{Look at that big truck going some place.} \quad [\text{Adam 3,0}] \\
(34) & \quad \text{No that one that couldn’t be snapped.} \quad [\text{Abe 3,6}] \\
(35) & \quad \text{I take the ones that fall out.} \quad [\text{Adam 4,0}]
\end{align*}
\]

Animacy is an important ontological category that is often reflected in linguistic structure (cf. Comrie 1989); but animacy is not the only semantic features that correlates with structural properties in children’s relative clauses. There are other, more fine-grained semantic categories that vary with the relativized syntactic role. In a second step, I divided animate and

\(^5\) The predominance of intransitive subject relative clauses has also been observed in adult language (Fox and Thompson 1990).
inanimate nouns into several semantic subclasses. Animate nouns were divided into humans and animals, and inanimate nouns were divided into things, machines, and locations. The five semantic categories of the head were crossed with four relativized syntactic roles: (i) transitive subject (i.e. A), (ii) intransitive subject (i.e. S), (iii) direct object (i.e. P), and (iv) adverbial (i.e. ADV). Figure 7 shows the proportion of the various structural types of relative clauses after different semantic types of nouns.

As can be seen in this figure, after human referents, subject relatives are predominant; after animals and machines, subject and object relatives are about equally frequent; after things (including abstract entities) object relatives are dominant; and after place nouns the relativized syntactic role typically functions as an adverbial.

How do we account for these relationships? Why are different structural types of children’s relative clauses associated with different semantic roles? I suggest that the semantic biases in children’s spontaneous relative clauses reflect the prototypical link between grammatical relations and semantic roles. One can think of the relationship between syntactic functions and semantic roles as an associative network that emerges from children’s experience with language: the more often a semantic role is expressed by a particular syntactic category, the stronger the associative link between form and meaning (cf. Figure 8).
Since the associations are largely independent of the clause type, it is a reasonable hypothesis that children acquire this network before they begin to produce relative clauses based on their experience with simple sentences (cf. Diessel and Tomasello 2005). When they begin to produce relative clauses, the network is so deeply entrenched that it is automatically transferred to complex sentences: a human referent in the main clause is associated with the subject in the relative clause; a noun denoting an object or thing in the main clause is automatically interpreted as the object in the relative clause; and a locative expression is automatically linked to an adverbial. This explains why children and adults have fewer difficulties with object relative clauses that are attached to inanimate nouns than with object relatives that modify animate nouns.

Note, however, that the association between grammatical relations and semantic roles is skewed in subject and object relative clauses. Disregarding intransitive subject relatives with a single referent, I examined the animacy features of subject and object (or adverbial) in children’s relative clauses with two nominal referents. Figure 9 shows that subject and non-subject relative clauses include very different pairings of animate and inanimate nouns.

![Figure 2. Animacy of subject and object (or adverbial) in subject and non-subject relatives](image)

While subject relatives occur with various combinations of animate and inanimate nouns, non-subject-relatives are strongly skewed in favour of one particular type: 87 percent of the children’s non-subject relative clauses contain an animate subject and an inanimate object (or adverbial); all other types are infrequent, suggesting that subject and object relatives tend to denote different types of situations. Specifically, object relatives denote situations in which the subject is higher on the animacy hierarchy than the object (or adverbial), whereas subject relatives are commonly used with two nominal referents that are equal in terms of their animacy features.

---

6 In relative clauses with more than two referents I concentrated on the core roles, i.e. subject and object, and disregarded adverbials.
Moreover, the two types of relative clauses occur with different types of subjects. The vast majority of the children’s non-subject relative clauses include a first or second pronoun as subject. As can be seen in Figure 10, 79.0 percent of the non-subject relatives are of this type (cf. Fox and Thompson 1990, 2007); the rest occur with third person pronouns (7.8 percent), definite NPs (10.4 percent), and a few indefinite nouns (1.9 percent).

The subjects of subject relative clauses are radically different. Subject relatives do not occur with first or second person pronouns, but are primarily used with lexical nouns as subjects: 76.1 percent of the children’s subject relatives are attached to a common noun functioning as subject inside of the relative clause. Note that almost half of the nouns are indefinite, introducing a new referent into the discourse.

Thus, the two types of relative clauses occur with very different types of subjects: non-subject relatives occur with highly accessible subjects referring to the speech participants or other well-known referents (e.g. expressed by proper names), whereas the subjects of subject relatives are third person referents that often introduced a new discourse referent.

Finally, subject and non-subject relatives include different types of verbs. As pointed out above, the majority of the children’s subject relatives include intransitive verbs, notably copula verbs are very common, whereas non-subject relatives are usually transitive; only some of the adverbial relatives are intransitive. What is more, even if we exclude intransitive relative clauses, there is a remarkable contrast between subject and non-subject relatives. As can be seen in Figure 11, non-subject relatives include a much larger proportion of activity verbs than subject relatives: 59.9 percent of the non-subject relatives occur with an activity verb such as make, do, or eat, whereas subject relatives are commonly used with stative verbs such as have, own, or belong. Moreover, while non-subject relatives include many cognition, perception, and communication verbs (e.g. say, know, see, want), these verbs rare in subject relatives.
In general, the verbs of non-subject relative clauses tend to be high on the transitivity scale; they often include a prototypical transitive verb selecting an agent and patient or else denote a cognitive or verbal activity. By contrast, subject relative clauses are low in transitivity. The majority of the children’s subject relatives include an intransitive verb and those that do occur with a transitive verb (i.e. a verb with two mandatory participants) are often stative selecting a non-agentive subject and a theme as object.

How do we account for these data? I suggest that the semantic contrast between subject and non-subject relatives reflects the fact that they are used with different pragmatic functions (cf. Fox and Thompson 1990). Subject relatives are low in transitivity because they are primarily used to characterize a discourse referent, which is often newly introduced in the preceding main clause. By contrast, object relatives are high in transitivity because they are commonly used to identify (or to retrieve) a referent that is defined by its relationship to one of the speech participants or some other well-known person (or object) that grounds the referent in the universe of discourse.

Note that while subject relatives involve the same order of grammatical relations as declarative sentences, they are semantically very different from simple sentences. As can be seen in Table 4, while non-subject relatives basically occur with the same types of nouns and verbs as simple sentences, subject relatives are radically different: they include more inanimate subjects, more lexical subjects, and more stative verbs than simple sentences, suggesting that subject relatives are semantically ‘unusual’ sentences. Both types of relative clauses are grammatical constructions with particular structural properties that are paired with specific semantic and pragmatic features.
Table 4. The meaning of children’s relative clauses in comparison to simple sentences

<table>
<thead>
<tr>
<th></th>
<th>Subject relatives</th>
<th>Non-subject relatives</th>
<th>Simple sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animacy of subject</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) Animate</td>
<td>55.6%</td>
<td>92.2%</td>
<td>96.1%</td>
</tr>
<tr>
<td>(ii) Inanimate</td>
<td>44.6%</td>
<td>7.8%</td>
<td>3.9%</td>
</tr>
<tr>
<td>NP-type of subject</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) pronominal</td>
<td>24.0%</td>
<td>87.6%</td>
<td>90.3%</td>
</tr>
<tr>
<td>(ii) lexical</td>
<td>76.0%</td>
<td>12.3%</td>
<td>9.7%</td>
</tr>
<tr>
<td>Meaning of transitive verbs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) Activity</td>
<td>28.7%</td>
<td>59.9%</td>
<td>57.9%</td>
</tr>
<tr>
<td>(ii) Cognition/communication</td>
<td>5.7%</td>
<td>29.8%</td>
<td>32.7%</td>
</tr>
<tr>
<td>(iii) States</td>
<td>65.6%</td>
<td>10.4%</td>
<td>9.4%</td>
</tr>
</tbody>
</table>

Conclusion

This paper has argued that relative clauses constitute a network of constructions that children acquire in a bottom-up way. Three studies have been discussed. The first study examined the structure and meaning of the external properties of children’s early relative clauses in spontaneous speech. The study showed that the earliest relative clauses are embedded in complex sentences with particular communicative functions in which the main clause is either propositionally empty or low in transitivity. Most of the constructions including early relative clauses are similar to simple sentences in that they denote a single state of affairs. The second study was an experimental study investigating how children process the internal properties of English and German relative clauses. In accordance with much previous research, the study showed that subject relatives cause fewer difficulties than object and adverbial relatives, which in turn are easier to process than genitive relatives. Examining the children’s errors, it was argued that, apart from input frequency, the similarity between simple sentences and the various types of relative clauses is an important determinant of the acquisition process. Since subject relatives involve the same sequence of subject, verb, and object than simple sentences, they tend to cause fewer problems than other types of relative clauses, suggesting that the development of relative clauses is influenced by the children’s prior knowledge of simple sentences. Finally, the third study examined the semantic and pragmatic properties of children’s spontaneous subject and non-subject relatives. The study showed that the two types of relative clauses tend to denote different types of situations. Non-subject relatives typically include a first or second person pronoun as subject that is involved in a dynamic activity, whereas subject relatives occur with lexical subjects that are commonly embedded in intransitive clauses or transitive clauses that are low in transitivity. It was argued that the semantic differences between subject and non-subject relatives reflect differences in their pragmatic functions. While subject relatives are commonly used to describe (or to characterize) a referent, which is often newly introduced in the preceding main clause, non-subject relatives are primarily used to identify (or to define) a referent by specifying its relationship to one of the speech participants or another well-known referent.

In sum, the acquisition of relative clauses is determined by multiple factors: the ambient language, the communicative interaction between parent and child, the similarity between constructions, and the complexity of the various types of relative clause. The earliest relative clauses that English-speaking children learn occur in particular constructions that share important properties with simple sentences: they contain a single proposition, involve the same sequence of grammatical relations, and the same associations between syntactic and semantic roles. Starting from such simple structures children acquire the network of relative
clause constructions by extending constructions they know to slightly different grammatical patterns, which gradually increase in complexity. From this perspective, linguistic complexity results from many small extensions giving rise to intricate grammatical patterns that share individual properties with other grammatical patterns in the gradually emerging network of constructions.

References

Bybee, Joan 2006. From usage to grammar: The mind’s response to repetition. Language 82: 711-733.
Bybee, Joan and Paul Hopper. 2001. Frequency and the Emergence of Linguistic Structure. Amsterdam: John Benjamins
Fitz, Hartmut and Franklin Chang. To appear. The role of the input in a connectionist model of the accessibility hierarchy in development. Proceedings of the Boston University Conference on Child Development.

Fox, Barbara A. 1987. The noun phrase accessibility hierarchy reinterpreted: Subject primacy or the absolutive hypothesis. *Language* 63: 856-870.


