PART II

Child language
On the role of frequency and similarity in the acquisition of subject and non-subject relative clauses

Holger Diessel
University of Jena

Frequency and similarity are important determinants for the acquisition of children's early item-based constructions. This paper argues that frequency and similarity are equally important for the development of more complex and intricate grammatical phenomena such as relative clauses. Specifically, the paper shows that the acquisition of relative clauses is crucially determined by the similarity between particular types of relative clauses and simple SVO constructions. Two specific hypotheses are proposed: First, since subject relatives have the same word order as ordinary SVO clauses, they usually cause fewer difficulties in comprehension studies than non-subject relatives. Second, while non-subject relatives are structurally distinct from SVO clauses, semantically they are expressed by prototypical transitive constructions, which arguably helps the child to learn this type of relative clause.

1. Introduction

In the usage-based approach to language acquisition grammatical development is crucially determined by the child's experience with language. Specifically, frequency of occurrence plays an important role in this approach. Every time a child encounters a linguistic expression in the ambient language it leaves a trace in memory reinforcing its mental representation, i.e., the level of entrenchment, which in turn facilitates the activation of the expression in future language use (cf. Tomasello 2003; Diessel 2007a).

Another learning mechanism that plays an important role in the usage-based approach is analogy, which has been characterized as the core of human cognition (Hofstadter 2001). Analogy involves a mapping of information from one particular entity, the source, to another particular entity, the target. The mapping is based on the recognition of similarity, which can be divided into two basic types: substantial (or object) similarity, which involves the recognition of shared attributes, and structural similarity, which involves the recognition of shared structures or relationships (cf. Gentner 1983; Gentner & Median 1998).
There is abundant evidence that frequency and (structural) similarity play important roles in the early stages of grammatical development. For instance, Bybee and Slobin (1982) showed that the acquisition of the English past tense is determined by the child’s experience with individual verbs and the phonetic similarity between the various past tense forms, and Tomasello and colleagues demonstrated that the development of children’s early item-based constructions is crucially affected by the frequency and similarity of verb-argument constructions in the ambient language (cf. Tomasello 2000, 2003; Lieven et al. 2003).

However, grammar does not only consist of words and simple item-based constructions, it also includes more complex and abstract grammatical phenomena such as relative clauses. It is the purpose of this paper to show that frequency and similarity are not only important determinants of the early stages of grammatical development but are equally important for the acquisition of complex grammatical structures such as relative clauses. Specifically, the paper argues that the acquisition of subject and non-subject relatives is determined by the frequency of the various relative clauses in the ambient language and their similarity to simple sentences.

The paper is divided into two parts. The first part considers the results of an experimental study that I conducted together with Michael Tomasello (cf. Diessel & Tomasello 2005). The study shows that children’s comprehension and production of relative clauses is influenced by the structural overlap between the various types of relative clauses and simple sentences. Specifically, the study suggests that subject relatives cause fewer difficulties than non-subject relatives because they involve the same constituent order as ordinary (in)transitive clauses. The second part reports the results of a new corpus study that examines the meaning and use of subject and non-subject relatives in spontaneous child language. The study shows that while non-subject relatives are structurally distinct from simple sentences, semantically they are expressed by prototypical transitive clauses, whereas subject relatives comprise a variety of constructions that do not always match the corresponding properties of simple (in)transitive clauses. The paper argues that the structural and semantic overlap of children’s subject and non-subject relatives with simple sentences helps the child to bootstrap into the intricate grammar of relative clauses. On this view, grammatical development is an incremental process whereby children acquire new grammatical constructions based on structures they already know.

Before I turn to these studies, I will briefly consider the external properties of children’s spontaneous relative clauses providing the background for the subsequent analyses.

2. The external properties of children’s spontaneous relative clauses

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 patient and a verb denoting a physical activity as in examples (1) and (2) (adopted from Tavakolian 1977).

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

The main clauses of children’s early relative clauses are very different. As shown in Diessel (2004) and Diessel and Tomasello (2000), the development of relative clauses originates from particular constructions that are much less complex than the relative-clause constructions that have been used in most experiments. Investigating corpus data from four English-speaking children between the ages 2.0 to 5.0, they found that most of the children’s early relative clauses are attached to the predicate nominal of a copular clause (cf. example 3) or, less frequently, to an isolated noun phrase (cf. example 4).

(3) Those are bugs [that I throw].
(4) This thing [you were watching].

In both types of constructions the relative clause functions to elaborate a referent in a pragmatically marked position. Relative clauses that are attached to the predicate nominal of a copular clause modify a referent that is presented in focus position of the copular clause (cf. examples 5–6), and relative clauses that are attached to an isolate noun phrase are commonly used to characterize a referent that is presented in response to a content question (cf. examples 7–8). Both types of constructions emerge in the communicative interaction between parent and child. They are associated with particular discourse patterns that constrain their early structure and use (cf. Diessel 2004: chap 6; see also Givón 2009).

(5) *CHI: Mommy … I want to eat now and then go to the park, okay?  Abe 3;5
   *CHI: Daddy … this is part of the egg [I ate for breakfast].
(6) *CHI: Hey … Paul … come and look at this machine.  Adam 4;3
   *CHI: Dis is a new machine [dat Paul likes].
(7) *FAT: No what did you eat?  Abe 3;6
   *CHI: Some apples [that were sweet].
(8) *MOT: What is that?  Adam 3;9
   *CHI: Place [where you put your things when you eat dem].

What is more, Diessel and Tomasello argue that although relative clauses are embedded clauses, initially they appear in constructions that are not really bi-clausal. Relative clauses that are attached to an isolate noun phrase do not occur with a complete main clause, and relative clauses that are attached to the predicate nominal of a copular clause occur with a main clause that is “propositionally empty” (Lambrecht 1988): Although...
the copular clause includes a verb it does not denote an independent state of affairs but functions to present a referent in focus position such that it can be elaborated by the relative clause. In other words, the development of relative clauses originates from constructions that contain only a single proposition. Starting from such simple sentences, children gradually acquire more complex relative-clause constructions in which main and relative clauses express two separate states of affairs (cf. Diessel 2004; Diessel & Tomasello 2000, 2005).

Interestingly, similar developmental pathways have been observed in several 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 early relative clauses in French are attached to the predicate nominal of a copular clause. Moreover, Hermon (2004) argued that there are 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 (2008) 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.1

In what follows I report the results of two studies that are concerned with the internal properties of children's relative clauses. The first study is an experimental study that was designed to examine the effect of the relativized role on the processing and acquisition of relative clauses (cf. Diessel & Tomasello 2005); the second study is an observational study that is concerned with the meaning and use of different structural types of relative clauses in spontaneous child language.

1. Ozeki and Shirai (2005) showed that children's early 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 (cf. Kim 1989, who found similar types of relative-clause constructions in the speech of Korean-speaking children). Since adjectives express properties rather than full propositions, Diessel (2007b) suggests that children's early relative clauses in Japanese (and other East Asian languages) are not really bi-clausal. Like English-speaking children, Japanese-speaking children begin to use relative clauses in particular constructions that denote only a single state of affairs. In English, relative clauses originate from focus and topicalization constructions in which the main clause is propositionally empty, and 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.

3. Study 1

The experimental literature on the acquisition of relative clauses has concentrated on two structural types: Subject relatives, i.e., relative clauses in which the subject is relativized, and object relatives, i.e., relative clauses in which the direct object is relativized (e.g., Sheldon 1974; Tavakolian 1977; Hamburger & Crain 1982; Corrêa 1995; Kidd & Bavin 2002). However, subject and object are not the only syntactic roles that can be relativized. As can be seen in (9) to (13), the relativized syntactic role can be the subject, the direct or indirect object, an adverbial, or a genitive attribute.

(9) The man who met the woman. subj-relative
(10) The woman who the man met. obj-relative
(11) The boy who the girl gave the ball to. io-relative
(12) The girl who the boy played with. adv-relative
(13) The man whose cat caught a mouse. gen-relative

In subject relatives, the relativized role is indicated by a relative pronoun (or particle) that immediately precedes the verb of the relative clause; but in object and adverbial relatives the relative pronoun/particle can be omitted and the relativized syntactic role is expressed by a 'gap' in the argument structure. For instance, the relative clause in (14) includes a gap after the preposition at the end of the sentence.

(14) The thing [I was looking for __].
indicates the relativized role right at the beginning of the relative clause, as in the following examples from German.

(15) Der Mann, der mich gesehen hat. subj-relative
(16) Der Mann, den ich gesehen habe. obj-relative
(17) Der Mann, dem ich das Buch gegeben habe. io-relative
(18) Der Mann, zu dem ich gegangen bin. adv-relative
(19) Der Mann, dessen Hund mich gebissen hat. gen-relative

Since the relative pronoun is the first word in the relative clause, the processor does not have to keep unintegrated information in working memory while processing the relative clause. As soon as the processor encounters the relative pronoun, the filler can be integrated into the relative clause, suggesting that the processing of German relative clauses involves a different strategy than the processing of English relative clauses, which may affect the acquisition process.

3.1 Methods

In order to test this hypothesis Michael Tomasello and I conducted a sentence repetition experiments in which 4-to-5 year old English- and German-speaking children had to repeat complex sentences including one of six different structural types of relative clauses:

1. Intransitive subject relatives, i.e., S-relatives (e.g., the man who slept on the couch).
2. Transitive subject relatives, i.e., A-relatives (e.g., the man who saw the woman).
3. Direct object relatives, i.e., P-relatives (e.g., the man who the woman met).
4. Indirect object relatives, i.e., io-relatives (e.g., the man who John gave the book to).
5. Adverbial relatives, i.e., adv-relatives (e.g., the man who the girl went to).
6. Genitive relatives, i.e., gen-relatives (e.g., the man whose dog chased the cat).

We distinguished between transitive and intransitive subject relatives because previous studies had shown that transitivity can affect children’s comprehension of relative clauses (cf. Hamburger & Crain 1982; Goodluck & Tavakolian 1982). The test items were all of the same length and tightly controlled for various semantic, pragmatic, and structural factors. After a warm-up phase, children had to repeat a total of 41 sentences including 24 relative clauses and 17 filler items. The sentences referred to a play situation in front of the child providing a pragmatically adequate context for the use of a (restrictive) relative clause (cf. Hamburger & Crain 1982).

3.2 Results

Figure 1 shows the mean proportions of the children’s correct responses to the various types of relative clauses.

As can be seen, in both studies subject relatives (i.e., A/S-relatives) caused fewer errors than direct object relatives (i.e., 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 in the English and German study; however, they differ 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 (cf. Figure 1); 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 (The difference between direct and indirect object relatives is only marginally significant in the German study; cf. Figure 1).

In order to understand the children’s difficulties with the various types of relative clauses, Diessel and Tomasello investigated the children’s errors. Interestingly, they found one very common type of mistake: When children were given an object or adverbial
subject relatives, and 7.9 percent are adverbial relatives; indirect object relatives and genitive relatives did not appear. Similar frequency distributions have been observed in the ambient language of a German-speaking boy (Brandt et al. 2008) suggesting that input frequency alone does not explain why subject relatives are so easily activated. But what then accounts for the ease of activation?

Diessel and Tomasello suggest that children activate subject relatives more easily than other types of relative clauses because subject relatives have the same word order as simple sentences. As can be seen in (24) to (26), in subject relatives the subject precedes all other semantic roles, whereas in object and adverbial relatives it follows the patient or some other semantic role.

This does not explain why intransitive subject relatives caused fewer errors than transitive subject relatives, but it does provide an explanation for children's good performance on subject relatives, which has also been observed in many other studies (Smith 1974; Tavakolian 1977; de Villiers et al. 1979; Hamburger & Crain 1982; Kidd & Bavin 2002).

That similarity with simple sentences plays an important role in the acquisition of relative clauses has also been proposed by Bever (1970; see also de Villiers et al. 1979), but in contrast to this work, Diessel and Tomasello argue that similarity is not only important to account for children's ease with subject relatives, it also plays an important role in the acquisition of other structural types of relative clauses. Specifically, they argue that relative clauses constitute a family of constructions that children acquire in an incremental fashion such that new relative clauses are learned based on structures the child already knows.

To begin, Diessel and Tomasello argue that the English-speaking children basically produced the same amount of errors in response to direct object relatives, indirect object relatives, and adverbial relatives because these three types of relative clauses involve the same word order. As can be seen in (27), direct object relatives, indirect object relatives, and adverbial relatives include the same sequence of constituents (i.e., ... NP NP V ...), which contrasts with the constituent order in subject and genitive relatives. This explains why the three types of relative clauses basically caused the same amount of errors although direct object relatives are much more frequent than adverbial relatives and indirect object relatives. In fact, the latter are so rare in the ambient language that children have very little experience with this type of relative clause. However, since indirect object relatives are structurally similar to direct object relatives they did not cause significantly more problems.

2. There were also a few subject relatives that were converted to object relatives, but this type of mistake was rare.

Diessel and Tomasello (2005) argue that the conversion errors occurred because subject relatives are more easily activated than other types of relative clauses. One factor determining the ease of activation is frequency of occurrence: Frequently used construction are deeply entrenched in memory and therefore easily activated in language use (cf. Bybee 2006; Bybee & Hopper 2001). However, although subject relatives are among the earliest relative clauses children produce (cf. Diessel 2004: chap 6; Brandt et al. 2008), in spoken discourse they do not seem to be more frequent than direct object relatives (cf. Fox & Thompson 1990; Roland et al. 2007). In fact, Diessel (2004: 146) reports that 4-to-5 year olds produce object relatives more frequently than subject relatives. Moreover, he notes that object relatives are dominant in the ambient language: 57.9 percent of the mothers' relative clauses in his data are direct object relatives, 34.3 percent are
• Direct object relatives, indirect object relatives, and adverbial relatives caused basically the same amount of problems in the English study because they share important word order properties with each other.
• Indirect object relatives caused fewer problems than genitive relatives despite the fact that both types of relative clauses are basically absent from the ambient language because indirect object relatives are similar to other types of relative clauses.
• And genitive relatives and German adverbial relatives caused by far the greatest problems because they are very different from all other types of relative clauses.

Why is similarity such an important factor? 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). Children acquire this network in a piecemeal, bottom-up fashion, relating new relative clauses to constructions they already know. The development starts with subject relatives in copular constructions, which are only little different from simple sentences—they contain a single proposition and involve the same word order as simple main clauses—and it ends with genitive relatives that are most distinct from all other types of relative clauses.

Inspired by this research, Fitz and Chang (2008) conducted a connectionist study in which a recurrent localist network (cf. Elman 1990) was trained to learn various types of relative clauses from a training sample of simple and complex sentences. Interestingly, the model performed in the same way as the children in the Diessel and Tomasello study. One of the factors determining the simulation was input frequency; but in addition, the network's performance was affected by the similarity between constructions. Manipulating the sentences 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, 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 relatives.

4. Study 2

The first study has shown that structural similarity is an important determinant for the processing and acquisition of relative clauses. Specifically, we have seen that children have fewer difficulties with subject relatives than with object and adverbial relatives because they have the same word order as simple sentences, whereas object and adverbial
relatives deviate from the SV(O) schema of ordinary (in)transitive clauses. However, a number of recent studies have shown that the processing load of object relatives is crucially affected by semantic and pragmatic aspects that have been ignored in earlier research on the processing and acquisition of relative clauses. Two findings are relevant.

First, several studies have demonstrated that the semantic feature of animacy is an important determinant for the comprehension of relative clauses in adult language (cf. Trueswell et al. 1994; Traxler et al. 2002, 2005; Mak et al. 2002, 2006; Gennari & MacDonald 2008). For instance, Mak et al. (2002) conducted a reading time experiment with Dutch-speaking adults in which animacy had a differential effect on the processing of subject and object relatives. Using various combinations of animate and inanimate nouns, they found no significant difference in reading times between subject and object relatives if the subject of the relative clause is animate and the object inanimate (cf. examples 29–30); it was only when both subject and object relatives were animate that object relatives caused longer reading times than subject relatives (cf. examples 31–32) (cf. Trueswell et al. 1994; Traxler et al. 2002; Mak et al. 2006; Gennari & MacDonald 2008).

(29) The burglars who stole the computer … animate-inanimate/subject-REL.
(30) The computer that the burglars robbed … animate-inanimate/object-REL.
(31) The burglars who robbed the occupant … animate/animate/subject-REL.
(32) The occupant who the burglars robbed … animate/animate/object-REL.

Second, several experimental studies have shown that the processing difficulty of an object relative is affected by the type of subject it includes. For instance, Warren and Gibson (2002) found that object relatives including a first or second person pronoun as subject (cf. example 33) have shorter reading times than object relatives including a proper name (example 34), which in turn cause shorter reading times than object relatives including a lexical subject (example 35), especially when the subject is indefinite (cf. example 36). Warren and Gibson argue that the NP type of the subject influences the processing of object relatives 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 (cf. Warren & Gibson 2005; see Gordon et al. 2001, 2004 and Reali & Christiansen 2007 for somewhat different explanations).

(33) The man I met. First person subject
(34) The man Peter met. Third person subject
(35) The man the woman met. Definite lexical subject
(36) The man a woman met. Indefinite lexical subject

Continuing this line of research, Kidd et al. (2007) conducted a sentence repetition experiment with 4-to-5 year old English- and German-speaking children in which they manipulated the animacy of the head noun and the NP-type of the subject in object relatives. In accordance with the studies in adult psycholinguistics, they found that an inanimate head and a pronominal subject reduce children's difficulties with object relatives. Like adult speakers, children find object relatives as easy (or difficult) to comprehend as subject relatives if they modify an inanimate noun and include a pronominal subject.

In what follows I will take a closer look at the meaning and function subject and non-subject relatives in spontaneous child language. The study shows that subject relatives comprise a variety of constructions including transitive, intransitive, and copular verbs that occur with various constellations of animate and inanimate nouns, whereas non-subject relatives are realized by prototypical transitive clauses. They usually contain a dynamic verb denoting a goal-directed activity and include a first or second person pronoun as subject functioning as agent or experiencer. It is argued that the prototypical meaning of non-subject relatives helps the child to bootstrap into this type of relative clause.

4.1 Data and coding

The investigation is based on observational data from two English-speaking children from the CHILDES database (MacWhinney 2000), Adam (Brown 1973) and Abe (Kuczaj 1976). Adam's data comprise 55 one-hour recordings that occurred at regular intervals of 2 to 3 weeks between the ages of 2;3 and 5;2; overall there are 180 finite relative clauses in Adam's corpus. Abe's data consist of 210 files containing transcripts of 30 minutes recordings that were collected every week between the ages of 2;4 and 5;0; overall Abe's corpus includes 309 finite relative clauses. In Diessel (2004), I also looked at the relative clauses of four other children, Sarah (Brown 1973), Peter (Bloom 1973), Nina (Suppes 1973), and Naomi (Sachs 1983); but since the corpora of these children are too small to investigate the interaction between semantic and syntactic features, they were excluded from the current investigation.

<table>
<thead>
<tr>
<th>Table 1. Data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Adam</td>
</tr>
<tr>
<td>Abe</td>
</tr>
</tbody>
</table>
The data were coded for various syntactic and semantic variables concerning aspects of both the main clause and the relative clause. The variables of the main clause include:
1. the meaning of the verb
2. the syntactic function of the head noun
3. the thematic role of the head noun
4. the animacy of the head noun

The variables of the relative clause include:
1. the meaning of the verb
2. the syntactic function of the relativized role
3. the thematic role of the relativized role
4. the animacy of subject and object
5. the NP type of subject and object

The levels of the variables will be defined in course of the analysis. In order to ensure coding reliability, 100 sentences were double coded by the author and a student assistant. For the syntactic variables, inter-coder reliability was almost 100 percent; for the semantic variables it was between 94.0 percent (thematic role) and 98.0 percent (animacy).

4.2 The animacy of the head and the relativized role

In a first step I examined the relationship between the animacy features of the head and the relativized syntactic role. The latter were divided into two basic types, subject relatives, i.e., relative clauses in which the subject is relativized, and non-subject relatives, i.e., relative clauses in which the direct object or an adverbial is relativized; indirect object relatives and genitive relatives did not occur in the data. Interestingly, although children have often difficulties with non-subject relatives in comprehension experiments, both children began to use non-subject relatives at around the same time as subject relatives and eventually produced more non-subject relatives than subject relatives: 64.5 percent (n = 116) of Adam's relatives and 57.3 percent (n = 177) of Abe's relatives are non-subject relatives.

The head of the relative clause was also divided into two basic types: animate head nouns denoting humans and animals, including toy animals, and inanimate head nouns denoting concrete objects and abstract entities. The vast majority of the children's relative clauses elaborate an inanimate noun: 78.3 percent (n = 141) of Adam's relatives and 79.0 percent (n = 244) of Abe's relatives are attached to an inanimate noun. Figure 2 shows that the animacy of the head correlates with the relativized syntactic role.

Subject relatives are common with both animate and inanimate nouns—there is only a small difference between them. On average, 42.4 percent of the children's subject relatives occur with animate nouns and 57.6 percent occur with inanimate nouns. However, non-subject relatives occur predominantly with inanimate nouns. As can be seen in Figure 2, on average 91.7 percent of Adam's and Abe's non-subject relatives are attached to an inanimate noun and only 8.3 percent occur with an animate noun. A 2 x 2 \( \chi^2 \)-analysis revealed a significant association between the animacy of the head and the syntactic function of the relativized role in the data of both children (Abe: \( \chi^2 = 52.67, df = 1, p < 0.001, \phi = 0.418 \); Adam: \( \chi^2 = 27.09, df = 1, p < 0.001, \phi = 0.393 \)) suggesting that the semantic feature of animacy is an important aspect of children's subject and non-subject relatives.

But why does the animacy of the head correlate with the relativized role? The answer appears to be straightforward: The correlation reflects the prototypical links between syntactic and semantic roles. Since animate referents are likely to function as agents, they are commonly modified by subject relatives, i.e., relative clauses in which the animate head functions as an agent or experiencer, and since inanimate referents are likely to function as patients, they are commonly modified by non-subject relatives, i.e., relative clauses in which the inanimate head functions as patient, theme, or location. In other words, I suggest that the semantic biases in children's subject and non-subject relatives arise from entrenched relationships between semantic and syntactic roles. One can think of these relationships as an associative network that emerges from children's experience with verb-argument constructions: The more often a semantic role is expressed by a particular syntactic category, the stronger the associative link between form and meaning. This explains why object relatives cause fewer difficulties in comprehension studies if they modify an inanimate noun rather than an animate noun. However, it
Holger Diessel

does not fully explain the animacy features of subject relatives: While subject relatives are more frequent with animate heads than non-subject relatives, most of the children’s subject relatives modify inanimate nouns (Figure 2), suggesting that subject relatives do not instantiate the prototypical links between syntactic and semantic roles.

In what follows I show that subject relatives comprise a variety of constructions that vary in terms of their semantic and pragmatic features, whereas non-subject relatives are commonly expressed by prototypical transitive clauses involving an animate agentive subject, a dynamic verb, and an inanimate object. Based on these data, I argue that the prototypical meaning of non-subject relatives helps the child to process (and learn) the particular structural properties of non-subject relatives.

4.3 The NP-type of the subject

The second aspect of non-subject relatives that has been studied intensively in the recent psycholinguistic literature on relative clauses is the NP-type of the subject (see above). In accordance with the Kidd et al. study, I found that non-subject relatives typically include a pronominal subject, notably a first and second person pronoun is very common. As can be seen in Figure 3, an average of 80.4 percent of Adam’s and Abe’s non-subject relatives include I, you, or we as subject (cf. Fox & Thompson 1990, 2007). The rest occur with third person pronouns (mean 7.7 percent), definite NPs (mean 10.0 percent), and a few indefinite nouns (mean 7.9 percent).

While the occurrence of the various NP-types is strongly skewed in favour of pronominal subjects, it must be emphasized that the frequency distribution of the various NP-types in non-subject relatives corresponds very closely with the distribution of the various NP-types in simple transitive clauses. Examining a random sample of 350 transitive (main) clauses from Adam and Abe’s corpus I found that an average 86.9 percent of all subjects are first or second person pronouns, 7.5 percent are third person pronouns, 3.7 percent are definite NPs, and only 1.8 percent are indefinite NPs. In other words, the predominant use of pronominal subjects is not a particular trait of non-subject relatives but is characteristic of ordinary (i.e., prototypical) transitive clauses.

By contrast, the subjects of subject relatives are radically different from the subjects of simple (in)transitive clauses. Since relative clauses do not modify first or second pronouns, they are exclusively used with third person subjects. Moreover, the subject of a subject relative is often a new discourse referent. As can be seen in Figure 4, the vast majority of Adam and Abe’s subject relatives are attached to a lexical noun functioning as subject inside of the relative clause. Note that more than half of the head nouns are indefinite, expressed by an indefinite lexical NP or an indefinite pronoun, introducing a new referent that functions as the subject of the relative clause.

Thus, the two types of relative clauses occur with very different types of subjects: non-subject relatives occur with highly accessible subjects that are characteristic of prototypical transitive clauses, whereas the subjects of subject relatives are third person referents that are often newly introduced in the preceding main clause.

---

**Figure 3.** Subjects of non-subject relatives.

<table>
<thead>
<tr>
<th>Subject Type</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>I/youd/we</td>
<td>80.4%</td>
</tr>
<tr>
<td>he/she/it/they</td>
<td>7.7%</td>
</tr>
<tr>
<td>definite NP</td>
<td>10.0%</td>
</tr>
<tr>
<td>Indefinite NP</td>
<td>1.9%</td>
</tr>
</tbody>
</table>

**Figure 4.** Subjects of subject relatives.

<table>
<thead>
<tr>
<th>Subject Type</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>this/that</td>
<td>43.30%</td>
</tr>
<tr>
<td>definite NP</td>
<td>31.70%</td>
</tr>
<tr>
<td>indefinite NP</td>
<td>19.70%</td>
</tr>
<tr>
<td>indefinite PRO</td>
<td>5.40%</td>
</tr>
</tbody>
</table>
4.4 The verb of the relative clause

Subject and non-subject relatives do not only occur with different subjects, they also occur with different types of verbs. The vast majority of the children's non-subject relatives include a transitive verb denoting a physical or cognitive activity (cf. examples 37–39); there are only a few intransitive non-subject relatives in the data in which an adverbial is relativized (cf. example 40).

(37) Where's the balloon [I made]? Abe 3;1
(38) Dose are bugs [that I throw]. Adam 3;7
(39) No the one [you found …]. Abe 3;9
(40) The inside of it [where we sit]. Abe 3;6

In contrast, subject relatives occur with both transitive and intransitive verbs. In fact, the majority of the children's subject relatives contain an intransitive verb or a copula, notably the copula be. As can be seen in Figure 5, an average of only 37.6 percent of the children's subject relatives are transitive, the rest include either an intransitive verb (43.8 percent) or a copula (18.6 percent), whereas non-subject relatives are mostly transitive: 95.1 percent of the children's non-subject relatives are transitive and only 4.9 percent occur with intransitive and copular verbs.

Interestingly, transitive and intransitive subject relatives are headed by different semantic types of nouns. Most of the copular and intransitive subject relatives are headed by inanimate nouns: On average 71.7 percent of the children's intransitive subject relatives and 52.9 percent of their copular subject relatives modify an inanimate noun (cf. Figure 6). Note that the intransitive subject relatives comprise unergative and unaccusative verbs (including transitive verbs in passive voice), which tend to occur with different semantic types of nouns: Unergative verbs occur with both animate and inanimate nouns (cf. examples 41 and 42), whereas unaccusative verbs (and passivized transitive verbs) are almost exclusively used with inanimate nouns (cf. examples 43).

(41) The doggie [that runs away]. [Adam 3;8]
(42) Look at that big truck [(dat) going some place]. [Adam 3;0]
(43) The wheel [that's broken]. [Abe 3;10]

Transitive subject relatives are more frequent with animate head nouns than intransitive and copular subject relatives. As can be seen in Figure 6, an average of 63.7 percent of the transitive subject relatives are attached to an animate noun and only 36.4 percent modify an inanimate noun. The difference between the three types of subject relatives is highly significant in the data of both children (Abe: \( \chi^2 = 14.53, df = 2, p < 0.001, \phi = 0.336 \); Adam: \( \chi^2 = 10.49, df = 2, p < 0.001, \phi = 0.441 \)).

Note, however, that the children's (transitive) non-subject relatives include a larger proportion of animate subjects than the children's transitive subject relatives. Averaged across the two children, 93.2 percent of Adam and Abe's non-subject relatives contain an animate subject, which corresponds quite closely to the proportion of animate subjects in the children's simple transitive clauses (96.1 percent); whereas only 63.7 percent of the transitive subject relatives occur with an animate subject. The difference between

---

Figure 5. Verbs of subject and non-subject relatives.

Figure 6. Animate and inanimate heads of transitive, copular, and intransitive subject relatives.
transitive subject and (transitive) non-subject relatives is highly significant (Abe: \( \chi^2 = 43.60, df = 1, p < 0.001, \varphi = 0.435 \); Adam: \( \chi^2 = 26.1, df = 1, p < 0.001, \varphi = 0.440 \)) suggesting that object and adverbial relatives include more prototypical subjects than subject relatives.

In order to test this hypothesis, I analyzed the thematic roles of the subjects in both subject and non-subject relatives, using the following definitions of semantic roles (cf. Huddleston & Pullum 2002):

1. agent (instigator of a physical activity)
2. patient (animate or inanimate being affected by a physical activity)
3. experiencer (animate being experiencing or perceiving an event)
4. recipient (animate being receiving an entity)
5. beneficary (animate being benefiting from an activity)
6. instrument (inanimate entity used to accomplish a goal)
7. location (place, destination, source)
8. theme (entity that is perceived, known, located, or possessed).

In accordance with the above hypothesis, I found that an average of 88.7 percent of Adam's and Abe's non-subject relatives include an agent or experincer as subject. In contrast, subject relatives include a much smaller proportion of agentive subjects: Only 48.1 percent of the children's subject relatives include a prototypical subject functioning as agent or experiencer of the activity expressed by the verb; the remaining 51.9 percent function as patient, theme, or location. This is of course partly due to the fact that the children's subject relatives include many intransitive and copular verbs, but even if we restrict the analysis to transitive verbs an average of only 55.2 percent of Adam and Abe's subject relatives occur with an agent or experencer as subject. In particular, experiencer subjects are extremely rare in subject relatives. Overall, there are only 5 subject relatives in the entire database in which the subject of a transitive subject relative clause functions as an experiencer, i.e., the mental agent of a cognitive or psychological state or activity.

This is reflected in the types of verbs that subject and non-subject relatives include. Concentrating on relative clauses with transitive verbs, I distinguished between three different semantic verb types:

1. Verbs denoting a physical activity (e.g., put, make)
2. Verbs denoting cognition, perception, or communication (e.g., know, see, say)
3. Verbs denoting a state or possession (e.g., be, have, belong)

While this classification surely has its problems, Figure 7 shows that there are some striking differences between the two types of relative clauses: Both subject and non-subject relatives are commonly used with (transitive) activity verbs, but while non-subject relatives are also commonly used with verbs of cognition, perception, or communication, subject relatives are only rarely used with these verbs, including instead many stative verbs and verbs denoting possession; in particular, the verb have is very frequent in transitive subject relatives. The difference between subject and non-subject relatives is again highly significant (Abe: \( \chi^2 = 82.38, df = 2, p < 0.001, \varphi = 0.519 \); Adam: \( \chi^2 = 27.46, df = 2, p < 0.001, \varphi = 0.391 \)).

### 4.5 The nominal referents of the relative clause

Finally, I examined the semantic differences between subject, object, and adverbials in the children's relative clauses with multiple referents. Disregarding intransitive subject relatives with a single referent (i.e., subject relatives including only a subject), I distinguished between four different types of relative clauses:

1. Relative clauses including an animate subject and an inanimate object (or adverial).
2. Relative clauses including an animate subject and an animate object (or adverial).

3. In relatives clauses with more than two referents I concentrated on the core roles, i.e., subject and object, and disregarded adverbials.
3. Relative clauses including an inanimate subject and an inanimate object (or adverbal).
4. Relative clauses including an inanimate subject and an animate object (or adverbal).

Figure 8 shows that while subject relatives occur with various combinations of animate and inanimate nouns, non-subject relatives typically include an animate subject and an inanimate object or adverbial. An average of 44.9 percent of the children’s subject relatives include an animate subject and an inanimate object (or adverbial) as in prototypical transitive clauses; but very often the two referents are both animate or both inanimate: 14.6 percent of the children’s subject relatives occur with two animate nouns and 36.8 percent occur with two inanimate nouns; that is, in more than half of the children’s subject relatives, subject and object/adverbial do not differ in terms of their animacy features. By contrast, non-subject relatives are strongly skewed in favour one particular type. As can be seen in Figure 8, 88.6 percent of the children’s non-subject relatives contain an animate subject and an inanimate object (or adverbial), all other types are infrequent, supporting the hypothesis that non-subject relatives are prototypical transitive clauses, whereas subject relatives are distributed across various types of constructions that are low on the transitivity scale.

5. Conclusion

In the usage-based approach to language acquisition, grammatical development reflects the child’s experience with language. But what exactly constitutes the child’s linguistic experience? Two aspects have been emphasized in recent studies on the acquisition of morphology and the emergence of item-based constructions: Frequency of occurrence, which determines the level of entrenchment, and structural similarity, which underlies the emergence of general grammatical patterns (or schemas). This study has shown that frequency and similarity are not only crucial for the early stages of grammatical development but are equally important for the acquisition of more complex grammatical patterns such as relative clauses. Specifically, the paper has argued that the acquisition of relative clauses is influenced by children’s prior knowledge of simple sentences.

The first study has shown that in comprehension, children have often fewer difficulties with subject relatives than with object, adverbial, and genitive relatives because subject relatives involve the same sequence of subject, verb, and object as simple sentences whereas non-subject relatives deviate from the familiar SV(O) pattern. Moreover, the study suggests that structural similarities between object, adverbial, and genitive relatives affect the processing and development of these constructions in the early stages of grammatical development.

The second study has shown that although non-subject relatives are structurally distinct from ordinary SV(O) clauses, semantically they are expressed by prototypical transitive clauses, whereas subject relatives comprise a variety of constructions including transitive, intransitive, and copular verbs. Interestingly, although subject relatives have the same word order as ordinary SV(O) clauses, semantically they often deviate from ordinary (in)transitive clauses. For instance, while the vast majority of the children’s simple SV(O) clauses include a pronominal subject, notably a first or second person pronoun, subject relatives are exclusively used with third person subjects that are often newly introduced into the discourse. In general, subject relatives comprise a diverse group of constructions with rather unusual semantic and pragmatic properties, whereas non-subject relatives are commonly expressed by prototypical transitive constructions including a pronominal animate subject functioning as agent or experiencer of a dynamic verb with an inanimate object functioning as patient or theme. The semantic and pragmatic properties of non-subject relatives are thus predictable from the properties of prototypical transitive clauses, which arguably may help the child to bootstrap into this type of relative clause. In other words, the difficulties that arise from the unusual structure of non-subject relatives are mitigated by the fact that non-subject relatives instantiate the semantic and pragmatic properties of prototypical transitive clauses.
References


