University of Massachusetts Occasional Papers in Linguistics

Volume 3 *Papers in the Structure and Development of Child Language -- UMOP Volume 4*

Article 3

1978

The Conjoined-Clause Analysis of Relative Clauses and Other Structures

Susan L. Tavakolian

Recommended Citation

Tavakolian, Susan L. (1978) "The Conjoined-Clause Analysis of Relative Clauses and Other Structures," *University of Massachusetts Occasional Papers in Linguistics*: Vol. 3 , Article 3.
Available at: https://scholarworks.umass.edu/umop/vol3/iss2/3

The Conjoined-Clause Analysis
of
Relative Clauses and Other Structures

Susan L. Tavakolian

0. Introduction

A theory of linguistic knowledge incorporates linguistic universals and ascribes implicit knowledge of these universals to the child. Universal aspects of language are assumed to result from the biological composition of the human mind and therefore to be an innately determined set of schemata and principles which limit the set of possible grammars. The existence of such schemata and principles greatly simplifies the task of a child learning the language of his community. The child approaches the data presented to him with a predetermined set of principles which restrict the nature of the hypotheses he can make about the structure of a given language. The child must select the hypothesis which is consistent with the data, and if more than one hypothesis is consistent, there must be a method of selecting the appropriate one. The latter problem is the problem of evaluating competing grammars, and in the conclusion to this paper a possible evaluation principle will be proposed. However, we will be dealing primarily with children's formulation of hypotheses concerning the structure of the linguistic data, and the implications of these hypotheses for universal grammar.[1]

The existence of universal aspects of language has been widely recognized by researchers in language acquisition, but the identification of specific universals and their relationship to other areas of cognitive development are unresolved issues which have been approached within a variety of theoretical models. (See Chomsky 1965; McNeill 1966, 1970; Bever 1970b; Slobin 1971, 1973; Bowerman 1973; Brown 1973; Roeper 1973; Lust 1974; Sinclair 1975). One general approach has emphasized the independence of linguistic structure from other cognitive systems. Other approaches have proposed that linguistic systems are extensions of more general cognitive properties. This paper will assume the first approach as a fruitful way of assessing the implications of the results presented in this paper, but a contrary assumption would not affect the results or their analysis.

The discovery of linguistic universals, regardless of their ultimate relationship to other cognitive structures, will tell us what kinds of assumptions a child brings to the task of language learning. These assumptions will place heavy restrictions on the hypotheses formulated by the child and on the form of his grammar.

1. The Hypotheses

Multiple-clause sentences are a challenge to children's parsing strategies and rules for interpreting missing elements. The child must determine the boundaries for each simplex sentence and determine a referent for any missing elements in the sentence. A great many of children's "errors" in interpreting complex sentences are systematic response patterns and provide us with information concerning the types of hypotheses a child makes in the process of language acquisition. This study focuses on two such hypotheses and accounts for their formulation in terms of fundamental concepts used in theories of adult grammars. Hypothesis A deals with rules of parsing, and Hypothesis B deals with the relationship of lexical entries and structural analyses.

An assumption underlying this study is that children will rely on the grammatical rules they already possess in an attempt to process difficult or unfamiliar structures, even though these already existing rules may be inappropriate to the data at hand. The types of rules children use to interpret sentences which are difficult for them will give us some insight into the linguistic processes a child has at his disposal in analyzing and interpreting material which his grammar does not yet generate.

Hypothesis A: When children are uncertain about the structure of a multiple-clause sentence, they attempt to parse the string as though it consisted of conjoined simplex sentences.

When a child is faced with a string of nouns and verbs to which he must give an interpretation and which his grammar does not yet generate, the child attempts to segment the string into simplex sentences. Given a string such as the following

1) NP...V...NP...V...NP

where the "..." indicate that material such as a relative pronoun or conjunction but not a verb or NP may intervene between the NP and verbal elements, we propose that a child's first hypothesis about the structure of the sentence is that it consists of two conjoined simplex sentences.

The string shown in (1) abstracts over a number of different structures. The intervening material may contain morphemes which indicate structural differences in an adult grammar. Schema (2) shows the different elements which may appear in the intervals marked by "..." and gives an example of the kind of structure in which each appears:

$$2)\quad \text{NP (a. }\underline{\text{that}}\text{) V NP} \left\{ \begin{array}{ll} \text{b.} & \underline{\text{that}} \\ \text{c.} & \underline{\text{in order to}} \\ \text{d.} & \underline{\text{and}} \\ \text{e.} & \underline{\text{to}} \end{array} \right\} \text{V NP}$$

a. The sheep that knocks down the rabbit stands on the lion.
b. The sheep knocks down the rabbit that stands on the lion.
c. The sheep knocks down the rabbit to stand on the lion.
d. The sheep knocks down the rabbit and stands on the lion.
e. The sheep tells the rabbit to stand on the lion.

Each sentence used in the experiments reported in this paper contained only one of the structural possibilities indicated above. That is, a sentence contained a relative clause, or an in order to clause or conjoined clauses, etc., but not more than one of these.

We propose that a child initially assigns the same structure to all of the above strings and analyzes them in the following way:

$$3)\quad {}_S[{}_S[\text{NP...V...NP}]\ {}_S[\Delta\text{...V...NP}]]$$

where Δ indicates that the subject of the second clause is empty. The empty subject is interpreted as being coreferential with the subject of the first clause. The string has been analyzed as containing two simplex sentences with the subject of the second clause missing.

It is not obvious that a child should segment a string as shown in (4); another possible way to parse the string would be to identify the second NP of the string as the subject of the second clause. This kind of segmentation would result in an analysis containing an embedded clause:

4) $${}_{S}[NP\ldots{}_{VP}[V\ldots\underset{\uparrow}{NP}\ {}_{S}[\underset{|}{\Delta}\ldots V\ldots NP]]]$$

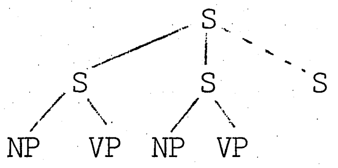
Children do use this kind of analysis with certain verbal complements, but it does not appear to be the hypothesis they try initially with unknown structures.

Tom Roeper has suggested a variation of the conjoined-clause hypothesis to me, namely, that attachment of new material occurs initially on the topmost S of a structure. He points out that this formulation would explain why children first put negatives before a sentence, why they misunderstand relative clauses and construe them with reference to the subject, why they put quantifiers in sentence-initial position (as in "Only I want this."), why they first learn questions with an invariant initial word ("Did you can come?"), and why some children misinterpret perception verbs and other verbs with participial complements as subject controlled (Goodluck and Roeper, this volume). This formulation is argued for in greater detail in Solan and Roeper (this volume).

The S-attachment formulation of the hypothesis is more general than the conjoined-clause hypothesis. It accounts for structural relationships between other items (such as negatives) and clauses as well as between clauses. The conjoined-clause hypothesis can be subsumed under the more general S-attachment hypothesis. However, since we are investigating only the relationship between clauses in this paper, we will continue to formulate the discussion in terms of the conjoined-clause analysis.

The primary difference between the analysis presented in Solan and Roeper and the present analysis is in the kind of tree structure that is postulated to account for children's responses. The present analysis claims that children assign the following structure to multiple-clause sentences:

5)

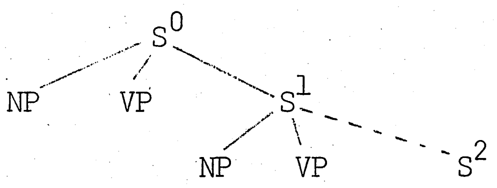


This structure would be generated by the phrase structure rule

6) $S \longrightarrow S^{*}$

The analysis presented by Solan and Roeper would assign multiple-clause sentences the structure shown in (7):

7)



which would be generated by the phrase structure (p.):

8) S $\longrightarrow$ NP VP (S)

In structure (5) each clause is directly dominated by the top-most S; however, in (7) only S^1 is directly dominated by the top-most S. Clause S^2 and any subsequent clauses would be directly dominated by the immediately preceding S, creating an embedded structure.

The second of our two hypotheses concerning the processing of complex sentences deals with how lexical information constrains the surface parsing.

Hypothesis B: Where the correlation of thematic relations and subcategorization features of a verb occurring in a surface string differs from the correlation of thematic relations and subcategorization features of the verb in deep structure, children will attempt the most direct projection of thematic relations onto the surface string when the verb occurs in a complicated sentence whose structure they are uncertain of.

We will suppose that a child has an internally represented lexicon which contains for each lexical entry subcategorization information concerning the deep structure trees into which the item may be inserted.[2] For verbs, the entry will also indicate the semantic relationship between the arguments of the verb and indicate the mapping of its thematic relations onto the subcategorization features. For example the entry for the verb hit might be the following:

9) hit
+V
+[NP^1 __ NP^2]
HIT (NP^1, NP^2)
NP^1 = agent
NP^2 = goal

The feature +[NP^1 __ NP^2] indicates that the verb is transitive; HIT(NP^1, NP^2) indicates that HIT is the relationship that holds between the two arguments of the verb. The last two lines of the entry indicate that NP^1 is the agent of the verb and that NP^2 is the goal of the verb. The use of superscripts correlates the thematic relationships with the subcategorization features. So for hit we know that the first NP in the deep structure tree is the agent and the second NP is the goal.[2] For all the verbs used in this study, NP^1 of the subcategorization feature is the agent and NP^2 is the goal.

Many verbs of English are like hit in that NP^1 will be the agent or experiencer of the action of the verb and NP^2 will be the patient or goal of the action. When the order of the atomic predicates (the noun phrases and the verbs) in surface structure is different from the order in deep structure, there is not a simple mapping of thematic relations onto structural relations. For example, in a string

10) NP_a...hit...NP_b

the most direct mapping of thematic relations onto this structure, given the lexical entry (9) for hit, is to assume that NP_a corresponds to NP^1 and is an agent and that NP_b corresponds to NP^2 and is a goal. If, in fact, the string is a passive,

11) NP_a...be+en hit...by NP_b

then the most direct mapping will produce an incorrect interpretation of the sentence.

We hypothesize that children will attempt the most direct mapping of thematic relations onto a surface string. For some structures, such as object clefts, it is not clear what the most direct mapping would be. In the string

12) It's NP_a that NP_b hit

it's not obvious which NP corresponds to NP^1 and which to NP^2 since the linear order of the NP's conflicts with their relationship to the verb. That is, NP_b, which immediately precedes hit and so might be construed as NP^1, is the second NP in the linear string and so might be construed as NP^2. Based on these kinds of difficulties we would predict that children would err in their assignments of thematic relationships to the surface string, sometimes assigning NP_a as the agent and at other times assigning NP_b as agent.

The problem of how children assign grammatical relationships in sentences where the deep and surface configurations are different has been discussed a great deal in the literature (Slobin 1966, Turner and Rometveit 1968, C. Chomsky 1969, Bever 1970a, Brown and Hanlon 1970, Roeper, 1973, Solan, 1975). Some proposals, such as those of Brown and Hanlon and of Roeper, have tended to emphasize the transformational component of the grammar as a locus of difficulty in comprehension and production. Transformational rules which alter the deep structure configurations of a string make the sentence more difficult for children to correctly comprehend or produce. In a different approach work by Bever, Sheldon, Legum and Solan has emphasized perceptual strategies, which can aid the child in interpreting structures or can be inappropriately applied to yield an incorrect response. Other studies have emphasized the importance of semantic cues in interpreting complicated sentences (Slobin, Turner and Rometveit) and have pointed out that lack of semantic cues increases the difficulty of a sentence for a child. The focus of the present study differs from that of other studies in that greater emphasis is placed on the role of lexical entries in children's comprehension of sentences which their grammars do not yet generate, and it is proposed that it is the mapping of thematic relationships onto surface trees which are not isomorphic with their deep structure trees which creates difficulty for children in their comprehension.

It is not unreasonable to assume that lexical information can aid a child in assigning structure to a sentence. If the child is attempting to parse a sentence whose structure he is uncertain of and which contains the verb hit, for example, he can determine from the subcategorization feature in the lexical entry for hit that it requires an object, and so he will not postulate a sentence boundary for the clause containing hit until he has processed an object for the verb. When he does process an appropriate object NP, then he closes off the current string as a simplex sentence.[3] If a string such as (1) above has hit as the first verb, the first clause can be closed with an S boundary as soon as the second NP is processed since the subcategorization feature, $+[NP^1 \_\_ NP^2]$, has been satisfied.

Fodor, Bever and Garrett (1974) have made a similar proposal, "the lexical analysis strategy," for adult sentence perception. They propose that "an early step in sentence recognition involves postulating any deep structure for a sentence that is compatible with the lexical analysis of its verbs."

(p. 350). They cite a number of experiments (Fodor, Bever and Garrett 1968, Holmes and Forster 1972, Hakes 1971) which indicate that the greater the number of deep structure configurations into which a verb can be inserted, the more difficult it will be to understand sentences containing the verb. The greater the number of subcategorization features in the lexical entry of a verb, the greater the difficulty in understanding it. The results of Bever, Lackner and Kirk (1968) and Fodor, Fodor, Garrett and Lackner (1974, reported by Fodor, Bever and Garrett (1974)) suggest that adults use lexical subcategorization information in determining sentence boundaries of embedded sentences. The available experimental evidence suggests that the lexical analysis strategy is actively employed by adults in postulating deep structures for a sentence and in parsing multiple clause sentences; so it is not at all surprising that we should find that children also exploit this information in parsing difficult sentences and in determining deep structures.

Hypothesis A and Hypothesis B have been discussed as though they are separate hypotheses, but obviously they are closely related. When presenting evidence, we will discuss the two together.

The data presented in the following sections support Hypothesis A that a child will attempt to parse a multiple-clause sentence into conjoined simplex sentences when he is uncertain of its structure. This data will show that the most commonly occurring response can be accounted for by the proposal that the string is parsed into two simplex sentences and that the subject of the second clause is interpreted as being coreferential with the subject of the first clause. This paper provides evidence primarily for Hypothesis A, and to a lesser extent, for Hypothesis B. The support for Hypothesis B is not a direct result of the experimental data presented here. (The studies cited earlier do provide such experimental evidence for adult grammars.) Rather, Hypothesis B is a result of plausible model construction and is a tentative and theoretical account, which is consistent with, though not directly supported by, the experimental evidence.

2. The Experiments

2.0 Introduction

The data to support the above hypotheses were collected from two comprehension experiments conducted with preschool children. One experiment tested children's comprehension of relative clauses, and the other experiment contained sentences with conjoined clauses, sentences with verbal complements and sentences with <u>in order to</u> clauses.

Each experiment contained three tokens of each sentence type. The relative clause experiment had 3 instances of each relative clause type, and the second experiment had three instances of conjoined sentences, three instances of sentences with verbal complements, etc. Both experiments contained declarative sentences as a control to see if the children understood the task, as a basis of comparison between the experiments, and to determine whether there was any tendency among the children to use all the animals put out on the floor even when only two were mentioned.

Each experiment consisted of four interview schedules. Each schedule was drawn up by randomly assigning the six animals to the NP slots in each structure and also randomizing the eight verbs in each sentence. The only restriction was that no verb or animal was repeated within the same sentence. The list of sentences obtained in this way was then randomized. Thus each experiment had four different lists of randomized sentences with each sentence constructed from a randomized list of animals and verbs.

2.1 The Population

Each experiment was administered to a different group of 24 children who ranged in age from 3 to 5 years. Eight children aged 3.0-3.6 (age reported in years and months), eight aged 4.0-4.6 and eight aged 5.0-5.6 participated in each experiment, and so a total of 24 children responded to each experiment.

There were four schedules of sentences for each experiment. These were distributed among the three age groups so that two 3 year olds, two 4 year olds, and two 5 year olds responded to each interview schedule.

The children were enrolled in various nursery schools in Western Massachusetts and represented a wide range of family incomes. Most children were from middle income families. However, some came from families receiving state aid to help pay for child care and others were children of high income families. All the children spoke standard English as a first language, and none were bilingual.

2.2 Administration Procedure

The sentences for the experiments were administered to each child individually. The child and tester both sat on the floor with the animals between them facing the child. Each child was asked to name the animals; a few children were unsure of the names of the sheep and the lion, calling the latter a tiger, but were prompted on the names. The animals were then discussed for a while so that initial hesitation about the names didn't confuse later acting out of the sentences with the animals.

The children were given the following instructions after having named the animals and talked about them briefly: "We're going to play a game with these animals. All of these animals like to play together. I'm going to tell you some of the things the animals like to do. I'll put three animals on the floor in front of you and then tell you something that the animals are doing. You make the animals move to do what I say. When you're done moving the animals then you put them back on their places on the board. I'll do the first one so you can see how the game goes and then you can do the rest of them."

There were two demonstration sentences:

13) The horse walks around the lion.
14) The pig jumps over the rabbit and the horse.

Three animals were placed in front of the child for each sentence, even when only two were named, as in (13). Sentence (13) was designed so that children would realize that not all three animals need be used to act out a sentence, and it was usually called to the children's attention that in (13) the third animal didn't do anything.

Children easily comprehended the nature of the task and enjoyed moving the animals.

2.3 Coding the Children's Responses

For ease of analysis the responses to the sentences are coded according to the occurrence of the noun phrase in the sentence.[4] Each NP is assigned a number corresponding to its occurrence in the linear order of the string. For example,

        1   2   3
15) The duck stands on the horse and jumps over the pig.

There are two clauses to be acted out in (15), and these are separated by a comma in the coded response. A 12,13 response with respect to (15) means that a child made the duck stand on the horse and the duck jump over the pig. A 12,23 response to the same sentence means that the duck stands on the horse and the horse jumps over the pig.

Some sentences, such as the simple active declaratives, have only two animals mentioned in the sentence. In this case the animal not mentioned in the sentence is listed to the right of the sentence and coded number three. For example,

          1   2   3
16) The rabbit kisses the lion.   duck

A 12 response indicates that the rabbit kisses the lion; a 13 response means that the rabbit kisses the duck.

The first number mentioned in each two-number sequence indicates the noun phrase functioning as the subject of the first verb, and the second number indicates the noun phrase functioning as the object of the first verb. The coded response of a simple active declarative sentence, such as (16), has one double-number sequence (for example, "12") since there is only one verb. The coded response of a multiple-clause sentence, such as (15), has two double-number sequences (for example "12,13"). The first sequence indicates the noun phrases functioning as subject and object of the first verb, and the second sequence indicates the noun phrases functioning as subject and object of the second verb.

3. The Conjoined-Clause Analysis

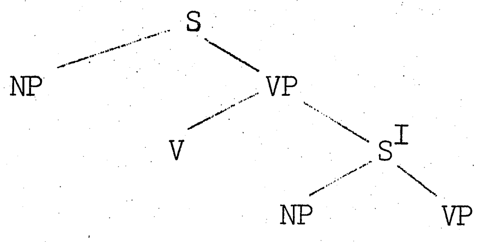
The proposed conjoined-clause analysis is a set of rules for assigning structure to a string and for interpreting missing noun phrases in a string. The rule for assigning structure converts an unanalyzed string such as (1) into a conjoined structure such as (3) with labelled bracketing. The subject of the second clause is lexically unfilled and an antecedent for the missing subject must be determined.

Williams (1975:256) has suggested a principle for determining the NP which is coreferential with a missing noun phrase. Following Solan and Roeper (this volume) we have phrased it as follows: [5]

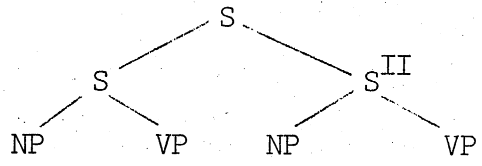
17) Universal Control Principle: A missing NP cannot be higher on the tree than the NP which is its antecedent.

Williams defines the notion "higher in a tree" in terms of four classes of items which may follow the verb in a sentence. We will tentatively use a simplified definition in terms of two classes of clauses: those which are attached to the VP and those which are attached to the top-most S.[6] The former we will call Class I clauses and the latter Class II clauses. Tree (18a) shows the structure of a sentence containing a Class I clause, and (18b) illustrates the structure of a Class II clause:

18) a.



b.



The Universal Control Principle restricts the choice of an antecedent for a missing NP. An NP in a Class I complement may not serve as the antecedent of a missing NP of a Class II clause. There are also other restrictions in a child's grammar on the choice of an antecedent (see, for example, Tavakolian (this volume)).

Thus the "conjoined-clause analysis" refers to the set of rules which assigns conjoined structure to a string and which determines an antecedent for a missing NP in accordance with the above control principle.

We propose that this set of rules is established in a child's grammar very early and is utilized in interpreting more difficult and unfamiliar constructions. If the linear order of noun phrases and verbs in a multiple-clause sentence varies from the order specified in the conjoined-clause analysis, the sentence will be eliminated from the conjoined-clause analysis earlier than sentences which have the same order of elements as conjoined clauses. But, if the linear order of elements in a sentence is the same as the order in conjoined sentences, then children will use the conjoined-clause analysis to assign clausal structure to it and to interpret missing elements. Until the child has mastered the adult system for a particular structure, he will continue to use the conjoined-clause analysis even when this results in an incorrect interpretation of the sentence. The evidence to be presented will support the proposal that children utilize a conjoined clause analysis to interpret multiple-clause sentences which their grammar does not yet generate and will also indicate that the conjoined-clause analysis is a very productive set of rules.

As expected, conjoined sentences themselves are quite easy for children to correctly interpret. In my data children correctly responded to 96% of the sentences with two conjuncts in which the subject of the second conjunct was missing. Each child was given three sentences such as

19) The lion hits the pig and stands on the sheep.

Only three children missed any conjoined sentences, and they missed only one sentence each. The other 21 children correctly responded to all three sentences. Given a schematized string (20a),

20) a. NP V NP <u>and</u> V NP

b. $_S[\,_S[\underset{\uparrow\!\_\_\_\_\_\_\_\_\_\_\_\_\_}{\text{NP}}\ \text{V NP}]\ \underline{\text{and}}\ _S[\underset{\_\_\lrcorner}{\Delta}\ \text{V NP}\ ]]$

it is parsed as (20b). The missing subject of the second conjunct, represented by Δ, is interpreted as being coreferential with the subject of the first conjunct.

Conjunction of elements in spontaneous speech is an operation which occurs very early (Menyuk 1969, 1971; Limber 1973; Brown 1973). By nursery school age clausal conjunctions with "and" are produced quite frequently (Lust 1974: 48), and at this age children also form relative clauses by generalizing the rule for conjunction (Menyuk 1969: 92, Limber 1973: 182).

Data from imitation studies also indicate that conjunctions are easy for children. Slobin and Welsh (1973) found that Echo at age 2.2.2 (age in years, months, weeks) would correctly imitate full sentential conjunctions. Beilin and Lust (1975) found that 2 and 3 year olds, when administered a sentence consisting of two conjoined sentences, such as

21) Give me the girls and give me the boys.

often reduced the full sentential form of the second conjunct to produce

22) Give me the girls and the boys.

Both studies indicate that yound children are able to imitate conjoined structures and convert them into related forms.

Both imitation and production data indicate that conjunction is an easy operation for children to acquire and imitate. Conjunctions are mastered early in a child's language development and the rules underlying their production and comprehension are in existence at an early stage and are available to be generalized to other multiple-clause structures.

3. Relative Clauses

3.0 Introduction

Relative clauses are a particularly difficult construction for children to produce, comprehend and imitate. For this reason a number of researchers in language acquisition have investigated them. The spontaneous production of the various types of relative clauses spans a wide age range depending on the internal structure of the relative clause. Limber (1973) reports that some relative clause types are produced quite early and that children under the age of 3 years produce relative clauses in which the head of the relative is the matrix sentence object and the object of the embedded clause is relativized. Other types of relative clauses are not produced until much later; Menyuk (1969: 16) reports that at age 7 children rarely produce utterances in which the head of the relative clause is the matrix subject.

Comprehension of relative clauses is also difficult for children. The results of Sheldon (1974a) and Legum (1975) indicate that the degree of difficulty varies depending on the function of the relativized NP in the embedded clause. Differences in the ease of comprehension depending on the structure of the relative clause have also been found by H. D. Brown (1971) and by Gaer (1969).

Studies of children's imitations of relative clauses by Slobin and Welsh (1973) indicate that children have difficulty correctly repeating stimulus sentences with center embedded relative clauses and often omit part of the sentence or restructure it into conjoined sentences. Other imitation studies of relative clauses (M. Smith 1974, C. Smith 1970, Brogan reported in Sheldon 1972) indicate that the position of the clause (whether it is right or center embedded) and the function of the relativized NP contribute to the difficulty of imitating a relative clause.

Given that relative clauses are quite difficult for children to produce, comprehend and imitate and that the degree of difficulty depends on their internal structure, they offer a rich source of material with which to elicit the rules that children use to comprehend sentences which are difficult for them to interpret. The most extensive evidence for children's use of the conjoined-clause analysis comes from them.

3.1 The Sentences Containing Relative Clauses

Four types of relative clauses were tested:

23) a. The head of the relative clause is subject of the matrix clause and the subject of the relative clause is relativized (SS);
b. The head of the relative clause is subject of the matrix clause and the object of the relative clause is relativized (SO);
c. The head of the relative clause is object of the matrix clause and the object of the relative clause is relativized (OO);
d. The head of the relative clause is object of the matrix clause and the subject of the relative clause is relativized (OS).

The following sentences are examples of each type:

24) SS: The sheep that knocks down the rabbit stands on the lion.
SO: The lion that the horse kisses knocks down the duck.
OO: The horse hits the sheep that the duck kisses.
OS: The lion stands on the duck that bumps into the pig.

3.2 Distribution of responses

Presentation of the results has the following form: first we establish the importance of the response in which the subject of the first clause is also the subject of the second clause (coded "12,13"). This response forms the basis of the argument that children use a conjoined-clause analysis in comprehending sentences they have not yet mastered, and so its pervasiveness is carefully documented. The documentation shows that 12,13 is the most commonly occurring response to all four types of relative clauses (section 3.2.0). The most frequently occurring response types are then discussed, and again we see that response 12,12 is the only response consistently used on all four types of relative clauses (section 3.2.1). Next we examine the percentage of errors accounted for by each response type, and find that response 12,13 accounts for the greatest number of errors (section 3.2.2). Thus the pervasiveness and importance of the response 12,13 to relative clauses is documented by evidence from the overall distribution of responses, from an examination of the predominant response types, and from the breakdown of errors.

Having established the importance of the 12,13 response to all four types of relative clauses, we show how application of the proposed conjoined-clause analysis to each relative clause type results in a 12,13 response (section 3.3). We propose a hierarchical arrangement of the four relative clause types according to children's maintenance of the conjoined-clause analysis for each type. The characteristics of each type which cause children to eliminate it from the conjoined-clause analysis are also discussed. Finally we show that the relative clause responses also provide support for Hypothesis B. The evidence indicates that when there is a divergence in the surface structure of the mapping of thematic relations onto subcategorization features from the correspondence in the lexical entry, children will have great difficulty assigning an interpretation to the structure.

The following tables indicate the distribution of responses for each type of relative clause. The correct response is tabulated in the first column of each table. Recall that the coded responses refer to the occurrence of the NP in the linear order of the string; so "1" refers to the first NP in the sentence, "2" to the second NP and "3" to the third NP. The two clauses of the sentence are separated by a comma. For example, a response of "12,13" means that the first NP and the second NP were subject and object, respectively, of the first verb and that the first NP and the third NP were subject and object, respectively, of the second verb in the string. A response such as "12,12" indicates that the first NP and the second NP were subject and object, respectively, of both the first and second verbs in the sentence. The category "Other" consists of all the response categories which had only one response in them. Any response type which had two or more responses is listed as a separate response category.

Table 1 shows the distribution of responses to SS relative clauses. Clearly, most children responded correctly to these relative clauses. The response 12,13 was the predominant response and there were very few errors.

TABLE 1

Distribution of Responses to SS Relative Clauses
"The sheep that knocks down the rabbit stands on the lion."

Age	Response Categories				
	Correct 12,13	12,23	21,23	12,32	Other
3.0-3.6	18	2	1	0	3
4.0-4.6	16	5	1	0	2
5.0-5.6	22	0	0	2	0
Totals	56	7	2	2	5
Percentage	78%	10%	3%	3%	7%

The high percentage of correct responses and the fact that the "Other" category is small indicate that the children have an effective set of rules for correctly interpreting SS relative clauses.

Table 2 shows the distribution of responses to SO relative clauses, and we find a very different distributional pattern from that of Table 1.

TABLE 2

Distribution of Responses to SO Relative Clauses
"The lion that the horse kisses knocks down the duck."

Age	Response Categories						
	Correct 21,13	21,23	12,13	13,23	21,32	31,13	Other
3.0-3.6	5	4	5	4	0	0	6
4.0-4.6	6	5	8	0	2	1	2
5.0-5.6	4	6	9	1	0	1	3
Totals	15	15	22	5	2	2	11
Percentages	21%	21%	31%	7%	3%	3%	15%

In Table 2 only 21% of the responses are correct, and there is no dominant response as there was in Table 1. The SO relative clauses are clearly more difficult for children to comprehend. The difficulty is reflected in the small percentage of correct responses, in the lack of any predominant response strategy, and in the large number of response categories. Each response in the "Other" category represents a unique response to the SO relative clauses. Summing the eleven different responses of the "Other" category with the remaining response categories, we get a total of seventeen different responses. In the case of SS relative clauses, there was a total of nine different responses, only slightly over half as many as for SO relatives. This difference in the number of response categories is an indication of the difficulty children had with SO relatives and also indicates that children had no one set of rules which was easily applicable to them. Although there was not a predominant response, response 12,13 accounts for the greatest percentage of reponses to SO relatives.

Table 3 presents the distribution of responses to OO relative clauses.

TABLE 3

Distribution of Responses to OO Relative Clauses
"The horse hits the sheep that the duck kisses."

Age	Response Categories						
	Correct 12,32	12,13	12,31	12,12	31,32	12,23	Other
3.0-3.6	8	4	5	2	1	0	4
4.0-4.6	9	5	6	0	1	1	2
5.0-5.6	10	5	5	1	0	2	1
Totals	27	14	16	3	2	3	7
Percentage	38%	19%	22%	4%	3%	4%	10%

Table 3 shows that OO relatives were more difficult than SS relative clauses, but not so difficult as SO relative clauses. As with Table 2 there is no predominant response type, but three main response categories account for 79% of the responses, and response 12,13 accounts for 19% of the responses. There is less scatter in the overall number of response categories than in SO relative clauses. There are thirteen categories for OO relative clauses versus seventeen for SO relatives. The greater percentage of correct responses and the smaller number of response categories indicate that children can more easily find a set of rules for interpreting OO relative clauses than is the case for SO relatives. Although response 12,13 accounts for a smaller percentage of responses to OO relatives than to SO and SS relatives, it is still one of the three main response categories for OO relatives.

In Table 4 we find a pattern for OS relatives similar to that for SS relatives and quite different from the distributional patterns of OO and SO relatives.

TABLE 4

Distribution of Responses to OS Relative Clauses
"The lion stands on the duck that bumps into the pig."

Age	Response Categories					
	Correct 12,23	12,13	12,31	12,32	21,23	Other
3.0-3.6	1	17	1	2	1	2
4.0-4.6	4	15	3	1	0	1
5.0-5.6	9	13	1	0	1	0
Totals	14	45	5	3	2	3
Percentage	19%	63%	7%	4%	3%	4%

In Table 4 the number of correct responses is lower than in any of the other tables, but it is clear that children have an effective set of rules for dealing with OS relative clauses. The response category 12,13 accounts for 63% of the responses; of the four types of relatives, this is the second highest percentage of responses falling in any one category and is surpassed only by the correct responses of Table 1. The overall number of response categories is low compared to the other types of relative clauses; there were eight response categories for OS relatives versus nine for the SS relatives, thirteen for OO relatives, and seventeen for SO relatives. The existence of a predominant response and the relatively low number of response categories indicate that children have a set of rules which is easily applicable to OS relatives even though it yields an incorrect interpretation in terms of an adult grammar.

From these four tables, we see that the response 12,13 was the predominant response to SS and OS relative clauses and that it was also a common response to SO and OO relatives. The OO and SO relative clauses did not have a predominant response; for each one, there were three response categories which accounted for most of the responses.

The emphasis in these tables has been on response patterns rather than on the percentage of correct and incorrect responses. In many ways the errors are the most interesting aspect of the research. Correct responses only tell us that the child responds as an adult would and tell us very little about the kinds of hypotheses a child makes to arrive at a correct response. By analyzing the predominant errors children make, we have an indication of the rules a child has at his disposal and of the way he goes about applying them to complicated structures.

3.2.1 Predominant Response Types

A very small number of response types account for nearly all of the relative clause responses. Six different responses account for 88% of the total number of responses to all four types of relative clauses. The remainder of the responses are distributed among nineteen different types of responses with an average of 1.8 responses per response category. In contrast, these six categories had an average of 42.2 responses per category; clearly, they are the most frequently used responses. Tables 5 through 10 show the distribution of these six most frequently used responses tabulated by relative clause type and by age. A comparison of the tables indicates that only one response type (12,13) was consistently used for all four relative clause types.

The most prevalent response by far is 12,13. Table 5 indicates the number of 12,13 responses for each relative clause type; 12,13 is a correct response for SS relatives and that fact is indicated at the top of the column.

TABLE 5

Distribution of 12,13 Responses

Age	Correct for SS	SO	OO	OS	Totals
3.0-3.6	18	5	4	17	44
4.0-4.6	16	7	5	15	43
5.0-5.6	22	9	5	13	49
Totals	56	21	14	45	137

The response 12,13 was used quite frequently for all types of relative clauses. It accounts for 48% of the total number of responses. If we consider just the patterned responses (that is, the six most commonly occurring ones), then the response 12,13 accounts for 54% of the total patterned responses. In particular, the response 12,13 was used for SS and OS relatives. These are relatives in which the subject of the embedded clause is relativized; it was a much less frequent response when the object is relativized, as in SO and OO relatives. We shall discuss this difference between relativized subjects and relativized objects in detail below.

Tables 6, 7 and 8 are tabulations of responses which were used only where they were correct; there is no over-generalization of these responses to other relative clause types. The response 12,32 was correct for OO relative clauses, and it was used almost exclusively for that type of relative. Similarly 12,23 was correct for OS relatives, and 21,13 was correct for SO relatives; each of these responses were used almost exclusively for the appropriate relative clauses.

TABLE 6

Distribution of 12,32 Responses

Age	SS	SO	Correct for OO	OS	Totals
3.0-3.6	0	1	8	2	11
4.0-4.6	0	0	9	1	10
5.0-5.6	2	0	10	0	12
Totals	2	1	27	3	33

TABLE 7

Distribution of 12,23 Responses

Age	SS	SO	OO	Correct for OS	Totals
3.0-3.6	2	0	0	1	3
4.0-4.6	5	0	1	4	10
5.0-5.6	0	1	2	9	12
Totals	7	1	3	14	25

TABLE 8

Distribution of 21,13 Responses

Age	SS	Correct for SO	OO	OS	Totals
3.0-3.6	0	5	0	0	5
4.0-4.6	1	6	1	0	8
5.0-5.6	0	4	0	0	4
Totals	1	15	1	0	17

These three response types were not overgeneralized to relative clause structures for which they were an incorrect response. This suggests that these particular responses are not the result of general rules which children attempt to use on complicated sentences which they do not fully comprehend. If these three responses were the result of a general hypothesis about English that a child brings to the comprehension task, we would expect some evidence of incorrect rule generalization such as we found in Table 5. Table 7, showing 12,23 responses, may provide some evidence of overgeneralization among the 4 year olds. There were 5 responses of 12,23 to SS relative clauses and 4 responses of 12,23 to OS relative clauses, but the pattern is clearly not as pronounced as the distribution in Table 5. These three responses were used primarily where they were correct and indicate a basically adult pattern of comprehension.

Tables 9 and 10 present the distribution of two response types which were not correct for any relative clause type, but which were frequently occurring errors for particular relative clauses.

TABLE 9

Distribution of 12,31 Responses

Age	SS	SO	OO	OS	Totals
3.0-3.6	1	0	5	1	7
4.0-4.6	0	0	6	3	9
5.0-5.6	0	0	5	1	6
Totals	1	0	16	5	22

TABLE 10

Distribution of 21,23 Responses

Age	SS	SO	OO	OS	Totals
3.0-3.6	1	4	0	1	6
4.0-4.6	1	5	0	0	6
5.0-5.6	0	6	0	1	7
Totals	2	15	0	2	19

The response 12,31 was used almost exclusively for OO relative clauses, and the response 21,23 was used primarily for SO relative clauses. Later in the chapter we will examine more closely why these particular errors are associated with these relative clause structures.

Two conclusions can be drawn from the above tables. First, they indicate that there are a very small number of response types (six) which account for most of the responses children gave to relative clauses, although the total number of actually occurring response types was rather large (25). Clearly, children are not just randomly associating noun phrases and verbs to assign a meaning to those relative clauses which they don't yet fully comprehend; instead, they have pre-existing rules which they attempt to use to arrive at an interpretation of the sentences. Secondly, they indicate that the response 12,13 was by far the most common response children gave; it is the only response type which was consistently used on all four types of relative clauses. The other five response types were associated primarily with only one type of relative clause. This heavy weighing in favor of 12,13 responses suggests that it is the result of a very productive system of rules.

3.2.2 Percentage of Errors

Tables 5 through 10 presented the distribution of the six most common response categories tabulated by relative clause type. The following table shows the percentage of errors tabulated by relative clause type that each response category accounts for. It also shows, in the last column, what percentage of the total number of errors each response category accounts for.

TABLE 11

Percentage of Incorrect Responses
Accounted for by Each Response Category

Response Category	SS (16)	SO (57)	OO (45)	OS (58)	Percentage of Total Number of Incorrect Responses
12,13	Correct	39%	31%	78%	46%
12,32	13%	2%	Correct	5%	3%
12,23	44%	2%	7%	Correct	6%
21,13	6%	Correct	2%	0	1%
12,31	6%	0	36%	9%	13%
21,23	13%	26%	0	3%	11%
Other	19%	32%	24%	5%	20%

The number in parentheses under each relative clause heading indicates the number of errors for that relative clause type. For example, there were 58 incorrect responses to OS relatives, but only 16 incorrect responses to SS relatives. Response type 12,13 accounts for 46% of the total number of incorrect responses, which is much more than any of the other response types. The next most frequent incorrect responses were 12,31 and 21,23 which accounted for 13% and 11% of the total number of incorrect responses, respectively. The other three response types, 12,32 and 12,23 and 21,13, which were correct responses to three of the four types of relative clauses, accounted for very few of the total number of incorrect responses. Response 12,23 accounted for a high percentage (44%) of the errors made on SS relative clauses. However, the 44% represents a total of only seven responses, and four of the seven responses were given by just two children. So until further research is conducted with a larger sample and with other structures, I hesitate to attach too much importance to this small number of 12,23 responses to SS relatives.

3.2.3 Variation by Age

An unexpected result of the research was the fact that an analysis of variance showed no significant variation in the total number of correct responses by age ($p > .468$). This may be partially due to the fact that younger children were selected to participate in the experiment who tended to be verbal and cooperative, and so children who were verbally sophisticated may constitute a large portion of the younger age groups. Since we were not particularly interested in discovering any variation by age but were primarily interested in eliciting responses which would indicate the rules children use, the lack of correlation between age and correct responses is not important for the current research. The absence of statistically significant variation by age indicates that there are wide differences across ages and that linguistic internal measures of development are more reliable indicators of linguistic progress than chronological age.

An analysis of variance indicates there was also no significant variation by age in the total number of 12,13 responses given ($p > .5$), indicating that the rules underlying this response type (that is, the conjoined-clause analysis) are productively used at all the age levels tested.

3.2.4 Summary of Distributions

The several ways of looking at the data, which were presented in the above Tables, all indicate that response 12,13 was the most important response in terms of giving us information about children's rules. The overall distributions (Tables 1 through 4) show that 12,13 was the predominant response to SS and OS relative clauses, indicating the presence of an easily applicable set of rules to these structures. The relatives for which 12,13 is a correct response (SS relatives) are easiest for children to correctly interpret. The fact that the set of rules was easily applicable to SS and OS relatives also reduced the total number of response categories for these two types of relatives. Children guess less frequently because they have an analysis which can easily give them an interpretation of the sentence. Response 12,13 was also a common response to SO and OO relatives, but was not a predominant response to them. There was no other response category which predominated for these relatives, and the lack of any single set of rules which was easily applicable to them shows up in a greater number of response categories for OO and SO relatives than for SS and OS relatives. Not only was 12,13 the predominant response to SS and OS relatives, and the only response which was used consistently for all four types, but it accounts for the greatest percentage of errors made on all of the relatives (Table 11). Thus the overall distributions of responses and the percentage of incorrect responses accounted for by each response category indicate the presence of a productive set of rules which manifests itself in 12,13 responses.

3.3 Parsing

The evidence presented in the preceding sections supports the hypothesis that children attempt to parse a multiple-clause sentence whose structure they are uncertain of into conjoined simplex sentences, resulting in a 12,13 response. This section discusses how the conjoined-clause analysis applies to each relative clause type to produce a 12,13 response, and how lexical information about the subcategorization of verbs may aid a child in establishing simplex sentence boundaries.

The linear order of NP's and V's in SS relatives is exactly the same as the order in conjoined sentences where the subject of the second clause is missing. An SS relative, schematically represented as (25a) would be parsed as (25b).

25. a. NP $\underline{\text{that}}$ V NP V NP

b. ${}_{S}[{}_{S}[\underset{\uparrow}{\text{NP}}\ \underline{\text{that}}\ \text{V NP}]\ {}_{S}[\triangle\ \text{V NP}]]$

The noun phrases and verbs are grouped into simplex sentences. The SS relative is parsed the same as conjoined clauses, and the same rule which identifies the subject of the first clause as the referent of the missing subject in conjoined clauses may also be used productively to determine the referent of the missing subject in (25b). So the missing subject of the second clause is interpreted as being coreferential with the subject of the first clause.

Lexical information about the verbs in both the matrix and relative clause can aid a child in determining sentence boundaries for each simplex sentence. As soon as the first verb in the string is processed and the child determines that it is transitive, he knows that the subcategorization feature +[NP___NP] will be satisfied when a second NP is processed. When the second NP is processed and the subcategorization feature is satisfied, an S boundary, closing off the substring, can be successfully postulated. Since the child has grouped the first NP V NP sequence into a simplex sentence, there is no overt NP available to fill the subject position in the subcategorization feature of the second verb. In the surface string the NP immediately preceding the second verb has been used to fill the object position of the subcategorization feature of the first verb. The child must analyze the remaining V NP sequence as a clause which lacks an overt subject, and must seek a rule which will provide an antecedent for the missing subject. The most accessible and easily applicable rule is the already existing one used to determine the referent of a missing subject in the second conjunct of conjoined sentences.

It is not clear what interpretation is assigned to the relative pronoun <u>that</u>. It may be that it is left uninterpreted. It is not uncommon for children to ignore those elements of a sentence which they don't fully comprehend. This often shows up in children's imitations, where they eliminate or restructure those elements they don't understand (cf. Slobin and Welsh 1973 and M. Smith 1974).

The evidence from children's comprehension of SS relatives presented here and the evidence from production of SS relatives is not consistent. Table 1 indicates that SS relatives are very easy for children to correctly comprehend; however, evidence from production indicates that sentences in which the matrix subject is the head of the relative clause are spontaneously produced later than sentences in which the head is the matrix object (Menyuk 1969:16). So we have a situation where SS relatives are correctly comprehended early in the language development of the child, but spontaneously produced rather late. Since SS production is late, we conclude that the child's grammar probably does not generate SS relatives at an early stage, and so the percentage of correct responses to SS relatives is not due to children's competence in relative clause formation.

We propose that the high percentage of correct responses to SS relatives is primarily due to the fact that a conjoined-clause analysis is easily applicable to them since the linear order of the elements is exactly the same in both SS relatives and conjoined sentences. There is a high percentage of correct responses because the conjoined-clause analysis happens to provide the correct semantic interpretation for SS relatives.

In the distribution of Table 4 we saw that the percentage of 12,13 responses to OS relatives is nearly as high as the percentage for SS relatives, and there is no significant difference between their means ($t=1.157$, $df=23$, $p>.15$). Children treat OS and SS relatives as though they were structurally identical. The schematic representation of OS relatives (26a),

26. a. NP V NP <u>that</u> V NP

    b. $_{S}[_{S}[\underset{\uparrow}{\text{NP}}\ \text{V NP}]\ \underline{\text{that}}\ _{S}[\triangle\ \text{V NP}]]$

shows that the linear order of NP's and V's is the same as the order of atomic predicates in SS relatives and conjoined simplex sentences. Using the conjoined-clause analysis, children parse the OS relative as shown in (26b), with the missing subject of the second clause interpreted as coreferential with the subject of the first conjunct. This parsing produces a structure which is semantically interpreted the same as SS relatives. Given sentences such as

27. a. The cow that kisses the pig hits the horse. (SS)
    b. The cow kisses the pig that hits the horse. (OS)

most children act out both of them with toy animals by having the cow kiss the pig and the cow hit the horse.

Lexical information can help a child determine simplex sentence boundaries in the same way that it could be utilized in interpreting SS relative clauses. The correspondence between the thematic relations and the subcategorization feature of the lexical entry is the same as the correspondence between the two in the surface structure. The only difference is that the subject of the second verb is missing in surface structure. As discussed above, children have an easily applicable rule which determines a referent for the missing subject.

We propose that SS and OS relatives are given identical interpretations by most children because the linear structure of both sentences is easily susceptible to interpretation by the conjoined-clause analysis. The high percentage of 12,13 responses to these two types of relatives provides evidence for children's use of this set of rules.

Response type 12,13 was also commonly given to OO and SO relative clauses, but less frequently than to the two types discussed above. We suggest that it was not a predominant response because the structure of OO and SO relatives makes it more difficult to easily apply a conjoined clause analysis to them. The lienar order of elements in OO relatives is not the same as the order in conjoined sentences:

28. NP V NP <u>that</u> NP V

In OO relatives the subject of the second clause is present, but the object is "missing." Despite this difference between OO relatives and conjunctions, many children still impose a conjoined-clause analysis on OO relatives. Table 3 indicates that 19% of the OO relatives received a 12,13 response from the children. The conjoined-clause analysis imposes the following constituent structure on OO relatives:

29. $_{S}[_{S}[\underset{\uparrow}{NP}\ \ _{VP}[V\ NP]]\ \underline{that}\ _{S}[\underset{|}{\triangle}\ _{VP}[NP\ V]]]$

Children still analyze the second clause as lacking a subject, even though one is present in the string, and interpret the missing subject as coreferential with the subject of the first conjunct. The NP immediately preceding the verb of the second clause is assigned object status. Clearly the linear structure of OO relatives makes them much less suitable for the conjoined-clause analysis.

The lexical entry for the verb in the first clause can help children in establishing sentence boundaries because the linear order of NP's in the subcategorization feature is isomorphic with the initial noun-verb-noun sequence. The first clause was easy for children to parse, and 89% of the responses correctly interpreted the first clause. However, the remaining noun-verb sequence will not satisfy the subcategorization requirements of the second verb; one of the nouns in the subcategorization feature is not overtly expressed in the string. If the child analyzes the existing NP as the subject, then he must find a rule which will determine a referent for the missing object; we will discuss these responses in detail later in this section. Our claim is that many children will interpret the NP preceding the second verb as the object instead of the subject, because the conjoined-clause analysis is a readily accessible set of rules which can provide an interpretation for missing subjects, and children do not have equally accessible rules for assigning a referent to missing objects. The difficulty of imposing the conjoined-clause analysis on OO relatives can account for the fact that they are the first relative clause type for which children drop the 12,13 response and replace it with a more appropriate set of rules.

The final type of relative clause, SO relatives, also have a linear structure which is different from the linear order of conjoined clauses:

30. NP $\underline{\text{that}}$ NP V V NP

The initial sequence of nouns and verbs does not directly correspond to the subcategorization feature for transitive verbs; there are two NP's which precede the first verb, and the second verb has an NP following it and no NP immediately preceding it. These features make it difficult to apply a conjoined-clause analysis to SO relatives; however, 31% of the SO relatives still received 12,13 as a response. The conjoined-clause analysis can be imposed on the linear structure as follows:

31. $$\begin{tikzcd}[column sep=small] {}_S[{}_S[\text{NP} & \underline{\text{that}}\ \text{NP V}]\ {}_S[& \triangle \arrow[ll, bend left] & \text{V NP}]] \end{tikzcd}$$

The sentence-initial NP is interpreted as the subject of the first clause, and the missing subject of the second clause is interpreted as coreferential with the subject of the first clause.

One of the problems that children have with SO relatives is determining which NP of the first clause is the subject; 51% of the responses to SO relatives had the NP immediately preceding the first verb as the subject, and 46% of the responses had the sentence-initial NP as the subject. There was no significant difference by age in the choice of NP as subject, contrary to what we might have expected from the results for cleft sentences reported by Bever (1970a).[7]

For SO relatives there is another pattern of responses besides 12,13 which provides evidence for the conjoined-clause analysis, and these are the 21,23 responses. 21% of the SO relatives were interpreted by the children as 21,23 responses. As discussed above, children who do not yet fully comprehend these relatives are unsure which NP in the first clause is the subject. The 12,13 responses discussed above were given by those children who chose the sentence-initial NP as the subject of the first clause; the 21,23 responses were given by children who chose the second NP as the subject of the first clause. Both groups parse the SO relative as indicated in (20), but those children giving 21,23 responses choose the NP immediately preceding the verb of the first clause rather than the sentence-initial NP as subject. Both groups interpreted the missing subject of the second clause as coreferential with the subject of the first clause. Combining the 12,13 and the 21,23 responses results in a total of 52% of the SO relatives which were interpreted using the conjoined-clause analysis.

We have shown how the conjoined-clause analysis applies to each type of relative clause and results in a 12,13 reponse to all the relatives as well as a 21,23 response to SO relatives. The conjoined-clause analysis accounts for 56% of the total number of patterned responses (the six main response types). For SS relatives, it gives the correct semantic interpretation, but for the other relative clause types, it yields an incorrect interpretation. We also saw that those relatives to which it was easily applicable because their linear order was identical to the linear order of conjoined sentences (i.e. SS and OS relatives) have a very high percentage of 12,13 responses and very little scatter in terms of the number of different responses given to them. Those relative clauses whose linear order of elements differs from that of conjoined sentences (OO and SO relatives) receive fewer 12,13 responses and have a greater number of different responses given, indicating that children still attempt to use the conjoined-clause analysis in comprehending them and that children do not have another easily applicable set of rules which can replace the conjoined-clause analysis.

3.4 A Hierarchy

In this section the relative clauses are arranged into a hierarchy according to children's retention of the conjoined-clause analysis for each type. This section also provides evidence supporting Hypothesis B that children attempt the most direct projection of thematic relations onto the NP's in the surface string and that when there is no direct correspondence between the thematic relations and the subcategorization features of the lexical entry and the thematic relations and the order of NP's in the surface string, children will have a great deal of difficulty imposing an interpretation on the string.

An interesting way to look at the responses of individual children to the four relative clause types is to group the children according to their use of the conjoined-clause strategy. If we stipulate that at least three children must perform similarly to form a group, then there are four main groups into which 18 of the children fall. The remaining six children are distributed among four different groups. The four main groups are presented in the following chart:

32.

	Response 12,13 Given to			
	SS	OS	SO	OO
Group I (N = 3)	X	X	X	X
Group II (N = 3)	X	X	X	
Group III (N = 6)	X	X		
Group IV (N = 6)	X			

Group I consists of children who gave at least 2/3 of their responses as 12,13 to each relative clause type. They did not discriminate among the different types of relative clauses in their responses and used the conjoined-clause analysis for all types. Group II consists of children who gave 2/3 of their responses as 12,13 to SS, SO, and OS relatives but not to OO relatives.

Group III consists of 6 children who gave 2/3 of their responses to SS and OS relatives as 12,13 but did not use this strategy for SO and OO relatives. The final group consists of 6 children who gave 2/3 of their responses as 12,13 to only the SS relatives and did not use the conjoined-clause analysis on any other relative clause types. In this last group there were 4 children who gave adult responses to all four relative clause types. They appear to have an adult grammar for relative clause interpretation. There were also 2 children who correctly responded to SS relative clauses, but gave incorrect responses to the other three types.

There were six children who did not fall into any of the above groups. Two children did not give 12,13 responses to any of the relative clauses. Both of these children had a great deal of difficulty with the relative clauses and did not give consistent answers to any of the four types. Two of the children gave 12,13 to OS relatives but not to SS relatives as the hierarchy would predict. They each gave one 12,13 response to SS relatives, but one child gave the other two responses as 12,23 and the other child gave them as 12,32. One child gave 12,13 only to OO relatives. The sixth child gave 12,31 as a response to both SS and SO relatives but not to OS relatives as predicted by the hierarchy. She did give one of the three responses to OS relatives as 12,13, but the other two responses were 12,23. These children's responses are different from those given by the rest of the children and do not represent a frequently-occurring pattern.

Chart (32) shows that if a child gives a 12,13 response to OO relatives, then he gives it to all the other relative clause types also; these are the children in Group I. If a child gives a 12,13 response to SO relatives, then he gives it also to SS and OS relatives; these are the children in Group II. Children who give 12,13 responses to OS relatives also gave them to SS relatives. The final group consists of children who gave 12,13 as a response only to SS relatives. These different patterns of responses can be used to arrange the relative clause types into a hierarchy:

33. SS
OS
SO
OO

If a child gives 12,13 responses to one particular type of relative clause, he will also give that response to any relative clause type above it in the hierarchy. For example, if a child gives 12,13 responses to OO relatives, then he also gives that response to SO, OS and SS relatives. Another way of looking at the hierarchy is to say that the relative clause type on the bottom of the hierarchy, the OO relatives, are the first ones eliminated from the conjoined-clause analysis. The SO relatives are eliminated next, and finally the OS relatives.

The characteristic which cause a particular relative clause type to be eliminated from the conjoined-clause analysis is the extent to which the relative clause deviates from the pattern of conjoined simplex sentences. As mentioned above, the linear order of NP's and V's in OO and SO relatives diverges from the order of these elements in conjoined clauses. The remainder of this section discusses the factors operating to eliminate SO and OO relative clauses from the conjoined-clause analysis and proposes an account of the (still incorrect) responses which replace the 12,13 response for OO and SO relatives.

In conjoined clauses and in SS and OS relatives there is a direct correspondence between the linear order of elements in surface structure and the order of elements in the lexical entry of each verb. The subcategorization of each verb can be directly superimposed on the surface string with the exception of the subject NP of the second verb. There is no NP in the string which can be matched up with this part of the second verb's subcategorization feature. This gap can be filled by using the conjoined-clause analysis which parses the string into conjoined simplex sentences and interprets the missing NP as coreferential with the subject of the first clause. Schema (34a-c) show how the subcategorization feature of a transitive verb maps directly onto conjoined sentences, SS relatives and OS relatives. The missing subject in conjoined clauses and the 'gaps' in relative clauses are represented by '$\triangle$'.

34. a. Conjoined Sentence: $_S[_S[$NP V NP] $\underline{\text{and}}$ $_S[\triangle$ V NP]]

Subcategorization features: +[NP ___ NP] +[NP ___ NP]

b. SS Relative: $_S[$NP $_S[\underline{\text{that}}$ V NP] $\triangle$ V NP]

Subcategorization features: +[NP ___ NP]+[NP ___ NP]

c. OS Relative: $_S[$NP V NP $_S[\underline{\text{that}}$ $\triangle$ V NP]]

Subcategorization features: +[NP ___ NP] +[NP ___ NP]

In OO and SO relatives this kind of correspondence does not hold:

35. a. SO Relative: $_S[$NP $_S[\underline{\text{that}}$ NP V $\triangle$] $\triangle$ V NP]

Subcategorization features: +[NP ___ NP] +[NP ___ NP]

b. OO Relative: $_S[$NP V NP $_S[\underline{\text{that}}$ NP V $\triangle$]]

Subcategorization features: +[NP ___ NP] +[NP ___ NP]

When the subcategorization feature of the first verb in SO relatives is matched with the surface string, there is no NP which matches the object NP in the subcategorization feature of the first verb. There is no easily applicable rule which will assign a referent for the missing element as there is for SS and OS relatives. The same problem arises with OO relatives. No NP in the surface string directly matches the object NP in the subcategorization of the second verb. For both SO and OO relatives there is no easily accessible strategy which will provide an interpretation for the "missing" object.

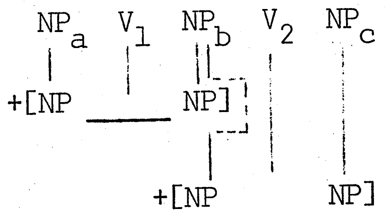
The claim that children have difficulty comprehending SO and OO relatives may seem to be contradicted by the fact that OO relatives received the second highest number of correct responses. Closer inspection of the responses to each relative clause type reveals that the ease of imposing an interpretation, whether correct or incorrect, is quite different for OS and SS relatives as opposed to SO and OO relatives. Response 12,13 accounts for 78% of the responses to SS relatives and 63% of the responses to OS relatives. In contrast, the highest percentage of OO and SO responses accounted for by a single response category is 38% for OO relatives and 31% for SO relatives. So it was relatively easy for children to provide an interpretation for SS and OS relatives, although it was an incorrect one for the latter, and more difficult for children to interpret OO and SO relatives.

Children's responses suggest that they do not simply superimpose lexical entries onto a string without any attempt to analyze the string structurally. The claim is that children do not just line up the NP's in the subcategorization frame with the NP's in the linear order of the string. Rather we propose that children analyze the string into simplex sentences, assigning structure to the string. They do not attempt to fill in the subcategorization frame without structurally analyzing the string. If they did do this, we would expect children to produce many more 12,23 responses. If a multiple clause sentence contains two transitive verbs, such as (36) does,

36. Linear String

Subcategorization of V_1:

Subcategorization of V_2:



and if children are merely trying to line up the subcategorization features with a surface string, we would expect NP_b to be the object of V_1 and also the subject of V_2. This response does occur as a correct response to certain verbal complements and to OS relatives. However, it is not a common response to other multiple-clause sentences, accounting for only 6% of the total number of responses to relative clauses other than OS relatives, and does not occur at all as a response to conjoined clauses.

In the lexical entries of the verbs used in these experiments the first NP of the subcategorization feature bears the thematic relation of agent to the verb, and the second NP is the goal. So in conjoined sentences, the first NP of each simplex sentence is the agent and the second NP is the goal.

In the linear order of OO relatives there is an NP which immediately precedes the second verb and which may be assigned to subject position in the subcategorization of the verb. We propose that this fact precipitates the rejection of the conjoined-clause analysis for OO relatives. The conjoined-clause analysis requires a missing subject in the second clause. In OO relatives the second clause does not have a missing subject, but instead has a "missing" object. The relative ease with which children identify the pre-verbal NP as the subject may be due to the integrity of the NP-V sequence. Sixty-eight percent of the responses correctly identified the NP immediately preceding the second verb as the subject of that verb. Of this 68%, 55% correctly identified the matrix object as the referent of the "missing" object (a 12,32 response), and 32% incorrectly interpreted the subject of the first clause as the referent of the missing object. Given a sentence such as

37. The horse hits the pig that the sheep kisses.

the latter children had the horse hit the pig and the sheep kiss the horse (a 12,31 response).

Tom Roeper has pointed out to me that children's 12,31 responses to 00 relatives can be accounted for by the Universal Control Principle (p. 14,(18)). If, as proposed, children analyze the sentence as conjoined clauses rather than as embedded clauses, the "missing" object of the second clause is in Class II because the clause as a whole is attached to the top-most S. The object of the first clause is lower than the missing NP because the object NP is attached to the VP and so is at the Class I level. So the subject NP is the only NP which satisfies the Control Principle. This point is also made by Solan and Roeper (this volume).

However, the pattern of responses is actually more complicated than the preceding discussion suggests. Nearly half of the children (42%) used a combination of 12,32 and 12,31 responses to 00 relatives, and there were no children who gave only 12,31 responses to these relatives. In terms of the analysis presented here, this pattern of responses can be accounted for in two ways. One way is to say that children are uncertain about the attachment node of the clause--whether it is attached to the top S or to the VP node. In the former case only the subject of the first clause would satisfy the Universal Control Principle; in the latter case the object of the first clause would also satisfy it. A second possible explanation is to say that children are analyzing the clause as attached to the VP, but because there are two NP's (the subject and the object of the first clause) which are higher than the missing object and because children have not yet mastered relative clause formation, they vacillate between the subject and object as an antecedent for the missing form.

There is some reason to prefer the first alternative because it imposes tighter restrictions on language learning theory. It has been suggested [8] that the theory be stated so that S-attachment of a clause implies subject control of missing NP's and VP-attachment implies object control. Thus once a child knows where to attach a clause, he has a rule for assigning an antecedent for missing NP's. This formulation would need to be refined, perhaps along the lines suggested in this study, to account for the fact that children find some clauses which are S-attached in their grammars (such as conjoined clauses) much easier to interpret than others (such as 00 relatives). However, at this point, the first alternative seems preferable as an account of children's vacillation between 12,31 and 12,32 responses to 00 relatives.

Sheldon (1972) has suggested that 12,31 and 12,13 responses to 00 relatives results from an extraposition strategy. "From their [children's] knowledge that in English some relative clauses modify the matrix subject they are overgeneralizing that all do." (p. 77, the emphasis is Sheldon's). There are two criticisms of this proposal. One is that we know that children produce object relatives earlier than subject relatives (Limber 1973), but there does not seem to be overgeneralization from object relatives to subject relatives. That is, because a child knows that some relative clauses modify the matrix object, he does not overgeneralize that all of them modify the matrix object. This suggests that some learning principle besides just generalization is operating in children's interpretation of 00 relatives.

Secondly, if children are in fact overgeneralizing subject modification, they should get SO relatives right. Under the Extraposition Hypothesis children who interpret (38a) as (38b)

38. a. The cow jumps over the horse that the dog kisses.
    b. The cow that the dog kisses jumps over the horse.

presumably are making the assumption that (38b) underlies (38a) and that the relative clause has been extraposed from its head, "the cow" to produce the surface string (38a). Thus, the interpretation of (38b) is one step in the interpretation of (38a). We would expect children who incorrectly comprehended OO relatives because of extraposition to get SO relatives correct. However, this is not the case with either Sheldon's data or with mine.

TABLE 12

Relationship of 12,31 Responses to OO Relatives
and Correct Responses to SO Relatives

Age Groups	Number of 12,31 Responses to OO Relative Clauses	Number of Correct Responses to SO Relative Clauses
3.0-3.6		
S1	1	0
S2	1	0
S3	1	1
S4	1	0
S6	1	0
4.0-4.6		
S12	1	1
S13	2	0
S15	1	0
S16	2	1
5.0-5.6		
S17	1	0
S19	1	0
S21	1	0
S22	1	2
S24	1	0

Column 1 lists the children who gave at least one 12,31 response to OO relatives by age group and by child number. There were 10 children who gave no 12,31 responses to OO relatives. Column 2 gives the number of 12,31 responses to OO relatives for each child. Column 3 shows the number of correct responses to SO relatives for these children who gave at least one 12,31 response to OO relatives.

A glance at Table 12 reveals that there is no similarity between the number of 12,31 responses and the number of correct SO responses, and this is confirmed by a t test of significance. The t test shows a significant difference between the two means (t=2.858, df=13 p=.01). These results indicate that children are not just treating SO relatives as a step in the interpretation of OO relatives.

In Sheldon's data there were 23 out of 33 children who made one or more errors on OO relatives which were labelled Extraposition errors (p. 83), but there were only 8 of these children who answered one or more SO relatives correctly. [9] Sheldon's data include 12,13 responses as instances of Extraposition, and so any correlation between the strategy of Extraposition and the number of correct responses to SO relatives should show up in her data since the number of response categories which could count as evidence of Extraposition was greater than mine. However, we find no such evidence in the data she presents of an interaction between OO responses and SO responses.

If the Extraposition explanation were correct, we would have a situation where children interpret an extraposed relative as modifying the matrix subject, but when the relative clause is actually present next to its head as an unextraposed relative, children incorrectly interpret it. That is, they incorrectly interpret SO relatives although SO relatives are the presumed underlying source for the interpretation of OO relatives and so should be correctly interpreted. For the above reasons we conclude that children do not use an Extraposition strategy.

We have discussed the way that the linear order of elements in OO relatives differs from the order in conjoined sentences. The presence of a preverbal NP which may fill the subcategorization slot for subject may account for the rejection of the conjoined-clause analysis. A child must then complete the subcategorization entry by finding a referent for the missing object. There is no generally applicable rule for determining the antecedent of a missing object, unlike the rule available for determining the antecedent of a missing subject, aside from an adult understanding of relative clauses. Children's uncertainty about the structure of the sentence and the nature of the applicable rules is reflected in the vacillation between 12,31 and 12,32 responses by individual children.

The other relative clause type which differs from the order of elements in conjoined clauses is the SO relative. As pointed out above, in these relative clauses the first clause has two NP's preceding the verb:

39. NP <u>that</u> NP V V NP

These were the most difficult relatives for children to comprehend. Not only did they correctly answer only 21% of them, but the scatter of the responses was quite great. One aspect of the interpretation was easy for children and that was the determination of the object of the second clause. Eighty-nine percent of the sentences correctly had the third NP of the sentence as the object of the second verb. This indicates that children have independent principles for assigning subjects and objects. They do not have to assign a subject with any degree of certainty in order to determine the object of a transitive verb. In SO relatives there is an NP immediately following the second verb, in the position corresponding to the post-verbal NP of the subcategorization feature. The correspondence between the subcategorization feature and the surface structure can account for the fact that it was easy for children to identify the object, although it is difficult for them to identify the subject.

As we would expect, children have problems determining which NP is the subject and which is the object of the first verb. In the lexical entry for a transitive verb the NP immediately preceding the verb will fill the subject position of a deep structure tree, and this NP will also be the first NP in the linear order of elements in the deep structure tree. In the surface order of elements in SO relatives, both the subject and object precede the verb. This creates a conflict between two ways of identifying subjects: as the first NP of the sentence and as the NP immediately-preceding the verb. In the deep structure tree these two strategies will select the same NP, but in SO relatives, they select different NP's. There is a great deal of individual inconsistency in the choice of subject. Only three children gave the same response to all three SO relatives, and everyone else gave inconsistent responses. Children do not exclusively choose either the sentence-initial NP or the NP immediately preceding the verb, but vacillate between the two in responding to SO relatives.

Unlike the case with OO relatives, there is not a strong impetus for eliminating SO relatives from the conjoined-clause analysis. In the OO relatives, the NP preceding the second verb satisfies two characteristics of subject status in English, clause-initial position and pre-verbal position, providing a strong impetus to choose it as the subject and reject the conjoined-clause analysis. However, the lack of correspondence between the order of elements in the lexical entry and their order in the surface structure of SO relatives makes it difficult to apply the conjoined-clause analysis to them. The large number of idiosyncratic responses (eleven) to these relatives suggests that many children realize that the conjoined-clause analysis is not appropriate but do not have an alternative set of rules which can easily provide an interpretation.

The last relative clause to be eliminated from the conjoined-clause analysis is the OS relative. We suggest that they are the last ones not because they are structurally more difficult than OO or SO relatives, but because it is so deceptively easy to apply the conjoined-clause analysis to them. As illustrated in (34) the linear order of elements in the lexical entries and surface strings of OS relatives and conjoined clauses are identical, and we propose that this isomorphism induces children to use the conjoined-clause analysis on OS relatives after they have rejected it for OO and SO relatives. Until a child totally masters the rules involved in correctly comprehending relative clauses, he will continue to use the conjoined-clause analysis on OS relatives because the order of elements in them is the same as the order specified by the conjoined-clause analysis.

By arranging children's responses in terms of exclusion from the conjoined-clause analysis, we have been able to suggest which relative clauses are excluded first and the reason for their exclusion. It is worth noting that the order of relative clauses established in this way is different from the order which results if the relative clauses are ordered according to the number of correct responses. The latter order, from most correct to least correct, is SS, OO, SO and OS. The relative clause which receives the fewest correct responses is the last to be excluded from the conjoined clause analysis. The hierarchy established in (33) reflects the ease of imposing an interpretation on sentences containing various relative clause types and does not correspond to the number of correct responses.

One further point should be made about the children's response to relatives. There is no direct evidence that children interpret a restrictive relative clause as a restriction of the head noun rather than a nonrestrictive comment about it. For example in OO relatives such as (40a) a better description than (40b) of the interpretation involved in children's 12,31 responses might be (40c):

40. a. The cow jumps over the horse that the dog kisses.
b. The cow that the dog kisses jumps over the horse.
c. The cow jumps over the horse and the dog kisses the cow.

Certainly from children's movement of the toy animals involved in this experiment we cannot infer any more than that.

3.5 Other Analyses of Relative Clauses

Sheldon's (1972) research was one of the first to systematically study all four types of relative clauses. She studied the comprehension of children aged 3.8 to 5.5 years and found an order of difficulty in correctly interpreting relative clauses which was nearly identical to that found in the present study. Sheldon found that SS relatives were easiest for children to correctly interpret, followed, in order of number correct, by OO relatives, then OS relatives and finally SO relatives. The order according to the number correct was the same for the present study except that there was one more correct response to SO relatives than there was to OS relatives, and so the order of the last two is interchanged. Sheldon proposed the notion of parallel function to account for the fact that SS and OO relatives were easiest for children to correctly comprehend. In both these types of relatives the function of the head of the relative and the function of the relativized item are the same in their respective clauses. In SS relatives they are both subjects and in OO relatives they are both objects. Sheldon suggests that this parallelism makes the sentences particularly easy for children to comprehend.

In the present study emphasis has been placed on the difference in children's responses between "easier" and "more correct". We have suggested that OS relatives are easier for children to interpret than OO relatives even though the interpretation assigned to OS relatives is incorrect. We distinguish between assigning an interpretation (correct or incorrect) and assigning a correct interpretation. There were more correct responses to OO relatives than to OS relatives, but the evidence suggests that the latter were easier for children to assign an interpretation to.

The conjoined-clause analysis provides an explanation for children's ability to easily comprehend SS relative clauses and also provides an explanation for the fact that OO relative clauses receive the second highest number of correct responses. SS relatives are easily comprehended because the conjoined-clause analysis is easily applicable to them and also results in the correct response. The conjoined-clause analysis is not easily applicable to OO relatives, as discussed in some detail in an earlier section, and OO relatives are the first type of relative clause eliminated from the conjoined-clause analysis. Children are forced to find another strategy for OO relatives and for most children this results in a mixture of correct (12,32) and incorrect responses. Thus the conjoined-clause analysis provides an account of children's responses without adopting the notion of parallel function.

In my data parallel function accounts for 66% of the total number of responses. But of these responses, 81% have parallel subjects while only 19%

have parallel objects. Most of the evidence supporting parallel function comes from parallel subjects rather than being equally divided between parallel subjects and parallel objects. The ease of interpreting parallel subjects can be accounted for if we suppose that children are using the conjoined-clause analysis.

Parallel function is not independently necessary in a child's grammar,[10] and there is no evidence that it is operating in adult grammars.[11] Relative clauses with parallel function do not seem any easier to a native adult speaker of English than relative clauses with nonparallel function. For example (41a) does not seem any easier to interpret than (41b):

41. a. The man saw the boy who the girl kissed.
    b. The man saw the boy who kissed the girl.

The conjoined clause analysis is independently necessary in both child and adult grammars to account for conjoined simplex sentences. The conjoined-clause analysis plus the learning principle of generalization, a well-documented psychological process, suffice to account for children's responses.

Legum (1975) studied comprehension of relative clauses by kindergarten, first grade and second grade children. He found the same general order of difficulty in comprehending relative clauses as the present study and as Sheldon's study. Legum found that the most common error in the second clause of a sentence containing a relative clause was what he calls "the bird-in-the-hand" strategy. He suggests that a child is using this strategy when he chooses the subject of the first clause as the subject of the second clause as well, that is, when he gives a 12,13 response or a 21,23 response to SO relatives. He points out that it is an empirical question whether or not the "bird-in-the-hand" strategy is incorporated into the grammar, but suggests that this strategy is "merely a task-related problem solving strategy and not a linguistic strategy per se" (p. 30). In fact, a task-oriented strategy, such as the "bird-in-the-hand" strategy, cannot account for the different degrees of difficulty of the four relative clause types, whereas a linguistic approach, such as the conjoined-clause analysis does account for this data.

The two approaches make different predictions about children's responses to multiple-clause sentences. A task-oriented strategy, which may be formulated as "take the subject of the first clause and use it as the subject of the second clause," cannot account for the differences in children's responses to relative clauses, which vary according to the linguistic structure of the relative clause. The problem-solving strategy predicts that there will be no differences in children's utilization of the "bird-in-the-hand" strategy related to the internal linguistic structure of the sentence since it is explicitly a nonlinguistic strategy. However, both Legum's data and my data do not support this prediction.

Table 5 shows that children gave fifty-six 12,13 responses to SS relatives, forty-five to OS relatives, twenty-one to SO relatives and fourteen to OO relatives. Including the 21,23 responses to SO relatives (Table 8), makes the total for SO relatives thirty-six. Clearly, the number of "bird-in-the-hand" responses varies greatly depending on the internal structure of the relative. Such differences are accounted for by the linguistic approach proposed here, the conjoined-clause analysis. We have argued that children initially utilize the conjoined-clause analysis to interpret multiple-clause sentences which their grammar does not generate, but that the analysis is rejected earliest for those relatives (SO and OO) whose order of elements is different from the order in conjoined-clauses and maintained for those (SS and OS) whose order is isomorphic with the order of conjoined clauses. Thus the greater number of 12,13 responses to SS and OS relatives (101 responses), compared with the number of such responses to OO and SO relatives (50, including the 21,23 responses to SO relatives), can be accounted for by the conjoined-clause analysis.

Legum's data also show differences in the use of the "bird-in-the-hand" strategy according to the structure of the relative clause. In his study of a total of 60 kindergarten, first and second grade children, he found 177 parallel subject responses to SS relatives, 112 to OS relatives, 107 to SO relatives and only 13 to OO relatives (pp. 9 and 13). Obviously, his data also indicate that children's use of the "bird-in-the-hand" strategy varies according to the linguistic structure of the sentence.

Sheldon (1972) also provides support for a linguistic account of the data. She included in her study conjunctions in which the subject of each conjunct was lexically filled. She tested children's comprehension of sentences such as

42. The horse stands on the pig and the duck kisses the horse.

If children utilize a task-oriented strategy of having the same animal as subject of both clauses, regardless of the actual linguistic structure of the sentence, we would expect children to have "the horse" as the subject of both clauses in (42) even though "the duck" is the expressed subject of the second clause. Sheldon does not give a breakdown of children's errors on these sentences; however, she does report that "the high scores on the conjunction experiment rule out the possibility that... a strategy of moving the same toy twice... was important in the relative clause results." (p. 61).

Thus we conclude that children are not utilizing a task-oriented strategy when they select the subject of the first clause as also the subject of the second clause and that these responses provide evidence for the operation of the conjoined-clause analysis and for its productivity as a linguistic set of rules.

3.6 Summary of Relative Clause Data

Relative clauses provide persuasive evidence for the existence and productivity of the conjoined clause analysis. The only response used across all four relative clause types is 12,13 and it accounts for the greatest percentage of errors for the three types to which it was an incorrect response. We have shown how this set of rules applies to each type of relative clause and what characteristics of the relative clause structure may lead to rejection of the conjoined-clause analysis. We suggest that the reason children find conjunction an easy operation is that there is a direct correspondence between the linear order of elements in the subcategorization feature of the lexical entry for the verb and the linear order of elements in the surface string.

We saw that relative clauses whose linear order of elements differs from the order of elements in conjoined clauses were more difficult for children to assign an interpretation to than relatives whose order conformed to the order in conjoined clauses. There was a much greater variety of responses to OO and SO relatives, which do not conform to the order in conjoined clauses, than to SS and OS relatives, which do conform. We have proposed that this difference in the relative difficulty of assigning a consistent interpretation to a string is due to the lack of direct correspondence between the correlation of thematic relations and subcategorization features in the lexical entry and the correlation of thematic relations and the order of NP's in the surface structure.

4. Sentences with Verbal Complements

4.0 Introduction

Sentences with verbal complements also provide evidence for children's use of the conjoined-clause analysis in interpreting structures which are not yet secure in their grammars. One of the most important studies of children's comprehension of sentences with verbal complements is C. Chomsky's (1969) study of children's responses to complements of "promise" and "tell." She found that many children interpret the missing subject of the complement to "promise" as the matrix object rather than the matrix subject. For example given her stimulus sentences (p. 40):

43. Bozo $\left\{\begin{array}{l}\text{promises}\\ \text{tells}\end{array}\right\}$ Donald to do a somersault. Can you make him do it?

the child makes Donald do the somersault in both cases. One child, while making Donald do the somersault said, "I promised you you could do a somersault." Chomsky analyzes this interpretation of the "promise" sentences by the children as an application of the Minimum Distance Principle (MDP)[12] to a structure containing a verb which is in the appropriate semantic class for application of the MDP but is an exception to it. She suggests that children use a general strategy in cases such as (43) that the understood subject of the complement sentence is the NP most closely preceding it. Children apply the same principle to "promise" to determine the complement subject as they apply to complements of "tell." Thus many children treat "promise" and "tell" alike in determining their complement subjects.

In my study I also found that children treat "promise" and "tell" the same, but, in addition to those children who selected the NP most closely preceding the gap as the implicit complement subject, there was also a group of children who chose the first NP of the matrix clause as the complement subject. In sentences such as Chomsky's example (43) above, they would select Bozo as the actor doing the somersault in both cases.

This is a surprising departure of the facts from the commonly held assumption that the MDP is a strategy acquired very early by children. Chomsky interviewed children who were somewhat older (5 to 10 years) than the children in the present study, and so this difference in responses may be accounted for by the fact that her analysis is based on responses from children at a more sophisticated level of development than the level of children in the present study.

The following sections show the distribution of responses for verbal complements and indicate how the results provide support for the conjoined-clause. The results also suggest a developmental sequance for the acquisition of the conjoined-clause analysis and the MDP strategy. Finally it is proposed that the MDP is not a surface structure principle but is the result of semantic features in the lexical entry of certain verbs.

4.1 The Sentences

As with the relative clause experiment, children were given three tokens of each sentence type to act out with toy animals. Examples of the sentences used are the following:

44. a. The lion tells the pig to stand on the horse.
    b. The pig promises the rabbit to jump over the duck.

Sentences such as (44a) will be called MDP sentences since a correct response selects the NP most closely preceding the missing complement subject as the referent of the subject.

In scoring the children's responses to the MDP and "promise" sentences, children were not required to act out the first clause of the sentence to have their responses counted as correct. "Telling" and "promising" were not prominent as actions which should be carried out by the animals, and many children did not act them out. Since the primary interest of the study was children's choice of a complement subject, the fact that only the complement clause was enacted in many cases was not considered important for the analysis of the results.

4.2 Distribution of Responses

The responses are coded in the same way as the relative clause responses; each NP is numbered according to its occurrence in the linear order of the string. The responses in the following tables show only the NP's that were used to act out the complement clause of each sentence. A response of "23" to (44a) means that the child had the pig stand on the horse. A response of "13" means that the lion stands on the horse.

Individual children's responses indicate that they were quite consistent in their choice of complement subjects for MDP complements and "promise" complements. Table 13 shows the four categories into which the children fall when classified on the basis of 2/3 of their responses to "promise" and MDP verbal complements.

TABLE 13

Cross Tabulation of Children
According to Choice of Complement Subject
for "Promise" and MDP Verbal Complements

Response Categories for MDP Complements	Response Categories for "Promise" Complements	
	Response 23	Response 13
Response 23 (N=13)	8	5
Response 13 (N=10)	2	8

There was one child who did not give 2/3 of his responses to MDP verbal complements as either 23 or 13, and he is not represented in Table 15. All the other children fell into one of the four categories in the table.

There were five children who correctly responded to at least 2/3 of both the "promise" and MDP verbal complements, and these children seem to have an adult grammar for both types. There were 8 children who chose the matrix indirect object as the subject of both the "promise" and MDP complements; these children responded as many of C. Chomsky's children did. Eight children chose the subject of the first clause as the implicit subject of both types of verbal complements. Two children gave at least 2/3 of their responses to MDP verbal complements as 13 and at least 2/3 of their responses to "promise" complements 23. They reversed the correct responses for the two types. Aside from these exceptional children, it is clear that those children who have not completely mastered the adult system for these verbs interpret the missing subject in both "promise" and MDP complements the same. A chi-square test of significance carried out on Table 15 indicates that the relationship between the two types of complements is statistically significant ($x^2=3.968$, df=1, p=.05). If a child who has not yet mastered the adult system responds to MDP complements in one way, we can predict with a high degree of reliability that he will respond to "promise" complements in the same way.

The breakdown by age of children in the four cells of Table 13 reveals that there is a tendency for 3 year olds to give 13 responses to both types and for 4 and 5 year olds to give 23 responses to both types:

TABLE 14

Breakdown by Age of Children
According to Choice of Complement Subject
for "Promise" and MDP Verbal Complements

Response Categories for MDP Complements	Response Categories for "Promise" Complements	
	Response 23	Response 13
Response 23	3 years - 2 4 years - 3 5 years - 3	3 years - 1 4 years - 1 5 years - 3
Response 13	3 years - 1 4 years - 1 5 years - 0	3 years - 4 4 years - 2 5 years - 2

These differences reflect the kind of distribution we would expect if the conjoined-clause analysis, which would select the subject of the first clause as the referent for the subject of the second clause, is a set of rules which is acquired earlier than the rule which selects the matrix indirect object as the referent.

Tables 15 and 16 show the distribution of responses to MDP sentences and "promise" complements, respectively.

TABLE 15

Distribution of Responses to MDP Verbal Complements

Age	Response Categories		
	23	13	Other
3.0-3.6	11	11	2
4.0-4.6	14	9	1
5.0-5.6	17	7	0
Totals	42	27	3
Percentage	58%	38%	4%

TABLE 16

Distribution of Responses to "Promise" Complements

Age	Response Categories		
	23	13	Other
3.0-3.6	10	12	2
4.0-4.6	13	11	0
5.0-5.6	9	14	1
Totals	32	37	3
Percentage	44%	51%	4%

A comparison of Tables 15 and 16 shows that the distribution of responses for "promise" and MDP complements is nearly identical for 3 and 4 year olds. They do not discriminate between "promise" and "tell." For 5 year olds, however, there is quite a difference between the two tables. Five year olds begin to differentiate between MDP verbs and "promise." A t test combining 3 and 4 year olds versus 5 year olds shows a significant difference in the number of correct responses to MDP verbal complements and "promise" complements (t=1.902, df=22, p=.05). Developmentally, it appears that children begin to be aware of the difference in interpretation of MDP complements and "promise" complements at about 5 years of age.

This age differentiation dovetails nicely with Chomsky's results with children 5 years and older who correctly determined the implicit subject in complements to "tell" but frequently misconstrued the implicit subject of "promise" complements as the matrix object rather than the matrix subject. The 5 year olds in her study correctly interpreted "tell" complements, which was also true of most of the 5 year olds in my study (6 out of 8 responded correctly), but the children in my study who were under 5 years of age frequently misinterpreted MDP verbal complements and chose the matrix subject as the referent of the missing complement subject. For example one 3 year old, while acting out (45) by having the lion talk to the pig and then bump into the horse,

45. The lion tells the pig to bump into the horse.

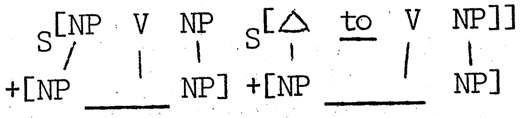
said, "The lion says, I'm going to bump into the horse.'"

The results of Chomsky's study and my study suggest a developmental trend in children's comprehension of verbal complements. Children first interpret missing complement subjects as coreferential with the subject of the preceding clause, and, as we have discussed in detail, this choice is the result of parsing the multiple clause string into conjoined clauses and using the same rule to determine the subject of the second clause as is used in determining the referent of the missing subject in conjoined clauses. We suggest that children correctly interpret "promise" complements at this stage, not because they understand them as lexical exceptions to the choice of the matrix indirect object as complement subject, but because the conjoined clause analysis provides the correct interpretation for them. The child next goes through a stage where he correctly interprets "tell" complements, but incorrectly interprets "promise" complements. Following Clark (1973) we account for this as follows: at this stage a child knows that a certain semantically based class of verbs, to which "tell" belongs, selects the matrix indirect object as the referent of the complement subject, but the child does not identify "promise," which is semantically a member of this class, as an exception to the general rule which selects the indirect object as the controlling NP. Many of the children in Chomsky's study are in this stage. This analysis of infinitival complements is at variance with the results found by Goodluck and Roeper for participial complements. The discrepancy between the two sets of results defines an area which clearly needs further research. In the final stage the child has an essentially adult grammar for both verbs and correctly interprets the verbal complements of both.

4.3 Parsing

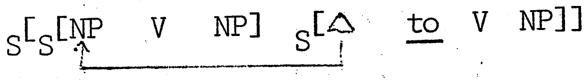
The parsing of these sentences containing verbal complements is quite straight-forward. The order of elements in them is the same as the order of elements in conjoined simplex sentences, and so the conjoined-clause analysis is easily applicable to them. There is also a direct correspondence between the order of NP's in the surface string and their order in the lexical entry of the matrix and complement verbs:

46.



So it is very easy for children to apply the conjoined-clause analysis and parse the sentence as conjoined simplex sentences:

47.



The missing subject of the second clause is interpreted as being coreferential with the subject of the first clause.

4.4 The Minimum Distance Principle

Although we have discussed complements to verbs in the tell class as though a principle such as MDP were operating in the choice of a complement subject, some reflection on how such a principle would operate and a consideration of children's responses suggest that it is a lexical semantic feature rather than a principle applicable to the surface string which is operating in the selection of an antecedent for a missing complement subject.

If the MDP were a productive strategy which was formulated as

48. Choose the closest preceding NP as the referent of a missing subject.

we would expect some evidence of overgeneralization of the principle to strings which are identical to MDP sentences in the linear order of their elements. If a child is using MDP to determine the implicit subject of (49a), we would expect him to overgeneralize to (49b) at some stage:

49. a. The lion tells the pig to stand on the horse.
    b. The lion jumps over the pig to stand on the horse.

In the following section we present data which shows that children do not overgeneralize from (49a) to (49b).

Generalization of rules to inappropriate forms is a common phenomenon in child language. In morphology Berko (1958) showed that children overgeneralize regular inflectional endings to classes which are exceptions to the general rule. In syntax the conjoined-clause analysis is overgeneralized to relative clauses and sentences with verbal complements. The fact that we find no evidence for overgeneralization of the MDP to strings which correspond to (49b) in surface structure suggests that the rules which determine the coreferent of the missing subject operate at a level other than surface structure.

Recent analyses (Bresnan 1976) of these verbal complements in adult grammars suggests that verbs such as "tell" and "promise" are marked in the lexicon as requiring the matrix indirect object or the matrix subject as the implicit subject of their complement clauses. If children's interpretation of these complement subjects is the result of lexical semantic features, the fact that we do not find any overgeneralization of the MDP would be accounted for. Features on individual lexical items would never apply to structures not containing that item. So we would not expect children to choose the closest preceding NP as the referent of the missing subject in (49b) because the matrix clause does not contain a verb which is lexically marked for such a controller.

Maratsos (1974b) also provides evidence that a surface structure MDP is not adequate as an account of children's choice of an antecedent for missing subjects. He found that children who correctly interpreted simple passives also correctly interpreted both of the following complex sentences:

50. a. Bill told John to get in.
    b. John was told by Bill to get in.

If children were utilizing a surface structure MDP, we would expect them to select Bill in (50b) as the antecedent of the missing complement subject. However, children did not give this kind of response.

Maratsos suggests a semantic role theory to account for the children's responses. He proposes that children utilize a rule which always selects the goal (John in (50)) as the referent of the missing NP for the class of verbs which includes tell. Children's misinterpretation of complements to promise would be accounted for by saying that the child has not learned that promise is an exception and selects the source rather than the goal as the antecedent of the missing NP.

Maratsos' proposal is consistent with the above suggestion that control is lexically marked. Thematic relations are established in the lexicon, and control could be marked either in the lexical entry of each verb or through a redundancy rule applying to a particular semantic class.

Both Bresnan's and Maratsos' proposals are examples of how control might be handled in the lexicon to replace the use of a surface structure principle.

4.5 Summary of Data from Verbal Complements

Children's responses to verbal complements indicate that the conjoined-clause analysis is used by all age groups but is particularly productive among 3 and 4 year olds. Children use it for both "promise" and "tell" complements to select the subject of the first clause as the referent of the missing subject of the second clause. The fact that 3 and 4 year ols have more incorrect responses to "tell" complements than 5 year olds indicates that the conjoined-clause analysis is developmentally prior to the MDP strategy. We proposed that the lack of generalization of the MDP strategy may be due to the fact that no such strategy exists but that children learn lexical features on the verbs which determine the complement subject chosen. If this proposal is correct, then the pattern of age-differentiated responses indicates that at an early stage of development of the child, verbs are not subcategorized for sentential complements and are not marked with semantic features which indicate the referent of missing subjects. All verbs may occur in simplex sentences, which may be conjoined to form multiple-clause sentences. If the subject of the second clause is missing, it is interpreted as coreferential with the subject of the first clause. Subcategorizing verbs and marking them for particular controllers in complement clauses develops later then the conjoined-clause analysis.

Goodluck and Roeper (this volume) have also found evidence consistent with the proposal that subcategorization develops later than the conjoined-clause analysis. They differentiated two groups of children on the basis of the children's interpretation of participial complements to perception and non-perception verbs. In sentences such as

51. a. John hit Mary wearing the watch.
    b. John saw Mary wearing the watch.

there was one group of children which selected John as the subject of the participial phrase in both (51a) and (51b). The second group of children subcategorized all verbs for participial complements and selected Mary as the subject of the participial phrase in both cases.

Thus Goodluck and Roeper found evidence for overgeneralization of the features which subcategorize a verb for a participial phrase--the stage of object control. In contrast to this pattern for participial phrases, evidence will be presented in the next section which suggests that children do not overgeneralize these features which subcategorize a verb for infinitival complements beyond the appropriate semantically based class of verbs.

These two sets of results lead us to conclude that after the initial period of subject control, children quickly develop a sophisticated subcategorization system which differentiates between participial complements and infinitival complements.

5. Sentences with "In Order To" Clauses

Children's responses to sentences with "in order to" clauses, such as

52. The lion jumped over the pig to stand on the horse.

also provide support for the existence of the conjoined-clause analysis. An adult response to (52) interprets the subject of the first clause as the implicit subject of the "in order to" clause, and 89% of the sentences were also interpreted in this way by the children.

In the initial stages of this experiment we expected some children to use the MDP on sentences such as (52). It was anticipated that a common error children might make would be to select the NP most closely preceding the missing subject of the second clause as the referent of the missing subject because of the high degree of similarity in the surface structure of sentences with MDP complements and sentences containing "in order to" clauses. Contrary to expectations, this was a very uncommon response. As discussed in the preceding section, this pattern of responses may be accounted for is the choice of a complement subject is not due to the operation of a surface strategy but results from application of lexical semantic features.

The following table shows the uniformity of children's responses to these sentences:

TABLE 17

Distribution of Responses to Sentences
Containing "In Order To" Clauses
'The sheep jumps over the horse to kiss the lion.'

	Response Categories		
Age	12,13	12,23	Other
3.0-3.6	19	2	3
4.0-4.6	22	1	1
5.0-5.6	23	0	1
Totals	64	3	5
Percentage	89%	4%	7%

Clearly, these sentences were very easy for children to interpret correctly. The 12,23 responses, which would provide evidence for the operation of a Minimum Distance Principle, represent an insignificant proportion of the total number of responses.

We propose that children did well on these sentences because the conjoined-clause analysis is so easily applicable to them. The order of elements in these sentences is the same as the order in conjoined clauses, and so they can be parsed as conjoined simplex sentences:

53. $_{S}[_{S}[\text{NP} \;\; \text{V} \;\; \text{NP}] \;\; _{S}[\triangle \;\; \underline{\text{to}} \;\; \text{V} \;\; \text{NP}]]$

There are, no doubt, children who correctly responded to sentences such as (52) because they interpret them semantically and syntactically as an adult would comprehend them, but we cannot attribute an adult grammar for these sentences to all the children who correctly responded since there was an alternative, and probably more basic, set of rules which also produced the correct response.

A comparison of the responses to these sentences and the responses to sentences with verbal complements shows no evidence that children subcategorize verbs other than those in the tell class for infinitival complements. There were eight children (Table 13) who selected the matrix indirect object as the NP coreferential with the missing subject for both tell and promise. These children have subcategorized a semantically based class of verbs (verbs depicting speaking actions) for infinitival complements. These complements are attached to the VP and so have object control. However, these same children gave a total of only two 23 responses (an object control response) to sentences containing "in order to" clauses, and these two responses were given by the same child. The other seven children gave no 23 responses to "in order to" clauses. This distribution supports the proposal that children do not overgeneralize subcategorization for infinitival complements to verbs which are not in the semantic class of speaking verbs. However, children do overgeneralize within the class to promise which is an exception in its selection of controller for the complement subject.

In summary, we have presented evidence supporting the hypothesis that children initially analyze multiple-clause sentences which they have not yet mastered as conjoined simplex sentences and interpret the missing subject of the second conjunct as coreferential with the subject of the first clause. Support for this hypothesis was found in children's responses to conjoined sentences, to four types of relative clauses, to sentences with verbal complements and to sentences with "in order to" clauses. For all these different sentences, the response 12,13 formed a substantial proportion of the responses, and we argued that this reponse was generated by the system of rules we have called the "conjoined-clause analysis." We have proposed that the conjoined-clause analysis consists of a set of rules for assigning a string conjoined structure and for interpreting missing noun phrases. A modified version of Williams' control principle was adopted as a restriction on children's choice of an antecedent for a missing noun phrase.

Conclusion

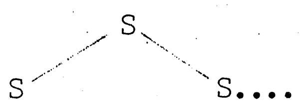
One of the implications of the conjoined-clause analysis is that the Language Acquisition Device is quite narrowly restricted in the early stages of development. We propose that the conjoined-clause analysis is a universal feature of language acquisition and that all children will initially analyze multiple-clause sentences as conjoined clauses. This proposal greatly limits the number and kind of possible hypotheses a child must entertain in constructing a grammar for his language.

A second implication of the conjoined-clause analysis is that it suggests an order of development for recursive rules in the grammar. We propose that recursion in a child's grammar occurs initially through iteration rather than embedding. The first recursive rule which introduces sentences would be

54. $S \longrightarrow S^*$

which would be structurally realized as

55.



When a child's grammar begins to generate multiple-clause sentences, they are generated by rule (54) and have the structure shown in (55). This proposal is consistent with various data (Menyuk 1969; 1971; Brown 1973; Limber 1973) concerning children's early production of multiple-clause sentences and is also corroborated by the evidence presented in this study.

Developmental priority of iterative rules would constrain the kinds of assumptions a child brings to the language learning task at an early stage and thus restrict the number of possible grammars. Such a restriction is desirable in a theory of language acquisition. Principles which limit the number of possible grammars at a particular stage can account for the fact that children learn their language at a phenomenally rapid rate. At each stage of development children can utilize far-reaching principles, to reduce the number and kind of hypotheses they make about new linguistic data. Although this kind of restriction on language learning will produce a great many incorrect interpretations, as we have seen, it is also a very powerful hypothesis with wide applicability and can be used to structurally analyze a great many strings.

Footnotes.

1. For a fuller discussion of language universals and their relationship to language acquisition see Chomsky (1965, 1968, 1976).

2. See Jackendoff (1972) for a more detailed discussion of the structure of lexical entries. Following Jackendoff I have included the subject as part of the subcategorization frame.

3. See Kimball (1973) for a discussion of the notion 'closing off a string.'

4. This method of coding is adopted from Sheldon (1972).

5. The use of this principle as an explanation for children's ease in determining the referent of a missing subject in conjoined clauses was called to my attention by Tom Roeper and is elaborated in Solan and Roeper (this volume).

6. This suggestion was also discussed and elaborated by the participants in the 1976 summer research group at the University of Massachusetts and by Solan and Roeper (this volume).

7. Bever found that 2 year olds analyzed a "noun-verb" sequence in cleft sentences as an "actor-action" unit. They did not just take the first noun of the sentence as the actor, but analyzed it as part of a gestalt. At 4 years of age, the child no longer relies on a "noun-verb" gestalt, but interprets the first NP of the sentence as the actor. Bever's results on age differentiation in the interpretation of clefts are not supported by the evidence from children's comprehension of relative clauses.

8. This suggestion was made by Tom Roeper and elaborated in discussions of the 1976 Summer Research Group at the University of Massachusetts.

9. On page 83 the heading for subject relatives indicates that the numbers represent the number of incorrect responses; however, in the paragraph at the head of the page, it says that the subject relative scores are the number of correct responses. Comparing the scores on pages 83-85 with the chart on page 86, it appears that the figures for subject relatives do represent the number of correct responses, and so I have taken them as that.

10. A study of children's comprehension of relative clauses by de Villiers, Flusberg, Hakuta and Cohen (n.d.) extended Sheldon's study of parallel function to include sentences in which the head and relativized NP were both indirect objects. These constructions were quite difficult for children to correctly interpret. Children's responses provide evidence that parallel function does not always operate to facilitate comprehension and casts further doubt on the viability of the parallel function hypothesis.

Solan (1975) provides evidence that parallel sequencing is a factor in children's interpretation of infinitival complements. This differs from Sheldon's parallel function in that the latter refers to the deep structure grammatical function of NP's while the former deals with surface structure ordering of NP's.

11. Sheldon (1974b) found no statistically significant effect of parallel function in adult responses to relative clauses. She found that interruption of the main clause by the relative clause (SS and SO) and relativizing the object NP (SO and OO) did result in statistically significant differences in performance. The increase in errors in sentences with relativized objects is also reported by Cook (forthcoming, cited in Keenan and Comrie, 1977). Thus evidence from adult experiments indicates that parallel function is not a factor in facilitating comprehension.

12. The Minimum Distance Principle was first proposed by Rosenbaum (1967) as a principle governing deletion of complement subjects in the surface structure of certain embedded sentences. C. Chomsky and others have used the terminology as a convenient way of referring to the recognition strategy used by children in interpreting such sentences. Following C. Chomsky, in our discussion of the MDP, we will mean a general recognition strategy rather than a principle governing deletion.