DOCUMENT RESUME

ED 253 097 FL 014 839

AUTHOR Smith, Carlota S.; van Kleeck, Anne
TITLE Linguistic Complexity and Performance.
PUB DATE 84
NOTE 37p.; Paper presented at the Annual Boston University Conference on Language Development (Boston, MA, October 12-14, 1984).
TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Child Language; *Difficulty Level; *Language Acquisition; *Language Processing; *Language Proficiency; Learning Processes; *Sentence Structure

ABSTRACT

An experimental study investigating the interaction of linguistic complexity and performance in child language acquisition tests the hypothesis that children learning a first language acquire relatively complex sentences somewhat later than less complex sentences. In one of three tests, the subjects, 44 children aged 3.6 to 6 years, were presented with a series of structurally diverse sentences about toys within reach. The children's comprehension of the sentences was observed in their reactions as measured by their movements of the toys. The second and third tests used two imitation tasks focusing on adverbial structures, using the same toy-moving sentences, varying in length, and toy-moving sentences carefully controlled for length, to find the subjects' critical sentence length for imitation. Analysis of the various responses to sentence types and complexity in all three tests supported the "weak complexity" hypothesis and suggest that the linguistic factors affecting performance are task-dependent and deny a simple one-to-one relationship between complexity and performance. (MSE)

Linguistic Complexity and Performance

Carlota S. Smith and Anne van Kleeck
University of Texas

1.

This paper discusses an experimental study investigating the interaction of linguistic complexity and performance in child language acquisition. Although the role of linguistic complexity in performance is not well understood, two quite plausible assumptions about it are often made. It is often assumed that linguistically complex sentences involve more effort in production and understanding than do less complex sentences; and that children learning a first language acquire relatively complex sentences somewhat later than less complex sentences[1]. The study reported here was designed to test the second of these ideas. It is based on a linguistic theory of complexity developed by the first author (Smith, forth- coming).

In their strongest form, the plausible assumptions in question make a rather simple hypothesis about the relation between linguistic complexity and performance - namely, that the one affects the other in a consistent manner that does not vary with the type of performance. One might expect a weaker, less simple relation. It seems likely that certain factors of complexity affect performance more than others; and that not all types of performance are affected in the same way by linguistic complexity. We will refer to these as the strong and weak forms of the Complexity Hypothesis.

In this study young children were presented with sentences that varied

[1] other things, particularly cognitive things, being equal.

in linguistic complexity using two tasks, toy moving and imitation. These tasks make quite different demands of a subject. If the strongest version of the Complexity Hypothesis is correct, the effect of linguistic complexity would be similar for both tasks; if a weaker version holds, different types of complexity might be dominant in the different tasks.

Our theory of complexity is a modular one, identifying several factors that contribute to the linguistic complexity of a sentence. The present focus is on syntactic and semantic structure: the factors of systematic complexity, surface structure complexity, and interpretive complexity are taken to be most important[2] . These factors are assessed for individual sentences, resulting in a complexity profile. The assessment is based on the derivation of a sentence in the grammar of the language. The basic idea is simply that sentences are complex when their derivation involves a relatively long path - in terms of operations, a large number of steps - through the grammar of a language.

The theory is not unrelated to the ill-fated Derivational Theory of Complexity (DTC); indeed, it is an attempt to use what was interesting about that theory while avoiding the shoals on which it foundered. An important different between our theory and the DTC is that we focus on surface and semantic structure rather than the rules and intermediate structures that were essential to the DTC(cf. Fodor & Garrett 1967). Further, our treatment of interpretive complexity adds an important dimension.

---

[2]In its present form the theory does not include morphology or phonology, but it can very naturally be extended to them

highest NP and PP nodes in surface structure.[5]

Interpretive complexity deals with the difference between surface syntactic and scopal semantic structure, or Logical Form. It arises when the two structures are not one-to-one: in such cases interpretive rules are involved in the grammatical derivation of the sentence[6] . Sentences with empty categories in surface structure, that must be interpreted, are relatively complex in this sense. Quantifiers and other elements that have wide scope also add to interpretive complexity; they involve rules relating surface structure to scopal semantic structure. Interpretive complexity is crudely assessed, at this stage of the theory, by the number of interpretive rules in the derivation of a sentence; a more sensitive method of assessment may be developed later.

A sentence may be relatively simple at one level, and relatively complex at another. Perhaps the most complex sentences are high on all levels. There is a certain trade-off between levels, in that a short sentence with a reduced surface structure - that is, a structure that involves empty categories - requires interpretation. Consider 3 for example, which contains an empty category indicated by [e][7]

---

[5]Bounding nodes in Government Binding theory. See Smith, forthcoming, for discussion.

[6]Assuming a linguistic theory that deals with semantic interpretation, such as Government Binding theory

[7]We assume an analysis of this sentence in which before introduces a reduced dependent clause.

3. John arrived before Mary [e]

This sentence has few terminal elements: in surface structure complexity it is relatively low. The empty category requires interpretation, so the sentence has some interpretive complexity.

1.2 Structures Used in the Study

Three related adverbial structures that differ in complexity, and whose linguistic properties are quite well understood[8] were chosen as the focus of our study. In these sentences adverbs introduce full and reduced surface clauses, allowing comparison between them. The examples illustrate.

4. John called Mary before Bill invited Sue

5. John called Mary before inviting Sue

6. John called Mary before Sue

These sentences vary in interpretive complexity: the first requires no interpretive rules besides the standard rules for tense and aspect, irrelevant here (see Smith, forthcoming, for discussion); the second requires that one empty category be interpreted; the third requires that two empty categories be interpreted. (We can say this much without committing ourselves as to how the interpretation is made.) We will refer to the different structural types according to the structure of the dependent clause: S, Ving, and NP.

---

[8]They are discussed at length in Smith 1978.

The sentences above also differ in length, of course. In our terms they differ in amount (number of terminal nodes). Since they have the same density the amount difference is decisive for surface structure complexity. In surface complexity the order of the structures is reversed: the full clause is most complex because it is longest, Ving shorter, the NP structure least complex in S-structure because it is shortest. They have the same systematic complexity.

The strong version of the complexity hypothesis does not differentiate between types of complexity, treating them in a blindly additive way. This version would predict little or no difference between sentences with reduced and full adverbial clauses. The reduced sentences have a higher degree of interpretive complexity, which is balanced by their lower amount, hence lower surface structure complexity. The NP structures are ambiguous, however, which may contribute to their complexity in a way not captured by our complexity measure[9].

Weaker versions of the hypothesis would predict different complexity effects in the toy-moving and imitation tasks, according to their demands on the subject. Very generally, toy-moving focuses on interpretation and might be expected to be sensitive to interpretive complexity; imitation requires production of a surface structure and might be sensitive to surface complexity.

Section 2: The Toy-Moving Experiment

---

[9]The question of ambiguity is discussed below, and in Appendix C.

Toy-moving tasks are frequently used in studies of linguistic competence, because they show how subjects understand the basic roles talked about in a sentence[10]. Toy-moving is most useful with structures requiring interpretation to fill these roles, e.g. control structures or sentences with pronouns (cf the work of Solan 1978, Carol Chomsky 1969, etc). Toy-moving tasks have an important limitation: they do not indicate the grammatical structure that a child ascribes to a sentence. This is clear from the extensive literature on sentences with relative clauses, in which different researchers give different interpretations to identical toy-moving responses. It is claimed that children ascribe to the same examples conjunction structure (Tavakolian 1981), extraposed relative clause structure (Sheldon 1974), and differently attached relative clauses (Solan & Roeper 1978). For the structures we are working with, toy-moving tasks are quite informative. Both Ving and NP dependent sentences have phonetically empty elements, requiring interpretation of just the type mentioned above. Children's reponses in a toy moving task indicate whether they know the adult interpretation to these structures; and what other interpretation, if any, they give to them.

Sentences with adverbials of the S, Ving, and NP type were presented to children in a toy-moving experiment. The main variable was the structure of the adverbial sentence. A second variable was transitivity, of main and dependent clause. The stimulus sentences varied in length due

---

[10]These roles are the theta roles of GB theory: agent, patient, etc.

to differences in both structure and transitivity.[11] In terms of the theory, they varied in both interpretive and surface structure complexity. The test sentences thus confound two factors, amount and structure. Note that confounding could not have been avoided by holding constant the length factor. If we had lengthened the single NP sentences, the lengthening material would have made some sentences different from others. Two temporal connectives (before and after) were used; this variation was lexical only and did not affect systematic complexity[12] . The specific types of sentences are indicated in the table below. This task has been used repeatedly in work on temporal connectives, varying the connectives and order of main and dependent clause: e.g. Amidon & Carey 1972, Barrie-Blackley 1973, Clark 1973, Coker 1978, Feagans 1974, French 1977, Keller-Cohen 1975, Johnson 1975, and others. The form of the dependent clause has not before been varied, however. This body of work gives a good background for interpreting our results: in particular with respect to order of mention and the different connectives.

If interpretive complexity strongly affects performance in this task, then we expect the NP structures to be most difficult. If amount is important, we expect NP structures to be least difficult; and we expect sentences with transitive clauses to be more difficult than sentences with

---

[11]These variables were chosen because of their naturalness in the test structures, where different forms of adverbial sentences appeared. The complexity factors of density and systematic complexity were held constant in order to keep the stimulus sentences to a manageable number.

[12]Systematic complexity is unaffected because these two connectives have the same syntactic constraints.

7.

Structural Variations in Toy-Moving Sentences

Dependent clause: Sentence

NVN before/after NVN

NV before/after NVN

NVN before/after NV

Dependent clause: Ving

NVN before/after Ving N

NV before/after Ving N

NVN before/after Ving

Dependent clause: NP

NVN before/after NP

intransitives.

The stimulus sentences were presented to 44 children, ranging in age from 3.6 to 6. The sentences talked about four animals (cow, horse, elephant, tiger) and actions that were easy to act out (jump up, sleep, tickle, chase, kiss). Throughout, the same four small toy animals were visible and within reach. Sentences with S-structure adverbials talked about three or four animals; the other types talked about two or three animals, depending on the transitivity variable. The S and Ving structures were tested six times each: thrice with before and thrice with after. The NP structures were tested twice with each connective. Two filler sentences with the connective when were also presented. The sentences are given in Appendix A. The sentences were presented in two blocs. The experimenter presented the sentences, and an assistant noted the child's actions. If

the child asked for the sentence to be repeated, the experimenter complied without comment.

In order to determine the influence of age on performance, we divided the children into three groups on the basis of age, as follows:[13]

8. Subjects grouped by age

Group III:	5;0 - 6;0 years;	18 children
Group II:	4;6 - 5;0 years;	11 children
Group I:	3;6 - 4;6 years;	15 children

2.2 Response categories

There were five initial response categories: Correct, Reverse, Single Clause, and Error. Reverse responses reverse the correct clause order and have no other error. For example, to the sentence The cow jumped after the horse fell a Correct response would have the horse fall and then the cow jump; a Reverse response would have the cow jump and then the horse fall. Reverse responses with additional errors were categorized as Errors; there were very few such responses. Single Clause reponses are those of single actions, whether or not mentioned in the stimulus sentence. (All sentences talked about two actions.) All responses with fewer than two elements from the stimulus sentence were considered Failures.

---

[13]The numbers are somewhat uneven. A few children dropped out of the experiment. Further, five subject groups were collapsed to three. There were originally five age groups consisting of children within each 6-month range between 3;6 and 6;0. When the data were analyzed there were no age differences between the two oldest and the two youngest groups. They were therefore collapsed together for the sake of both statistical power and simplicity of presentation.

Correct and Reverse Reponses were combined into a single category, Correct. This step was taken because Reverse responses apparently indicated only the effect of the after connective and reverse order of mention; the experimental variables did not affect the pattern of Reverse responses. It is known that order of mention, and temporal connective, affect performance (Clark 1973, Keller-Cohen 1975). Children and adults find it easier to deal with sentences in which the order of mention mirrors the actual order of events: sentences with before are easier than sentences with after. In fact the latter are often misinterpreted; generally before is acquired earlier.

2.3 Results

Our main finding is that the children knew something about all three structures, but the degree of control differed. Interpretive complexity strongly affected the results, as predicted by the weaker, task-sensitive version of the complexity hypothesis. S structures were best controlled, Ving structures next, NP structures least. We discuss here the pattern of responses; the statistical treatment is presented in Appendix B.

2.3.1 Correct Responses.

As shown in the table below, the percentage of Correct responses was strikingly similar for S and Ving sentences, a little over 50% for the group as a whole. The NP structures received significantly fewer correct responses, 37%. (This is true for a Correct response category that includes both Correct and Simple Reverse responses. Otherwise there is a consistent difference between sentences with before and after, the former receiving many more Correct responses than the latter.)

9. Toy Moving: Responses by Structure (% of all Responses)

	Correct	Single Clause	Error
S	58	22	20
Ving	55	22	23
NP	37	45	18

2.3.2 Single Clause Responses

The pattern of Single Clause responses gives additional information. We take it that such responses indicate difficulty with the stimulus sentence, pace Hamburger and Crain 1982[14] . The children generally responded to the stimulus sentences with two actions, although all structures have the same presuppositions. The Single Clause response was a minimal response in this experiment; only rarely did the children refuse to act out anything at all. NP sentences tended to elicit many more Single Clause responses, especially from the younger children (see below for age-related data). This suggests that S and Ving structures were about the same in difficulty, and that NP structures were more difficult. In discussing the error types we will also argue that the S structures were

[14]The work of Hamburger and Crain with relative clauses suggests that pragmatic and age factors may make an important difference in children's responses to sentences with embedded clauses involving presuppositions.

better controlled than the Ving structures.

The structures that elicited relatively many Correct responses elicited few Single Clause reponses, and vice versa, as the table above indicates.

There was a significant relation between age and success. An analysis of variance indicated that the oldest children gave a significantly greater number of correct responses than the two younger groups (see Appendix B for discussion of the statistical analysis). The oldest children also gave significantly fewer Single-Clause responses than the other two groups. The percentages of correct and Single-Clause responses are given in the tables below.

10. Correct Responses

(Correct category combines Correct & Reverse, as noted above)

	Group III	Group II	Group I
	5;0-6;0	4;6-5;0	3;6-4;6
S	74	38	35
Ving	70	36	36
NP	56	32	20

11. Single Clause Responses

S	8	43	27
Ving	2	39	40
NP	24	52	57

Amount, another variable in the stimulus sentences, did not affect the responses in a consistent manner. Shorter S and Ving structures, which contained intransitives, did not pattern with any particular type of response. The shortest sentences were NP structures, which were the most difficult.

2.3.2 We turn now to the category of Error responses. The most frequent errors differed for each structural type, so the structural types were considered separately. Looking at the responses in this way, we found evidence that the children's knowledge of the Ving and NP structures was less solid than their knowledge of the S structure. We develop this evidence below.

S structure errors. To appreciate the error patterns it is useful to recall the main characteristics of the S structure sentences. They have complete main and dependent clauses, involving three or four animals and two different actions. All the material of the S structure is overt: subjects must remember the nouns and verbs and act them out in the correct relation to each other. There is nothing that corresponds to the interpretation of one or more empty categories required for the other two structures.

Most of the error responses correctly presented both actions talked about in the stimulus sentences; but one or more of the NPs was incorrect. Two types of errors predominated: Doubling, in which one animal does double duty, participating in two actions where the sentence mentioned two different animals; and Interchange, in which roles are changed, within or across clauses. Many responses had multiple errors; almost all errors

involved the animals rather than the actions.

Doubling errors all have the effect of reducing the number of animals involved in the action: one animal is used for two roles. Four different Doubling strategies were identified: A) the Subject of Clause 1 is the same as the Subject of Clause 2, noted Subj 1 = Subj 2; B) Obj 1 = Subj 2; C) Obj 1 = Obj 2; D) Subj 1 = Obj 2. Examples follow, using one of the test sentences.

12. Sentence: The cow chases the walrus before
the horse tickles the tiger

Response: Cow chases Walrus, Horse tickles Tiger

Doubling Errors

A: Cow chases Walrus, Cow tickles Tiger*

B: Cow chases Walrus, Walrus tickles Riger

C: Cow chases Tiger, Horse tickles Tiger*

D: Cow chases Walrus, Horse tickles Cow

*The structurally parallel NP might also be doubled

The most frequent strategy by far was A, in which either Subj1 or Subj2 is the agent of both actions. B is an Adjacency strategy : an NP that plays one role, an object, is given the adjacent subject role as well.

The other frequent response type was Interchange: e.g. Subj1 interchanged with Obj1, Subj1 with Subj2, etc. Finally there were Outside

errors (using an animal not mentioned in the sentence[15]), and Verb errors. Multiple errors were coded according to main error types: Doubling, Interchange, Verb, Outside.

13. Error Responses to S structure

% of all Error responses

Doubling-types		All Doubling	- 64 %
A	- 38 %	Interchange	- 22 %
	- 9.5 %	Out	- 9.5 %
C	- 7 %	Verb	- 4.5 %
D	- 9.5 %		

There were few differences in the responses to sentences with transitive and intransitive clauses. It is particularly striking that the intransitives did not elicit many Outside responses. Of 38 error responses to intransitive sentences, only 5 were of the Outside type. This suggests that the Doubling and Interchange responses to the S stimuli were due to confusion about the NPs mentioned in the sentences and not the lack of other animals in the situation.

We interpret these errors to mean that the children had reasonably good control of the S structure, but that the sentences tended to overload their memory. In particular it seems likely that the large number of NPs and Vs was difficult for the children to hold in memory; see the discussion below of the second imitation task, where it is shown that the presence of

---

[15]These were possible for four of the six S sentences, those with one intransitive clause.

many nouns and verbs in a sentence make it difficult to imitate. Note that the task itself did not impose any constraints on which NP played which role. Indeed, in the various sentences all the animals are sometimes agents and sometimes patients. Since there was no particular reason for a given animal to play one role rather than another, and since memory capacity was taxed, the children tended to mix up the animals.

The most frequent response, the A type, has one animal playing the agent role in both the main and adverbial clause. Such a response may be due partly to the toy-moving task itself. Several investigators have noticed that, in acting out multiple actions, children prefer to use a single toy as agent (Huttenlocher, Eisenberg, & Straus 1968, Legum 1975, Hamburger 1979).

To complete the argument for this interpretation of the S error responses, we consider what a structure-violating response to these sentences would be like. Most drastic would be a single clause response; and indeed such responses were taken to indicate real difficulty with the structure. If the response has two actions, then a structure-violating error might make an intransitive clause transitive, or vice versa. There were very few such errors. Note that it would be rather implausible to say that the A-doubling responses reflect a different interpretation of S structures, namely as Ving structures or some other type of reduced clause. Such an explanation is implausible because it involves an extra load on memory. On the usual assumptions about memory, remembering a reduced structure is more costly than remembering an unreduced one: a

representation of both structure and reduction must be retained[16]. The imitation data suggest that the children's memory spans were already taxed by the S sentences, and it is unlikely that a strategy that also burdens memory would be adopted. Morever, such an explanation cannot be extended to the other Doubling responses.

We conclude, then, that the error pattern showed the children to have fairly good knowledge of the S structure, and some memory difficulties with the stimulus sentences.

Ving structure errors. The characteristic errors to Ving structures were of two types. The Ving structure lacks an overt subject in the second clause; indeed, Ving sentences are ungrammatical with an overt subject (*John called before his/Mary's leaving). The second clause subject is unambiguously taken to be the same as the subject of the main clause. So, to a sentence The cow chases the horse before tickling the sheep the correct response is one in which the cow chases the horse, and then the cow tickles the sheep. The children gave many responses in which the subject of the dependent clause was not the subject of the main clause: most frequent were an Adjacency response and and an Outside response.

In the Adjacency response the NP object1, adjacent to subject2 position, is the actor for clause 2. So to the sentence The cow chases the horse before tickling the sheep, an Adjacency response has the cow chasing the horse, and the horse tickling the sheep. This response corresponds to

---

[16]See the discussion of memory, in Daneman and Case 1981, for example; and Smith, 1970.

the Doubling Strategy B above[17] In the Outside response, an animal not mentioned in the sentence participates as subject or object of the second action. These two error types, in which the subjects of the two clauses are different, accounted for over 50% of the Error responses to Ving structures.

The other Error responses preserved the equal subject interpretation but showed confusion elsewhere: the children used an animal not mentioned as subject, interchanged the animals mentioned, or confused or interchanged the verbs. The frequencies of the different responses are shown in the table below.

14. Error Responses to Ving Structures

	% of Error Responses		
I: Subj1 = Subj 2		II: Subj1 diff Subj2	
Out	21 %	Adjacency	37 %
Int	12 %	Out	17 %
Verb	13 %		

We return to the interpretation of the error responses to S and Ving structures. One possibility is that the children had difficulty keeping straight the animals and roles in the Ving structures, as we have argued

---

[17]Adjacency responses were scored as such if the NPs in question were adjacent in the stimulus sentence; or in the response (this second proviso applied to Reverse responses to sentences with after.

that they did for the S structures. In this event we would expect them to respond similarly to both structures, since acting them out required the same numbers of nouns and verbs. But the response patterns were actually quite different, for same-subject and adjacency interpretations. The children responded to S structures with a same-subject interpretation (Doubling strategy A) 38% of the time. For Ving structures this response, which is the correct one, occurred 55% of the time. The difference in use of the adjacency strategy is even more striking: it was used 37% of the time with Ving structures and 9.5% of the time with S structures.

The different responses to S and Ving structures, then, are due to their differences in interpretive complexity. The Ving structure requires interpretation; the S structure does not. The Adjacency strategy was essentially a response to the empty subject of the Ving structure; its frequency, together with the frequency of Outside responses, show that the children did not fully control the Ving structure.

NP structure errors. The most frequent response to the NP structure was the Single Clause; there were relatively few Error responses, and few Correct responses, as noted above. The few Error responses were Interchanges and actions using Outside animals. We take both as indicating poor understanding of the way the lone NP figures in the second clause.

What is interesting is the almost total absence of errors involving the type of action of the second clause. There is only one such error. This shows that the children who attempted to act out the second clause understood the basic structure of these sentences. Recall that there were many Single Clause responses to these structures. We suppose that those

children who did not understand the structure were those who acted out only one clause - recall that 45% of the responses to NP structures were Single Clause responses. Responses in the Error category, then, indicate some control of the structure in question.

It is possible that responses to the NP structures were affected by the fact that they are ambiguous. This is an effect of the greatly reduced surface structure of such sentences, a feature of their relatively high interpretive complexity. We present in Appendix C the evidence for our subjects' awareness of ambiguity in the NP sentences.

2.6 Conclusions: Toy-Moving

The predictions of the weak complexity theory were borne out. The Correct and Single Clause response categories separated the NP structures as most difficult: there were fewer Correct responses, and many more Single Clause responses, to the NP sentences than to the other two. The Error responses showed that the Ving structure was somewhat less well understood than the S structure, in that many responses supplied an incorrect subject for the dependent clause. In contrast, responses to the S did not indicate consistent structural misunderstanding.

The factor of interpretive complexity predicts the apparently paradoxical order of control indicated in this experiment. Sentences that are relatively long and explicit are controlled better than shorter, elliptical sentences. One might take difference in performance on the toy moving task to indicate different degree of acquisition. We now discuss the Imitation experiments.

3. Imitation

Imitation tasks are based on the interaction between structure and memory. The basic strategy is to present for imitation sentences that overload a subject's memory to the degree where knowledge of structure makes a difference. It is important to find the critical length for subjects in imitation tasks. Short sentences can be imitated by rote, long sentences cannot be imitated even with structural knowledge.

Imitation requires the subject to reproduce the surface structure of a sentence. It is thus quite different from the toy moving task. If linguistic complexity plays the same role in both tasks, that would be dramatic evidence favoring the strong complexity hypothesis. The weak complexity hypothesis would predict that factors of surface structure would be of paramount importance for this task. It is not clear how fully subjects interpret sentences presented in imitation tasks; errors are often semantic in nature, suggesting partial interpretation[18] . The task in itself does not require that sentences be interpreted.

Two imitation studies were done, focussing on the adverbial structures used in the toy-moving study. The first imitation study presented the same sentences used in the toy-moving experiment, defying the caveats above -the sentences varied in length. The second study controlled carefully for length.

---

[18]Smith 1970, Keller-Cohen 1975

3.2.2 Results

Length of stimulus sentence predicted the distribution of correct responses. Correct responses were quite frequent to sentences of 8 words or less; they were quite infrequent to sentences of more than 9 words. Since the NP structures are shorter than the others, this meant that the factor of length overrode the other factors. The predictions of the weak complexity hypothesis were thus borne out. We present the percentage of correct responses, organized by subject groups, below. The oldest children, Group III, did very well with all three structures; variation occurred for the two younger groups.

16. Percent Correct Responses to Imitation Study A

Structure	Group III	Group II	Group I
S	80%	36%	25%
Ving	90%	33%	36%
NP	92%	71%	54%

Although length alone determined success on the imitation task, the error data is suggestive for the S and Ving structures. The data, while sparse, indicates difference in the children's control of the two structures. The comparable responses were too few for statistical analysis, so our discussion will be informal. Error responses to NP sentences yielded little information.

S-structures: the great majority of error responses - 83.5% - were

difficult. This result is hardly surprising: length is known to affect imitation, and the stimulus sentences varied quite widely in length.

3.3 Imitation study B

In order to investigate more closely the imitation results and the effect of the task, a second imitation study was run. This study focused on the variable of structure and controlled for length; there was no variation in transitivity or connective. The intention was to explore as fully as possible children's imitation responses to the basic structures -full and reduced adverbial clauses- of the experiment. The sentences all had 9 words, transitive clauses, and the connective before; they differed in the structure of the adverbial clause. In order to keep length constant, we were forced to vary the nounphrases and verbs: nounphrases were complex [determiner + common noun] or single [proper noun]; verbs were complex [verb + preposition] or single [verb]. They were presented with fillers to avoid predictability and boredom. For example:

16. S: Michael caught the cat before Pete tickled the dog
17. Ving: The clown kissed the giraffe before petting the dog
18. NP: The monkey waved to the cow before the elephant

There were four exemplars of each structure; the test sentences appear in Appendix A. The sentences were presented in one session, interspersed with filler sentences of varying structure. There was at least one filler before each stimulus sentence, and three fillers before the first test sentence.

The subjects were 10 children, aged 4-5. Our earlier imitation study had determined that this age group was most responsive to the imitation task: children older than 5 generally imitated most or all of the sentences correctly, children younger than 4 tended to be unable to imitate any of them correctly.

3.3.1 Results

The results of this experiment were surprising: they supported the previous imitation study, even though length was held constant. The NP sentence were easiest, Ving sentences less easy, S sentences most difficult.

20. - % Correct Imitation
Responses

S	Ving	NP
25%	45%	75%

The children who were less advanced linguistically tended to have more difficulty with the S and Ving sentences, but there was little significant variation in the group as a whole.

3.3.2 Discussion

Since length was held constant we must look to some other factor to explain this rather surprising result. Continuing with the hypothesis that

imitation tasks focus on surface structure, we consider carefully the surface structures of the test sentences.

The sentences differ in the number of surface nouns and verbs that they contain, and this difference patterns with the imitation results. The sentences that were easier to imitate have fewer nouns and verbs than the others:

22. - Nouns & Verbs in Test Structures

S: Michael caught the cat before Pete tickled the dog (6)

Ving: Michael caught the cat before tickling the dog (5)

NP: Michael caught the cat before the dog (4)

Since the test sentences of each structure are entirely consistent with regard to length, this factor differentiates between the structures. S structures have 4 nouns and 2 verbs; Ving structures have 3 nouns and 2 verbs, S structures have 4 nouns and 2 verbs.[19]

We can only speculate as to the steps required in uttering nouns and verbs, especially in comparison to those required for determiners and prepositions (and perhaps other function words). Nouns and verbs have more semantic content, in some sense, than do function words: the latter are partially automatic and/or partially redundant. In order to explore this question more fully one would have to understand the processes tapped by the task of imitation.

---

[19]Earlier work with imitation tasks has also suggested that high-content words make a difference, cf Smith 1970.

But the task of elicited imitation, straightforward as it may seem, is poorly understood. We know that syntactic structure makes a difference, and that subjects comprehend the stimulus sentences to some degree; but it is not clear what level or kind of processing is involved. To see this, consider the typical error responses to imitation tasks.

Semantic substitution errors are very common: subjects tend to change articles -e.g. the for a, names - e.g. John or the boy for Jim, even verbs -e.g. go for run. See Smith 1970 for discussion and examples of such errors. These would be serious errors in the world, or in certain tasks involving interaction with a toy world. But for the imitation task they seem trivial. Typically in this task a subject hears a succession of sentences, one after the other; no response other than imitation is called for. In particular there is no requirement that a subject relate the stimulus sentences to anything in a real or imagined world. In this task it really doesn't matter whether a sentence mentions John or Jim. There is nothing about the task, or subjects' responses, suggesting more processing than that of short-term memory.

Indeed, these considerations support the general view that an imitation task mainly taps short-term memory. The fact that subjects indicate some degree of comprehension is not surprising: studies of word recognition suggest that people comprehend automatically words presented to them[20] . Returning to the results of our experiment, it seems that high semantic content demands more -units? time? we do not know- of short-

---

[20]For example, Forster 1970.

memory than does lower semantic content. Syntax is not the only factor of surface structure that affects short-term memory, and therefore not the only factor that affects success in an imitation task. Concerning the difference between linguistic recognition and production, Daneman & Case 1981 suggest that the latter is more difficult because it requires attention to all the features of a sentence, whereas the former does not. "Semantic and syntactic features must be dealt with as independent units for production...these features needn't be treated as independent units for recognition." Daneman & Case claim that production, and not comprehension, requires that each part of an utterance be processed sequentially and placed briefly in short term memory.

From this point of view the results of the imitation experiment are consonant with the other results and predictions of the weak complexity hypothesis. That hypothesis denies a simple one-to-one relation between linguistic complexity and performance: rather, the linguistic factors affecting performance are claimed to be task-dependent. We have shown that the task of imitation involves at least three factors of surface structure, length, syntax, and number of nouns and verbs.

4. Conclusions

We conclude that the weak complexity hypothesis is correct and the strong complexity hypothesis incorrect. The relation between linguistic complexity and performance depends on the requirements of performance, which vary widely. This is hardly a surprising result, but it is an important one to document.

The results of our two experiments are, of course, very different; so

different that the relation between the tasks is called into question. It seems clear that different abilities are tapped by the two tasks used, and that success in one does not necessarily predict success in the other[21]. This fact is of some interest for the field of language development. It means that the results of elicited imitation cannot be taken as more than partial evidence of children's language knowledge. Tasks of comprehension and elicited imitation may sometimes have similar patterns of results (as in Lust, Solan, Flynn, Cross, & Shuetz 1981, for example); it is plausible to assume that such similiarity indicates good knowledge of the material tested. But the patterns of results may differ, as shown in the experiments reported here. This shows that different tasks are needed for assessing children's linguistic competence as a good deal more than checks or safety measures. They tap different abilities and are therefore complementary in a very basic sense.

The modular theory of complexity has been useful, indeed essential, in construing these results. The identification of separate factors in linguistic complexity has allowed us to vary them and to discover some interesting interactions of complexity and performance. The theory with which we have been working has proved useful enough to merit further study. Our results suggest that other structures involving empty categories might usefully be investigated in this framework. Of direct relevance are other reduced structures, such as empty categories keyed to particular verbs; such structures involve systematic complexity and interpretive complexity. Also relevant are other types of interpretively complex sentences involving

---

[21]Similar results are reported in Keller-Cohen 1975, (p. 107).

quantification, complex temporal reference, and adverbials. And more generally, the three types of linguistic complexity need further investigation as they affect the linguistic performance of children and adults.

Appendix A List of stimulus sentences

Toy-Moving and Imitation A

These sentences are in blocs according to structure rather than in the order of presentation.

1. The cow chases the horse before the tiger sleeps.
2. The horse sleeps before the walrus kisses the cow.
3. The cow kisses the tiger before the horse chases the walrus.
4. The horse tickles the walrus after the tiger kisses the cow.
5. The tiger jumps up after the horse tickles the cow.
6. The horse kisses the tiger after the walrus jumps up.
7. The horse sleeps when the cow tickles the walrus.
8. The walrus sleeps when the tiger chases the cow.
9. The tiger jumps up before chasing the cow.
10. The cow tickles the walrus before sleeping.
11. The walrus tickles the horse before kissing the cow.
12. The horse chases the tiger after jumping up.
13. The tiger chases the horse after tickling the cow.
14. The tiger jumps up after kissing the horse.
15. The tiger chases the horse before the walrus.
16. The walrus tickles the tiger before the horse.
17. The walrus chases the cow after the horse.
18. The cow kisses the walrus after the tiger.

Imitation B (without fillers)

1. The horse chased Tommy before the lion pushed Susan.
2. Harry patted the dog before Mary kissed the cat.
3. Michael caught the cat before Pete tickled the dog.
4. Robert fed the chickens before Janie rode the donkey.
5. The horse kicked the dog before chasing the cow.
6. The tiger patted the lamb before riding the horse.
7. The lion pushed the tiger before chasing the zebra.
8. The clown kissed the giraffe before patting the dog.
9. The elephant knocked down the giraffe before the camel.
10. The girl talked to the lady before the boy.
11. The snake looked at the turkey before the rabbit.
12. The monkey waved to the cow before the elephant.

Appendix B Statistical analysis of Toy-Moving results

The children acted out a total of 18 sentences. Two were filler sentences with <u>when</u> as connective. The S and Ving structures were tested a total of six times each, three times with <u>before</u> and three times with <u>after</u>. The NP structure was also tested with both connectives: there were

two sentences with before and two with after.

In order to statistically determine whether the children responded similarly to sentences of the same structure, a Kuder-Richardson (KR-20) reliability procedure was used (this procedure is appropriate for dichotomous variables). The program Reliability-Model Alpha in the Statistical Package for the Social Sciences was used (Hull and Nie, 1981). We combined responses to the two connectives since they did not differ. As noted above, all sentences tended to be acted out in the order of mention, and this meant that the before sentences were correctly acted out whereas the after sentences were acted out in reverse order.

In using this procedure we created a correct category (including correct and reversed clause order responses) and an incorrect category (which included the various error responses). The standard alpha levels obtained for the structures having six test sentences exemplars (S and Ving) were .73 and .74, respectively. We considered this acceptable for a six item subtest. Furthermore, in all cases except one the analyses indicated that the alpha levels were not improved by subsequent elimination of test sentences. Sentence 13, an S after sentence, was the exception; it was eliminated from subsequent analyses since the reliability procedure indicated that the children treated it differently from the other two S after sentences.

The alpha level for the NP structures, where there were only four sentences, was .62. This was achieved by dropping sentence 2, an NP after sentence. Again, further elimination of sentences did not improve internal reliability. Using only the reliable items (that is, eliminating sentences

2 and 13), we were able to assume that responses were sufficiently homogeneous to allow deriving a single overall index of performance by using the composite score of the percentage of correct responses to each structural type. These scores were entered into the analyses of variance, discussed below.

Two analyses of variance were conducted using a two factor (3 x 3) design with repeated measures on one factor (using the biomed program BMDP2V-Analysis of variance and covariance with repeated measures). The factors were age (with 3 levels), and structure type (also with 3 levels). The structural type was the repeated factor, since each child acted out all structural types of sentences.

The first analysis determined the effect of age and structural type on the total number of sentences acted out correctly (recall that order of mention, notionally an order reversal, was considered a correct response). The results indicated a significant main effect for the age factor (F=9.82, p =.0003), and a significant main effect for type of structure (F=6.20, p.=.0003). The interaction effect between age and structural type was not significant. A follow-up multiple comparison on the age factor showed that the oldest children performed significantly better than the two younger age groups, and that the differences between the middle and youngest group was not significant. For structure, the results on the NP structure were significantly poorer than for the S and Ving structure; S and Ving were not significantly different.

The second analysis considered single clause error responses. These were the only erro· responses numerous enough to subject to statistical

analysis. Internal consistency was determined, using the KR-20 reliability procedure. The alpha levels obtained were .80 for the S structures, .71 for the Ving structures, and .76 for the NP structures. These levels indicated that the children treated sentences of each structural type with sufficient similarity to allow conducting an analysis of variance on the data. Again, the age factor was significant ($F = 8.62$; $p=.0007$) Follow-up comparison showed that the youngest children's performance was significantly poorer than the middle and oldest groups; the difference between the latter two was not significant. Structural type was also significant ($F = 19.91$; $p = .0000$). The children's performance was significantly poorer with the NP structures than with the S and Ving structures; performance on the latter two did not differ significantly.

Appendix C: Ambiguity in NP Structures

The NP sentences have two interpretations, one with the adverbial NP as subject and one with the NP as object of the adverbial clause. For instance, the adverbial clause in The horse tickled the sheep before the cow may have the horse tickling the cow; or, the cow tickling the sheep. Although there is some evidence in the literature that children of this age are unaware of syntactic ambiguity (Kessel, 1970), the children's responses suggest that they were uncertain and aware that more than one interpretation was appropriate.

To investigate the question of whether the NP sentences were ambiguous for our subjects, we look at how the lone NP was interpreted and whether the children as a group were consistent in their interpretations. The NPs were interpreted as both Subject and Object. The older children slightly favored the Object interpretation and the younger children strongly favored

the Subject interpretation.

NP Subject and Object Interpretations

% of Correct Responses by Age Group

	Subject	Object
Group III	40	60
Group II	71	29
Group I	66 2/3	33 1/3

Finally, we ask whether individual subjects consistently gave a single interpretation of the lone NP; if most children were consistent, that would be evidence that they were unaware that the NP sentences had another interpretation. Of the 44 subjects, 36 were not consistent: that is, they interpreted the lone NP as both Subject and Object. This inconsistency suggests that the children were aware of both interpretations of the NP structure. (Of the 8 who were consistent, 5 treated the NP as Subject and 3 treated it as Object; the 8 did not belong to one particular age group.)

References

Amidon, A. & Carey, P. 1972. Why 5-year-olds cannot understand before and after. Journal of Verbal Learning and Verbal Behavior. 11:417-423.

Barrie-Blackley, S. 1973. Six-year-olds' understanding of sentences with time adverbs. Journal of Psycholinguistic Research. 2:153-185.

Chomsky, Carol. 1969. The acquisition of syntax in children from 5 to 10. Cambridge, Mass; MIT Press.

Clark, E. 1973. How children describe time and order. In Ferguson C.A, & Slobin, D.I. (eds.), Studies of Child Language Development. New York; Holt, Rinehart and Winston, 1973.

Coker, P. 1978. Syntactic factors in the acquisition of before and after. Journal of Child Language 5:261-277.

Daneman, M. & Case, Robbie. 1981. Developmental Psychology. Syntactic form, semantic complexity, and short-term memory: influences on children's acquisition of new linguistic structures. 17.4: 367-379.

Feagans, L. 1974. Children's comprehension of some temporal and spatial structures. Papers from the tenth regional meeting of the Chicago Linguistic Society. Chicago, Illinois.

Fodor, Jerry, & Garrett, M. 1967. Some reflections on competence and performance. Lyons, J. & Wales, R., eds. Psycholinguistics Papers. Edinburgh; University Press.

Forster, K.I. 1970. Visual perception of rapidly presented word sequences of varying complexity. Perception and Psychophysics, 8:215-221.

French, L.A. & Brown, A.L. 1977. Comprehension of before and after in logical and arbitrary sequences. Journal of Child Language. 4:247-256.

Hamburger, H. and Crain, S. 1982. Relative Acquisition. In S. Kuczaj, ed. Language Development, Volume I: Syntax and Semantics. Hillsdale, New Jersey; Lawrence Erlbaum Associates.

Hull, C.H., and Nie, N.H. 1981. Statistical Package for the Social Sciences Update 7-9: New procedures and facilities for releases 7-9. New York; McGraw-Hill Book Company.

Huttonlocher, J., Eisenberg, K., & Straus, S. 1968. Comprehension: Relation between perceived actor and logical subject. Journal of Verbal Learning and Verbal Behavior, 71:527-530.

Johnson, H.L. 1975. The meaning of before and after for preschool children. Journal of Experimental Child Psychology. 19:88-99.

Keller-Cohen, D. 1975. The acquisition of temporal reference. Indiana University Linguistics Club.

Kessel. F. 1970. The role of syntax in children's comprehension from ages six to twelve. Monographs of the Society for Research in Child Develoment, 35;6.

Legum, S. 1975. Strategies in the acquisition of relative clauses. Southwest Regional Laboratory Technical Note, TN2-75-19. Austin, Texas; Southwest Regional Laboratory.

Lust, B., Solan, L., Flynn, S., Cross, C., & Shuetz. 1981. A comparison of null and pronominal analysis in first language acquisition. Proceedings of the New England Linguistics Society, 1981.

Sheldon, A. 1974. The role of parallel function in the acquisition of relative clauses in English. Journal of Verbal Learning and Verbal Behavior 13, 272-281.

Slobin, D., & C. Welsh. 1973. Elicited imitation as a research tool in developmental psycholinguistics. In Ferguson, C. & Slobin, D.I. (eds), Studies of Child Language Development. New York; Holt, Rinehart, Winston.

Smith, C.S. 1970. An experimental approach to children's linguistic competence. In Hayes, J. (ed.), The acquisition of language. New York; John Wiley.

Smith, C.S. 1978. The syntax and interpretation of temporal expressions in English. Linguistics and Philosophy 2.1, pp 43-100.

Smith, C.S. forthcoming. A theory of linguistic complexity.

Solan, L. 1978. Anaphora in child language. Graduate Students of Linguistics Association, University of Massachusetts at Amherst.

Solan, Larry, & T. Roeper. 1978. Children's use of syntactic structures in interpreting relative clauses. In Goodluck, H., & L. Solan, eds. Papers in the structure and development of child language. University of Massachusetts Occasional Papers, volume 4. University of Massachusetts at Amherst.

Tavakolian, S. 1981. The conjoined clause analysis of relative clauses. In Tavakolian, S., ed. Language Acquisition and Linguistic Theory. Cambridge, Mass: MIT Press.