DECLARATION BY AUTHOR OF THE THESIS

Name of Author in Full (Block Letters) DOROTHY VERA MARGARET BISHOP

Faculty Board Clinical medicine

Title of thesis as approved by the Faculty Board

Note

(*a*) A candidate wishing to amend the title previously approved by the faculty board must apply to the faculty board for permission to do so.

(*b*) The title stated here must be precisely the same as that stated on the title page of the thesis submitted.

Comprehension of grammar: normal & abnormal development.

1. I understand that I am the owner of the copyright of this thesis and that the copyright rests with me unless I specifically transfer it to another person.

2. I understand that the University requires that I shall deposit one copy of my thesis in the Bodleian Library where it shall be available for consultation and that photocopies of it may be made available to those who wish to consult it elsewhere.* I understand that the Library, before allowing the thesis to be consulted either in the original or in a photocopy, will require each person wishing to consult it to sign a declaration that he recognizes that the copyright of the thesis belongs to me and that no quotation from it and no information derived from it may be published without my prior written consent.

3. I agree that my thesis shall be available for consultation in accordance with paragraph 2 above.

SIGNED D Bishop DATE 17-11-77

* (*a*) A candidate who, for good reason, does not wish his thesis to be made available for consultation or to be photocopied, may apply to the relevant faculty board for dispensation from this requirement. Any such application should state, with full reasons, both the nature and the period of the dispensation requested, and should be sent with this form to the Faculties Office, University Offices, Wellington Square, Oxford.

Note: Dispensation will always be granted in cases where confidentiality has been made a condition of access to material which is subsequently incorporated in a thesis.

(*b*) A candidate who wishes to be informed of the names of persons consulting his thesis should write separately to Bodley's Librarian, The Bodleian Library, Oxford.

3,000: 1.76

D25194/79

Microfilmed by NLL

Ms. D. Phil d 6392 = R.S.L.

COMPREHENSION OF GRAMMAR: NORMAL AND ABNORMAL DEVELOPMENT

by

D.V.M. Bishop

BODLEIAN LIBRARY OXFORD
DEPOSITED THESIS
17.7.78.

Submitted at the University of Oxford
for the Degree of
Doctor of Philosophy

1977

COMPREHENSION OF GRAMMAR: NORMAL AND ABNORMAL DEVELOPMENT

D.V.M. Bishop

St Anne's College

submitted at the University of Oxford for the Degree of Doctor of Philosophy, Michaelmas Term, 1977

Abstract

This thesis is in two parts. Part 1 reviews work on comprehension of grammatical structure in children and describes the development of a new Test for the Reception of Grammar (TROG). TROG assesses understanding of a variety of grammatical contrasts using a format where the child is required to select from an array a picture to correspond to a phrase or sentence spoken by the tester. Results from 280 normal children aged from 3-6 to 13-2 are presented. It is shown that several grammatical structures are not completely mastered by age 13, and in some instances it appears that the accepted rule that comprehension precedes production is contravened. Data on reliability and validity of TROG are also presented.

Part 2 is concerned with the application of TROG to 73 children with receptive and expressive developmental language disorders. It is shown that the majority of these children perform below age level both on TROG and on a test of receptive vocabulary. Further, there is a strong relationship between performance on TROG and grammatical complexity of expressive speech. Structures involving decoding of word order appear to give these children particular difficulty, over and above their general comprehension deficits. Results are discussed in relation to psychological theories of developmental language disorders, and implications for the classification of these disorders are considered.

CONTENTS

page

Abstract ii

List of tables iv

List of figures v

Acknowledgements vii

PART 1. NORMAL DEVELOPMENT

Chapter 1. Comprehension of grammar: experimental work 1

Chapter 2. A review of language comprehension tests for children 16

Chapter 3. Methodological considerations 33

Chapter 4. Experimental study: normal children 41

PART 2. DEVELOPMENTAL LANGUAGE DISORDERS

Chapter 5. Developmental language disorders: a review 92

Chapter 6. Experimental study: language disordered children 141

Appendix 1. Examples of TROG items 162

Appendix 2. Possible biases in samples of preschool children 171

Appendix 3. Calculation of %-scores for TROG 172

References 174

LIST OF TABLES

		page*
1.	ITPA subtests	27
2.	Intercorrelations for a subset of ITPA subtests	29
3.	Contrasts tested in TROG	43
4.	Schools used in the study	51
5.	Vocabulary pretest	55
6.	Prepositions: errors on individual items	63
7.	Reversible active and passive sentences: errors on individual items	67
8.	Rank order of difficulty of structures	79
9.	Test-retest: Pearson correlation coefficients	83
10.	Picture Completion and PPVT: Mean scaled scores	83/84
11.	Correlation of TROG with PPVT and Picture Completion	84
12.	Correlation of TROG with Reynell Comprehension Scale A	85
13.	Delay vs. deviance	107
14.	Comprehension scores: sex differences in distribution	147
15.	Language disordered children: mean age and Picture Completion scores	148
16.	Criteria for selection of control groups	151/152
17.	Analysis of variance: summary tables	151/152
18.	Outcomes for matched pairs on grammatical contrasts	154
19.	Neurological abnormality: allocation of children to categories	157
20.	WISC-R at age $8\frac{1}{2}$	171
21.	IQ and preschool experience	171
22.	Incidence of missing data	173

*Where a single page number is given, the table will be found facing that page. Where two numbers are given, the table will be found between those pages.

LIST OF FIGURES

		page*
1.	comparative/superlative	56
2.	comparative terms	57
3.	negative	58
4.	noun: singular/plural	59
5.	prepositions: session 1	60
6.	prepositions: session 2	61
7.	"the hat under the knife" from a different view	62
8.	pronouns	64
9.	pronouns: individual items	65
10.	reversible actives and passives	66
11.	reversible passive: mean number of items correct	68
12.	tense	69
13.	embedded	73
14.	verb: singular/plural	75
15.	2-word combinations	77
16.	3-word combinations	78
17.	median scores on TROG scales	82
18.	TROG total: medians	82/83
19.	TROG total %-score: boys and girls	82/83
20.	PPVT and Picture Completion raw scores	83/84
21.	Theoretical distributions: brain weight vs. body height	93
22.	Discriminating power of tests	138
23.	Expected distributions of comprehension scores	141
24.	Language comprehension: experimental group in relation to controls	145
25.	Distribution of PPVT scaled scores	146
26.	Models of sex differences in incidence of disorders	149
27.	Grammatical development: expressive vs. receptive language	150
28.	TROG scores in relation to PPVT	151
29.	Grammatical scales	152
30.	"Expressive" and "receptive" groups	153

continued

List of figures, ctd.

		page*
31.	Relation of hearing loss to comprehension scores	156
32.	Positive family history in relation to comprehension scores	156/157
33.	Children with late onset: comprehension scores	156/157
34.	Neurological abnormality in relation to comprehension scores	158

*Where a single page number is given, the table will be found facing that page. Where two numbers are given, the table will be found between those pages.

Acknowledgements

I would like to thank my supervisor, Dr Freda Newcombe, who first aroused my interest in this area, and who provided me with the facilities for carrying out this research.

Testing of Oxfordshire children was made possible through the co-operation of Oxfordshire Education Services, particular thanks being due to Miss K. Jenkins. I would also like to express my appreciation to the following head teachers and their staff, many of whom not only allowed me to test their pupils, but also joined in constructive discussion about my work: Miss T. Lawrence, Miss H. Chapman, Miss M. Baker, Miss R. Archer, Mrs S. Wilkinson, Miss K. White, Mr P. Jones and Mr B. Mulliner.

Miss M. Palmer of John Horniman School, Worthing, and Mrs S. Naidoo of Dawn House School, Mansfield, provided facilities for testing language disordered children, and were most helpful and tolerant throughout.

Dr R. Hiorns of the Department of Biomathematics, Oxford University gave me his valuable time and statistical advice, and Mrs N. Amery of the same department was most helpful with data punching and analysis.

Part 1

NORMAL DEVELOPMENT

Chapter 1. Comprehension of grammar: experimental work

The vast majority of research on language development is concerned with the analysis of expressive speech. In Brown's (1973) review of this field, "A first language, the early stages" only 13 out of 413 pages of text are directly concerned with research on language comprehension. We can safely assume that this paucity reflects the lack of work in this area rather than lacunae in Brown's knowledge.

This overwhelming preponderance of research on expressive speech must be largely due to methodological considerations. Collecting speech samples is not difficult, and there is little question of the validity of such raw data. More importantly, the child is not required to co-operate beyond tolerating a taperecorder and uttering speech. Analysing the data is a more formidable enterprise - the issues in this area have been ably discussed by Brown (1973) and Crystal, Fletcher & Garman (1976). We may contrast this situation with the study of language comprehension, where ability is inferred from the child's response to the tester's utterances. For this procedure to be valid, the child must attend to the task and be motivated to respond. It is not surprising that few experimenters have obtained reliable results with children younger than two years who tend to be shy and disinclined to obey instructions. Quite apart from this difficulty, various task situations bring with them additional problems, as outlined in more detail on pages 33 to 40. In brief, a child may fail to respond correctly for reasons other than lack of comprehension, e.g. inability to make sense of pictures, or to scan an array, or to manipulate toys, or to remember a sequence of words, and so on. Thus the validity of

tasks designed to test comprehension is dubious. On the other hand, data analysis is much simpler than for expressive speech. The experimenter can design his task so as to investigate variables in a systematic manner rather than waiting for the child to produce structures of interest.

review of research

This chapter will give a chronological account of development of comprehension in general terms, leaving more specific discussion of particular grammatical structures to chapter 4.

Before the age of 9 months, few children appear to understand single words. Research on what they do understand in this period falls outside the domain of the chapter and has been reviewed by Friedlander (1970) and Menyuk (1974). Few experimenters have tackled the topic of comprehension in children who are just beginning to understand language. An exception is Benedict (1976) who carried out a longitudinal study of 8 children aged from 0-9[1], when little or no linguistic comprehension was evident, to 1-4, at the beginning of productive language. Three data sources were used: written diaries kept by the mothers on a daily basis, observations of mother-child interaction, and observation of the child's responses to specially formulated commands. Benedict's data confirmed the widely held impression that comprehension precedes production at the lexical level - the children understood an average of 60 words before they produced 10 words. Benedict claimed that several types of word were learned from the beginning, so that by the time the child understood 50 words, all major grammatical categories were included, but in the absence of further detail about the criteria used to assess comprehension one must

[1]Throughout this thesis, ages in years and months will be denoted in the form N-M, where N = years and M = months.

remain rather sceptical of this claim. Of particular interest are the findings from the structured comprehension task. These consisted of a series of commands, constructed by the experimenter from words known to be familiar to the child, and presented by the mother in a free play situation. Commands were analysed in terms of semantic units, defined as any meaning unit within the phrase which allowed a discrete response by the child, e.g. "give the dog to Mummy" has three semantic units, corresponding to the underlined words. Each unit was scored independently, e.g. in the above example, if the child gave the shoe to the experimenter, he would be scored as responding correctly to one unit - give. Benedict found she could describe a chronological sequence of comprehension in 5 stages. At stage 0 there is no evidence of understanding the meaning of any specific words. Most of the children were beyond this level at the start of the study. At stage 1A the child demonstrates understanding of a few stereotyped words in highly specific circumstances, and by stage 1B (age 1-0 to 1-3) single words are understood in a variety of situations. By age 1-3, most children entered stage 2, and started responding to two units within an utterance. By the end of the study a few children had entered stage 3, and were responding to three units in an utterance. Benedict noted that comprehension at stages 2 and 3 was additive rather than grammatical, i.e.

> "he understands the various units singly and puts them together additively on the basis of prior conceptual knowledge of contextual information, rather than deriving the information on how the units go together from the relational information contained in the language itself."

Evidence for this came from instances where children responded correctly to more than one semantic unit in an utterance, while putting them in an incorrect relationship to one another.

A similar conclusion was reached by de Villiers & de Villiers (1973) who asked children to act out reversible active and passive sentences[1] with toys. They found that:

> "There was no evidence that children with an MLU[2] between 1.0 and 1.5 morphemes per utterance could use word order information in this comprehension task, since they showed no preference for either noun as agent."
>
> p. 335

There was a marked improvement in ability to use word order as MLU increased above 1.5. Chapman & Miller (1975) used a similar task with slightly older children. Subjects were asked to act out reversible sentences with small toys and to describe their actions. Once again, many children failed to use word-order information correctly in the comprehension task, although they did better in their descriptions of what had happened. Chapman & Miller drew attention to this as an instance of production preceding comprehension.

Several experiments have looked at the child's sensitivity to grammar from a different standpoint. Rather than scrutinizing responses to normal sentences for evidence of appreciation of word order, etc, the child is presented with deviant sentences. At the simplest level this method may be used to discover whether children can discriminate well-formed from ungrammatical sequences, irrespective of whether the content of the sentences is understood. This procedure was adopted by Starr (1974) who observed the responses of children aged from 0-8 to 2-6 (all of whom spoke 1 or 2 words per utterance) in a situation where they could choose to hear grammatical or ungrammatical

[1] Reversible sentences are those where subject and object can be interchanged to produce a sensible sentence: e.g. the boy pushes the girl.

[2] MLU refers to mean length of utterance, a widely used index of grammatical development in preschool children. The rationale for this index, together with rules for calculating MLU are described by Brown (1973) pp. 53-55.

sequences from a taperecorder. She found that children responded differently to the two types of sequence, most subjects demonstrating a preference for grammatical sentences. This held for children at all levels of expressive ability.

Other experimenters have examined the number of relevant responses made by children to various sorts of input. The best known study in this area is by Shipley, Smith & Gleitman (1969) who divided 11 children aged from 1-6 to 2-9 into two groups on the basis of expressive speech. The telegraphic[1] group had median length of utterance ranging from 1.40 to 1.85, the holophrastic[2] group from .06 to 1.16. Children were given a variety of commands concerning a set of toys. In one condition, well-formed sentences were used, e.g. "Throw me the ball"; in another, telegraphic commands were given, e.g. "Throw ball". These commands were presented by mothers in a free response situation. Shipley et al are widely quoted as having demonstrated that the telegraphic group responded more to well-formed than to telegraphic commands, whereas for the holophrastic group the converse was found. However, the latter half of this conclusion is not supported by their data. In their analysis, Shipley et al lay most emphasis on "touch" responses by the child, which include any interaction with the named toy, i.e. any response to the noun of the command, regardless of the verb used. By this criterion, the telegraphic group do indeed respond consistently more often to well-formed than telegraphic commands. However, the holophrastic group

[1] word-combinations used by young children have been described as "telegraphic" because they are characterised by omission of nonreferent words such as articles and prepositions. See Brown (1973) pp. 74-78.

[2] The first single-word utterances of children learning language are termed "holophrases".

are far less consistent:- 2 of the 4 children in this group do not appear to respond differently to the two types of command, whereas the remainder respond more to telegraphic than to well-formed commands. If we consider another category of response, replies to commands, we find that both telegraphic and holophrastic subjects are more likely to respond to well-formed than to telegraphic commands. Similar experiments performed by Duchan & Erickson (1976) and Petretic & Tweney (1977) find that both holophrastic and telegraphic children respond more often to well-formed commands. The effect in the latter experiment was particularly strong when children were required to perform appropriate actions, rather than merely making some response to the object, in order to be credited with a correct response. The differential responsiveness to well-formed and telegraphic sentences demonstrated by this research may reflect attentional factors. Brown (1973) has suggested that the oddity of the telegraphic commands from mother may make the child pause and so interfere with his response. It cannot be argued that telegraphic sentences are less effective simply because they are shorter. Shipley et al included a condition where the child was presented with lengthened telegraphic commands, e.g. "Please Johnny, throw ball". These were no more effective than other telegraphic commands.

Wetstone & Friedlander (1973) investigated children's responsiveness to three types of sentences: i) well-formed, e.g. "where is the truck?" ii) misplaced, formed by reducing sentences to telegraphic form and arbitrarily reinserting nonreferent words out of order, e.g. "where the is truck?" and iii) scrambled order, e.g. "truck the where is?" Twenty 2- to 3-year-olds were divided into a nonfluent group of holophrastic and telegraphic speakers and a fluent group who

used well-formed sentences. The nonfluent group responded appropriately to all three types of sentences (i.e. looked at, touched the object, etc), but the fluent children did poorly on scrambled sentences. There was a significant negative correlation between number of relevant responses to scrambled sentences and MLU. The most fluent child did not respond to scrambled sentences at all and gave concrete evidence of her awareness of grammar when she exclaimed: "That lady who comes to play is nice, but she sure talks funny." The fluent and nonfluent groups did not differ in response to misplaced sentences, indicating that the order of nonreferent words was irrelevant to their response, despite the fact that the fluent group used such words correctly and consistently in their own productions. Wetstone & Friedlander interpreted the lack of sensitivity to word order shown by the nonfluent group in terms similar to Benedict:

> "without access to the subtle meanings and relationships that are established by means of linguistic structure, young children rely very heavily on their ability to integrate the meanings of familiar words they hear with the objects and events immediately before them."
>
> P. 739

Gleitman, Gleitman & Shipley (1972) measured children's ability to detect syntactic anomaly directly by asking them to make judgements of acceptability. Their subjects were three $2\frac{1}{2}$ year old girls, all of whom had been subjects in the study by Shipley et al (1969) and all of whom had responded more accurately to well-formed than to telegraphic sentences in that study. However, when asked to judge whether sentences were "silly" or not, only one child judged telegraphic sentences with word-order preserved as sillier than well-formed sentences, and in general, Gleitman et aldescribed the metalinguistic skills of these children as "feeble". Nevertheless, all subjects judged reversed word-order sentences as sillier than well-formed sentences.

De Villiers & de Villiers (1972) conducted a larger study of the same kind. 8 children aged from 2-4 to 3-9 (MLU 2.87 to 4.67) were presented with hand puppets and told that one puppet (P1) said things all the wrong way round and that another puppet (P2) would teach P1 to speak properly. P1 then gave examples of correct and reversed word-order sentences, and P2 judged these as right or wrong, correcting the aberrant sentences. Most children could be encouraged to join in the game, anticipating P1. A third puppet speaking semantically anomalous sentences was finally introduced. It was found that only the linguistically most advanced children (MLU of 4 or more) were able to make a significant number of correct judgements of ungrammatical word-order, and no child was consistently correct in detecting reversed word-order. Performance on the semantic task was much better. Indeed, correction of syntactically anomalous sentences often consisted of semantic alterations; e.g. "man a draw" was corrected by one child to "keep the man out"! Nevertheless, most of the children who did not judge reversed word-order sentences as silly responded consistently to word-order in a subsidiary task where they were required to act out reversible active and passive sentences (although their performance on the latter tended to be consistently wrong).

Research on children older than 3 years can utilize new methods of assessing comprehension which are not feasible with younger subjects. An early study by Brown (1957) investigated children's ability to infer part of speech of a word from its syntactic context. 16 children aged from 3 to 5 years from a Harvard preschool group were presented with nonsense words fulfilling the role of count noun, mass noun, or verb in a sentence while being shown corresponding pictures of nonsense activities. For example, the child might be shown a picture of a man

cutting a mass of cloth with a strange tool. He would be told that the picture showed either sibbing, some sib or a sib. A further set of three pictures showing the action (cutting), count noun (tool) and mass noun (cloth) respectively were then presented, and the child had to find the one corresponding to "sib". Brown showed that his subjects would pick the action for "sibbing", the count noun for "a sib" and the mass noun for "some sib", demonstrating good ability to infer part of speech from linguistic context. However, Lieber & Spitz (1976) using a more representative sample of children found a much later age of acquisition. Performance remained constant just above chance level for children aged from 6 to 8 years, with scores comparable to Brown's sample being attained only by 12-year-olds, the oldest group tested. These different results could have been due to the fact that the same pictures were not used by Brown (1957) and Lieber & Spitz (1976) but as the latter authors point out, if the outcome is so dependent on materials, then the results are not sufficiently general to be of great interest.

A classical study by Fraser, Bellugi & Brown (1963) popularized the picture-pointing procedure for studying comprehension. Pairs of pictures were devised differing only in terms of the grammatical contrast to be tested (e.g. active/passive: "the boy chases the dog" versus "the boy is chased by the dog"). 10 grammatical contrasts were used. Three parallel tasks were devised. In the Imitation task, the child did not see the pictures, but merely had to repeat pairs of sentences after the tester. In the Comprehension task, the experimenter showed the child a pair of pictures while saying the corresponding pair of sentences. The two sentences were then repeated in turn, and the child was required to select the picture corresponding

to each one. In the Production task the child was again shown both pictures while two sentences were spoken. He was then shown a single picture and asked to produce the correct sentence for that picture. More detailed discussion of the findings on individual structures will be presented in chapter 4. The main interest of Fraser et al, however, was not in the relative difficulty of different structures, but in the relationship between Imitation, Comprehension and Production. Their conclusion that Imitation precedes Comprehension which precedes Production has been replicated (Lovell & Dixon, 1967) and is widely quoted, but it has been challenged on methodological grounds. The difficulty is inherent in the comparison between three tasks of different format. It is difficult to determine whether the probability of correct performance by chance (P_c) differs between tasks. Fraser et al point out that it might seem that P_c would be .5 on the Comprehension task if children treated pairs of contrasts as such, always pointing first to one picture then to the other. However, this did not occur. Within one trial, many children would point twice to the same picture for the two contrasting sentences. Thus P_c is lower than .5, making comparison of the three tasks difficult. Fraser et al argue that nevertheless they are justified in analysing raw scores since this yields the same pattern of results as a secondary analysis in which only trials where children point first to one picture and then to another are considered. Baird (1972) has criticized this procedure, arguing convincingly that it cannot be assumed that P_c will be .5 for this subset of items. Presumably a better method would be to use different pictures to test the two contrasting sentences so that the child would be required to make only one response to each pair of pictures. However, difficulty also arises in estimating P_c for the

Production task. Fernald (1972) points out that there are three response possibilities in this task: the child may say the correct sentence, the contrasting sentence, or something different, which may be partially correct. The last alternative was considered to be an error by Fraser et al. Fernald performed a replication of Fraser et al's study and showed that Comprehension did precede Production if the original scoring system was used for the Production task, but if children were credited with a correct response for producing only the critical distinction in a sentence rather than the whole sentence, then there was no difference between Comprehension and Production.

Fraser et al's claim that children older than 4-0 approached ceiling performance on their task, together with oft-quoted statements to the effect that the child has mastered grammar by 4-0 (see pages 85 to 86) appear to have discouraged most researchers from investigating grammatical development beyond that age. An exception is found in the work of Chomsky (1969) who tested comprehension of sentences whose deep and surface structures did not correspond in a simple manner. Such sentences may be illustrated by the famous pair:

"John is easy to please"
"John is eager to please"

which have the same surface structure but different deep structures. When adjectives such as "eager" are used, deep and surface structures coincide: John is the subject of the sentence and the person who does the pleasing. When adjectives such as "easy" are used, deep and surface structures no longer coincide: although John is the subject of the surface structure, he is no longer deep structure subject, for it is someone else who is doing the pleasing. Comprehension of such

structures was tested by Chomsky by presenting children with a blindfolded doll and asking "Is the doll easy to see or hard to see?" Chomsky found that most 5-year-olds were consistently wrong on this task, replying that the doll was hard to see and could be made easier to see by removing the blindfold. Comprehension of the correct meaning was not mastered by most children until age 9-0. Chomsky also tested the verbs "ask" and "tell". Children were given commands such as "Ask/tell John what colour this is". When "ask" is used in such a sentence frame, surface structure is misleading, since it is John who must provide the information, whereas with most other verbs inserted in the same sentence frame, John would receive the information. Chomsky found that "tell" was easier to comprehend than "ask" in such a context, the latter not being fully mastered until children were 10-0.

However, interpretation of these results is complicated by several methodological considerations. Chomsky herself pointed out that the blindfolded doll in the "easy to see" task could be misleading, in that in order to perform correctly, the child must realize that the prominent blindfold is irrelevant. Further, the ask/tell difference might result because the "tell" command is a more natural one for the child to receive. Also the "ask" condition requires the child to formulate a more complex response than does the "tell" command. Chomsky attempted to overcome these difficulties by using a picture-selection task to test for comprehension of ask and tell. She found the differential difficulty of ask and tell remained although overall performance improved in this situation. This approach was extended by Kessel (1970) who devised sets of pictures from which the child had to select the one corresponding to sentences such as "the boy asks/tells the girl which pencil to sharpen". Kessel found that "tell" was easier than

"ask", but he found an earlier age of acquisition than did Chomsky, with all but one child performing perfectly on "ask" sentences by age 8-0. The easy/eager type of distinction was tested using two dolls. The child was familiarized with the dolls' names and then asked to make them play a hide-and-seek game in accord with instructions such as "Lucy is easy to find", etc. Kessel found that most children of all ages performed correctly with "eager" and similar adjectives. "Easy" and other adjectives of the same type were more difficult, and four 6-year-olds were consistently wrong on structures of this kind. The oldest children tested, 12-year-olds, showed complete mastery of the task.

Cromer (1970) divided the adjectives which can occur in the "John is ... to see" sentence frame into three types. S-types such as "eager" always indicate that surface and deep structure subjects coincide; O-types, such as "easy" indicate that surface and deep structures subjects do not coincide; and A-types, such as "nice" are ambiguous. Cromer studied 41 children aged from 5-3 to 7-5 who were asked to act out with a wolf and a duck puppet sentences with frames such as "the wolf is ... to bite", where S-, O- and A-type adjectives could occur. Many children consistently interpreted all adjectives as S-types. Other children would make either the named puppet or the other do the biting, but commit several errors. Finally, some children mastered the distinction. Classification of children by mental age (MA) as assessed by the Peabody Picture Vocabulary Test showed that those with MA below 5-7 never demonstrated correct comprehension, whereas almost all those with MA above 6-8 were consistently correct. Cromer concluded that children initially interpret all adjectives as S-types, then learn this is wrong in some cases but remain confused as to which instances disobey the rule, and eventually master the rule. Cromer

(1972) extended his study of these structures by investigating children's interpretation of nonsense words which were introduced in S- or O-type sentence frames. For example, the child would hear the sentence "I am skelgious that you read me the title" in one context, and then be asked to act out the sentence "the wolf is skelgious to bite" with puppets. Only those children who had performed correctly with real words were able to master this task, the remainder performing at chance level.

Cromer (1970) had raised some interesting points about Chomsky's original task with the doll. He pointed out that not only was the irrelevant blindfold misleadingly prominent, but some children might respond that the doll was not easy to see, while comprehending perfectly, because they regard her as difficult to see when her face is obscured. This point was taken up by Cambon & Sinclair (1974) who replicated Chomsky's experiment with French-speaking children, asking: "La poupée est-elle facile ou difficile à voir?" Results corresponded closely to Chomsky's: only after age 8-0 did the majority of subjects answer correctly. However, unlike Chomsky, Cambon & Sinclair found that 5-year-olds did better than 6- or 7-year-olds. They argue that Chomsky might have falsely classified some of her young subjects as failing the task because she failed to question them about their response. They give examples of subjects who, as Cromer suggested, reply that the doll is difficult to see because her face is obscured. They propose an ingenious explanation for the dip in performance that occurs at around 6 to 7 years. They suggest that the 5-year-old regards the verb "to see" as having only one agent, himself, i.e. he cannot "decentre" and conceive of the doll as seeing. His correct response is dependent not on full comprehension of grammatical structure

of the command, but on failure to realize that the verb could have the doll as subject. Performance declines around ages 6 to 7 years as the child becomes aware that others can "see". Thus the correct responses of the younger children are of poorer quality than those of older children. This view has little evidence to support it. It will be remembered that Cromer found that children of mental age 5-7 or less never performed correctly on O-adjectives, even though "see" was not used.

Conclusions

The most compelling conclusion reached by this brief review is that we cannot afford to ignore methodological issues in the study of comprehension. Consider our assessment of the child's mastery of word-order distinctions. If, as in Starr's (1974) experiment, the child is only required to make a discriminative response to grammatical versus scrambled word sequences, we find that children as young as O-8 can master the task. However, the ability to use consistent word-order in expressive speech does not emerge until later. Techniques which use liberal criteria for comprehension of sentences, such as that used by Wetstone & Friedlander (1973) will yield earlier estimates of appreciation of word-order than those which require the child to indicate that he can understand the distinction between two acceptable sentences differing only in word-order (Chapman & Miller, 1975). Finally, de Villiers & de Villiers (1972) showed that the ability to make metalinguistic judgements develops very late, and children may not be able to make such judgements about distinctions even though they appear to comprehend them. The last result is perhaps not surprising. The child must hear a great deal which he does not fully understand. If he is to learn the language, he must try to make sense of what he hears, rather than rejecting it as "wrong".

Chapter 2. A review of language comprehension tests for children

Recognition of the need for tests of language functioning as distinct from general intellectual ability has been a relatively late development in the field of psychological testing. Some of the earliest tests of language comprehension were concerned with receptive vocabulary. Only the best-known of these, the Peabody Picture Vocabulary Test, will be described here.

Peabody Picture Vocabulary Test (PPVT) (Dunn, 1965)

rationale

This is not described by its author as a language comprehension test, but rather as a tool for estimating verbal intelligence. However, the validity of this role has been questioned (Shaw, Mathews & Klove, 1966; Carr, Brown & Rice, 1967; Taylor, 1975) and Dunn himself remarks:

> "in light of the growing body of literature on the many facets of intellect, one must concede that the PPVT is not providing a comprehensive measure of intellectual functioning."
>
> p.33

format & content

Each item consists of four pictures. The tester says a word and the subject must select the picture matching that word. The content was selected from an initial pool of all words from a dictionary that could be represented pictorially. Test items which showed good linear growth curves with age, and did not show sex differences were then selected. Two parallel forms are available. Some items are unsuitable for British subjects, and a shortened English version of the test, the English Picture Vocabulary Test (EPVT) (Brimer & Dunn, 1962), has been developed. Standardization of the EPVT is less

comprehensive than of the PPVT.

statistical considerations

Norms are available for ages $2\frac{1}{2}$ to 18, and are based on 4,012 subjects. Alternate-form and test-retest reliabilities are appreciably high (Dunn, 1965). As far as validity is concerned, in view of the conceptualization of the PPVT as an intelligence test, the author concentrates on demonstrating acceptable correlations with IQ tests. Such values depend on the IQ test used and the population studied, ranging from .3 to .9. Dunn remarks that in the few studies of predictive validity carried out, correlations with later tests of academic achievement are not very high.

evaluation

The test seems to be a useful tool for estimating receptive vocabulary, but its value as a measure of verbal intelligence is questionable.

*****

A number of tests use a similar format to the PPVT to investigate comprehension of grammatical structure.

Northwestern Syntax Screening Test (NSST) (Lee, 1970; 1971)

rationale

This test is intended to be used as a screening instrument to pick out children whose syntactic development is sufficiently deviant to warrant further study.

format & content

NSST measures both receptive and expressive use of syntactic forms. These tasks resemble the Comprehension and Production tests used by Fraser, Bellugi & Brown (1963) described in chapter 1, except

that the receptive task uses arrays of 4 rather than 2 pictures. A variety of grammatical contrasts is tested.

statistical considerations

Norms are based on 344 children aged from 3-0 to 7-11 from middle-income and upper-middle-income homes. Details of selection of these children are not given. Lee (1971) provides no data on reliability of NSST. As far as validity is concerned, data from 18 children classified as "language delayed" are given. These children seem to do poorly on NSST, particularly on the expressive scale, but detailed analyses are not provided. Ratusnik & Koenigsknecht (1975) gave NSST to 20 normal, 20 language delayed and 20 mentally retarded children, all with mental ages around $4\frac{1}{2}$ years, and calculated the internal consistency of the test for these groups. The Hoyt coefficient for the expressive section was around .8, and for the receptive section around .6.

evaluation

Lee (1971; 1977) is careful to stress that NSST is a screening test and should not be regarded as an adequate estimator of linguistic functioning. However, its utility even in this limited role is questionable in the absence of data on reliability and validity. Arndt (1977) has also pointed out that NSST norms, based on a middle-class sample, are not satisfactory. Larson & Summers (1976) gave NSST to normal children of low socioeconomic status and found that both expressive and receptive mean scores fell below the 25th percentile on Lee's norms. Lee (1977) replies that it should be "common practice" among clinicians to establish local norms. This seems an unrealistic expectation. The average clinician has neither time nor facilities to test a large representative sample, and relies on the author of a

test published for general use to provide detailed normative data. It is particularly important that some estimate of the influence of socio-economic status on scores be given in view of research demonstrating that this variable can have a large effect on performance in tests of this kind (Parisi, 1971; Nurss & Day, 1971). Arndt's (1977) criticisms extend to the size of Lee's normative sample as well as its representativeness. At some age levels, norms are based on less than 50 children.

Another objection raised by Larson & Summers (1976), Arndt (1977) and Byrne (1977) seems less tenable. These authors have suggested that the validity of NSST is brought into question by the finding that analysis of results of pairs of receptive and expressive items testing the same structure fails to support the accepted view that comprehension precedes production in language development. Such a criticism is invalid in view of the fact that the probability of correct performance by chance (P_c) is likely to be lower for receptive items, where the subject has to choose one of 4 pictures, than for expressive items, where he selects one of two statements. Given that children treat pairs of contrasting sentences as independent (Fraser, Bellugi & Brown, 1963), P_c in the receptive task is .25 for each sentence. Fernald's (1972) argument that P_c in expressive tasks of this kind is less than .5 is applicable to the NSST, but nevertheless it seems unlikely that P_c will be equal for both tasks, so direct comparison of scores is not meaningful.

Lee (1977) points out that many of the criticisms levelled against NSST assume that the test has more pretensions than is the case. She likens administration of NSST to a physician taking a

patient's temperature to discover if he is in need of further investigation. This analogy, though persuasive, is not perfect. We know the normal range of variation of temperature in the general population and the significance of a reading outside that range. The same cannot be said of NSST scores. It must be asked whether NSST contributes useful information that is not already evident. In a different context, Compton (1970) compares the task of the speech pathologist to the TV repairman who does not need to spend time identifying malfunctioning sets - these are all too obvious. Rather he must find out what is defective and fix it. Lee (1977) admits that NSST is not suitable for detailed analysis of language problems, but only for identifying whether a problem exists. It can only therefore be justified if it can be shown that NSST picks out children who would not otherwise be recognized as being in need of attention. Unless it can do this, demonstrations of its ability to discriminate language-disordered from normal children are not sufficient to convince us of its usefulness.

Carrow Test of Auditory Comprehension (Carrow, 1975)[1]

rationale

The test is based on Chomsky's (1957) generative grammar, and is designed to measure the auditory comprehension of grammatical distinctions. Carrow states that as well as assigning the child to a developmental level of comprehension, the test is designed to pinpoint specific areas of linguistic deficiency.

format & content

The test consists of 101 plates, each of three line drawings. The child has to select the picture corresponding to the tester's

[1]An early version of the test is described by Carrow (1968)

utterance. Four main types of contrast are tested: 1. form classes and function words: nouns, adjectives, adverbs and prepositions. 2. morphological contructions: e.g. "er" and "ist" endings added to nouns, verbs and adjectives. 3. grammatical categories: contrasts of case, number, gender, etc. 4. syntactic structures: predication, modification and complementation. Items are ordered according to type of contrast rather than level of difficulty.

A Spanish version is also available, but has been attacked by Rueda & Perozzi (1977) for being based on direct word for word translation from English, resulting in unnatural and ambiguous items.

statistical considerations

Norms are available for a selected sample of middle-class Black, Anglo and Mexican American children aged from 3 to 6 years. Carrow (1975) reports high test-retest correlations for the total score over a two week interval. The criterion for selecting test items was content validity, rather than statistical considerations. Evidence for other forms of validity presented by Carrow is that a) scores increase with age; b) the test discriminates language disordered from normal children and c) scores improve with clinical improvement in the language of language disordered children.

evaluation

This test is already being widely used in the USA. It is evident that much thought and ingenuity has gone into its design. However, there are both practical and theoretical objections to the currently available version.

First, the norms are open to the same criticism as the NSST, in

that a middle-class sample cannot be regarded as representative of the general population. Second, a large number of British speech therapists who tried out the test found it unacceptably long (Renfrew, personal communication). The whole test must be given and this takes about 20 minutes. Personal experience with TROG (see chapter 4) suggests that the average preschool child cannot be expected to concentrate on this type of task for more than 10 minutes.

More serious objections are concerned with item validity. Failure on some test items could result from the child not knowing component words of a sentence, rather than being unable to handle the grammatical structure. Carrow does attempt to avoid this problem by keeping the vocabulary reasonably straightforward, and the majority of the items contain words which are first tested in the "form classes and function words" section. Words which have not been pretested are however also included, and it cannot be assumed that all children will know these (they include relatively unusual words such as "bump" and "show"). Furthermore, Carrow gives no advice on how to interpret scores if a child does fail a lexical item early on in the test; e.g. if a child fails item 12 ("black") he is likely to also fail item 86 ("it's not black"), and such failure could not be attributed to lack of comprehension of the negative.

Criticism of individual items is always easy, while suggesting how they may be improved upon can be more difficult. Nevertheless, some items do have unnecessarily ambiguous pictures (e.g. items 40, 66 and 76) while others require the child to adjust his point of view by putting himself "in" the picture (e.g. imperatives, and "we": items 89, 90 and 65). Other items are so long that one suspects that memory failure could be responsible for errors, particularly

since only one administration of each item is allowed; e.g. Item 99: "If you're the teacher point to the dog, if not, point to the bear." Other criticisms of items are put forward by Rueda & Perozzi (1977). Rondal (1976) has made the general point that the test has an important visual component, so failure may result from visual perceptual rather than linguistic complexity. This difficulty applies to all tests using this format, and must be recognized. However, no test format is completely free from the influence of nonlinguistic factors and it is difficult to see how this can be overcome. These issues are discussed at greater length in chapter 3.

Sentence Comprehension Test (SCT) (Hobsbaum & Mittler, 1971)

This test is in an experimental stage and not generally available.[1]

rationale

The authors aimed to devise a test which minimized situational and contextual cues and made few demands on the child's response repetoire.

format & content

Format is similar to the PPVT: multiple choice black and white pictures, from which the child must select one to correspond to the tester's sentence. Comparative/superlative, negative/affirmative, active/passive, and other distinctions are tested. Four examples of each of 15 sentence types are given.

statistical considerations

The test is relevant for normal children aged between 3-0 and 4-6, above which ceiling effects occur. Mittler et al (1974) have

[1] I would like to thank Dr Mittler for allowing our department to have a copy of this test.

however found SCT useful for investigating comprehension in older mentally subnormal populations. Test-retest reliability for 14 severely subnormal children over a period of 8 days was .77. Mittler et al report satisfactory correlations between SCT and the Reynell comprehension scale (.89) and the EPVT (.82) and low correlations with nonverbal tests. A high correlation (.91) was found between teachers' predictions and actual performance on SCT for 16 language disordered children.

evaluation

A very clear critique of the advantages and disadvantages of SCT has been provided by Mittler et al. They discuss at length the problem of validity of this type of test format, and note that some syntactic structures are difficult to illustrate in a satisfactory manner. Overall, the test is a carefully constructed tool, but its range of applicability would need to be extended in order for it to be of much use in a clinical setting.

*****

The remaining tests described here do not make a specific distinction between comprehension of grammatical structure and other aspects of language functioning.

Reynell Developmental Language Scales (Reynell, 1969a)

These scales are widely used in Great Britain. One expressive test and two comprehension tests are available. This description is concerned with the Verbal Comprehension Scale A (used unless the child is physically handicapped).

rationale

The theoretical background is fully discussed by Reynell (1969b) and appears to have developed independently from current psycholinguistic research. Given that the child has adequate hearing and experience of language, two processes are seen as crucial to development of comprehension of verbal language: ability to appreciate meaningful patterns in speech, and ability to incorporate these patterns into existing concepts. Initially words are only understood in specific contexts, then gradually the child becomes able to understand their meaning independently of context. By the age of 18 months, the child starts to internalize language as a vehicle for thought. Reynell's test attempts to tap these different levels of development.

format & content

The first section is flexible. The examiner has to demonstrate whether the child can consistently recognize simple words or phrases. The remaining sections require the child to make appropriate responses to a standard set of experimental objects. In sections 2 and 3, the child is asked to point to a named object from an array. In later sections, he must obey more complicated commands, e.g. "put all the pigs in the box and give me a brown horse." In the final section, the child must go beyond a presented situation and evaluate what will happen next, e.g. "This little girl is nearly late for school. What must she do?" Reynell & Huntley (1971) state that:

> "the scale is so designed that increasing difficulty depends mainly on the arrangement of the words used, rather than increasing the difficulty of the vocabulary."
> p.12

However, many factors other than word order are obviously implicated, including comprehension of word endings and nonlinguistic factors such as memory and ability to use contextual cues.

statistical considerations

The published version of the test is still in an experimental stage, with standardization so far limited to children in London and South-East England. Data from 636 children aged from 0-6 to 6-0 are given. Separate norms are presented for boys and girls, as small but consistent sex differences are found in the standardization sample. Figures for split-half reliability seem reasonable, but data on other types of reliability, or on validity, are not presented. Inspection of the developmental curves in the manual shows that the scale's discriminating power is low for children aged above 5-0.

evaluation

The test appears to be useful for evaluating comprehension in young children, particularly those below about 4-0, where many formal tests are not feasible. It is likely that the test gives a good estimate of the child's level of understanding in day-to-day life. However, Reynell & Huntley's (1971) claim that the scale allows qualitative as well as quantitative assessment of verbal comprehension does not appear to be justified. Reynell's model of development of language comprehension describes clear stages through which the child passes, but these are outlined in very general terms. A low score could result for a variety of reasons such as limited vocabulary, inability to make use of syntax, poor short-term memory, and so on, and no guidance is given on how to disentangle these factors. Finally, predictive validity studies are needed in order to be able to estimate the clinical significance of a poor score.

Illinois Test of Psycholinguistic Abilities (ITPA) (Kirk, McCarthy & Kirk, 1968)

rationale

The ITPA is a battery of 12 subtests, each designed to measure a

Table 1

ITPA subtests

SUBTEST	PROCESS	LEVEL	CHANNEL	FORMAT
auditory reception	decoding	representation	auditory	yes/no questions: e.g. "do trees fly?", "do logs burn?"
visual reception	decoding	representation	visual	child finds picture from array sharing same concept as sample picture.
auditory association	association	representation	auditory-vocal	verbal analogies: e.g. "cotton is soft, stones are ..."
visual association	association	representation	visuo-motor	child finds picture showing related concept to sample picture
verbal expression	encoding	representation	vocal	child describes 5 objects; scoring is concerned with number of concepts expressed rather than grammaticality
manual expression	encoding	representation	motor	child demonstrates the use of pictured objects by gesture
grammatic closure		automatic	auditory	grammatical constructions are elicited in standard contexts: e.g."here is a bed; here are two... ."
auditory closure		automatic	auditory	child completes words with sounds omitted: e.g. tester says "bo..le", child must say "bottle"
sound blending		automatic	auditory	real and nonsense words divided into successive sounds must be recombined: e.g. d-og becomes dog
visual closure		automatic	visual	exemplars of a named category are hidden in a standard picture and must be found
auditory sequential memory		automatic	auditory	similar to "Digit Span" in Wechsler scales
visual sequential memory		automatic	visual	child reproduces sequence of tiles showing abstract symbols

specific psych
be combined to
intended the t
differences.
"... the
logically
number of
education
compared
developme
an assess
since it
The ITPA is b
1957b). Thi
communication
channels of c
a specific p
format (
The 12
measure psyc
representati
that isolat
theoretical
they divide
statis
962 ch
Socioecono
States as
carefully

specific psycholinguistic function. Results from these subtests can be combined to give an overall "psycholinguistic age", but the authors intended the test to be used primarily to investigate intra-individual differences.

> "... the concept of intraindividual differences led logically to psychometric tests that could measure a number of specific and discrete areas of psycho-educational development. These areas could then be compared to determine discrepancies in growth and developmental imbalances in the child himself. Such an assessment is diagnostic rather than classificatory, since it pinpoints the deficient areas needing remediation."
>
> (Paraskevopoulos & Kirk, 1969; p.6)

The ITPA is based on Osgood's model of communication (Osgood, 1957a; 1957b). This model distinguishes three dimensions in terms of which communication can be defined: processes, levels of organization and channels of communication. A psycholinguistic ability is defined as a specific process at a specific level via a specific channel.

format & content

The 12 subtests of the ITPA are described in table 1. Tests measure psycholinguistic abilities at two levels of organization: representative and automatic. Paraskevopoulos & Kirk (1969) claim that isolating processes at the automatic level is "tenuous and theoretically problematic" and therefore use "whole level" tests which they divide into two types: closure and sequential memory.

statistical considerations

962 children aged from 2-7 to 10-1 comprised the normative group. Socioeconomic status of this group approximated that of the United States as a whole. It is important to note that this sample was carefully selected to include only children with IQs between 80 and

auditory sequential memory
visual sequential memory
auditory automatic
visual automatic
picture and must be found
similar to "Digit Span" in Wechsler
child reproduces sequence of tiles s
symbols

120. Most of the preschool children were siblings of school-age members of the normative group. Mittler & Ward (1970) gave the ITPA to a sample of English nursery-school pupils and obtained results which closely resembled those of the American standardization sample.

Paraskevopoulos & Kirk (1969) give detailed information about reliability of the ITPA, but little is said about validity. Internal consistency estimates range from .87 to .93 for the test as a whole at different age levels. Values for the subtests are lower, but never fall below .49 for the main subtests. Test-retest reliability for samples of restricted age over a 5 month interval ranges from .70 to .83 for the whole test. As is to be expected, scores on individual subtests are less stable over time. Test-retest reliability for some of the visual subtests is very poor.

evaluation

It is unfortunate for the authors of the ITPA that Osgood's model of communication has been largely disregarded by modern psycholinguists. The reasons for this neglect are manifold. The model is cumbersome and its three dimensional approach does little to clarify the communication process. Distinctions between "levels of organization" and "processes" are not at all clear. The model does not deal with central issues such as the nature of grammar, and, as Weener, Barritt & Semmel (1967) have pointed out, some of the subtests seem to be measuring cognitive factors only tangentially related to language functioning.

There is cause to be dubious about the content validity of some subtests. Consider, for example, Auditory Reception, the subtest which appears most relevant to this thesis. This subtest is designed to "assess the ability of the child to derive meaning from verbally

Table 2

Intercorrelations for a subset of ITPA subtests
(from Paraskevopoulos & Kirk, 1969)

		AC	SB	ASM	VR	VA	ME
subtests assessing automatic auditory-vocal functions	GC	.35	.32	.23	.39	.41	.31
	AC		.31	.18	.28	.23	.20
	SB			.17	.23	.23	.10
	ASM				.26	.27	.16
subtests assessing representational visuo-motor functions	VR					.38	.24
	VA						.26
	ME						

GC grammatic closure
AC auditory closure
SB sound blending
ASM auditory sequential memory
VR visual reception
VA visual association
ME manual expression

presented material" (Kirk et al, 1968) and takes the form of a series of yes/no questions, such as "do ants crawl?", "do logs burn?", "do sphinxes gallop?". Apart from the problems raised by the use of the yes/no paradigm (see chapter 3) the difficulty here is that failure might result if, although knowing the meanings of the words used, the child does not know the answer to the question: e.g. he may know what a log is and what burn means, but not be certain whether logs do in fact burn. This subtest would seem more appropriately described in terms of Osgood's model as a test of auditory association than auditory reception. This is not the only subtest with such difficulties. Rondal (1976) has discussed how factors other than knowledge of grammatical rules could affect performance in the Grammatic Closure subtest. Consider also Visual Reception, which uses a match-to-sample paradigm with a delay between sample and array. Surely here visual memory as well as conceptual ability is implicated.

Another approach to content validity is to examine the inter-correlations of subtests to see if they conform to the expected pattern. By Osgood's model, abilities at the same level of organization, or involving the same process or channel, should be more highly correlated than those which do not share these characteristics. Data relevant to this question are shown in table 2. These figures are not encouraging, since there is no evidence that tests purporting to measure the same type of psycholinguistic ability have higher inter-correlations than those measuring quite different abilities.

Weener, Barritt & Semmel (1967) have put forward similar criticisms of the ITPA, and have also argued that the use of a standardization sample of restricted IQ range can be misleading. It is most important

that the tester realizes that a score at the 15th percentile of the ITPA is not analagous to a score at the 15th percentile on the Binet or Wechsler intelligence tests.

The most important aspect of the ITPA is its claim to be able to go beyond the classification of children, and to specify educational and therapeutic aims. To this end, Paraskevopoulos & Kirk provide detailed statistical data to enable the tester to assess the child's score profile as well as his overall level of functioning. Final judgement of the ITPA will depend on its predictive validity. If it proves to be the case that isolation of particular strengths and weaknesses of a child can predict which type of therapy will be maximally effective, then the ITPA will fulfil its authors' aims and be a most useful addition to the psychologist's test battery.

The Token Test (de Renzi & Vignolo, 1962)

This test was designed to assess the comprehension of adult neurological patients, but has also been used with children (Whitaker & Noll, 1972; Noll & Lass, 1972; Cartwright & Lass, 1974; Fusilier & Lass, unpublished; Tallal, 1975). Only those aspects relevant to use of the Token Test with children will be discussed here.

rationale

The aim of de Renzi & Vignolo was to devise a test that was quick to administer and required no special apparatus, and that posed considerable difficulty at the linguistic level, without involving memory or general intellectual ability or a verbal response.

format & content

Shapes (tokens) varying on three dimensions (two sizes, two

shapes and five colours) are presented to the subject in an array, and he is given instructions concerning these. In parts 1 to 4, difficulty is increased by adding nonredundant words to instructions. For example, in part 1 the subject is given instructions such as "Point to the red circle" whereas by part 4 he is asked to "Point to the large green circle and the large blue rectangle." Part 5 introduces syntactic complexity as a relevant factor, with instructions such as "Except for the green one, touch the cirles." "Square" is usually substituted for "rectangle" when the test is used with children.

statistical considerations

Whitaker & Noll (1972) gave the Token Test to 252 children aged from 5-0 to 11-11, all with IQ scores between 95 and 110. They found a pattern of increasing test scores with age, with the curve beginning to plateau at around age 8-0. Most errors were made on part 5. This is in contrast to Tallal (1975) who gave the Token Test to a group of normal children in the course of a study of language disordered children and found that part 4 was more difficult than part 5. It is possible that differences in administration may account for this discrepancy. Noll & Lass (1972) reported that children of low socioeconomic status did more poorly on the Token Test than middle class children, even though the test items would appear to be relatively culture-fair. Lass et al (1975) found that a Short Form of the Token Test devised by Spellacy & Spreen (1969) was not useful with children, despite its demonstrated ability to discriminate dysphasic from nondysphasic neurological patients. Short Form scores failed to increase significantly between ages 6-0 and 11-11 in their sample.

Fusilier & Lass (unpublished) found that children's performance on the Token Test correlated significantly with only two of the 12

subtests of the ITPA (namely Grammatic Closure and Sound Blending). They concluded that these different tests may measure distinct aspects of linguistic functioning. Cartwright & Lass (1974) compared children's performance on the Token Test with the NSST and PPVT. Statistically significant correlations were found between the Token Test and total NSST score, but correlations between the Token Test and PPVT, and NSST and PPVT were close to zero. This is perhaps not surprising, for, unlike the PPVT, the Token Test is restricted to a very simple vocabulary, difficulty being introduced by the combination and arrangement of words.

evaluation

The Token Test fulfils its authors' aims in enabling the tester to assess complex linguistic functioning with simple and easily available materials. However, its suitability for children is questionable. It cannot be assumed that young children will know the colour and shape terms. Anyan & Quillian (1971) found that 14% of 4-year-olds and 8% of 5-year-olds were unable to name any colour from an array of red, blue or yellow. Newcombe (unpublished) found that 4 out of 55 children (7%) in the first class of a primary school (nearly all 5-year-olds) could not point correctly to five named colours (red, green, blue, yellow and black). Even if children do know the colours and shapes, it is unlikely that these abstract materials will hold their interest for the ten minutes or so that it takes to complete the test. Although one aim of de Renzi & Vignolo was to minimize errors due to memory failure, some of the instructions in parts 4 and 5 seem all too likely to exceed the memory span of many children.

Chapter 3. Methodological considerations

It can be seen from chapter 2 that although tests of language comprehension are available, none of these is ideal for studying understanding of grammatical structure in a language disordered population. It was therefore decided to design a new test for this purpose, attempting to avoid some of the drawbacks encountered by previous workers. A major decision concerned the type of task to be used. Methods which require the child to formulate or imitate sentences will not be considered here, since these are likely to confound comprehension and production. Potentially suitable paradigms will now be described, and their advantages and disadvantages briefly discussed.

paradigm A: obeying commands

This paradigm encompasses a wide range of test situations. The commands may be concerned with the natural environment (as in clinical neurological examination, when a patient may be asked: "shut your eyes", "pick up my pen", etc) or with a standard array of objects, as in the Reynell Comprehension Scale and the Token Test. This method can be very effective with small children when a set of toy materials is used, and seems to have been the preferred method of experimental psychologists working with subjects under 3-0. However, as Mittler (1972a) has noted, toys may prove to be too interesting, so that children become absorbed in playing with them. More fundamental objections also exist. An important result obtained by experimental workers using this method has been the demonstration that young children have strong tendencies to do what seems sensible with the objects provided, rather than obeying instructions (see account of work by Clark, 1973; Wilcox & Palermo, 1974; Grieve, Hoogenraad &

Murray, 1977, on pp 60 to 61 for examples). This difficulty can be circumvented by the use of meaningless materials, as in the Token Test, but such methods introduce new difficulties, since children may not know the words used to designate colours and abstract shapes, and meaningless materials are unlikely to hold their interest for long. Finally, this approach assumes that the subject is capable of carrying out the necessary motor movements to perform the commands. Subjects with dyspraxic difficulties might fail items despite adequate comprehension. Even for normal subjects, item difficulty may depend on the complexity of the motor response, rather than linguistic factors: e.g. "make the pig chase the mouse" may be more difficult than "make the pig jump over the table" because the former requires the child to manipulate two toy animals, the latter only one.

paradigm B: multiple choice

The subject is presented with an array of stimuli and is required to select the one which corresponds to the examiner's utterance. Pictured stimuli are typically used, presumably because they are less cumbersome than arrays of real objects, and allow more flexibility when designing stimuli. This method overcomes some of the difficulties raised by paradigm A. The tendency to select the most likely response regardless of instructions can be circumvented to some extent by using pictured situations which do not produce strong response biases. The motor response is reduced to a minimum and kept constant for all stimuli. However, young children may fail such tasks because they do not recognize pictures as representations of objects. Beckwith & Thompson (1976) reported success with children as young as 1-5 using an adaptation of this method where the child pointed to projected coloured slides of real objects and was reinforced for responding by

illumination of a rabbit face. The tester could stand behind the child unobtrusively, so that shyness on the part of the child was less likely to interfere with responding. However, the pictured situations used by Beckwith & Thompson were representations of single objects. Sleigh (1970) developed a nonverbal test of symbolic function in which the child is required to match toy objects with a pictured situation which is of related content but perceptually dissimilar. She found that not only were pictures more difficult for children to comprehend than toys, but also that complexity of pictures was important:

> "A child of mental age below two and a half years cannot be expected to assimilate more than one object from a picture containing several objects without adult help: a child with mental age below three years cannot be expected to understand a composite picture as such although he will be able to pick out some objects from it."
>
> p.173

Thus, if complex pictures are used in a comprehension task, it seems unlikely that a valid estimate of linguistic ability will be obtained for children younger than 3-0. Mittler (1972a) has enumerated several other potential sources of errors in tasks using this type of format. First, he points out that children may simply point to the most attractive picture without regard for the tester's utterance. Second, children with visual problems may fail to discriminate perceptually between test pictures. For example, in the Sentence Comprehension Test, the item "the boy is reading the paper" includes the distractors "the girl is reading the paper" and "the boy is reading the book". Here the subject might fail to discriminate between boy and girl or book and newspaper on a visual basis. Mittler also draws attention to the related and more subtle problems that visual discriminability introduces when we wish to compare the

relative difficulty of a set of items. If we found, for example, that a set of items testing comprehension of singular/plural distinctions was more difficult than a set testing the negative/affirmative distinction, we might be tempted to conclude that the former is the later function to develop. However, the result might be due to the fact that the distractors used in singular/plural items happen to be more visually confusable than those used to test comprehension of negative/affirmative. The reverse result might be obtained with a different set of items. Finally, this method is unsuitable for subjects with visual perceptual difficulties such as unilateral visual neglect, or scanning problems.

paradigm C: yes/no questions

Two forms of this paradigm are possible. The child may be presented with a picture plus a statement and have to indicate whether or not the statement agrees with the picture. Another approach uses no visual stimuli: the subject is given a sentence and must indicate whether it is true or false. This is the method used in the Auditory Reception subtest of the ITPA. This method overcomes some of the disadvantages of paradigms A and B. In the true/false version, no equipment is required, no visual abilities and only minimal motor abilities are assumed on the part of the subject. However, the method could underestimate comprehension if the subject had difficulty with the concepts of "yes" and "no". Personal experience with adult dysphasic patients illustrates this. One such patient perseverated with a "yes" response when presented with pictured nouns and asked questions such as "is it a dog?". However, when the correct answer should have been "no", the response was of much longer latency and was an explosive ejaculation, unlike correct responses. This patient

could identify the same nouns very easily by pointing to one out of two pictures. Furthermore, she learned after a few trials to make a pencil-tapping response accurately, where one tap signalled "yes" and two taps "no". It would appear that she had adequate comprehension of the questions asked her, and of the concepts of affirmation and negation, but could not code these into words. A second patient would reply to yes/no questions while nodding or shaking her head, but the verbal and nonverbal responses did not correspond, with the latter tending to be more accurate. Here again there is evidence of comprehension that would go undetected if a verbal response were relied upon. Finally, a third patient gave evidence of good comprehension of nouns in a pointing task, but could not respond accurately when asked yes/no questions. This patient did not learn to use a tapping response correctly and it is possible that she no longer understood the concepts of "yes" and "no". These patients demonstrate that the yes/no method can underestimate comprehension in some cases.

paradigm D: metalinguistic judgements

In this paradigm, the child does not necessarily have to understand the content of a sentence, but rather has to appreciate the distinction between acceptable and unacceptable sentences. This is the method used by de Villiers & de Villiers (1972) and Gleitman et al (1973), whose work is described in chapter 1. Like paradigm C, this method has the advantage that it does not require the use of equipment, and does not assume any visual or manipulatory skills on the part of the subject. Further, it can be used to test appreciation of grammatical rules which do not affect meaning: e.g. noun and verb number agreement; "the boy eat/the boy eats". However, this method

has the great drawback that young children may not understand what is required of them. This comes across most clearly from experiments where children are asked to correct ungrammatical sequences, as well as judge them. Menyuk (1969) describes her experience with this sort of task as follows:

> "The task was put to (the subjects) as 'I'm saying something wrong. How should I say that?' Their concept of 'wrong' was not always concerned with grammaticality....The younger children, in their efforts to correct the sentences sometimes carefully articulated each word in the utterance, sometimes said it as loudly as they could or generated a new utterance.....These younger children obviously had the complete set of rules in their grammar for such structures as tense restriction in conjunction as evidenced by their spontaneous corrections, but when specifically asked to apply them they did not seem to understand what was required or could not do so."
>
> pp 121-122

Finally, this method would seem to put language disordered children at a particular disadvantage. Such children are used to having adults tell them what is right or wrong about sentences, and may well lack the confidence to judge any adult utterances as "wrong".

paradigm E: sentence pairs

Waryas & Ruder (1974) proposed a new technique in which the child is presented with a single picture while two sentences are spoken, and then must select the sentence that best describes the picture. Like paradigm D, this method has the advantage that it can be used to test appreciation of grammatical contrasts that do not signal changes in meaning. Further, problems of visual discrimination and scanning are minimized. However, a language disordered child would be placed at a disadvantage if asked to respond verbally. The alternative would be to show the child the picture while presenting the pair of sentences once, and then to repeat the sentences, giving the child a

chance to indicate nonverbally which he preferred. But this raises other difficulties, in that it is assumed that the child can remember the pair of sentences and the order in which they were spoken. Limited personal experience with this approach indicates that it is not an easy paradigm to explain to subjects, and young children tend to always respond to the first sentence, or to both sentences.

paradigm F: multiple choice with nonsense materials

An imaginative technique using nonsense words was devised by Scholnick & Adams (1973) who presented children with coloured cartoons each showing one nonsense figure hiding, painting, washing or bumping another in the presence of a third nonsense figure. The scene was described to the child using nonsense words: e.g. "The miv was painted by the yud". He then had to identify the miv or yud. This method has the advantage that children's prior expectations about how objects behave are overcome. However, nonsense materials can raise new problems. Typically some children simply will not respond to them, saying that they don't know what the words mean. Other children may point to the nonsense object that they find most attractive (Lieber & Spitz, 1976; Brown, 1973, pp 334-335), or base their responses on similarities that they see between the nonsense stimuli and real words or objects.

*****

It can be seen that all the paradigms considered here have attendant complications, and none is free from contamination by task specific factors, so that the choice between methods involves selection of the least evil. It was eventually decided to use

paradigm B. Paradigm A was rejected because of the difficulty of controlling for a priori response tendencies except by using meaningless materials. Paradigms D, E and F were deemed too difficult to explain to subjects who might have comprehension problems.

Paradigm C was rejected after a pilot study showed that some children, expecially those with language disorders, were reticent, or if they did respond, seemed reluctant to disagree with the tester, and said "yes" to all items. There may be more to this than a social difficulty. Bloom (1970) demonstrated that although "no" is used by children from a very early stage of language development, the semantic functions expressed by it are initially limited to Nonexistence (e.g. "no sock" when an expected sock is missing) and Rejection (as when a child refuses something). Use of "no" to express Denial (the semantic function required by paradigm C) develops later, and although it should be acquired by age 4-0, this cannot be assumed for language disordered children. The main problem with paradigm B, apart from the difficulties discussed above, proved to be in the design of unambiguous pictures. Some grammatical contrasts are by their very nature difficult to represent pictorially. Specific difficulties which arose will be discussed in more detail in the next chapter.

Chapter 4. Experimental study: normal children

Introduction

This chapter describes the development of a new comprehension test, named the Test for Reception of Grammar (TROG), and discusses results obtained with normal children. The rationale for selecting the multiple choice paradigm has been discussed in chapter 3. Other considerations affecting test design were as follows:-

1. The test should yield an estimate of grammatical comprehension as far as possible uncontaminated by a) knowledge of vocabulary; b) fluctuations in attention, and c) a priori probabilities of sentences.

In order to minimize the effect of vocabulary on performance, all test items were constructed from a small set of common words, knowledge of which was assessed before giving the grammatical test. To reduce the likelihood of poor performance resulting from momentary fluctuations in attention, subjects were allowed to have the whole test item repeated if necessary. A main objective in designing sets of pictures was to ensure that target and distractors did not obviously differ in the a priori probability of the situations they depicted.

2. The test should be applicable to as wide an age range as possible.

Grammatical contrasts covering a wide range of difficulty were included in TROG, although structures which are not usually encountered in spoken language were avoided.

3. The test should give a qualitative as well as a quantitative estimate of grammatical comprehension

It is useful to know whether a child's comprehension is in keeping with his age, but it is of greater interest to both teacher and

theoretician to find out whether the child has particular difficulty with some grammatical contrasts rather than others. It was hoped to provide data on this point by including in TROG sufficient items so that for each grammatical contrast the probability of a subject getting all items right by chance was very low. This would enable us to describe the child's performance in terms of a list of contrasts passed (i.e. all items correct) and failed (one or more errors). In practice, this method generated interesting group data, but proved of less value in assessment of the individual than had been hoped. The difficulties which were encountered will be discussed below.

Design of the study

Two main test sessions were used to gather normative data. In the first session, an initial version of the test was administered to 220 children aged from 3-6 to 9-0. The test was then modified and readministered to the original sample 3 to 5 months later. A further sample of children aged 9-3 to 13-2 was also tested. The second test session was decided upon for three reasons. First, it enabled new items to be added to TROG, and existing items to be reordered and improved upon. Second, the results from session 1 proved sufficiently interesting for the author to decide to compare it with other tests. Finally, as there was considerable overlap between TROG items used in the first and second test sessions, by retesting the original sample, a minimum estimate of test-retest reliability could be obtained for TROG.

Materials

Test for Reception of Grammar (TROG)

The set of pictures for each item were drawn on a sheet of paper

Table 3

Contrasts tested in TROG

comparative/superlative inflections

negative

noun singular/plural inflections

active/passive

prepositions

pronouns

tense inflections

verb singular/plural inflections

embedded sentences

more/less

word combinations (lexical contrasts)

and arranged in a file. Brightly coloured pictures were used as black and white drawings which had been used in pilot testing had failed to capture children's interest.[1] In session 1, 70 grammatical items and 26 vocabulary pretest items were used. For session 2, 56 of the same items were used, with other items being added or modified to give a final set of 76 grammatical items and 27 vocabulary pretest items.

The final version of the vocabulary pretest consisted of 13 nouns, 9 verbs and 5 adjectives. Two or three-choice items were used to test comprehension of these words. Some examples are given in appendix 1.

The grammatical contrasts tested are listed in table 3. With the exception of "more" and "less", where a single item was used to test each word, sufficient items were included for each contrast so that the likelihood of getting all items right by chance was .05 or less. This enabled the child's performance on each structure to be scored dichotomously, a pass being awarded if all items testing that structure were correct, a fail otherwise.

comparative/superlative

This contrast was tested by one 2-choice comparative item ("spottier") and two 3-choice superlatives ("biggest" and "longest"). An example is shown in appendix 1, fig xxi. A difficulty encountered with these items was that adjectives which can be depicted and to which comparative and superlative inflections can be added are usually relative terms regardless of endings. This is true of "big" and "long"

[1]Mittler et al (1974) regard black and white pictures as more suitable in this sort of test as they maintain coloured pictures could impede fine discriminations. The author did not feel that the evidence they cited was adequate to justify rejecting coloured pictures.

but not "spotty". This means that children could respond correctly to the superlative terms without understanding the inflectional ending, because, e.g. the picture which best describes "big" will be the biggest one. This possibility will be considered when discussing the results.

In session 1, one item testing "most" and one testing "least" were included. Both items showed 2 trees with differing numbers of apples. In session 2 the same pictures were used, but the more appropriate terms "more" and "less" were used. The wording was of the form "which has more apples?", etc. These items were included in view of the extensive literature on the comprehension of such comparative terms (Donaldson & Balfour, 1968; Donaldson & Wales, 1970; Clark, 1970; Holland & Palermo, 1975; Carey, 1976).

negative

The negative was tested by five 2-choice items, where "not" preceded a verb or adjective used in the pretest. An example is shown in appendix 1, fig v.

singular/plural noun inflections

Four regular and one irregular (foot/feet) singular/plural noun inflections were tested. An example is shown in appendix 1, fig x. This distinction is notoriously difficult to test unambiguously as Waryas & Ruder (1974) have remarked. For example, if asked to point to "a dog" children might with justification point to one member of the pair of dogs shown in D (appendix 1, fig x). Conversely, when asked to point to "the dogs" children tend to point to both A and D. When this occurred the child was asked to point to one picture, and the item was then repeated, but it could be argued that the use of

the word "one" in this additional instruction is misleading. It should be noted that the two items with wording in the singular had an additional cue to the correct response, in that the child was asked to point to e.g. "a dog", whereas in the plural items the definite article was used. The author now feels it would have been better to use the definite article in all items (e.g. the dog/dogs).

prepositions

Different items were used in sessions 1 and 2. In session 1, a total of 6 items, two each for "in", "on" and "under" were used. Each item contained three pictures: one corresponding to the correct arrangement (C), one corresponding to reversed word-order (R) and one containing the objects mentioned in a different spatial relationship (X). An example is shown in appendix 1, fig xvii. All pictured situations were selected on the basis that the objects used would not be expected to have any particular relationship with one another in the real world. In session 2, for reasons that will be discussed, it was decided to use different items to test comprehension of the word-order aspect of prepositions (C versus R distractors) and the spatial relationship aspect (C versus X distractors). In this session, six 2-choice items testing "prepositions + word order" were used, two each for "in", "on" and "under", and four 2-choice items for "standard prepositions", one each for "in", "on" and "under" plus a new item testing "among". Examples are shown in appendix 1, figs xviii to xx.

pronouns

Four personal pronouns (he, she, they and it) and five possessive pronouns (his, her, their, my and your) were tested in session 1. Examples are shown in appendix 1, figs vi and xvi. "My" and "your" were omitted from session 2. By their very nature, first and second

person pronouns are difficult to test using the picture-pointing format. In session 1, "show me your feet" and "show me my feet" were interjected among the pictured items, but failure frequently resulted from the child having developed a set for pointing to pictures, and therefore pointing hesitantly to a picture in the book when given these items. Carrow (1975) does use a pictured item to test "we" (item 65: "we're eating apples") but here the child must realize that the tester is no longer offering a description of the picture, but is indicating what the characters in the picture might be saying. It was felt that items which demand that children shift their point of view in this way probably underestimate comprehension, and therefore such items were avoided in TROG.

The item testing "her" proved to be ambiguous, as illustrated by the response of a 4-year-old:

Tester: "show me her ball"

Child: (pointing to the boy holding a ball): "He's got it".

active/passive

The items used to test this contrast were similar to those devised by Fraser, Bellugi & Brown (1963) for their original study. A total of ten 2-choice items were used; in five the sentence was in the active voice (reversible actives) and in five it was in the passive (reversible passives). Construction of items was constrained by the limited number of reversible verbs that young children might be expected to know. Here "chase", "push" and "see" were used. An example is shown in appendix 1, fig vii. For all these items, the two nouns in the sentence could be logically interchanged.

tense

Limitation on the number of items it was feasible to test led to irregular and regular verbs referring to present, past and future time being

included together under the heading of "tense". All past time items involved the complex present perfective, e.g. "she has eaten the apple". Like singular/plural inflections, tense distinctions proved difficult to test unambiguously. This is best illustrated by an example, such as fig xi in appendix 1. If given the sentence "she is eating the apple" neither A nor C could be correct. However, given "she will eat the apple", A seems the best choice, but B is also justifiable, especially if we interpret this as "she will eat up the apple (completely)". The present perfective can also be ambiguous: one could make a case for selecting B for "she has eaten the apple" on the grounds that this might be interpreted as "she has eaten (some of) the apple". These responses seem particularly justifiable for children who do not realize that, unlike for any of the other items, the three pictures form a sequence. In session 1, an attempt was made to overcome this problem by introducing tense items as follows: "These three pictures tell a story. First this happens (pointing to A), then this (B), then this (C)." The test sentence was then presented. However, this modification raised new problems. Suppose a child with poor comprehension failed to understand this introduction, then he would be put at a disadvantage. For this reason, the introductory phrase was omitted in session 2. Another difference between the two sessions lay in the wording of future tense items. In session 1, each sentence was spoken twice, first using "going to" and then using "will", e.g. "she is going to paint; she will paint". In session 2, "will" was used alone.

embedded sentences

Embedded sentences were contrasted with situations that could be described by the same words in a different order. An example is shown

in appendix 1, fig xiv. Five 2-choice items were used, all with the relative pronoun deleted. In four of the items, the subject of the main clause was the object of the embedded clause (e.g. "the cat the cow chases is black"). In the remaining item, the subject of the main clause was also the subject of the embedded clause ("the boy chasing the dog is spotty").

verb singular/plural inflections

The irregular noun "sheep" was used in all items testing comprehension of the verb inflection for number. Three items tested the distinction between "is" and "are" (e.g. "the sheep is/are spotty"), and two items tested the regular noun inflection (e.g. "the sheep chase/chases the cat"). An example is shown in appendix 1, fig xv. The same difficulty arose as for singular/plural nouns, i.e. children would point to both pictures, and needed to be reminded "point to one picture". Another difficulty was that some children did not know that "sheep" was an irregular noun and would ask "sheep or sheeps?" If this occurred, the tester simply repeated the item but obviously this is unsatisfactory, as the child who thinks that "sheep" can only be singular is placed at a disadvantage.

Word combinations: lexical distinctions

Tests of comprehension tend to either assess ability to understand single words, or grammatical contrasts, but no test sets out to test the ability to understand combinations of words when grammatical considerations are not relevant, although both the Token Test and the Sentence Comprehension Test include items that fit this description. Tallal (1975) showed that the long but grammatically simple items in part 4 of the Token Test were particularly difficult for both normal

and language disordered children. One obvious explanation would be that part 4 puts a heavy load on short-term memory. Alternatively, some subjects might have difficulty in combining the meanings of concepts, despite adequate memory. It seems logically a trivial step to understand the phrase "black cat", given that one can understand "black" and "cat" in isolation, and can remember both words, but perhaps it is not so trivial for language disordered groups. In order to see if some subjects did have particular difficulty with long but grammatically simple word strings, "word combinations" items were included in TROG. All items of this kind were 4-choice. The simplest ones consisted of adjective + noun, or noun + verb; (see appendix 1, fig viii). Other items used sentences of subject-verb-object format, in the active or passive voice (see appendix 1, fig ix). It should be noted that although some of these sentences are in the passive voice, the contrast between active and passive is not tested by these items, and it should be possible to select the correct picture in the absence of comprehension of grammatical structure, by simply selecting the one corresponding to the correct combination of content words. In session 2, several longer word combinations were added, including adjectives and prepositional phrases as well as subject, verb and object (see appendix 1, fig xiii).

Other tests

In session 2, two additional tests were given. The Peabody Picture Vocabulary Test (PPVT) form A[1,2] provided an estimate of

[1]see pp 16 to 17.

[2]Two items that were incomprehensible to most British subjects (item 25, "wiener" and item 32, "caboose") were altered (to "sausage" and "tank" respectively).

receptive vocabulary. It was felt that a rough estimate of nonverbal ability would also be of value. It was decided to use Picture Completion items from the WPPSI and WISC-R intelligence scales (Wechsler, 1967; 1974) combined to form a new 40-item scale that was quick to administer, portable, and applicable to the age range 3-6 to 9-0. Each item requires the child to identify what is missing in a picture. Picture Completion is the second most reliable performance subtest in both Wechsler scales (split-half reliability = .83 for WPPSI and .77 for WISC-R) and correlation with Performance IQ is reasonable (.55 for WPPSI and .72 for WISC-R). Combining items from both tests is likely to increase these values, since these types of correlation increase with test length. Ordering of items was determined by using items common to WPPSI and WISC-R as anchor points, so that all items from both tests which preceded such items continued to do so. Of course this short test, which measures only one type of ability, should be regarded as giving only a very crude index of nonverbal skill.

It should be noted that for both these tests, raw scores were used in all statistical analyses. Scaled scores can be calculated for Picture Completion by considering only WPPSI items for those aged below 6-0, and only WISC-R items for those aged 6-0 or more, and using the tables provided by Wechsler (1967; 1974) to convert raw scores into standard scores with mean of 10 and standard deviation of 3. Scaled scores ("IQ" scores), with mean of 100 and standard deviation of 15, can be derived from raw scores for PPVT. However, it was felt that such scores should be used cautiously, since in both tests administration had differed in some respects from standard form, as described above, and further, both tests are based on American normative samples.

Table 4

Schools used in the study

	n	age range	intake
Nursery school A	19	3-6 to 5-0	half middle class, half working class
Nursery school B	17	3-6 to 5-1	wide range of backgrounds, mainly working class
Nursery school C	18	3-6 to 5-1	wide range of backgrounds with over-representation of middle class
Nursery class D	8	4-3 to 5-1	council estate with many "problem" families
Nursery class E	7	4-4 to 5-1	council estate
Primary school F	48	5-4 to 8-11	same as nursery class E
Primary school G	50	5-2 to 8-11	wide range of backgrounds without over-representation of any social group
Primary school H	53	5-3 to 8-11	same as nursery school A

subjects

In session 1, TROG was given to 220 children aged between 3-6 and 9-0, selected from 3 state nursery schools, 2 nursery classes in primary schools and 3 state primary schools. An attempt was made to use schools from a wide range of areas, so that pupils from diverse social backgrounds were included. Detailed information about socio-economic status was not available, but brief descriptions of the intake of the schools, arrived at after discussion with the head teachers, are provided in table 4. Lists of children attending the schools were obtained, and for each 6-month age group 10 boys and 10 girls were selected at random. The youngest group spanned the range 3-6 to 3-11, the next 4-0 to 4-5 and so on. At the time of this selection, the tester had no background information about any of the children. Teachers were then asked if any of the children selected had a parent who was not a native Briton, and these children were excluded and substitutes selected. One child with Down's syndrome was excluded. Information volunteered by teachers to the effect that children were thought to be educationally subnormal, came from deprived backgrounds or were receiving speech therapy was noted but was not regarded as grounds for excluding children from the sample. It had been hoped to extend the lower age limit to 3-0, but this was not possible for two reasons. First, only a few children under 3-6 attended nursery school, and these were often atypical, being sent to nursery school because of some problem such as failure to develop language, or parental assault. Second, attempts to test a few children in the 3-0 to 3-5 age range were frequently unsuccessful because the children would not complete the test. Taking only co-operative children is likely to bias the sample in the direction of

over-representing bright children.[1] Three children aged between 3-6 to 5-0 refused to be tested, but this seemed unlikely to appreciably bias the sample. When considering the results on preschool children, it should be noted that nursery school attenders might differ from the general population in important respects, although data from another study, reported in appendix 2, suggests not, at least as far as general intelligence is concerned.

In session 2 it was possible to retest 197 of the original 220 children, three to five months after the first test session. Most of the missing children were of nursery school age. Seven new children were selected to make up the numbers in this age range. A further 77 children aged from 9-4 to 13-2 were given TROG only. These children came from one middle school, situated next to primary school G and with the same sort of intake. Time did not permit giving these children the PPVT and Picture Completion tests. Children up to the age of 9-2 were again divided into 6-month age blocks, this time with the youngest block spanning the range 3-9 to 4-2 (mean age 4-0). Older children were divided into 12-month blocks (i.e. 9-3 to 10-2 and so on). Session 2 blocks did not have equal sex ratios, although each block contained at least six children of either sex. Mean block size was 19.13 (range 16 to 22).

Method

TROG

Administration was the same for both sessions. The vocabulary pretest was given first. The tester said "I am going to show you some pictures. I want you to point to the one I say." Book 1 (vocabulary items) was opened and the child was asked "show me the boy", and so on.

[1] Appendix 2 gives some data on this point.

The introductory phrase was not standardized: "which one ...", "point to ...", "show me ...", were all used. The instruction was repeated if necessary. If the child pointed to more than one picture, he was told: "point to one picture" and the item was repeated. No child in the normative group had any difficulty in quickly understanding what was required. The tester gave frequent praise and encouragement regardless of the response. After completion of the pretest, the grammatical items were given. The child was shown each set of pictures, and while looking at them was told the test phrase. For most items this phrase was introduced by the words "show me ...", "point to ...", etc, although this introduction was omitted for long items testing tense and embedded sentences. Important contrasts were stressed, even if this sounded rather unnatural: e.g. "which one is not sitting?" If no response occurred after a few seconds, or if the child requested it, the whole item was repeated. No limit was imposed on the number of repetitions allowed, although in practice more than one repetition was never needed. Children sometimes pointed to more than one picture, in which case they were told to point to just one picture, before the item was repeated. If a child said that more than one picture was right, or that no picture was right, the item was repeated and the child given time to change his mind, without the tester insisting on a response. If a child spontaneously altered his response, the last response was counted. In test session 1, nursery school children were not given the ten items testing embedded sentences and verb singular/plural inflections, as it was already evident that older children performed at chance level on these structures. In session 2, all children were given the entire test.

Some nursery school children became restless and inattentive during the test. In these cases, a "buzzy box" was used to maintain responding. The child was allowed to press a button on a special box to make a buzzing noise after every 3 to 5 responses. This reinforcement was not contingent on correctness of response, so did not provide feedback. For the majority of the children, praise and encouragement from the tester were adequate to maintain responding.

PPVT

Except for two alterations in wording (see foctnote on page 49) this test was administered according to the instructions in the manual (Dunn, 1965).

Picture Completion

The examiner noted which items had originated from the WPPSI (P-items), which from the WISC-R (C-items), and which were common to both test (PC-items). General administration was according to the WISC-R manual. The time limit for a response was 15 seconds for P-items and 20 seconds for C-items and PC-items. Criteria for discontinuing testing were applied according to the WPPSI manual for children aged under 6-0 (ignoring C-items) and according to the WISC-R manual for children aged 6-0 or more (ignoring P-items).

Results & discussion

1. TROG

vocabulary pretest

The majority of children made no errors on the vocabulary pretest. Five errors were made on nouns in session 1 (on "apple", "box" and "sheep") and one in session 2 ("knife"). Verbs and adjectives were

Table 5

Vocabulary Pretest

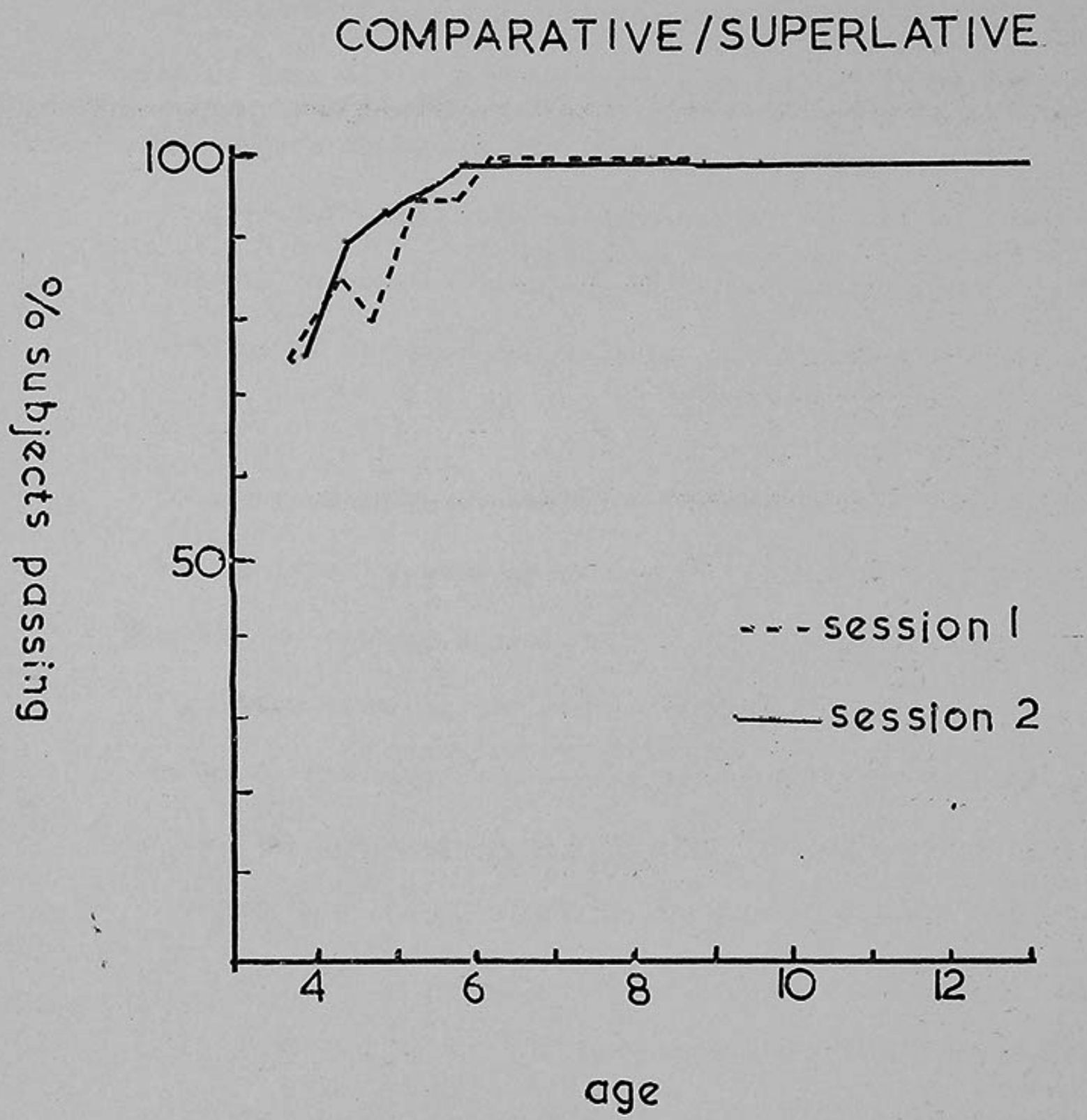
	age at which 95% correct performance attained (session 2)
all nouns	4-0
long	5-0
white	5-6
black	5-6
spotty	4-0
big	4-0
painting	4-0
eating	4-0
going	6-0
cutting	6-0
running	4-0
chasing	4-0
sitting	4-0
seeing	7-6

more difficult, with 95% correct performance not being reached until age 7-6. The pattern of performance was essentially the same in both test sessions. The relative difficulty of various words can be found in table 5. Some results are rather surprising. Many teachers expressed doubts about using "spotty" in the test because of its rarity, but in fact "spotty" was less difficult than the commoner adjectives "long", "black" and "white".

In both test sessions, at least 88% of the youngest children tested performed correctly on all pretest words except "cutting" and "seeing". Only 80% of the youngest children were correct on "cutting" (see appendix 1, fig ii), which may have been particularly difficult because it was the only verb used with an inanimate subject (i.e. an instrument rather than an agent). If this account is correct, we would predict that "cutting" would be easier if the picture showed a boy cutting something with a knife. "Seeing" (see appendix 1, fig iii) was the most difficult word used, with only 53% of the youngest children performing correctly. In the light of Chomsky's (1969) work[1] we may suggest that "see" is difficult either because attention is directed to the unusual blindfolded boy, rather than to the other boy who does not appear to be doing anything in particular, or because comprehension of "see" requires the child to "decentre", as suggested by Cambon & Sinclair (1974). In either case, this result suggests that the poor performance on Chomsky's (1969) original task with the blindfolded doll may have less to do with the complexity of the grammatical structure used, than with the difficulty of the verb. However, Cromer's (1970) work with a different verb (bite) suggests this is not the whole answer.

[1] see chapter 1.

Fig. 1


COMPARATIVE/SUPERLATIVE
100
50
% subjects passing
---session 1
——session 2
4
6
8
10
12
age

As mentioned above, failure on vocabulary pretest items meant that failure on later items using those words could not be assumed to be due to grammatical inability. Therefore errors on grammatical items were disregarded if they could have resulted from the child's lack of comprehension of a component word, as evidenced by his performance on the pretest.

grammatical contrasts

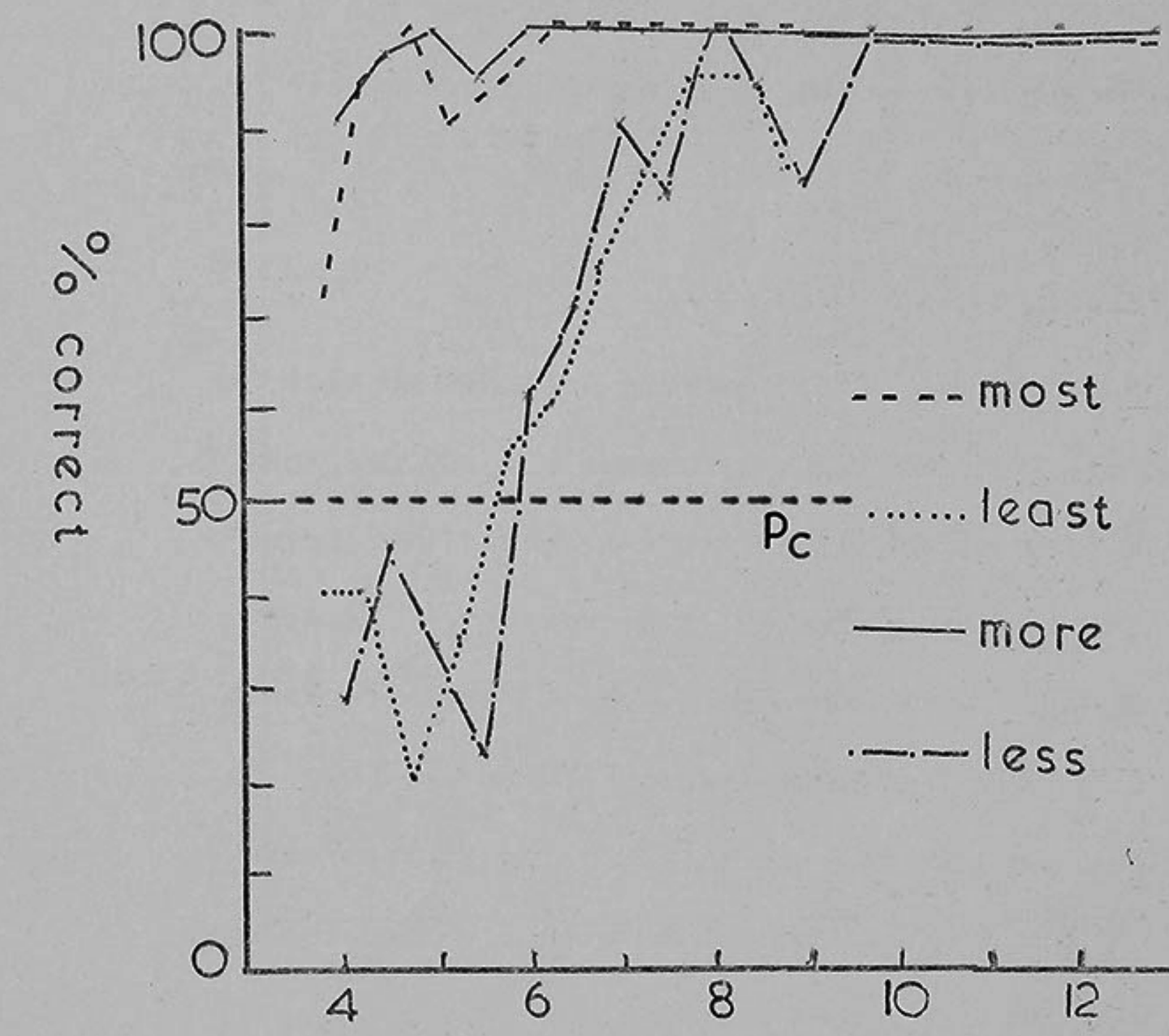
For each contrast, sufficient items were included so that the likelihood of getting all items right by chance was .05 or less. Thus for each child it was possible to award a "pass" (all items correct) or "fail" (one or more errors) on a contrast. Unless otherwise indicated, the data discussed below will be in terms of this dichotomous scoring. Instances where children persisted in pointing to more than one picture, or pointed to no picture were counted as errors. Errors which could have resulted from the child not knowing the component words were discounted.

comparative/superlative

Percentages of subjects passing the items testing comparative/superlative inflections are shown in figure 1. This proved to be the easiest set of items tested, with no errors occurring beyond age 6-0. However, as mentioned on page 44, on the two superlative items, correct performance was not critically dependent on the presence of the inflection, and indeed in both test sessions slightly fewer errors occurred to "biggest" and "longest" than to "big" and "long".[1] Similar results were obtained by Nelson & Benedict (1974) who presented

[1] This is not a statistically significant difference. Very few errors occurred to either inflected or uninflected adjectives.

Fig. 2

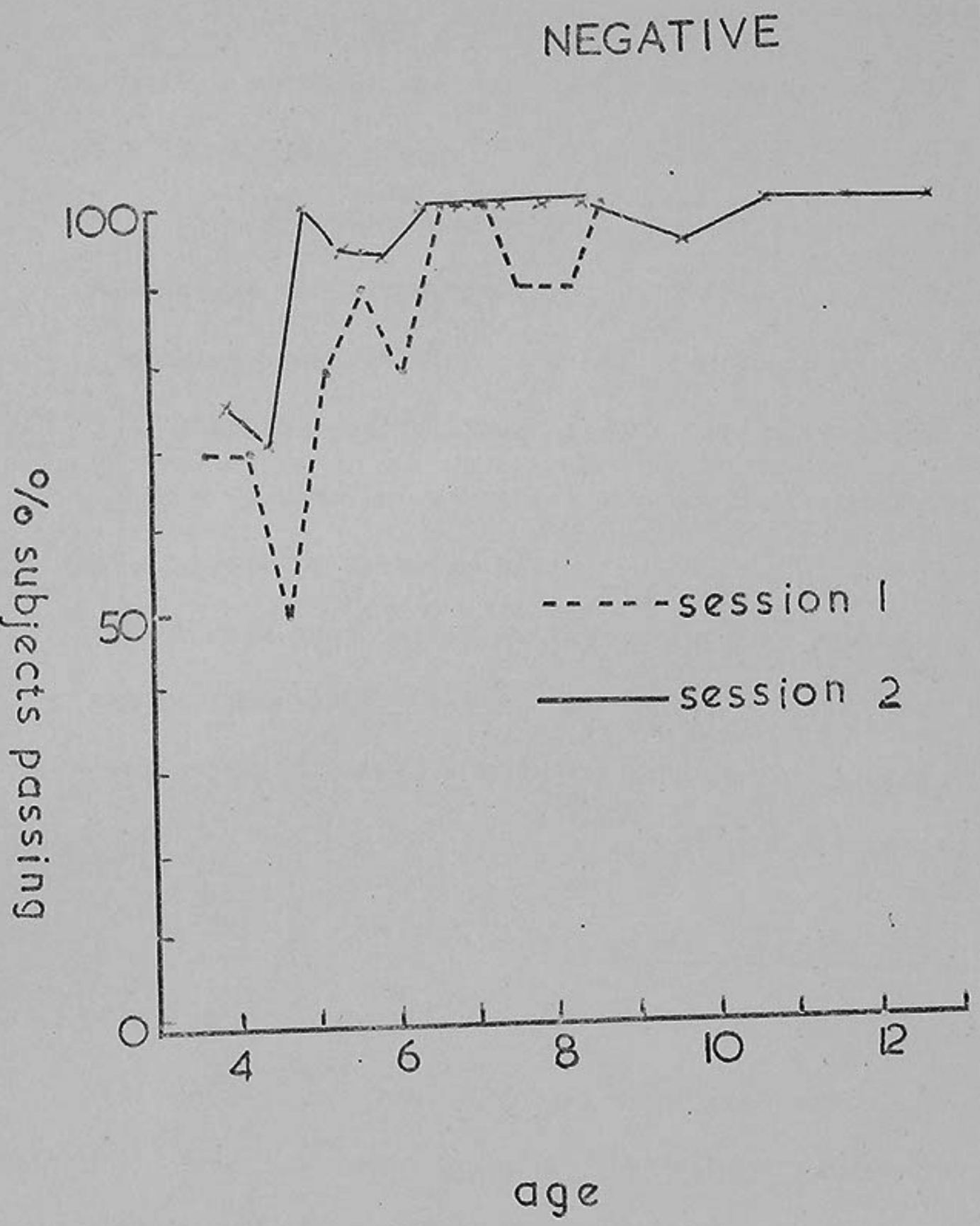

COMPARATIVE TERMS
100
50
0
% correct
Pc
most
least
more
less
4
6
8
10
12
age

adjectives in uninflected and comparative forms to 4- to 6-year-olds. Some of the adjectives they used were relative whether inflected or not (e.g. tall, fat, etc), whereas others did not indicate a contrast unless inflected (e.g. cloudy, round, etc). Their analysis of data is not at all clear, but it appears that comparatives are more difficult than uninflected adjectives in most conditions except when relative adjectives are used, when there is no difference between inflected and uninflected forms. They conclude that relative adjectives are processed in a comparative sense, whether or not inflections are added. This means that in order to be valid, a test of comprehension of comparative/superlative inflections should use only adjectives like "spotty", which do not imply a contrast unless inflected. Surprisingly few pictureable adjectives in the vocabularies of young children are of this type.

comparative terms: more, less, etc

In considering these results, it should be noted that unlike the other contrasts for which data is presented, there is a probability of .5 of correct performance resulting by chance, since only one item was used for each term. Figure 2 shows that results on TROG replicate the finding by other workers (Donaldson & Balfour, 1968; Donaldson & Wales, 1970) that "less" is more difficult than "more", with the former not being fully acquired by age 7-0 (Palermo, 1973). "Least" and "most" follow an identical pattern. Further, as remarked by previous researchers, performance on "less" is below chance level for the youngest children. This has been interpreted as meaning that for young children both "less" and "more" are understood as referring to quantity, but are thought to be synonymous, both meaning "more". This interpretation has been questioned by Carey (1976), who

Fig. 3


NEGATIVE
100
50
0
% subjects passing
4
6
8
10
12
age
session 1
session 2

demonstrated that children's responses to a meaningless word, "tiv", resembled their resporses to "less". This seems to indicate that the below chance level performance obtained with "less" is not so much due to children assigning the wrong meaning to "less" as to their having a strong response bias toward the alternative corresponding to "more", which is particularly evident when they are given words which they do not understand.

negatives

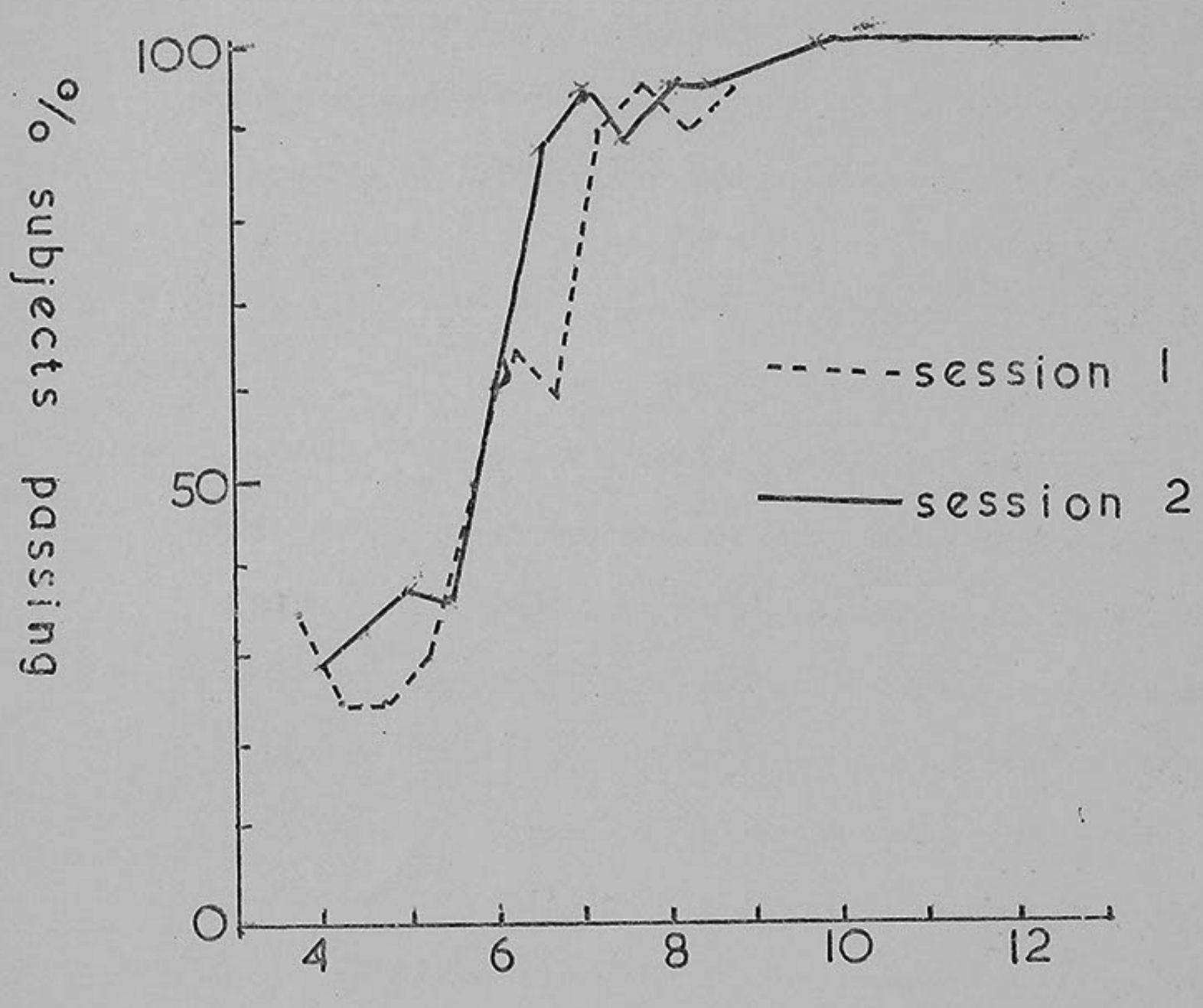
Results for sessions 1 and 2 are shown in figure 3. It can be seen that 90% of children understand "not" by age 6-6. Error rates did not differ significantly between the five items. The author had predicted that children would either understand the negative and get all items right, or fail to comprehend and get all items wrong. Although the latter pattern did occur for a few nursery school children, it was commoner to find children who made some mistakes without getting all items wrong. For example, in session 1, out of a total of 220 children, 3 made errors on all five negative items, 12 made errors on between 2 and 4 items, and 21 made one error on a negative. The remainder were correct on all items. This indicates that an "all-or-none" view of comprehension is wrong. There appears to be a stage between no comprehension and complete understanding, where "not" is understood, but only with difficulty and not consistently.

singular/plural nouns

Results are shown in figure 4. 90% correct performance is attained by age 7-0[1]. The five items did not differ significantly

[1]Errors always involved selection of the distractor showing the same noun (e.g. A in fig x, appendix 1), i.e. singular/plural items were effectively 2-choice items, and were treated as such in later analyses (see appendix 3).

Fig. 4

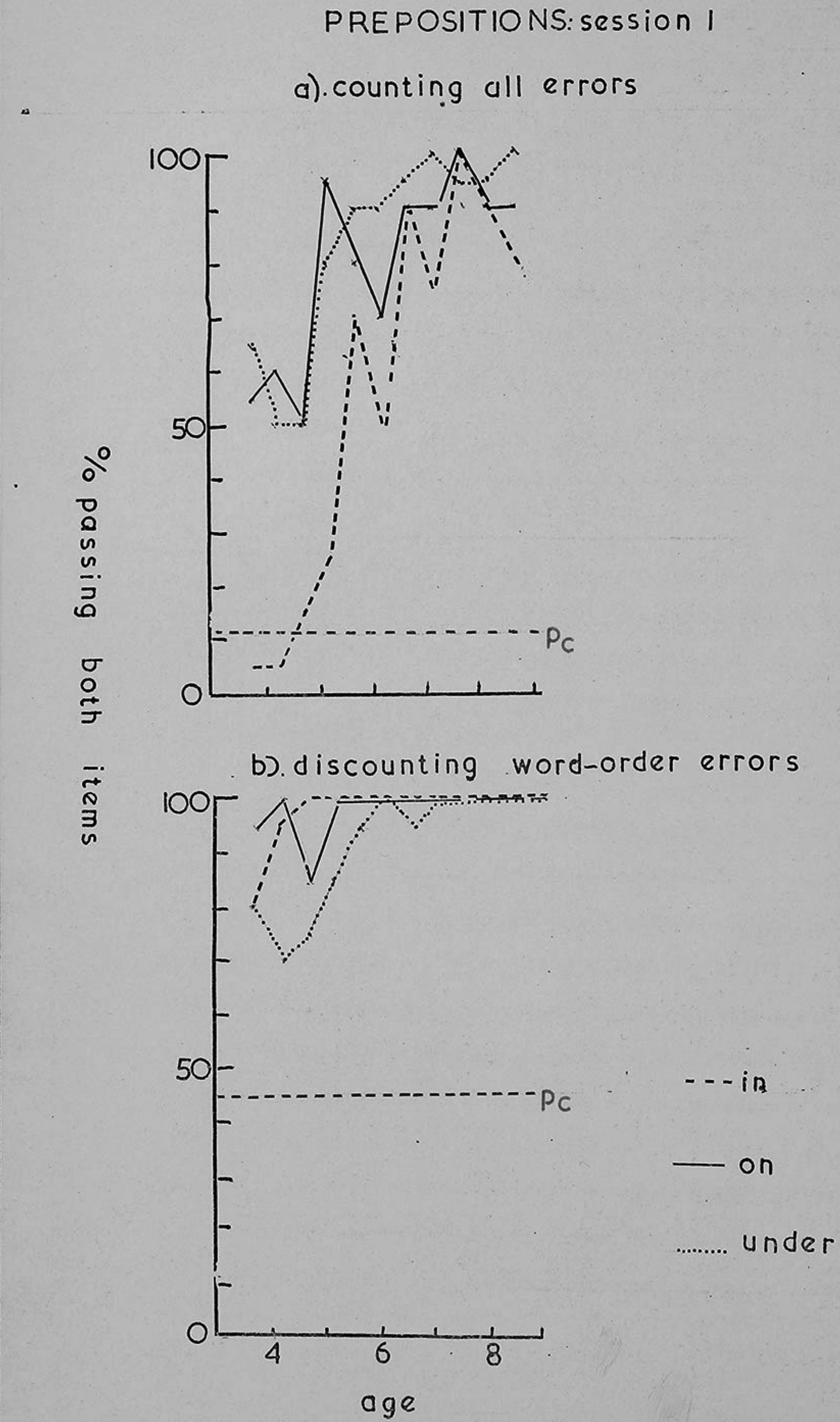

NOUN: SINGULAR/PLURAL
% subjects passing
100
50
0
4
6
8
10
12
age
session 1
session 2

in difficulty. It will be noted that the age of acquisition found here is much later than would be expected on the basis of Brown's (1973) work, where children were found to use plural inflections correctly in 90% of obligatory contexts by around 2½ years. This adds to our doubts about the validity of these items. However, it should be noted that Brown himself (page 331) reports that children who used these inflections correctly in their speech did not respond consistently to less ambiguous tests of comprehension, such as commands like "give me the pencil(s)" when the child has equal opportunity to present one or more pencils. Brown suggested that children may be unused to processing sentences differing in a single feature, and that in normal speech to the child the plural would be doubly marked, e.g. "give me some pencils". This raises the possibility that the accepted pattern of comprehension preceding production may be contravened in some circumstances. This notion will be discussed in detail later.

prepositions

In session 1, results for the three prepositions proved to be very different, so are shown separately in figure 5a, which gives the proportion of children responding correctly to both items for a given preposition. Probability of correct performance resulting by chance on any pair is .11. Many more errors occurred to "in" than to "on" or "under". More detailed analysis revealed that virtually all these errors involved selection of the R-distractor. If we count only selection of the X-distractor as an error, the order of difficulty for the prepositions alters, with "under" being the most difficult. In session 2, separate items were used to test the distinction between correct response and X- and R-distractors.

Fig. 5

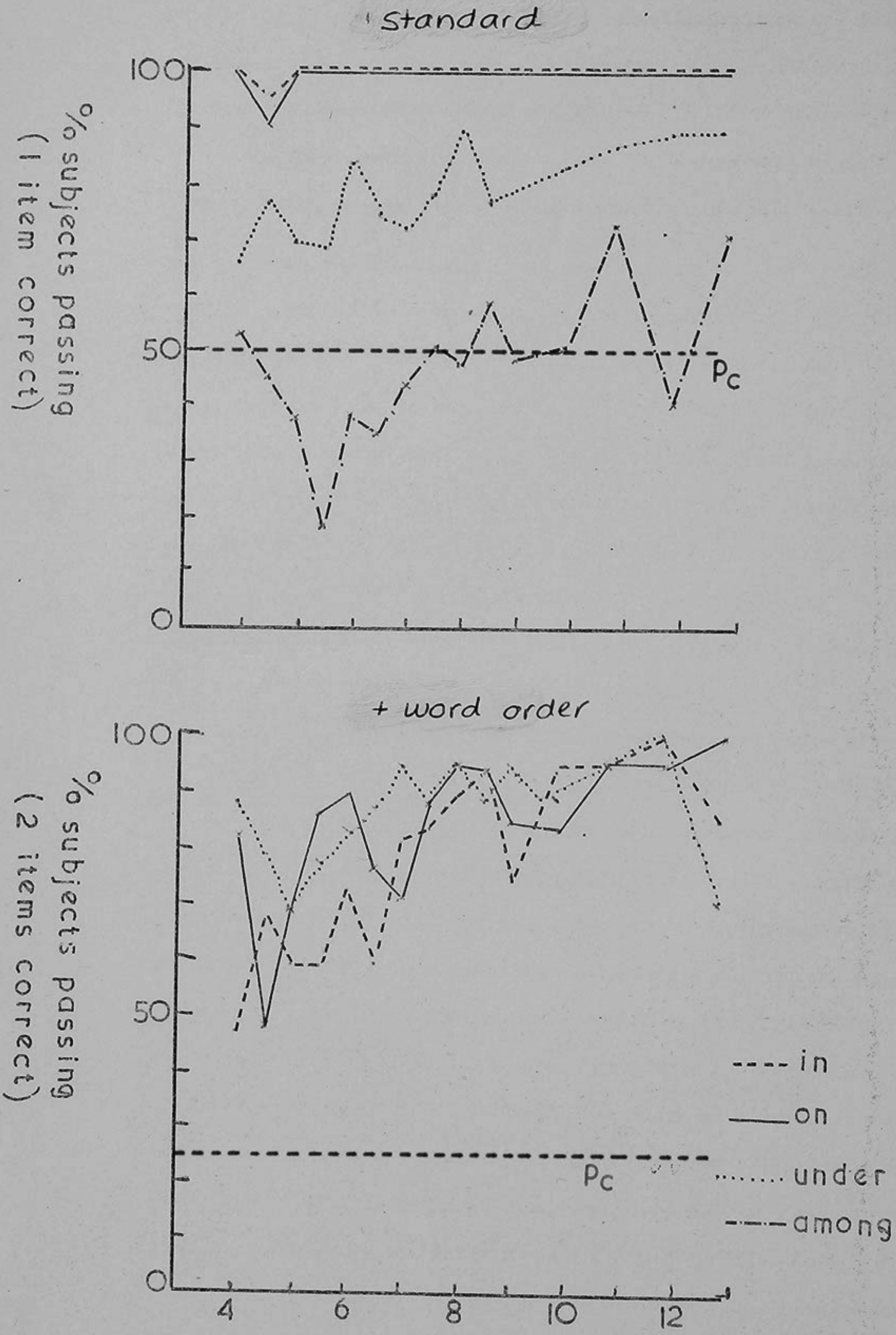

PREPOSITIONS: session I
a). counting all errors
100
50
0
% passing both items
Pc
b). discounting word-order errors
100
50
0
Pc
--- in
—— on
..... under
4
6
8
age

Results for standard prepositions (with X-distractors) and prepositions + word order (with R-distractors) are shown in figure 6. Standard preposition items testing "in" and "on" are nearly always correct, but "under" and "among" give more difficulty. Less difference between prepositions was fourd for items involving word order.

"In", "on" and "under" were used in TROG because previous workers had obtained interesting results with these prepositions. Clark (1973) used a task where children were presented with toy objects and animals and asked to carry out instructions of the kind "put the x in/on/under the y" where x refers to an animal and y to a reference point object. Clark showed that simple analysis of correct responses showed clearcut developmental trends, with "in" being easiest, and "under" most difficult. However, on investigating the types of errors which occurred, Clark found that these followed a regular pattern, namely, children tended to put animals <u>in</u> containers and <u>on</u> objects with flat surfaces regardless of the instructions. She pointed out that her original results which seemed to reflect differential acquisition of the three prepositions might in fact have very little to do with linguistic comprehension. The data could be the result of the child obeying nonlinguistic rules, whereby if the instruction concerned a container, the animal was put in it, and if it contained an object with a flat surface, then the animal was put on it. A high proportion of children's correct responses and errors could be accounted for in terms of these rules. Wilcox & Palermo (1974) proposed that the physical characteristics of the objects used were of less importance than the child's expectations of how objects behave. Using a new set of objects, they devised instructions which were either congruent or incongruent with the

Fig.6

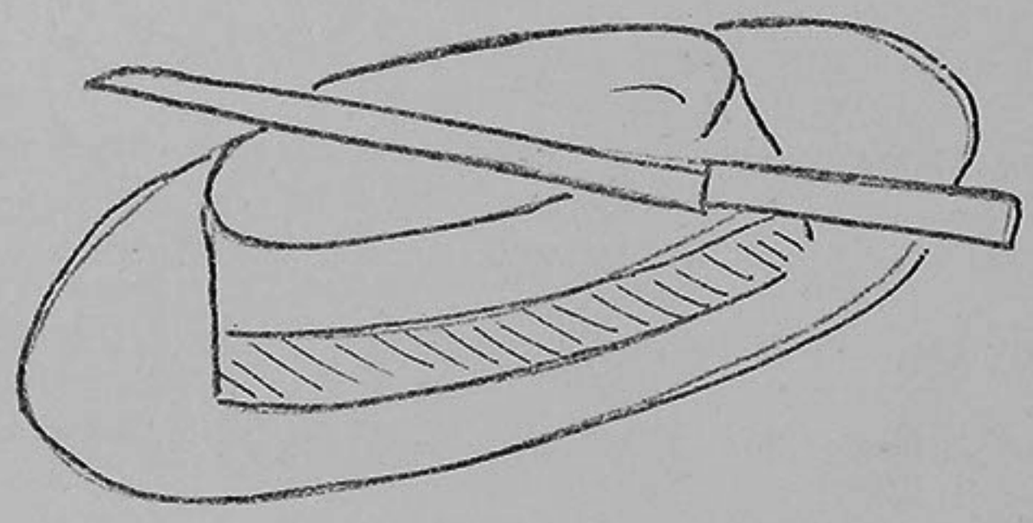
PREPOSITIONS : session 2



typical behaviour of objects in the real world; e.g. "put the boat under the bridge" is congruent, whereas "put the bridge under the boat" is incongruent. They showed that congruence had an important effect on response. An even clearer demonstration of such an effect was given by Grieve, Hoogenraad & Murray (1977) who gave children a variety of instructions involving "in" "on" and "under" using a big and little box as materials. In one condition, the boxes were referred to as such. In other conditions, children were told that the boxes were a) a table and chair; b) a baby and bath and c) a cup and table. Very few children had difficulty in identifying the two boxes as these imaginary items, and their responses to the boxes were shown to be highly dependent on the labelling of the boxes. Grieve et al also included a further condition similar to the picture-pointing task used here, where children were presented with three arrangements of the boxes and had to select the one corresponding to the experimenter's utterance. Error rates were much lower in this task than in the other conditions for "under" instructions. Few errors were made to "in" or "on" in either condition. Grieve et al propose that instructions to "put x under y" are particularly difficult in the manipulation conditions because the child is prone to start off by picking up x, the first named object, when in fact it is y that must be moved. A similar argument is advanced by Huttenlocher & Strauss (1968) who found that subjects' responses to instructions such as "x is on top of y" and "x is under y", when x and y were blocks in a ladder, depended crucially on whether x or y was moveable. It was easier to place the item that was the grammatical subject of the sentence relative to its object than vice versa.

Fig.7

"the hat under the knife" from a different view



The TROG paradigm was intended to minimize the effects of a priori expectations, ease of motor response and other biases inherent in the "obeying commands" paradigm. The results obtained here indicate that when word order is made critical, this will have a dramatic effect on accuracy of comprehension. A preposition may be regarded as having both lexical and syntactic aspects. The lexical aspect is the spatial or other relationship denoted by the word, whereas the syntactic aspect is concerned with meaning conveyed by word order. Items using X-distractors test knowledge of the former, those using R-distractors the latter. Further, pragmatic constraints will limit the nouns which can follow a preposition, e.g. the noun following "in" must be a container. In Clark's experiments, word order was often redundant because of pragmatic constraints: e.g. the cat may be in the bed, but the bed cannot be in the cat. TROG results show that when word-order is not redundant, it can pose considerable difficulties for children, long after they have learnt to comprehend the lexical aspect of a preposition. Analysis of responses to individual items proved interesting. The standard item: "the hat under the knife" (see appendix 1, fig xix) was surprising, because many errors occurred on this item, and 14 children replied "there isn't one", a response which was very rare to other items. Perhaps children have a narrower concept of "under" than the experimenter, who used it here synonymously with "below". Possibly children interpret "under" as meaning "hidden from view" or perhaps they expect to see something which is "under" from a different viewpoint (see e.g. figure 7). We may compare work by Parisi & Antinucci (1970) who found that prepositions whose meaning depended on the particular point of view of the subject gave more difficulty than those in which only topological space is represented. This

Table 6

Prepositions: errors on individual items

	no. of children selecting R-distractor in session 1 (n = 220)	no. of children making errors on "+ word order" items in session 2 (n = 284)
the apple on the hat	6	4
the ball on the cup	37	44
the knife under the apple	9	27
the knife under the hat	8	15
the cup in the box	37	25
the box in the cup	58	48

interpretation could account for the other surprising finding that the "+ word order" prepositions using "under" (see appendix 1, fig xviii) gave less difficulty. Presumably here the uncertainty about which spatial viewpoint to adopt was resolved. Asking children about why they found the standard "under" item difficult did not elucidate matters much, except to suggest that "the knife under the hat" is a more probable statement than the converse - some children said "I thought you meant the knife under the hat" before responding. It would be interesting to examine children's responses if asked to draw a hat under a knife.

A second point of interest was that error rates of pairs of "+ word order" items testing the same preposition differed significantly for "in" and "on", as shown in table 6. "The apple on the hat" was much easier than "the ball on the cup", and "the cup in the box" easier than "the box in the cup". Post hoc reflection on these differences in difficulty gave rise to the following speculations. "The ball on the cup" (see appendix 1, fig xx) is a bad item in that the situations depicted in the C-picture does not clearly represent the dictionary definition of "on":

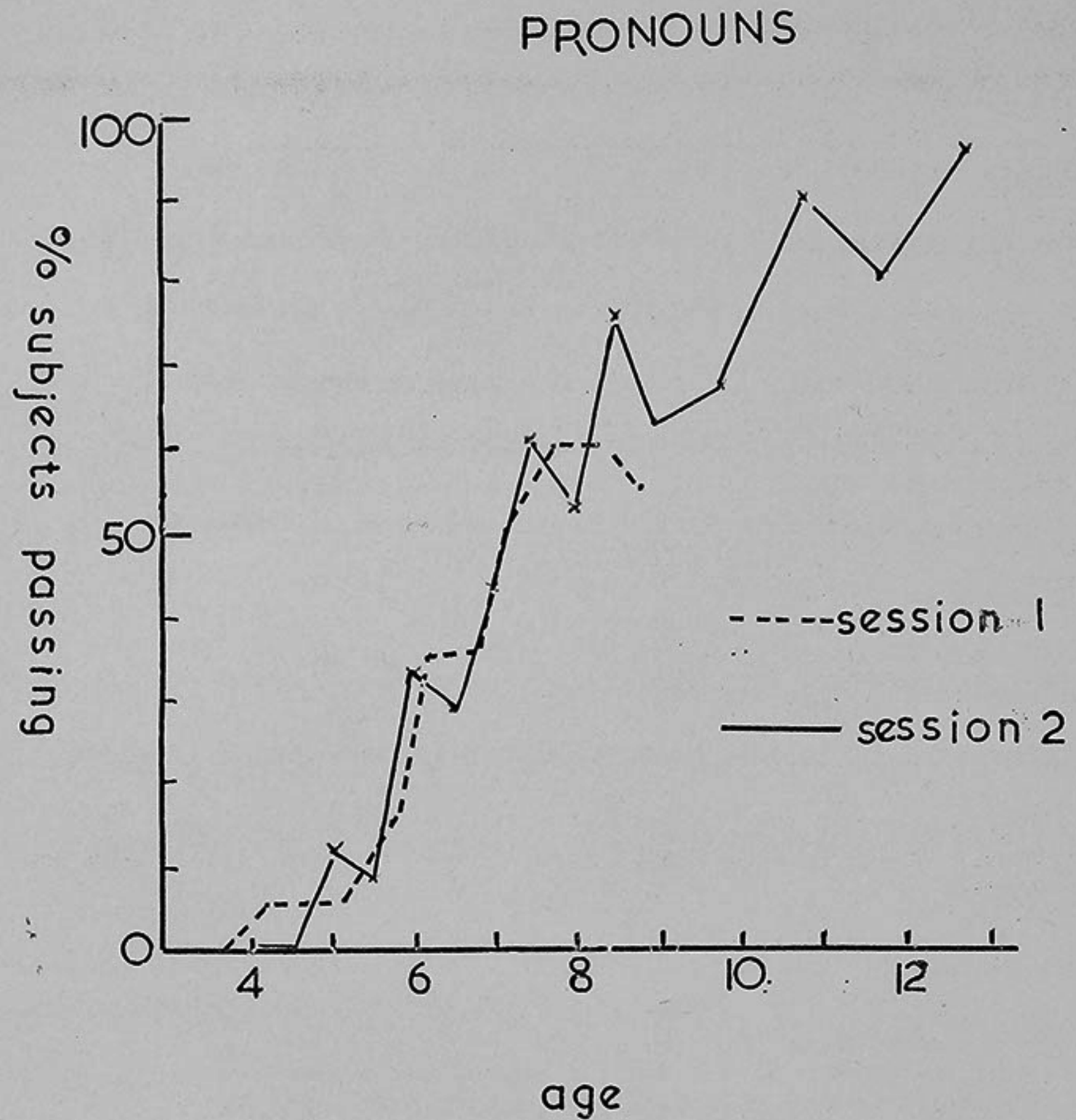
> "Above and in contact with, above and <u>supported by</u>"
> (my underlining)
>
> Onions (1956) Shorter Oxford English Dictionary

whereas the R-picture does. Differences between "the cup in the box" and "the box in the cup" might result because cups typically contain liquids and not solid objects, or perhaps the former situation is more probable than the latter.

<u>pronouns</u>

Overall results are shown in figure 8. The different pronouns

Fig. 8


PRONOUNS
100
50
0
% subjects passing
4
6
8
10
12
age
----session 1
—— session 2

were of differential difficulty, as shown in figure 9. The most surprising finding was the difficulty of "it" (see appendix 1, fig vi). Despite the fact that they used "it" correctly in expressive speech[1] and would be unlikely to call a boy "it" or a ball "he"[2], children as old as 9-0 still made errors in this pronoun, and children younger than 6-0 performed below chance level. Several post hoc explanations may be considered.

First, the questions formed to test personal pronouns ("show me he/she/it" or "which one is he/she/it") sounded rather odd. Indeed, because some children in pilot sessions had looked confused, the variant "which one would you call 'it'?" was allowed for the inanimate pronoun if the initial question produced puzzlement. However, children did not appear to be helped by such rephrasing. Second, although "it" is very frequent in expressive speech, it only rarely signals an important contrast in meaning with "he", "she" or "they", since, as Tanz (1977) has pointed out, pronouns are normally used by a speaker only when he believes that the referent is fully and uniquely recoverable by the hearer, because of previous mention or uniqueness in a context. Thus in most sentences, it is evident from the context that an inanimate referent is involved. Indeed, if the distinction between animate and inanimate pronouns were not

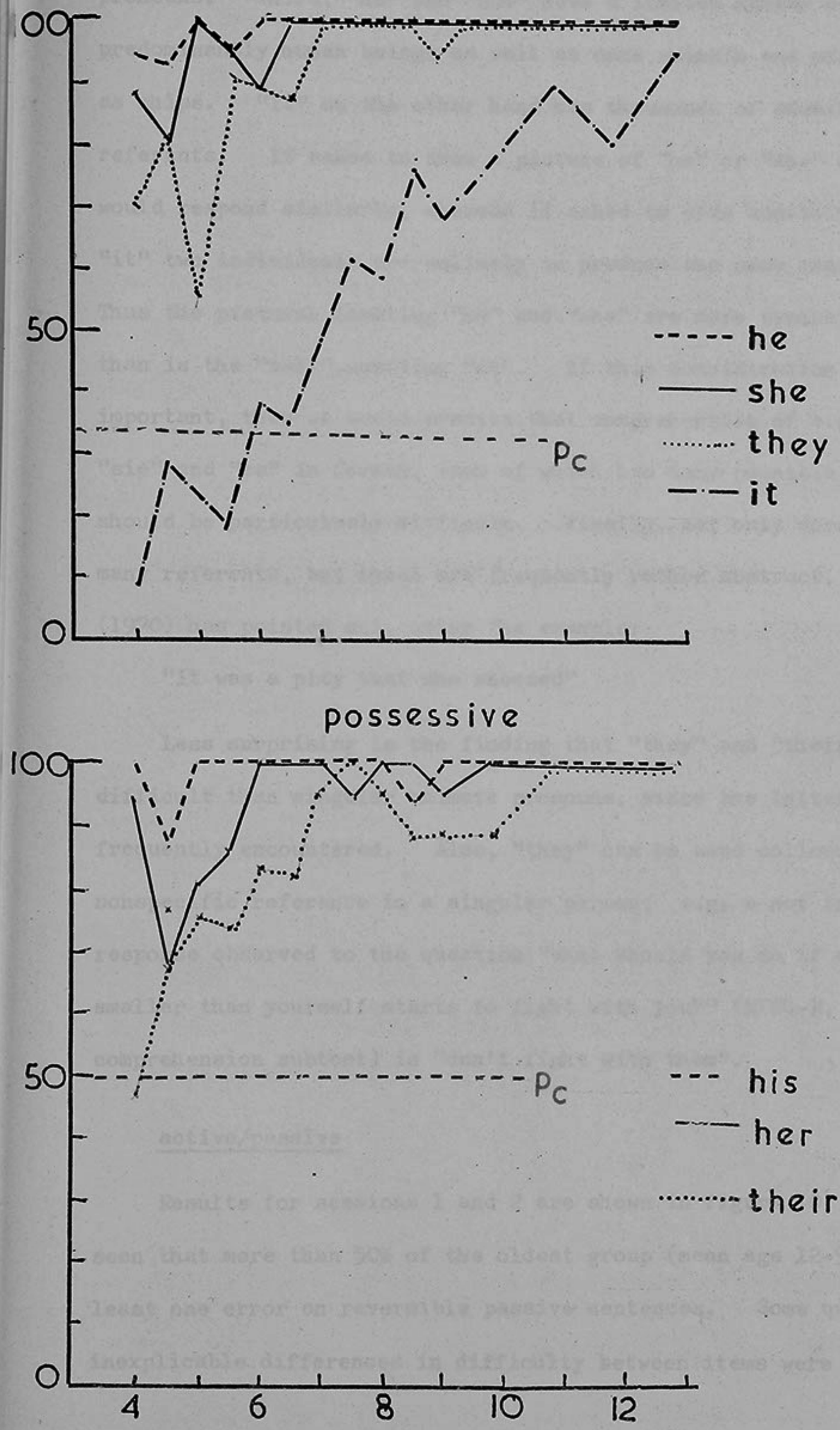
[1]Expressive speech was not evaluated in this sample, but it is hoped that few would regard this statement as unreasonable. Certainly casual observation of the children and discussion with their teachers supported it. Huxley (1970) working with much younger children reported instances of confusion between "he" and "him", and "she" and "her", but does not mention any confusion between "he", "she" and "it".

[2]Newcombe (personal communication) has noted that some varieties of Oxfordshire dialect do use animate pronouns to refer to inanimate objects, but this cannot account for the principal error occurring here, which involved selection of a boy or girl when asked for "it".

ig. 9

PRONOUNS: individual items

2nd session



usually redundant, it would be difficult to understand languages such as French which do not have a spearate term for inanimate pronouns. Third, "he" and "she" have a limited number of referents predominantly human beings as well as some animals and oddities such as ships. "It" on the other hand has thousands of possible referents. If asked to draw a picture of "he" or "she" most people would respond similarly, whereas if asked to draw something called "it" two individuals are unlikely to produce the same response. Thus the pictures denoting "he" and "she" are more typical exemplars than is the "ball" denoting "it". If this consideration is important, then we would predict that comprehension of e.g. "er" "sie" and "es" in German, each of which has many possible referents, should be particularly difficult. Finally, not only does "it" have many referents, but these are frequently rather abstract, as Huxley (1970) has pointed out, using the example:

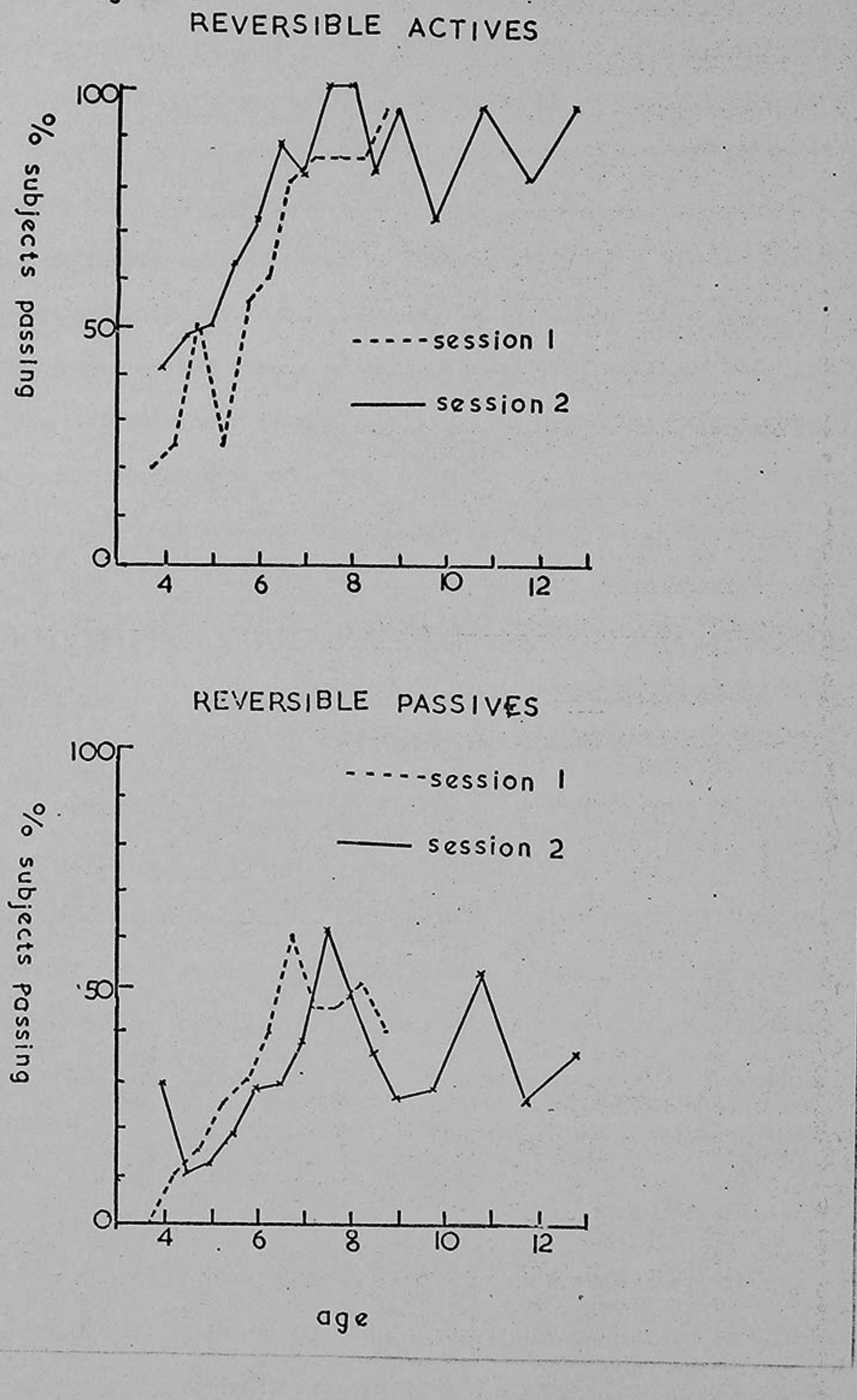
"It was a pity that she sneezed"

Less surprising is the finding that "they" and "their" are more difficult than singular animate pronouns, since the latter are more frequently encountered. Also, "they" can be used colloquially for nonspecific reference to a singular person; e.g. a not infrequent response observed to the question "What should you do if a boy much smaller than yourself starts to fight with you?" (WISC-R, comprehension subtest) is "don't fight with them".

active/passive

Results for sessions 1 and 2 are shown in figure 10. It can be seen that more than 50% of the oldest group (mean age 12-9) made at least one error on reversible passive sentences. Some quite inexplicable differences in difficulty between items were found, as

Fig. 10


REVERSIBLE ACTIVES
% subjects passing
100
50
0
4
6
8
10
12
session 1
session 2
REVERSIBLE PASSIVES
% subjects passing
100
50
0
4
6
8
10
12
session 1
session 2
age

shown in table 7. "The dog is chased by the boy" was much more difficult than the other sentences. There seemed no reason to regard the picture for this item as unsatisfactory, and left-to-right ordering of elements in this picture could be shown not to be critical. It could be said that it is more usual for dogs to chase boys than the converse, but "the boy chases the dog" was the easiest reversible active sentence. The even more unlikely sequence "the cat chases the dog" gave no particular difficulty, which makes one suspicious of post hoc explanations in terms of differential a priori probabilities.

Much research has been conducted on comprehension of active/passive distinctions. Bever (1970) worked with children as young as 2 years, whom he asked to act out sequences such as "the horse kisses the cow" and "the horse is kissed by the cow", using toy animals. On the basis of his results he proposed three stages through which children pass. Initially, performance on reversible actives is good, but reversible passives are not understood, with performance remaining at chance level. Bever points out that this difference indicates that children can discriminate active and passive sentences but cannot comprehend the latter. Slightly older children continue to perform well on actives, but their performance on reversible passives drops to below chance level. Bever argues that this drop indicates that the child has started applying a regular but incorrect strategy to passive sentences, in which the first-named object is taken as the agent of the sentence and the last named as the object. Use of this strategy (the agent-action-object strategy) would lead to consistently incorrect performance on passives. Finally, performance rises above chance level as the child starts to apply the correct strategy. There has been some

Table 7

Reversible active and passive sentences: errors on individual items

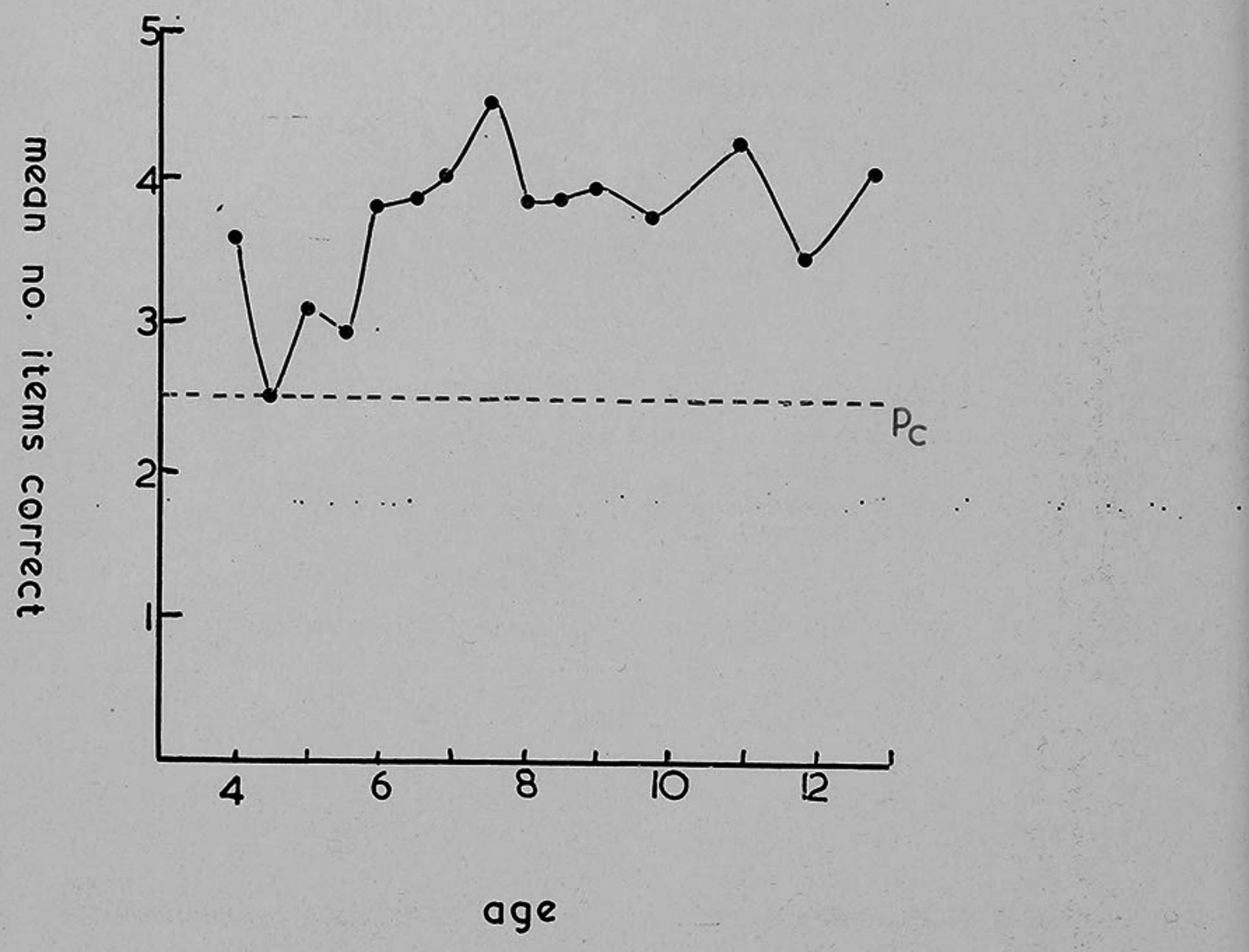
	no. of children making errors in session 1 (n = 220)	no. of children making errors in session 2 (n = 284)
the boy chases the dog	18	18
the girl chases the cat	21	12
the cat chases the dog	32	19
the boy sees the girl	34	21
the cow pushes the boy	39	28
the cow is chased by the sheep	51	56
the cat is chased by the cow	52	67
the boy is pushed by the girl	57	75
the boy is seen by the girl	64	65
the dog is chased by the boy	92	112

debate as to whether the dip in performance on reversible passives is replicable. Maratsos (1974) did find that 4-year-olds did more poorly on reversible passives than 3-year-olds. Strohner & Nelson (1974) studied comprehension of reversible sentences in 2 to 5-year-olds in a task where children were required to act out sentences using puppets. They claimed not to replicate the dip in performance found by Bever, but inspection of their data suggests that there is a dip, but at 3 years rather than 4 years. 2-year-olds scored at chance level on reversible passives (42% correct) whereas 3-year-olds performed below chance (27% correct). However, the 2-year-olds were a specially selected co-operative sample so this result could be an artifact. De Villiers & de Villiers (1973) pointed out that Bever only counted correct or reversed responses in his data analysis; irrelevant or idiosyncratic responses were ignored. In a replication of Bever's study, using 33 children aged from 1-7 to 3-1 from volunteer student parents, they showed that the most common response of young children to commands such as "make the horse kiss the cow" was for the child to kiss either the horse or the cow. If these Child-as-Agent responses were scored as errors, then performance of the youngest group on reversible actives was correct on only 70% of trials. The dip in performance on reversible passives observed by Bever was replicated, although this was clearer when children were grouped by MLU rather than age. Tendency to reverse passive sentences was particularly evident at MLU 3.0 to 3.5. Turner & Rommetveit (1967) used a different procedure where children were presented with a single picture and two sentences and were required to select the sentence corresponding to the picture. They too found below chance performance on passive sentences for 4-year-olds.

Fig·11

REVERSIBLE PASSIVE

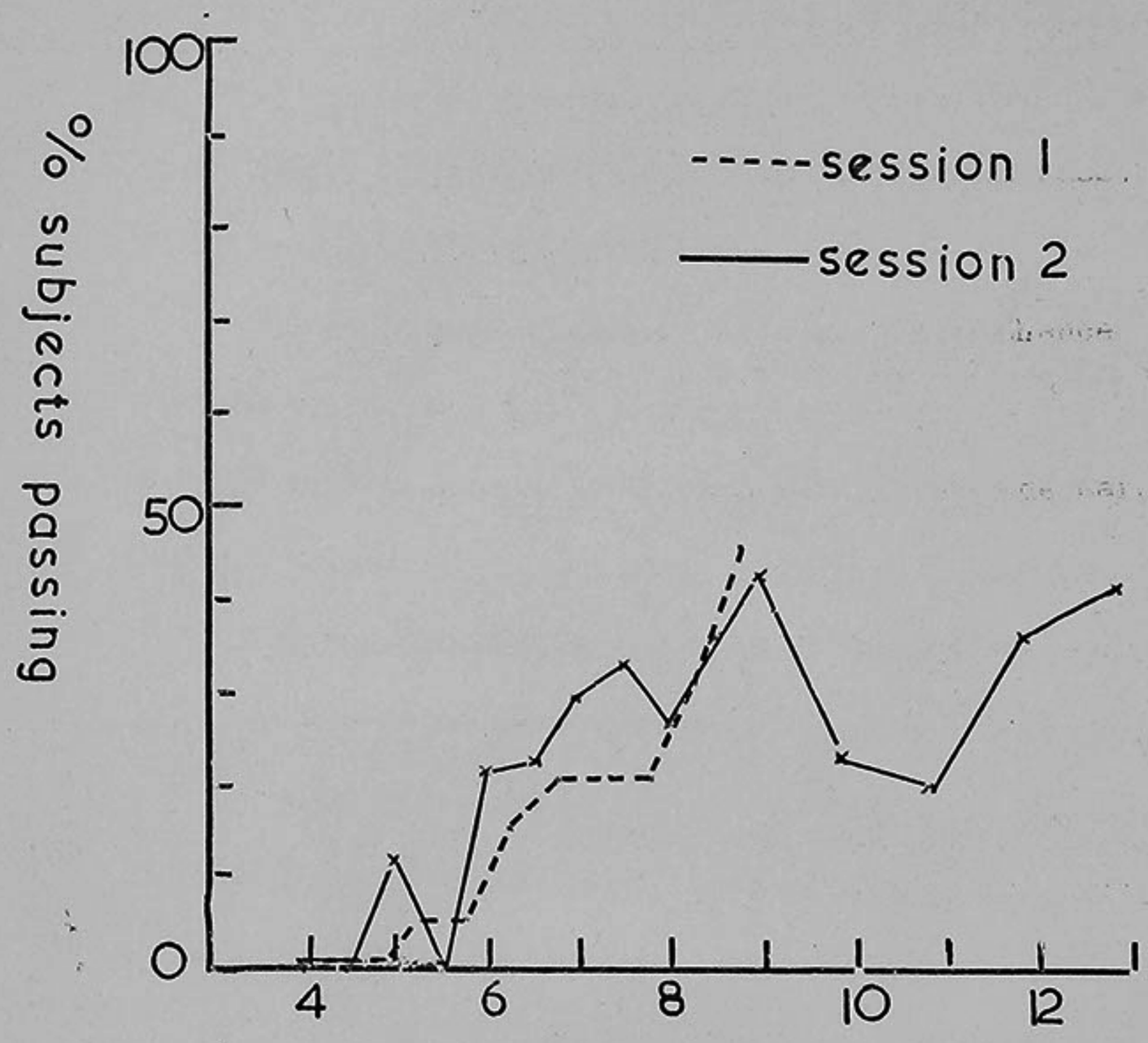
mean number of items correct: session 2



In all of these studies, performance is assessed in terms of total percentage of errors in a population of children. If the agent-action-object strategy does operate, we would expect to find evidence of this in the performance of individual children, some of whom should be consistently wrong in a series of reversible passive sentences. The number of children who scored 0 out of 5 in reversible passives in TROG was not greater than that expected by chance (6 out of 220 children in session 1) and these did not tend to be the youngest. Thus, as also found by Sinclair, Sinclair & de Marcellus (1971), children who make errors on reversible passives do not behave as we might expect if they were following Bever's agent-action-object strategy, but rather make a mixture of correct and incorrect responses. The second point to note is that when mean number of correct responses is plotted by age (see figure 11) troughs and peaks in performance are common, even though each data point is based on around 100 observations. It would seem hazardous to attach much importance to these fluctuations which probably reflect random noise. Obviously, a longitudinal study is needed to convincingly demonstrate a true decrease in performance on reversible passives at around age 4-0.

The difference between results found in this study and Bever's might be a consequence of the different test paradigms used. Huttenlocher, Eisenberg & Strauss (1968) tested children's ability to comprehend active and passive sentences in a paradigm where the subject was presented with a moveable (red) truck and shown a fixed (green) truck and was told to make it so that "the red/green truck is pushing/pushed by the green/red truck". The verb "pulling" was also used. Reaction time rather than accuracy was the dependent variable. It proved easier to place the truck that was the grammatical subject

Fig. 12


TENSE
session 1
session 2
% subjects passing
100
50
0
4
6
8
10
12
age

with respect to its object than vice versa. Thus the difficulty with acting out passive sentences may be similar to the difficulty with acting out sentences involving "under", discussed on p. 61 . The child may have a tendency to pick up the first-named object, which will lead him into difficulties when it is the second named that need to be moved.

A final point to note is that Sinclair, Sinclair & de Marcellus (1971) propose that the particular semantic and syntactic restraints attaching to the verbs used in this sort of task will be critical in determining comprehension. These workers used a variety of verbs in reversible structures, and found that even when it is demonstrated that children understand these verbs in simple lexical constructions, they may give differential difficulty in reversible sentences. They suggest that some verbs are more familiar in this type of syntactic and semantic context than others.

Tense

Results from sessions 1 and 2 are shown in figure 12. It can be seen that differences in administration between the two sessions had no obvious effect on performance. As might be expected in view of the diverse items used, individual items showed significant variation in difficulty. Analysis of errors revealed that B, the present tense picture (see appendix 1, fig xi), was by far the most popular choice for all items, comprising 89% of errors on present perfective items and 74% of errors on future tense items. It is unlikely that this was a simple position preference, since no such preferences were found for other items. It seems more likely that children picked B as the best exemplar of the verb they heard, with semantic factors outweighing syntactic ones. E.g. in the sentence "she has eaten the apple" the picture most obviously concerned with "eat" is B.

In both test sessions, many more errors were made to "the knife has cut the apple" than to "she has eaten the apple". This could have been because the picture for the latter was less ambiguous than the former, or because "cut" is irregular in that the past participle is not inflected. "Gone", an irregular inflected past participle, was the easiest past participle , a not surprising result when we consider that "gone" appears very early in the development of expressive speech (Brown, 1973) and may well be learned as a separate lexical item unrelated to the verb from which it is derived, equivalent to "absent".

By far the most difficult item was the future tense item "the boy will run" (see appendix 1, fig xii), which produced more errors than "she will paint", performance being below chance level. Perhaps the boy in picture B was not regarded as yet running, as no movement was evident - children might feel he will run when unfrozen from the picture. As mentioned above (p.47) there was reason to doubt the validity of this set of items, because of the difficulty of designing unambiguous pictures. It is therefore of interest to contrast these results with those obtained by Herriot (1969) using a more satisfactory technique. Herriot argued that a true measure of comprehension of tense inflections should use nonsense words to avoid accuracy resulting from children's familiarity with specific instances. 24 subjects aged from 3-0 to 4-0 were presented with 2 identical toys, woodpeckers attached to wooden poles. The woodpecker would remain stationary at the top of the pole unless tapped, when he would descend in a series of jerky movements. Future, present and past time contrasts were tested in pairs; e.g. the technique for the future/ present distinction was to show both woodpeckers initially at the top

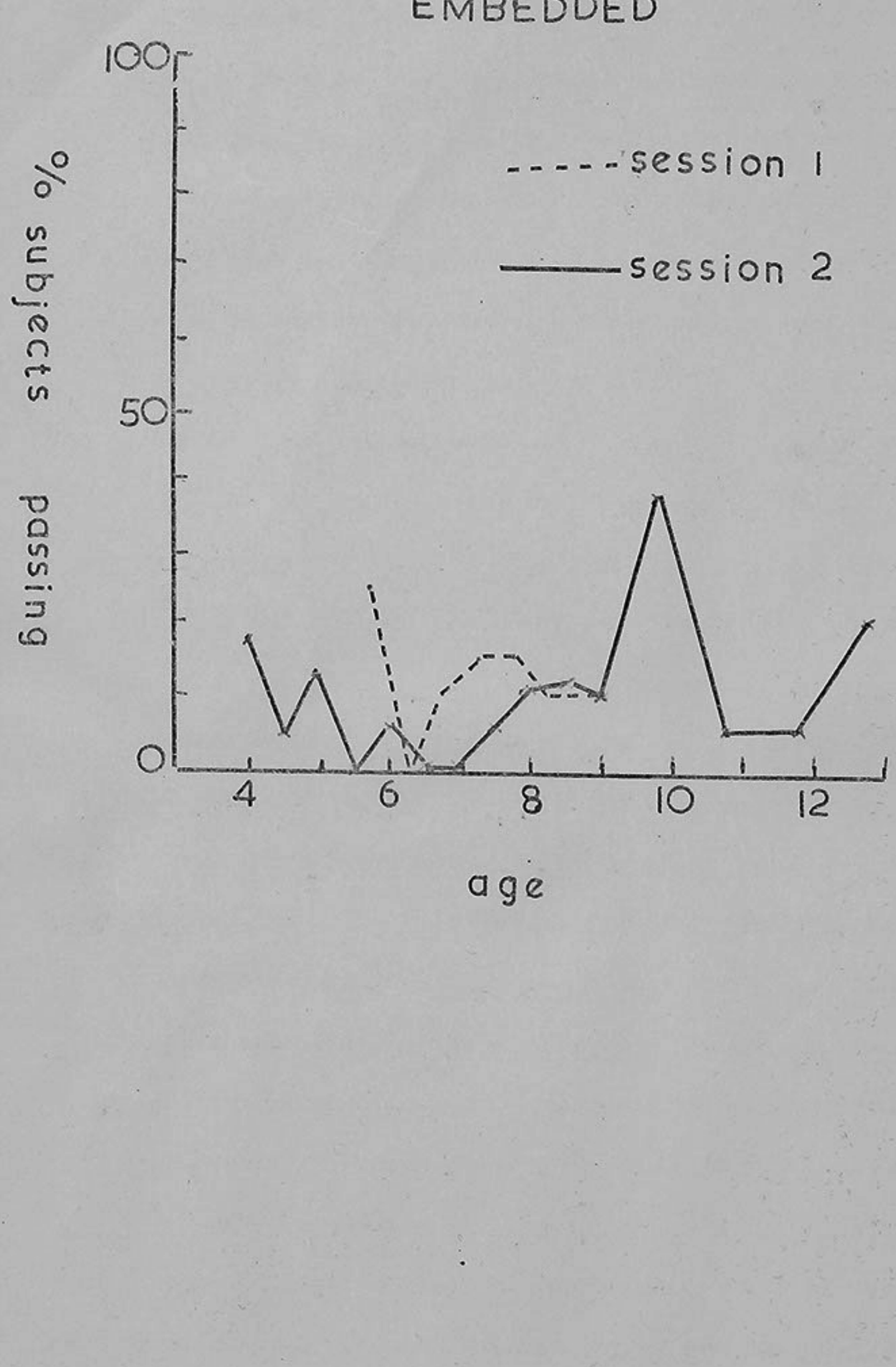
of the pole, then to set one in motion and ask "which one is going to gling?" and "which one is glinging?" Future/past and present/past distinctions used a similar method, with the past (has glinged) being indicated by the woodpecker reaching the bottom of the pole. Herriot found that subjects showed a strong preference for selecting the moving (present tense) woodpecker in all contrasts, so that future and past performance was very poor when these were contrasted with the present. However, in the future/past condition, performance was well above chance level. Herriot suggested that the preference for the present tense exemplar could be attributed to attentional factors - a descending woodpecker is more engaging than a stationary one. However, the results obtained with TROG showing a similar pattern could not be accounted for in these terms. In a second experiment testing comprehension of the present/future distinction, Herriot found that above chance performance on the future tense was not achieved until age 5-0 to 6-0, and even then about one third of his subjects still selected the present tense exemplar when given future sentences. Thus, as in TROG, comprehension of tense inflections appears to develop far later than might have been expected.

Harner (1976) obtained much higher levels of accuracy with 2- to 4-year-olds in a task where comprehension of the distinction between future and past was tested (the boy will fall/the boy has fallen) using pairs of toys carrying out various actions. This is in keeping with Herriot's finding that when the present tense is not used, performance on future and past improves. In a second experiment, children were presented with three boxes of toys, one they had played with previously, another to play with on the test day, and another they were told they could play with the next time. The

previous day's toys were then contrasted with those the child had not yet played with and he was asked to show the experimenter either "a toy you will play with" or "a toy you have played with". It was found that in this situation, comprehension of present perfective was as good as in the first experiment, but comprehension of future tense was less good. Harner concluded that the child can understand linguistic reference to past events which are a day removed from the present as well as he can understand immediate past reference. Reference to future events which are a day removed is more difficult to comprehend than reference to events which are imminent. However, as Harner points out, in the second experiment, it is possible that children are not certain that they would have the opportunity to play with the "future" toys, whereas they know about their past experiences with more certainty.

Brown (1973) claims that the present perfective and future "will" are late to appear in expressive speech. Even so, the finding that under 50% of 12-year-olds fail to get all tense items correct in TROG is surprising, and appears to contradict the principle of comprehension preceding production. Since all methods of testing comprehension of tense are somewhat ambiguous, this could be no more than a methodological artifact. . However, there are some grounds for expecting that comprehension of the distinction between past and present and present and future would be difficult, since the meaning carried by these distinctions is usually redundant in everyday speech. Although the contrast between past and future tense frequently does disambiguate meaning (e.g. I saw John Smith/ I am going to see John Smith), it is less easy to conceive of situations where contrasts with the present tense carry important meaning that is not usually evident from the context of the speaker. If a person says "I am eating my dinner" this is usually evident, unless for some reason we cannot see him.

Fig. 13


EMBEDDED
100
50
0
% subjects passing
session 1
session 2
4
6
8
10
12
age

Indeed, future tense "will" may be omitted altogether without affecting comprehensibility or grammaticality in semantic contexts where the time of the action is evident:

e.g. "I am playing football tomorrow"

or where the situation makes this obvious:

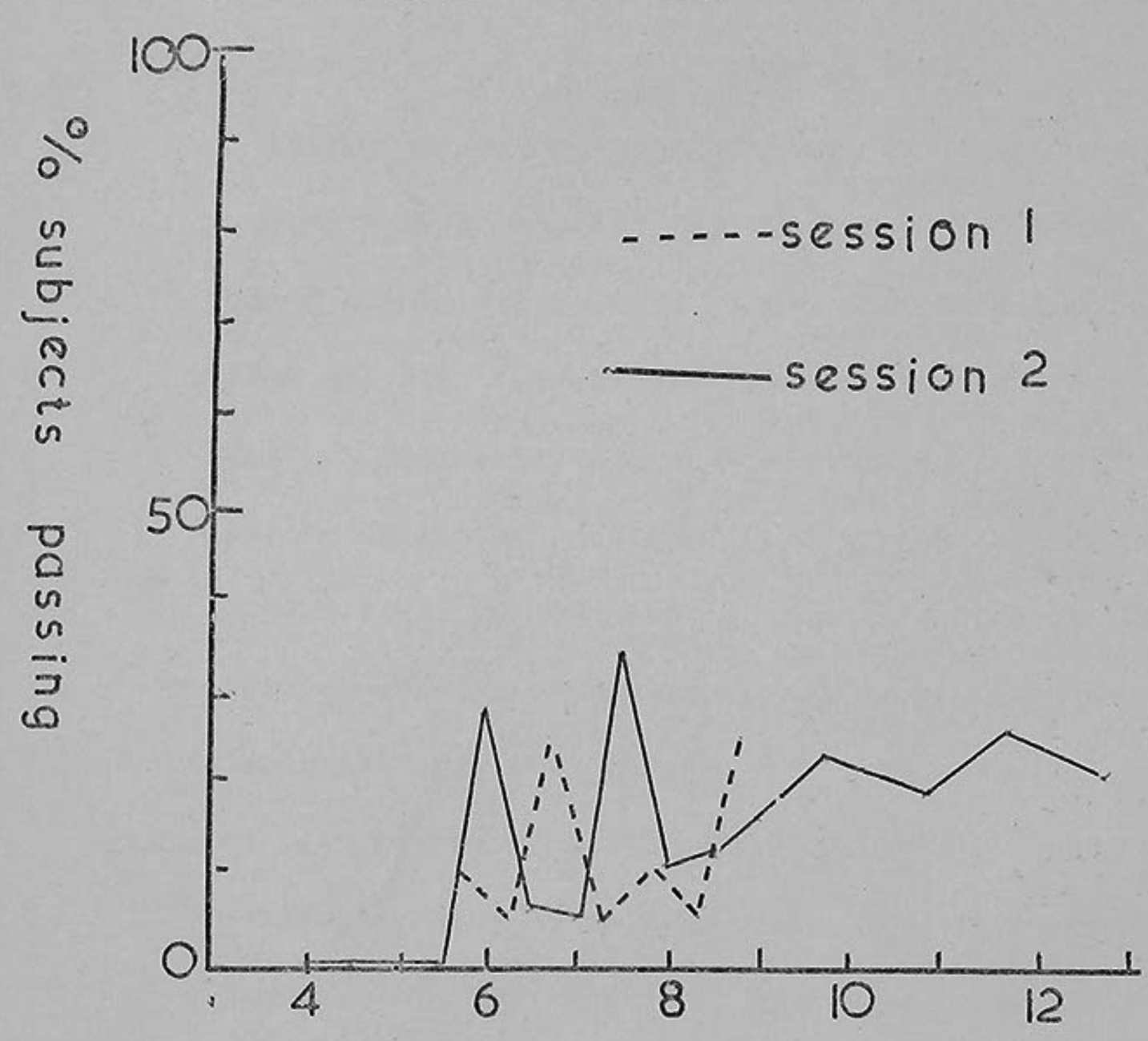
"We can't come to your party. Harry is driving to Newcastle" (with Harry present).

embedded sentences

These proved to be the most difficult structures tested. Results are shown in figure 13. It can be seen that performance remained constant at just above chance level over the entire age range tested. Individual items varied in difficulty with "the boy the cat sees eats" and "the cat the cow chases is black" being much more difficult than the other three items. The crucial difference between these two items and the others seems to be that the two nouns are potentially reversible in the former. The exception is "the boy chasing the dog is spotty" which is also potentially reversible, but which has a different structure, in that the noun in the relative clause is the object rather than the subject of the verb of that clause. Brown (1971) found such sentences easier than those where the noun of the relative clause is the subject of the verb. Other workers studying these types of embedded sentences also find that they are extremely difficult. Gaer (1969) used a test where children were presented with a sentence and a picture and had to respond "yes" if they corresponded and "no" otherwise. For centre-embedded sentences there was no evidence for a developmental trend between ages 3 to 6 years, with performance remaining just above chance level. Children were also given a production task and had considerable difficulty in

producing such sentences. Cook (1975) studied comprehension of a variety of embedded sentences in children aged 4 to 9 years and adults. The children were asked to act out sentences with toys, whereas the adults were asked to demonstrate the relationship described graphically by marking subject and object of the verb with arrows on an answer sheet (the task being explained in simple language). Not only did Cook find that children made a large number of errors, but also that one in five of the adults tested misinterpreted some types of embedded sentences. All the adults were native speakers taking courses in higher education. Cook cautions against overestimating the linguistic abilities of the naive adult speaker, a view thoroughly endorsed by the author who has found that many normal adults cannot decode embedded sentences of the type used in TROG when a 4-choice picture pointing task is used. Work by Lahey (1974) suggests that subjects may fail to make use of the syntactic information provided in embedded sentences. She found that 4- and 5-year-old children performed no differently on sentences such as "the cow that hit the horse chased the sheep" as on the same sentence with syntactic markers omitted, i.e. "cow hit horse chase sheep". The scoring procedure gave credit for each relation between subject and verb, verb and object, and subject and object which the child demonstrated in a task where he was required to act out sentences with toys. Lahey proposed that children do not make use of syntactic markers in comprehending embedded sentences, but adopt an ordering strategy, taking the first noun as agent of both actions, and the noun following the verb as the subject of the action.

Fig. 14


VERB: SINGULAR / PLURAL
100
50
0
% subjects passing
-----session 1
session 2
4
6
8
10
12
age

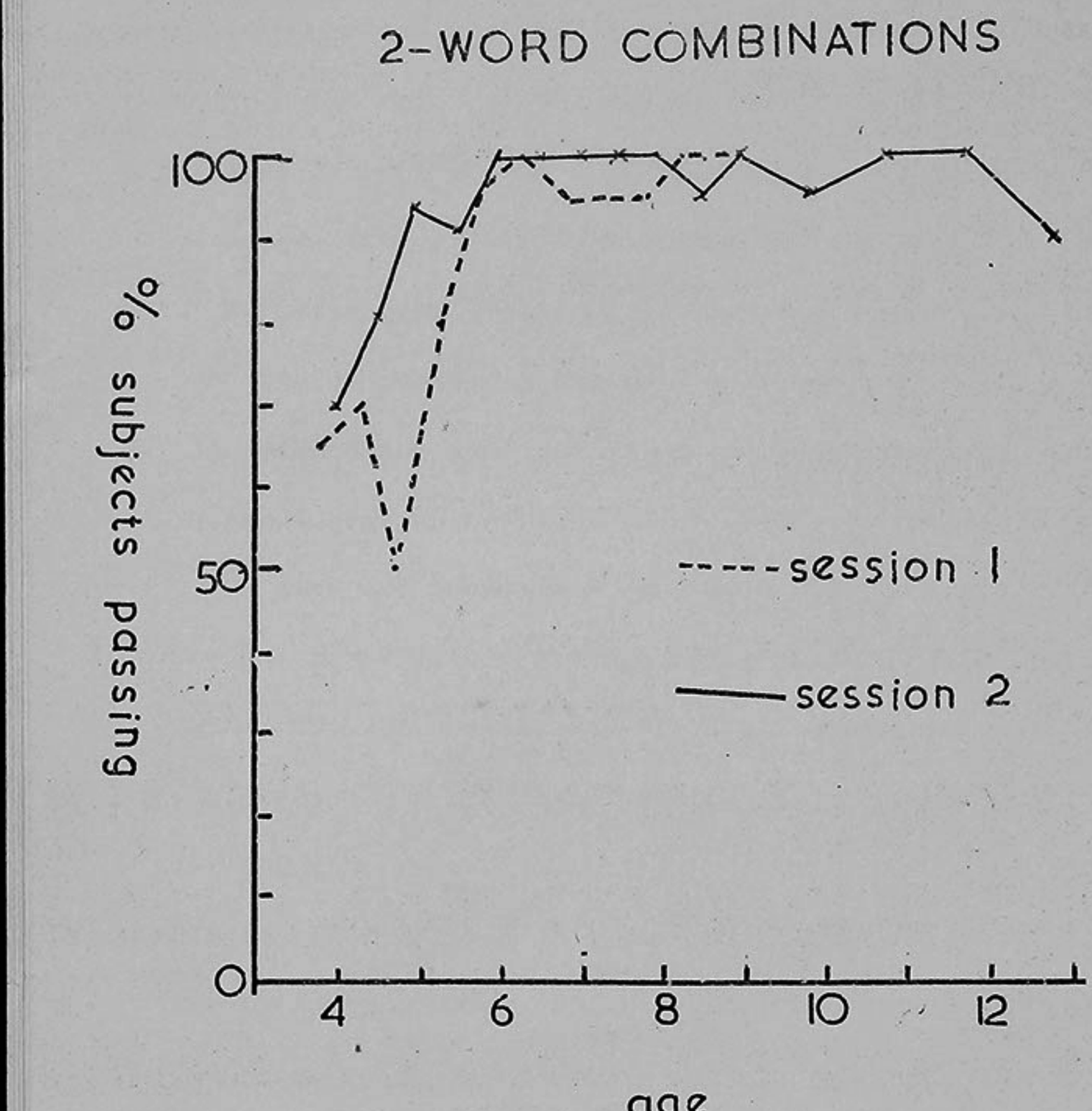
verb singular/plural inflections

Results are shown in figure 14. It can be seen that the majority of children in the oldest age group tested made errors on these items. This result is in contrast to that obtained by Fraser, Bellugi & Brown (1963) who found that comprehension for the distinction between is/are was above chance level for 3-year-olds. This might be partly due to differences in procedure. Fraser et al told children both sentences for a pair of pictures before presenting the test sentence, thereby drawing attention to the relevant contrast. In TROG the stress on the verb was the only clue that children had as to its importance, and this seemed inadequate for many subjects, who would look puzzled and say "they all are" when, for example, given the item "the sheep are black".

Keeney & Smith (1971) have suggested a number of reasons why singular/plural inflections may be acquired relatively late, compared to noun inflections. The noun inflection can occur in the absence of the verb, whereas the converse is not the case. Further, it is the referent of the noun and not the verb that actually is singular or plural. The rule for marking number in nouns (-s added to the plural) is the opposite for that in verbs (-s added in the singular). Keeney & Smith note that the irregular verbs such as "to be" should be learnt earlier as this latter difficulty does not apply. Perhaps most importantly, they point out that the marking of the verb designating number is nearly always completely predictable from the noun, so conveys redundant information. Keeney & Smith investigated imitation and comprehension of noun and verb singular/plural inflections in children aged from 4-2 to 5-3 using pictures of two nonsense animals called "glits" and "snups". Children were presented with pictures of

one or two of these beasts performing various actions and were asked to select the one corresponding to the experimenter's utterance, which was of the form "the glit(s) sit(s) in the tree". For some sentences, the verb only was inflected for plurality. Keeney & Smith found that children attended more to the inflection of the noun than the verb, i.e. when asked to find "the glit sit in the tree" they selected the single animal, even if they had been previously trained to say "one glit, two glit". This tendency to ignore the verb was reduced if the irregular real verb "are" was used. Nevertheless, Keeney & Smith noted that subjects who did not appear to comprehend the verb inflection successfully imitated sentences containing verbs inflected for number, and also used these in their spontaneous speech. Keeney & Wolfe (1972) tested a group of children aged from 3-0 to 4-11 on a new set of tasks. In the first task, the child heard only the verb of a sentence, e.g. "is singing" and had to respond verbally with the phrase "one bird" or "two birds". Other verbs were also used, with the noun being kept constant. In the second task, the same stimulus was presented, but the child had to select a picture of one or two birds. In the third, the child was given the whole sentence "the bird is singing" and was required to select the correct picture. All sentences had correct noun/verb number agreement. Finally, an imitation test was given where the subject was asked to report sentences which followed or violated rules of subject/verb number agreement. Spontaneous speech samples from these children were also analyzed, demonstrating that 94% of their utterances were correct with respect to subject/verb number agreement. In the imitation test there was a strong tendency for children to correct ungrammatical strings despite instructions to repeat exactly what the tester said. Analysis of results on the comprehension tests showed that in the

Fig. 15


2-WORD COMBINATIONS
100
50
0
% subjects passing
4
6
8
10
12
session 1
session 2
age

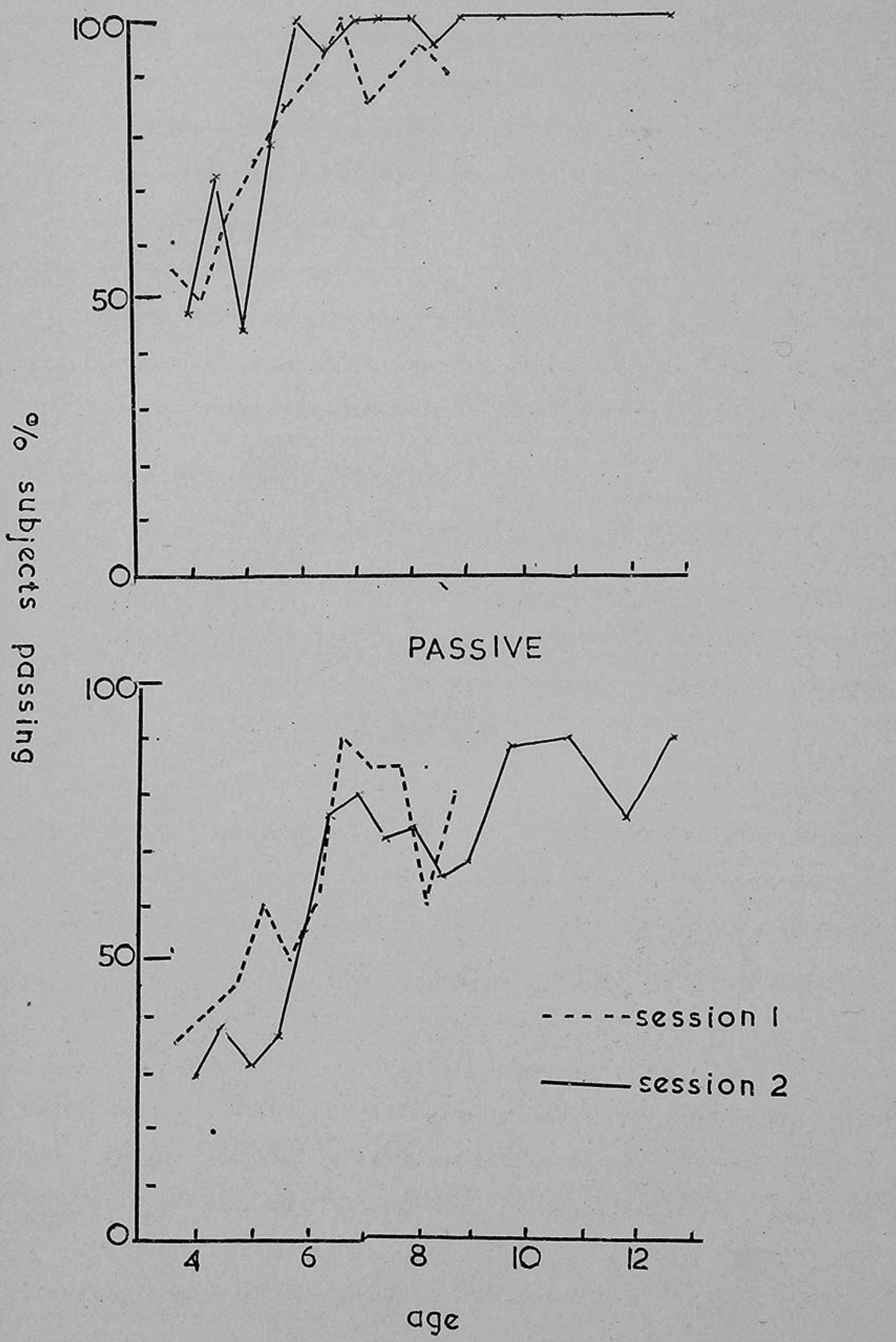
picture-pointing task there was no evidence that the verb inflection was controlling the response. When the inflected noun as well as verb was given, slight improvement in performance was found, but on the whole accuracy levels were barely above chance level. However, in the first test, where children were required to give a verbal response, responding was more accurate. Keeney & Wolfe point out that the latter test requires the child to utter two words that sound right in the context, whereas in picture-pointing tasks an external referent must be identified. They conclude that production of verb singular/plural inflections does precede comprehension, the former being controlled by a purely syntactic rule without comprehension of meaning.

word combinations

Figure 15 shows the results for the six two-word combinations (3 noun + verb and 3 adjective + noun). It can be seen that 90% correct performance is reached around age 5-0.

Word combinations with three content words (3-word combinations) were of two kinds: active and passive. As mentioned above, the comprehension of the distinction between active and passive forms is not tested by these items, and the subject and object can be deduced from the semantic context. In order to perform correctly, the child need only understand the three referent words of the sentence, as nonreferent words and word order are redundant. This makes it rather surprising that passive sentences proved more difficult than actives, as shown in figure 16. This indicates that children are processing the structure of these items, even though it is redundant, and are hampered by encountering the difficult passive transformation. Baldie (1976) compared children's ability to understand different types of passive construction in a 5-choice picture-pointing task

Fig. 16


3-WORD COMBINATIONS
ACTIVE
100
50
0
% subjects passing
PASSIVE
100
50
0
4
6
8
10
12
age
session 1
session 2

using items similar to TROG. He found that only the youngest children tested, aged 3-0 to 3-5, failed to achieve perfect performance on such sentences, indicating an earlier age of acquisition than that found in this study. However, Baldie's subjects each received only one 5-choice item, so were more likely to perform correctly by chance than in this study, where they were given three 4-choice items.

Grammatical contrasts: general comments

sex differences

For none of the contrasts tested was there a significant sex difference in accuracy.

relative difficulty of contrasts

Differences in testing and scoring procedure make direct comparisons with other tests difficult. However, general comparisons of order of difficulty were made between TROG and the work of Fraser, Bellugi & Brown (1963), Mittler et al (1974), Carrow (1975) and Parisi (1971) with whom there was a fair degree of overlap in contrasts tested. For each of these authors, contrasts were rank ordered according to difficulty. For TROG, rank ordering was determined by the age at which 90% of children performed correctly. For Carrow's test, the age at which 90% of children passed an item was used. When several items tested the same contrast, the latest age of acquisition was used. Where tied ranks occurred, the age at which 75% of children passed was used to order contrasts. Rank ordering was determined for the remaining studies according to the total number of errors reported for each contrast. Table 8 shows the rank orderings obtained for these studies. Inspection of this table shows that rather close

Table 8

Rank order of difficulty of structures:
comparison with other investigators

	TROG	Parisi	Carrow	Mittler et al	Fraser et al
more	2		2.5		
in)	2		1		
) (standard)					
on)	2		7.5		
comparative/ superlative	4.5		2.5	1	
he/she	4.5	2	9		
negative	6	1	4	3	1
they	7		6		
singular/plural nouns	9	3	7.5	4	5
active word combinations	9			2	
reversible active	9		11.5		2
under	11		5		
reversible passive	12	4	10	5	6
tense	13	5	13	6	3
verb singular/plural	14		11.5		4

agreement is found between TROG and Carrow's test for most contrasts, exceptions being prepositions, where as has been demonstrated by TROG, performance seems very dependent on the particular pictured situation used, and "he" and "she" where Carrow's items seem to have included some rather ambiguous distractors (a monkey for "he", and a boy and girl together for "she"). Agreement of TROG with the order of difficulty found by Parisi (1971) and Mittler et al (1974) seems reasonable, but there is a notable difference in difficulty of noun and verb singular/plural inflections for TROG and Fraser et al. As has already been mentioned, the relative difficulty of verb singular/plural items may be decreased if the child is told the contrasting sentences, as in Fraser et al's procedure, but it is hard to see why such a procedure should result in noun inflections becoming more difficult.

Fraser et al

Analysis in terms of scales

Analysis in terms of contrasts passed and failed is a useful approach for investigating developmental trends in a population. However, interpretation of the results for individual children in a clinical setting is less straightforward. Although it is possible to use the dichotomous scoring described above to list those contrasts passed and failed by an individual, assessment of the abnormality of any particular profile would pose considerable difficulty in all but extreme cases, since the number of potential patterns runs into thousands. Furthermore, in some cases this dichotomous scoring seems rather artificial, since it involves grouping together sets of items that differ quite substantially in difficulty level. An alternative approach is to calculate the total number of items correct for a child and compare this to the normative sample. This gives us a

quantative picture of performance, but does not enable us to judge whether the comprehension of a language disordered child is merely delayed, resembling that of a younger child, or if it is deviant, unlike a normal child of any age. An intermediate state between the detailed qualitative picture offered by grammatical contrast analysis, and the easily interpretable but limited quantitative result obtained by using total number of items correct was therefore devised. On the basis of their content, items were grouped together to form 4 scales, each encompassing a wide range of difficulty. It should be noted that this grouping was done on the basis of the content of the items, rather than being dictated by statistical analysis. Techniques such as factor or cluster analysis would be inappropriate for a test such as TROG where items cover a wide range of difficulty. As Guilford & Fruchter (1973) point out:

> "two items very far apart in difficulty might correlate less than .20, even when each measures the same thing and measures it well."
>
> p. 413

The decision to use scales was made between sessions 1 and 2 and affected some of the item changes made between these two sessions. It should be noted that in terms of content of the items included in scales, Word Combinations and Word Order seem more consistent than Function words and Inflections, where the individual items have less in common with one another. The scales will now be described.

Word Combinations

These items have been fully described above (see p.48). The scale consists of all items testing lexical rather than grammatical contrasts. New items were added between sessions 1 and 2 to

increase the difficulty of this scale, which includes some of the longest items in the test. Bloom (1970) and Slobin (1973) have suggested that there is a cognitive limitation on the number of elements young children can produce in a single utterance. Cromer (1976) suggested that a limitation in this production span capacity might be responsible for deficient language development in some children. Word Combinations enable us to test the analagous hypothesis for comprehension by asking whether it is the inability to handle a large number of referent words that gives some children problems, rather than difficulty in decoding structure.

Word Order

In all of the sentences included in this scale, word-order was critical to comprehension. All items were 2-choice, and distractors always showed a sentence that could be described by the same words in a different order. The scale consisted of all items testing "+ word order" prepositions, reversible active and passive sentences, and embedded sentences.

Inflections

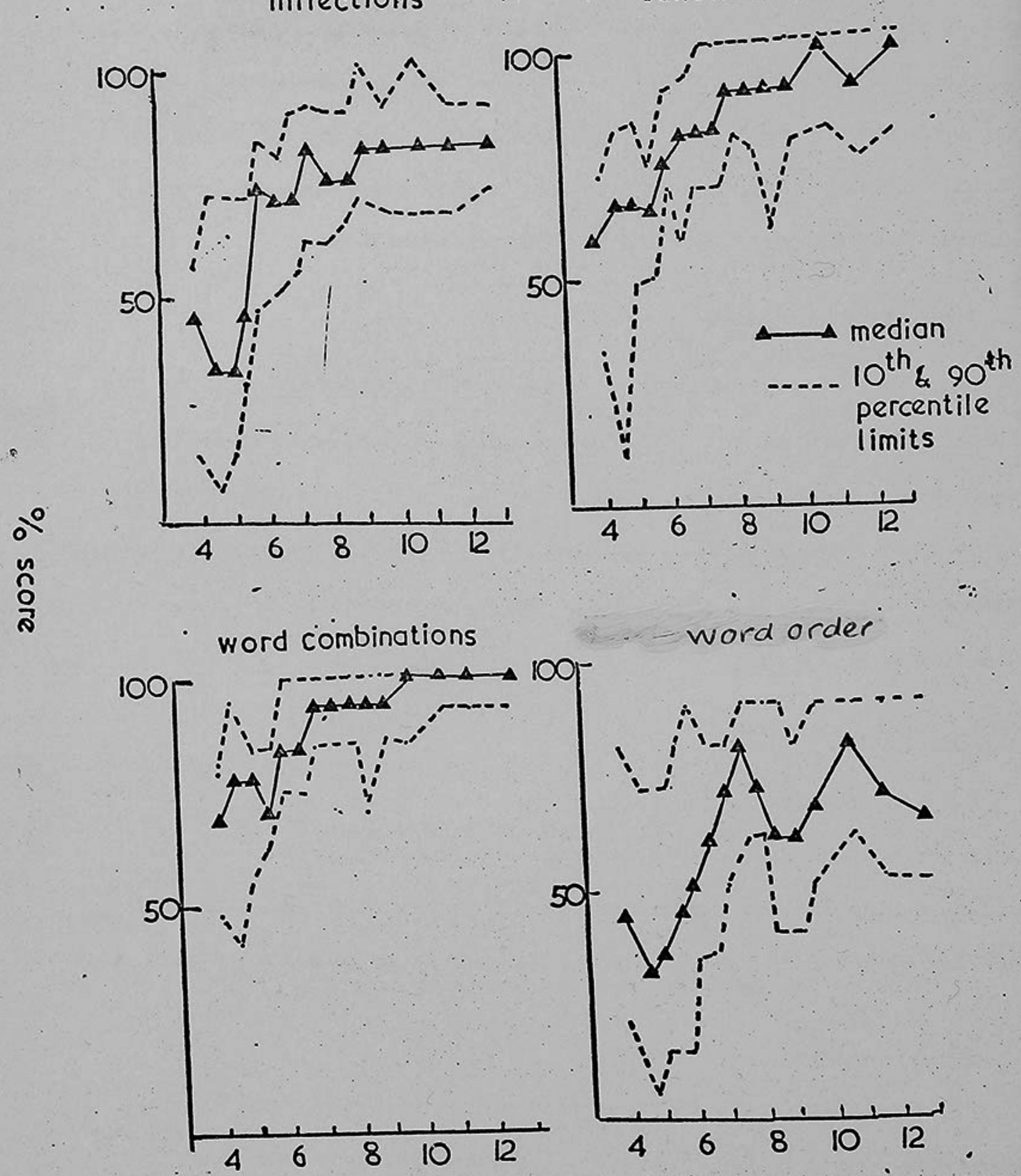
Comparative/superlative inflections, and noun and verb singular/plural inflections obviously belong to this category. It was decided to also include tense items in this scale, although tense was signalled by auxiliary verbs as well as inflection of the main verb.

Function words

The remaining items, those testing comprehension of "not", standard prepositions, "more" and "less", and pronouns were grouped together in a scale termed "function words". It can be seen that this

Fig. 17

median scores on TROG scales



is a heterogeneous collection of items, including some, e.g. "not" and "more" and "less" that might not be regarded as "function words" by many linguists. What these words have in common is that they do not in themselves convey referent meaning, their function being to refer to, specify relations between, or modify the meaning of those referent words with which they occur.

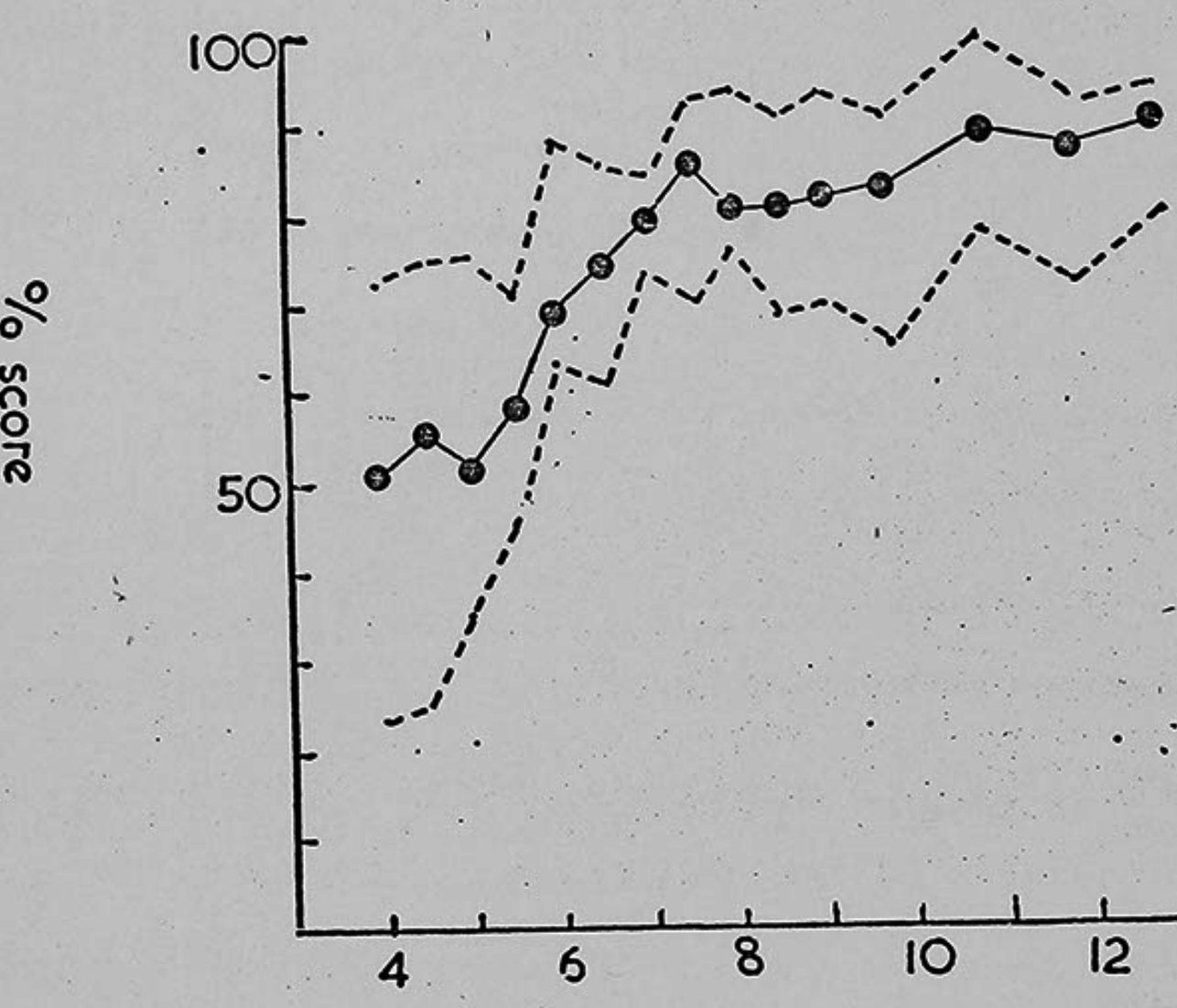
Since items contain differing number of distractor pictures, in order for scales to be comparable, the correction for guessing suggested by Guilford (1936) was applied to the proportion of items correct for each scale, and to the total test. Further details of the scoring procedure are outlined in appendix 3. For each child a corrected percentage correct score (henceforth %-score) for each of the 4 scales, and for the test as a whole was calculated. Median values from test session 2 are shown in figures 17 and 18, together with 10th and 90th percentile limits.[1]

It can be seen that score shows a steady increase up to age 8-0, when the curve begins to plateau, a result which is strikingly similar to that obtained by Noll & Lass (1972) with the Token Test. It should also be noted that, even for the oldest children total median scores are not at ceiling level. TROG thus appears useful for discriminating between children aged well above 5-0. Figure 17 shows that the Function Words and Word Combinations scales are easier than

[1]Medians were used in conjunction with 10th and 90th percentile limits rather than means and standard deviations for several reasons. The percentile limits are more suitable for defining the normal range of variation in scales such as these where scores approach ceiling for older children. Further, in samples as small as these (mean of around 19 children per age group) scores of one or two deviant children can give rise to very large fluctuations in standard deviation, making the setting of normal limits extremely unreliable.

Fig. 18

TROG total: medians



median

10th & 90th percentile limits

Fig. 19

TR

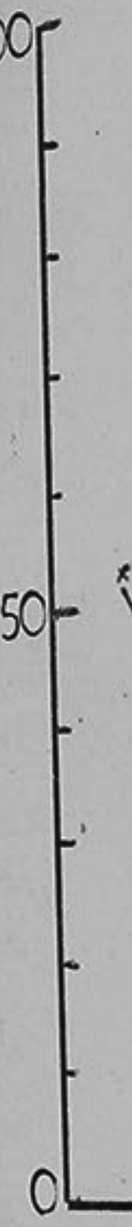
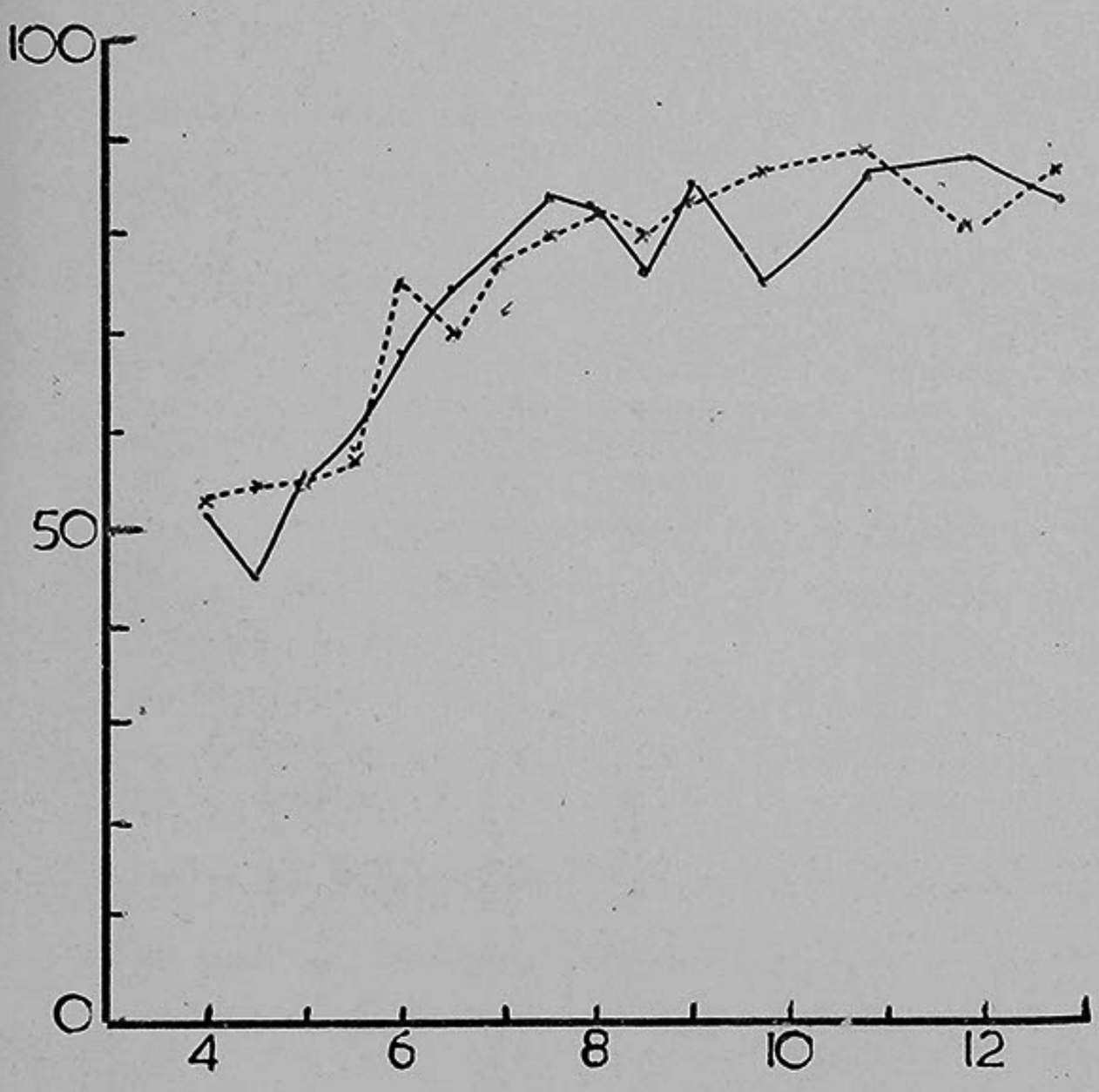


Fig. 19

TROG total % score: boys & girls



---------- boys

——— girls

Table 9

Test-retest: Pearson correlation coefficients

age	n	Inflections	Function Words	Word Combinations	word order	Total
3-9 to 4-2	13	.55	.69	.81	.60	.83
4-3 to 5-2	14	.02	.47	.63	.34	.60
5-3 to 6-2	29	.43	.35	*	.56	.80
6-3 to 7-2	37	.35	.69	*	.30	.56
7-3 to 8-2	36	.33	*	*	.46	.63
8-3 to 9-2	36	.53	*	*	.63	.69

Standard error of measurement

	Inflections	Function Words	Word Combinations	word order	Total
3-9 to 4-2	13.0	14.8	8.0	22.5	8.2
4-3 to 5-2	17.4	22.2	11.4	27.7	11.8
5-3 to 6-2	14.9	8.9	*	18.3	5.5
6-3 to 7-2	10.1	11.2	*	17.3	5.2
7-3 to 8-2	14.6	*	*	15.2	6.2
8-3 to 9-2	14.2	*	*	16.9	8.8

* value not calculated: many children at ceiling

inflections and
more variable
other scales.
Figure 19
similar for bo

Correlations b

It was po
shortened form
overlapping it
include the p
sessions, nor
2 in order to
for the total
word o
values for th
length.

The int
months, but
original ses
children tes
expected to
interval, lo
obtained fo
is consider
from 2 to 1
restricted
correlation

Inflections and Word Order, scores on the latter being much more variable both within and between age-groups than for any of the other scales.

Figure 19 shows that the pattern of performance is essentially similar for boys and girls.

Correlations between sessions 1 and 2

It was possible to estimate test-retest reliability for a shortened form of the test by calculating %-scores for the 56 overlapping items from sessions 1 and 2. These items did not include the preposition items which had been altered between sessions, nor six long Word Combinations items introduced for session 2 in order to increase the difficulty of this scale. Thus estimates for the total test, and for the Function Words, Word Combinations and Word Order scales are likely to be lower than the true values for the final version, since reliability increases with test length.

The interval between test and retest had varied from 3 to 5 months, but only 165 children tested 4 months or less after the original session are included in this analysis, as most of the children tested at longer intervals were under 5-0, and might be expected to show quite substantial changes in comprehension over this interval, lowering the correlations. High correlations can be obtained for almost any pair of measures if a wide enough age range is considered - e.g. shoe size and vocabulary level in children aged from 2 to 16! The test-retest figures given here therefore are for restricted age ranges. Values are presented in table 9. Some correlations for Word Combinations and Function Words are missing

Table 10

Picture completion and PPVT: mean scaled scores

age	n	Picture completion scaled score	Peabody "IQ"
3-9 to 4-2	17*	10.78	94.6
4-3 to 4-8	12	13.00	103.4
4-9 to 5-2	16	11.00	90.8
5-3 to 5-8	22	13.68	93.9
5-9 to 6-2	18	11.50	109.2
6-3 to 6-8	17	9.77	100.6
6-9 to 7-2	21	10.71	102.9
7-3 to 7-8	18	10.72	111.0
7-9 to 8-2	19	10.47	102.9
8-3 to 8-8	17	10.23	105.2
8-9 to 9-2	18	9.28	105.9
9-3 to 9-8	2	11.50	110.5
mean	197	11.03	101.89

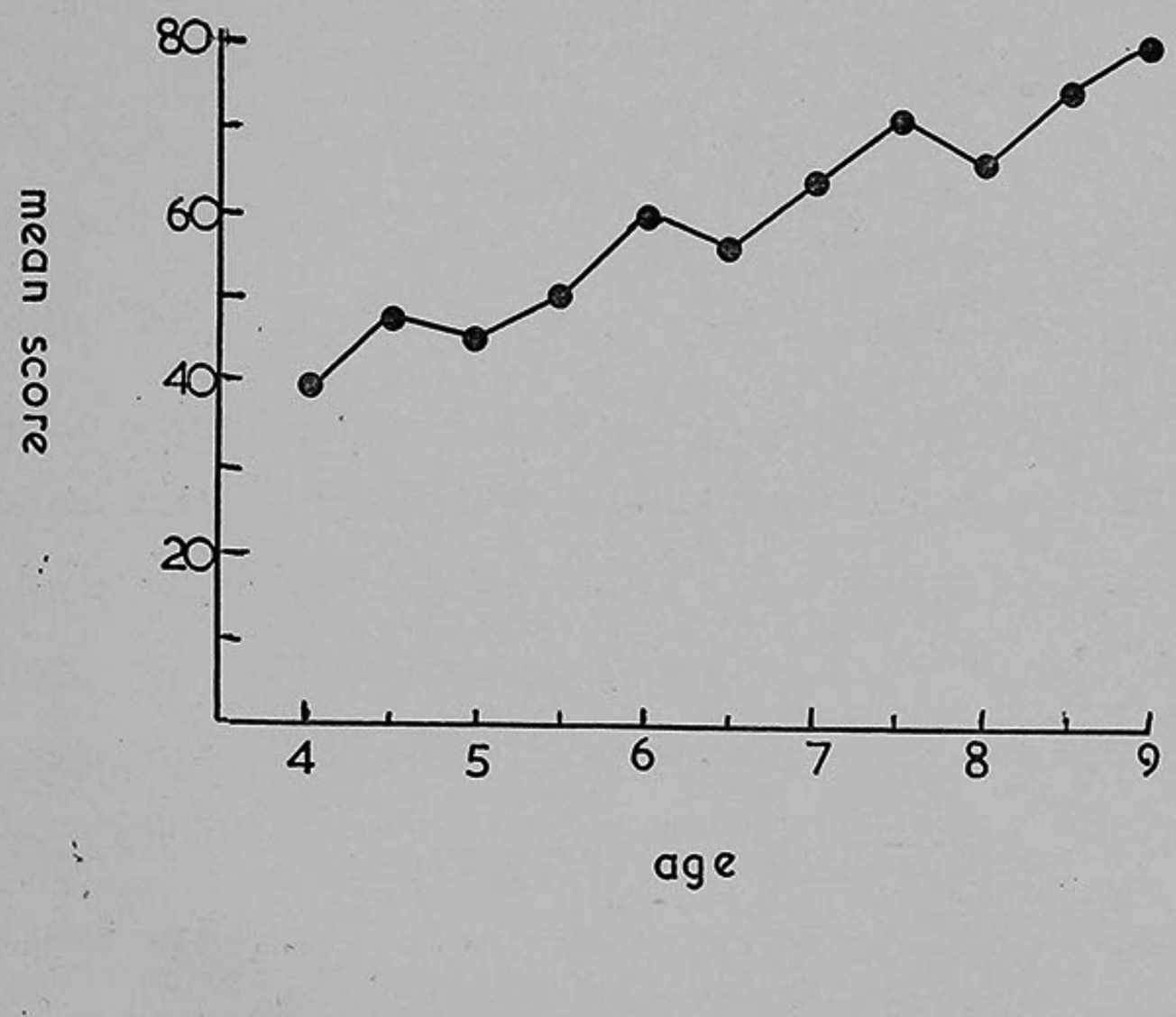
*Picture completion scaled scores could not be calculated for 8 children aged under 4-0

mean score

mean score

Fig. 20

PPVT: raw score



Picture Completion: raw score

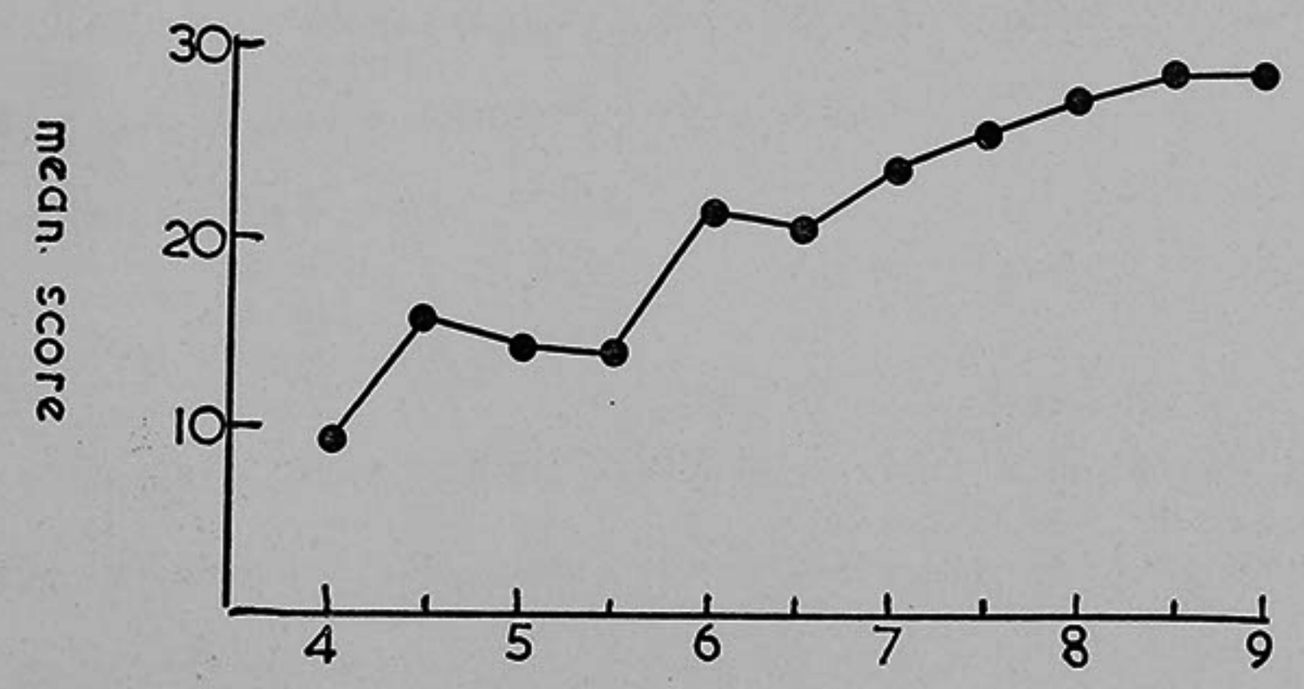


Table 11

Correlation of TROG with PPVT

age	Inflections	Function Words	Word Combinations	Word Order	Total
3-9 to 4-2	.57	.68	.73	.50	.75
4-3 to 4-8	.50	.33	.35	.41	.53
4-9 to 5-2	.18	.29	.42	.53	.50
5-3 to 5-8	.07	.19	.22	.34	.33
5-9 to 6-2	.50	.65	.42	.64	.76
6-3 to 6-8	-.10	.18	.35	.17	.19
6-9 to 7-2	.51	.36	.05	.02	.46
7-3 to 7-8	.40	.69	.33	.16	.54
7-9 to 8-2	.51	.48	.39	.59	.70
8-3 to 8-8	.74	.72	.53	.73	.76
8-9 to 9-2	.65	.51	.59	.70	.77
mean	.42	.46	.40	.44	.57

Correlation of TROG with Picture Completion

age	Inflections	Function Words	Word Combinations	Word Order	Total
3-9 to 4-2	.23	.36	.04	.00	.43
4-3 to 4-8	.65	.07	.39	.33	.46
4-9 to 5-2	.46	.28	.17	.45	.48
5-3 to 5-8	.29	-.13	.00	.20	.20
5-9 to 6-2	.53	.52	.65	.54	.73
6-3 to 6-8	.57	.12	-.15	.26	.34
6-9 to 7-2	.56	.27	-.02	-.03	.45
7-3 to 7-8	.27	.51	.13	.31	.48
7-9 to 8-2	.41	.10	.55	.06	.36
8-3 to 8-8	.11	.14	.36	.35	.26
8-9 to 9-2	.30	.51	.17	.50	.47
mean	.40	.25	.21	.27	.42

because when only overlapping items are considered, most older children score at ceiling, depressing the correlations. Standard errors of measurement are also given in table 9. Bearing in mind the restricted set of items, on the whole the values appear reasonable for this length of interval between test and retest, although the Inflections scale seems undesirably unstable over time.

2. Peabody Picture Vocabulary Test

Mean PPVT scores from session 2 for each age are shown in figure 20. Calculation of IQ scores using the norms from the manual provided a rough check that the population tested here was an average one (see table 10). PPVT IQs are best interpreted not as intelligence scores but as standard scores of receptive vocabulary, with population mean of 100 (s.d. = 15).

3. Picture completion test

Mean raw scores by age group are shown in figure 20. As for the PPVT, calculation of standard scores acted as a rough check on the representativeness of this population. These data are shown in table 10. Standard scores have a mean of 10 (s.d. = 3) in the standardization population.

Intercorrelations between tests

Correlations between TROG total score, PPVT raw score and Picture Completion raw score (all items) were calculated for each age group. Values are shown in table 11. As would be hoped, correlations between the two language tests are higher than between TROG and Picture Completion.

Table 12

Correlation of TROG with Reynell Comprehension Scale A*

	Spearman rho
Inflections	.69
Function Words	.63
Word Combinations	.40
word order	.67
Total	.80

*values for 19 children aged from 4-0 to 5-1

Postscript: a validity study with TROG

In the nursery schools, some eager volunteers came forward, and as it seemed unfair to turn them away, but not possible to include them in the normative sample, it was decided to run a small validation study. These children were given both TROG (final version) and the Reynell Language Comprehension Scale A. Some of the normative sample who were keen for more also completed the Reynell after the other tests. This gave a sample of 19 children aged from 4-0 to 5-1 who had been given both tests within 2 days of one another. Correlations of TROG scales and total score are shown in table 12. These values are highly significant, with the exception of the Word Combinations score. This is particularly gratifying since both tests purport to measure language comprehension with an emphasis on grammatical structure, but very different formats are used.

GENERAL DISCUSSION

Perhaps the most striking finding from this study is the late age of comprehension found for many distinctions. In some instances, notably items testing tense, there is reason to doubt that the distinction tested was adequately represented by the pictorial stimuli, but even discounting these, there are many instances where comprehension does not seem to be fully acquired until well after 4 years. We may contrast this finding with the impression conveyed by most writers on language acquisition, who are prone to statements of the following kind:

> "Of course, linguistic advancement continues long after age four, but these late developments involve such matters as style and possibly semantics, but apparently not syntax."
>
> McNeill (1971) p. 17

or

"In slightly more than two years, children acquire full knowledge of the grammatical system of their native tongue."

Morley (1972) p. 6

This impression has developed largely from studies of expressive speech. Certainly it seems to be true that the speech of the average 4-year-old is grammatically well-formed in terms of the adult rules. This does not mean that he is necessarily using the full range of potential structures defined by the adult grammar. The difference between results from this study and those on expressive speech may result from different methodologies - the person analyzing expressive output is limited to considering departures from well-formed grammar in the corpus presented by the child. It might be that if the child could be made to produce more complex structures that errors would become apparent. This is indeed what is found when the sentence repetition method is used (Menyuk, 1963).

However, this is not the whole answer. As demonstrated with the pronoun "it", there are contrasts which the child uses regularly and correctly but for which he does not demonstrate comprehension, and it does not seem to be arguable that all these instances are artifacts of the test procedure. These structures are surprising not only because comprehension develops late, but also because they appear to contravene the principle established by Fraser, Bellugi & Brown (1963) that comprehension precedes production. The wide acceptance of this principle depends not only on Fraser et al's experimental demonstration, nor on early longitudinal accounts of development (see Bloom, 1974 for a review) but also on logical considerations. How, it is asked (other than by parroting, which is easily detected), could a child produce a word or structure which he does not understand? Slobin (1973) states categorically

"a child cannot begin to use a given linguistic form meaningfully until he is able to understand what it means."

p. 180

However, the discussion above on the comprehension of "it", tense and verb singular/plural inflections suggests that this might not be true. It would appear that a grammatical structure can be used correctly without full comprehension of its meaning if that meaning is already evident from the semantic or syntactic context in most of the sentences in which it occurs. The hypothesis to be proposed here is that grammatical contrasts that do not typically disambiguate the meaning of sentences in which they occur may be produced before they are comprehended.

To clarify this point we may consider an imaginary world where colour terms are only applied to animals, and where colour of an animal is highly predictable, because 99% of cats are white and 99% of dogs are black. A child growing up in such a world may learn to talk about "black dogs" and "white cats", so that the external observer would think that the child understood the meaning of the colour terms. However, when presented with the test case of a black cat or white dog, he may prove unable to comprehend the adjectives. The meaning carried by "black" and "white" is redundant because in most instances the colour of an animal is evident from the distinction between cat and dog. Similarly, as discussed on p. 64 , the inanimacy of the referent of "it" is already evident in contexts in which it occurs, since that referent will have already been more completely specified. The same argument applies to tense distinctions, as discussed on pp 72 to 73. The distinction between past and future does convey information, and is comprehended relatively early. However, the distinctions between past and present, or present and future are seldom used to disambiguate a situation, so

comprehension of their meaning is acquired late. Except in very few irregular instances, verb singular-plural inflections are always predictable from the noun inflections and carry no extra meaning. The word order of sentences containing prepositions, especially "in", usually conveys redundant information, because pragmatic constraints limit the nouns that can follow prepositions, therefore comprehension of the syntactic aspect of prepositions may follow correct productive use of word order with prepositions. The same is true for reversible active and passive sentences. In everyday speech, word order is usually redundant, and it is obvious without knowing the word order which noun is subject and which is object.

We can predict from this argument that other grammatical distinctions whose meaning is usually redundant will be comprehended late. We need not rely solely upon our judgement as to whether a grammatical distinction fits this description or not. First we can see whether native speakers of a language are easily able to understand sentences where that structure is missing. Second, we can compare other languages and see whether they satisfactorily manage without that distinction in the absence of alternative means of signalling its special meaning.

An example of such a structure would be adjective gender agreement in languages such as French. The child learning French has many cues as to which inflection to apply: the gender of the definite or indefinite article preceding the noun, in many cases the phonological ending of the noun, and, less reliably, the sex of the referent of the noun. This latter cue is a poor one, since many referents are inanimate, and actual gender does not always concur

with linguistic gender. Karmiloff-Smith (1976) investigated the influence of these cues on children's productive use of gender agreement. Children were asked to supply an adjective to describe a pictured nonsense creature, given a nonsense name. The nonsense name had a typically male or female ending (e.g. "bicron/bicronne"). Portrayed gender of the depicted creature, gender of article, and gender of noun ending were pitted against one another, giving the child contradictory information, e.g. he would be shown a masculine picture which would be described as "un bicronne." Karmiloff-Smith found that children aged 3 to 4 years relied strongly on the phonological word-endings when selecting gender of the adjective. As they grew older, the article became a more important determinant of response. The results from this artificial situation suggest that in normal language, when cues from different sources are not discordant, the French-speaking child's correct use of productive gender agreement in adjectives may depend very little on the gender of the referent of that noun, or the article preceding the noun but more on the phonological context in which it occurs. Correct productive use might occur without appreciation of the gender distinction conveyed by that contrast. One way of testing comprehension of such a contrast would be to present the child with a picture of e.g. cows and dogs and then to ask him "Quel(le)s sont blanc(he)s?" (having first ascertained that the child was aware that one cow is "une vache" and one dog "un chien".) It is predicted that children would not demonstrate comprehension of the distinction between masculine and feminine adjectives in such a test until a period well after they were using such a distinction consistently and correctly in their own productive speech. A difficulty with the experiment suggested above would be that children might perform

correctly by first naming the picture to themselves, and then applying the adjective to correspond with the ending of the noun. This could be circumvented by using homonyms of different gender; e.g. the child might be shown a group of veils ("le voile") and a group of sails ("la voile") and be asked to show the experimenter "les voiles blanc(he)s". Of course, it would be necessary to first ensure that children were familiar with such homonymous pairs.

It should be noted that in the above discussion comprehension is taken to mean appreciation of the difference in meaning conveyed by a particular grammatical contrast. It should be evident that this is quite different from ability to detect well-formedness. It would be predicted that the latter ability should correspond more closely with correct use of a contrast in expressive speech than would the former. Children who do not give evidence of comprehending the distinction in meaning between "is" and "are" or "blanc" and "blanche" but who use these correctly in their own speech would be predicted to be able to detect that "the boy are running" and "le chien blanche" are incorrect. Failure to recognize that comprehension of meaning of grammatical structure and detection of well-formedness are different abilities can result in confusion. This is particularly evident in the arguments proposed by Ingram (1974).

Conclusions

1. This normative study with TROG has shown that it is a sensitive test which can be used with children aged from 3-6 to 13-2. Although doubts have been raised as to the content validity of some items, overall reliability and validity appear adequate to justify further research with a language-disordered population.

2. Results from this normal population demonstrate that the notion that grammatical structure is mastered by age 4-0 is not true for comprehension of many of the structures tested here. In some cases it seems that comprehension does not precede production as is usually thought.

Part 2

DEVELOPMENTAL LANGUAGE DISORDERS

Chapter 5. Developmental language disorders: a review

Definition

The defining characteristics of developmental language disorder consist more of negative than positive features. Abnormal language development, other than dysrhythmia (stuttering) or dysarthria, is termed developmental language disorder if acquired brain damage, physical abnormality of the speech apparatus, subnormal intelligence, deafness, gross emotional disturbance, or environmental deprivation can be excluded as primary causal factors.

Virtually all authors assume that developmental language disorder constitutes one or more specific syndromes of distinct, but as yet undiscovered etiology. While this seems probable, it is not the only possibility. Developmental language disorder could represent an extreme of normal variation, as suggested by Dozier (1937). Morley (1972) argues that this is unlikely. In a study of 15 language disordered[1] children she concluded:

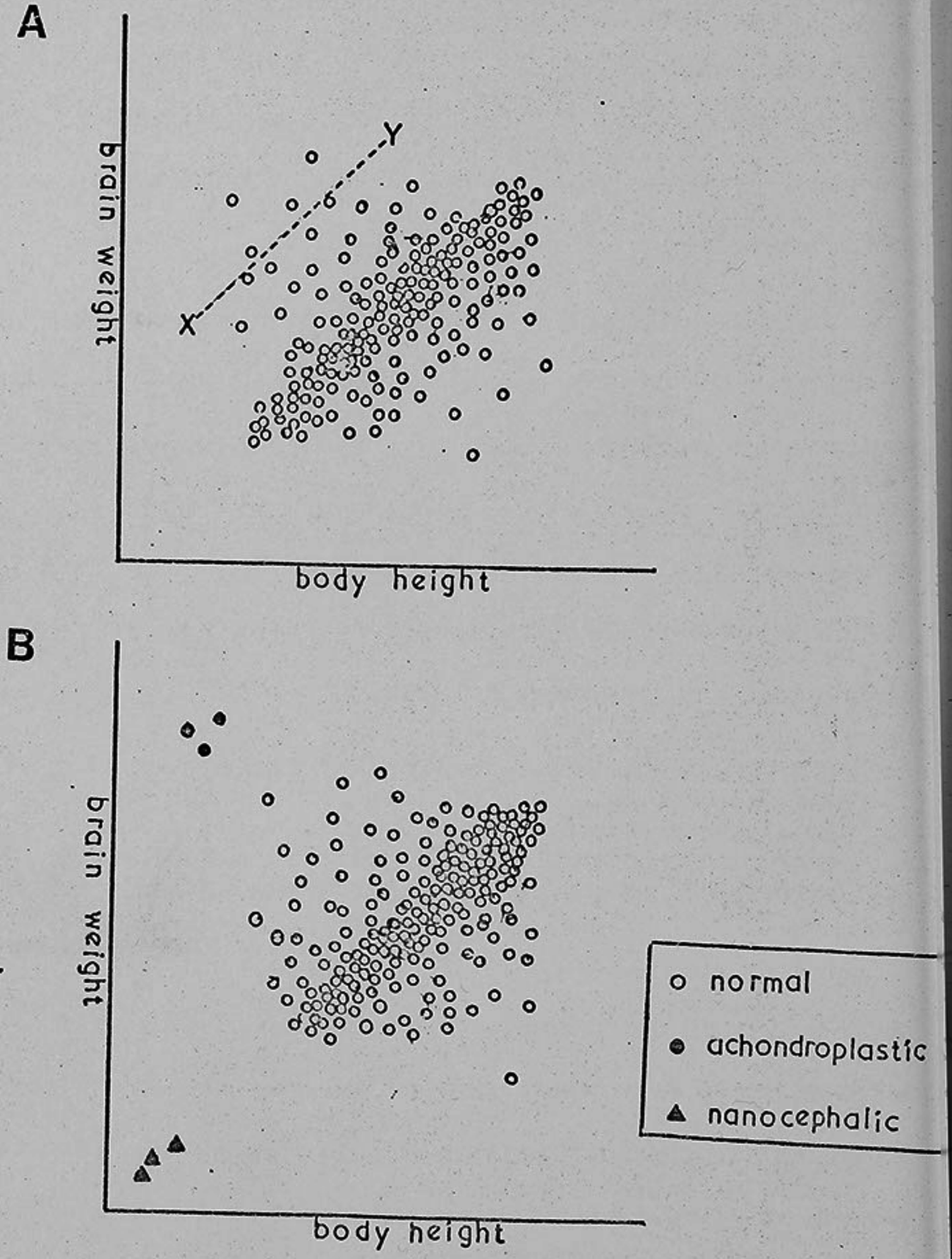
> "In view of the considerable delay in starting to speak and the continuing difficulties with language which these 15 children experienced, we feel that they probably express a true cerebral disorder and not merely the extreme end of the normal range of speech development."
>
> pp 166-167

The severity of the language problem is not the only factor which makes one suspect a qualitative difference from normality. Language disordered children are notable for the dissociation they demonstrate between linguistic and other cognitive skills in contrast to normal children in whom verbal and nonverbal skills are positively correlated. However, in any normal population, some children will show sizeable

[1] Unless otherwise qualified, "language disordered" will be used to mean "suffering from developmental language disorder".

Fig. 21

Theoretical distributions: brain wt. vs body ht.



verbal/nonverbal discrepancies and the possibility remains that children with developmental language disorders represent the most extreme examples of such dissociation.

An analogy might be useful here. Figure 21A shows a theoretical scattergram relating brain weight to body height in normal 8-year-olds. We could select extreme examples of low body height for brain weight by setting a cutoff line XY and taking individuals falling above this line as our sample. However, in a very large unselected population the scattergram would look slightly different (see figure 21B) owing to some individuals suffering from forms of dwarfism. Achondroplastic dwarfs will fall well above XY since their bodies are small in relation to their brain weight. Here a single dominant gene is responsible for the condition (Scott, 1976), and these individuals do not shade into the normal distribution. Another very rare type of dwarfism, nanocephalic dwarfism, results in lowered brain weight and body height without altering the relationship between them. Once again, a single gene, in this case recessive, is responsible for the condition (Seckel, 1960).

The analogy with developmental language disorders becomes clear if we substitute "nonverbal abilities" for brain weight and "linguistic abilities" for body height. It could be that children diagnosed as suffering from developmental language disorders represent an extreme end of the normal bivariate distribution relating nonverbal and verbal skills, corresponding to the children falling above XY in figure 21A. On the other hand, it may be that a specific etiological factor results in selective disturbance of linguistic functioning, much as a specific factor disturbs body growth in achondroplasia. To complete the analogy, we know that

other specific factors (e.g. the chromosomal abnormalities of Down's syndrome) appear to lower verbal and nonverbal skills, without disturbing the balance between them (c.f. nanocephalic dwarfism).

The most convincing way to resolve this issue would be to isolate a specific etiological factor responsible for developmental language disorder.

Terminology

The condition described here is also termed congenital or developmental dysphasia (or aphasia) by some authors.[1] The term "developmental language disorder" seems preferable despite its length since it avoids confusion with acquired disorders in adults and children (Critchley, 1967). Further, it does not carry the neurological connotations of the other terms. This is not to say that developmental language disorders cannot result from neurological dysfunction, but rather that the evidence for this is as yet unclear (see pp. 99 to 106).

Classification

One point about which there is no dispute is the heterogeneity of language disordered children. This has led to a variety of attempts at subclassification. Most classifications are based on the child's language, although, as Ingram (1969) points out, clinical impression rather than detailed phonetic or syntactic analyses must be relied upon, in the absence of adequate assessment techniques.

Many authors distinguish between expressive and receptive

[1]The picture is further complicated by the fact that a few authors use congenital and developmental dysphasia to refer to different conditions (Cohen, Caparulo & Shaywitz, 1976).

disorders (Orton, 1937; McGinnis, 1963; Worster-Drought, 1968; Johnson & Myklebust, 1967; Reynell, 1969b; Morley, 1972). The implication is that these are two clearcut distinct types: in the former, language comprehension is relatively intact, whereas in the latter both comprehension and expressive speech are disturbed. The receptive type corresponds to traditional "word-deafness" or "congenital auditory imperception" (Worster-Drought, 1943). Ingram (1969, 1972) and Griffiths (1972) prefer to regard developmental language disorders as constituting a spectrum ranging from mild articulatory disturbances at one end to severe receptive disorders at the other. Common to all these authors is the view that children with expressive disorders have relatively normal comprehension. Morley (1972) goes so far as to claim:

> "Failure of normal understanding for speech in the child with developmental expressive aphasia is rare, except where there is a partial hearing loss for some of the frequencies of sound required for speech."
>
> p. 159

In contrast, de Ajuriaguerra et al (1976) remark:

> "The verbal comprehension of (developmental) dysphasics is almost never normal."
>
> p. 347

and Eisenson (1968) goes even further, claiming that impaired comprehension is the basis of nearly all developmental language disorders:

> "in regard to manifestly expressive disturbances ... unless an oral apraxia or dysarthric impairment can be demonstrated, all of a child's language difficulties are products of his receptive capacity for language."
>
> p. 196

A similar view is proposed by Cohen, Caparulo & Shaywitz (1976). Some tentative experimental support comes from a study by Weiner (1969a) who found that a group of children with mild language disorders did more poorly than controls on both PPVT and Carrow's (1975) test. Furthermore, Garvey & Gordon (1973) found a high positive correlation between scores on the Renfrew Action Picture Test (Renfrew, 1975), which provides an estimate of complexity of expressive speech, and the PPVT in 12 children with long-term language problems.

This brief account of the receptive/expressive distinction reveals that even within this classificatory framework two important areas of disagreement exist. First, there is the question of whether receptive and expressive types are qualitatively distinct, or whether they differ only in terms of severity. Second, there is the related issue of whether children diagnosed as suffering from expressive disorders have normal language comprehension.

Eisenson (1968) rejects the traditional receptive/expressive distinction in favour of a unitary model of developmental language disorder. At the other extreme, Aram & Nation (1975) propose no fewer than 6 types of language disorder. This classification is arrived at in a study where 14 language tasks were administered to 47 language disordered children and the results subjected to factor analysis. It is unlikely that this result is reliable. Everitt (1975) has pointed out that spuriously highly correlations are likely to result from multivariate analyses if the number of subjects (N) is of the same order of magnitude as the number of measurements (P). He recommends that N should be at least 10 times as great as P and recommends that the conclusions of any analysis performed on data where N is less than 5P should be viewed with a certain scepticism.

Nevertheless, there may be some justification for distinguishing between different types of expressive disorder. Eisenson's distinction between apraxic and other expressive disorders has already been mentioned, and Edwards (1973) and Crookes & Greene (1963) claim that similar distinctions are useful. The difficulty here is that good objective tests for dyspraxia of speech are lacking.

Other classificatory schemes are based on presumptive etiology. Rose (1972) distinguishes children with no history of speech development from those who start to develop speech normally and then regress. It could be argued that the latter group should not be regarded as instances of developmental language disorder, since the disorder appears to be acquired. Cohen, Caparulo & Shaywitz (1976) have suggested that the latter group be termed idiopathic acquired aphasia. It is interesting to note that where cases of this type are reported, there is frequent mention of EEG abnormality (Landau & Kleffner, 1957; Worster-Drought, 1971; Huskisson, 1973; Gascon et al, 1973; Shoumaker et al, 1974) and Byers & McLean (1962) report on a series of children in whom language disorder of unknown origin developed accompanied by frank hemiplegia. All these reports mention a high incidence of receptive disorders. It may be that symptomatology of congenital and idiopathic acquired disorders does differ, but more empirical data is needed on this question.

Incidence

It is generally agreed that developmental language disorders of severity sufficient to warrant special education are rare. Ingram (1963) estimates the incidence of expressive disorder at .71 per 1000. No figure is given for receptive disorders, which are known to be even more uncommon. Highly unintelligible speech affected 0.7% of 6½-

year-olds attending normal schools in a survey by Morley (1972). Both this study and reports from the National Child Development Study do not distinguish between children with developmental language disorder, and those whose poor intelligibility is due to structural abnormalities or dyspraxia. The latter survey reported an incidence of 2.4% for unintelligible speech in a large population of 7-year-olds (Butler, Peckham & Sheridan, 1973), with 1% of the same sample being rated by their teachers as having "certainly poor speech" by the age of 11 years (Calnan & Richardson, 1976).

Associated factors

It is established that the incidence of developmental language disorders is higher in boys than girls, with a sex ratio of around 2:1 to 3:1 (Ingram, 1959). Similar sex ratios are reported for "unintelligible speech" (Butler, Peckham & Sheridan, 1973). Rutter, Bartak & Newman (1971) noted that the sex ratio appeared to be equal for receptive disorders. However, Worster-Drought (1943) collected 36 cases of "congenital auditory imperception" from the literature and found only 6 females among them.

A raised incidence of a positive family history for speech disorders is also frequently reported (Orton, 1937; Worster-Drought, 1943; Ingram, 1959; Bartak, Rutter & Cox, 1975) but it is not established whether this indicates a genetic contribution.

Association with left-handedness or ambidexterity has been remarked by some authors (Orton, 1937; Ingram, 1959).

Retrospective data on pre and perinatal factors must be treated with caution, for mothers of abnormal children may be more prone to report such influences than other mothers because they are searching

for reasons to explain their child's defect (Brown et al, 1973). Ingram (1959) did find that 15 children out of his sample of 80 language disordered children had a definitely abnormal pregnancy or delivery, but in the absence of control data the significance of this finding remains uncertain. Sheridan (1973) reported prospective data from the National Child Development Survey showing that children with marked speech defects did not differ from controls in birth weight, maternal age, incidence of haemmorhage in pregnancy, abnormal delivery, or foetal disorders, but were significantly more likely to be premature.

Sheridan also reported that fewer children with marked speech defects than controls had walked by 18 months and many fewer had spoken in phrases by 2 years. 46% of speech defective children compared to 13% of control children were considered clumsy by their teachers.

Neurological aspects

Goldstein, Landau & Kleffner (1958) compared 114 deaf children with 69 children from the same school who were diagnosed as aphasic. Half of the aphasic group did have hearing losses of some severity. An etiological classification of children revealed some interesting differences between these groups. Meningitis and other severe infections in infancy were far more likely to result in deafness than aphasia. In contrast, four cases with a history of rhesus incompatibility were all diagnosed as aphasic. Six out of seven cases with convulsive disorders without apparent precipitating cause were aphasic. These figures suggest that neurological disorder is associated with developmental language disorder, but this is of a

different kind from the neurological disorders found with deafness. However, 35 of the 69 aphasic children had no evidence of pre-, peri- or postnatal factors of etiological significance, although nine of these did have a positive family history for hearing loss or language disorder.

Benton's (1964) review of the relationship between brain damage and developmental language disorder considered only receptive disorders, and was of necessity limited by the dearth of research in this area. The only post-mortem study was a single case reported by Landau, Goldstein & Kleffner (1960). This child had suffered from a receptive language disorder before dying from a cardiac complication at the age of 10. Autopsy revealed extensive bilateral damage involving the temporal lobes with degeneration of the primary auditory projection areas and medial geniculate nuclei. Both auditory nerves were intact. The child's linguistic difficulties seem hardly surprising in the face of such damage, and indeed, Landau et al pointed out that it was remarkable that the boy had made as much progress as he did. With special instruction he had learned to communicate in simple sentences. The remainder of the evidence for an association between brain damage and developmental language disorders presented by Benton (1964) is circumstantial. He notes the occurrence of language disorders in children with erythroblastosis fetalis, and points out that this condition has been associated with lesions in the cochlear nuclei. He also remarks on the occurrence of language disorders after cerebral anoxia, severe head injury, meningitis and encephalitis, but does not give estimates of the proportion of affected children with such histories. Finally, he points out that many symptoms found in language disordered children

are characteristic of patients with demonstrated brain damage. Benton concludes that cerebral damage is probably an essential causative factor in developmental receptive disorders, even though such damage is frequently not demonstrable.

Since Benton's review, evidence has accumulated showing that a substantial proportion of language disordered children are neurologically abnormal. Forrest, Eisenson & Stark (1967) studied 73 language disordered children and found that 36 had EEG abnormalities. Further, 19 out of 22 children with localized abnormalities had a focus in the left hemisphere. However, interpretation of these data must be cautious, as a control group was not tested. As mentioned above, EEG abnormalities are frequently noted in cases whose language disorder follows a period of normal language development (Landau & Kleffner, 1957; Worster-Drought, 1971; Huskisson, 1973; Gascon et al, 1973; Shoumaker et al, 1974). However, demonstration of EEG abnormalities does not entail gross structural lesions. These children may rather have functionally abnormal nervous systems. This description would fit the children studied by Shoumaker et al (1974) and Landau & Kleffner (1957) in whom the severity of EEG abnormality was positively correlated with language disorder - as improvement occurred, EEG abnormalities died down. This led Shoumaker et al to suggest that it is the abnormal electrical activity itself which causes the symptoms.

However, there remains a considerable proportion of children without detectable neurological abnormality of any kind. These children might have lesions or functional disturbances which remain undetected by current neurological techniques. A popular hypothesis is that in some children neural connections fail to develop normally,

perhaps because of genetic factors (Worster-Drought, 1943, 1968; Carrell & Bangs, 1957; Gordon, 1964; Johnson & Myklebust, 1967; Morley, 1972), but such an abnormality remains to be demonstrated.

It is also possible that a metabolic disorder might be responsible for abnormal functioning of the central nervous system resulting in language abnormalities. Indeed, early reports of one such disorder, histidinaemia, noted an association between this abnormality of histidine metabolism and specific language retardation (Ghadimi, Partington & Hunter, 1961, 1962; Auerbach et al, 1962; La Du et al, 1963) but later reports are sceptical of these claims, noting that many children with histidinaemia are normal (Gordon, 1970). Neville et al (1972) collated all reported cases of histidinaemia, and noted that out of 42 subjects, 28 were of normal intelligence, and seven of these had defective speech. They point out that children are not normally tested for histidinaemia unless they are cognitively impaired, or have a sib with the disorder, so it is possible that the incidence of cognitively normal children with histidinaemia is much higher than it appears. An epidemiological study would be of help in resolving the question of whether a true association exists.

If the nature of the neurological abnormality poses problems, its location is even more troublesome. The obvious candidate would seem to be the left hemisphere, damage to which is associated with dysphasia in adults. Forrest et al's (1967) finding of a high incidence of EEG foci in the left hemisphere would support this notion. However, simple parallels with adult dysphasia must be rejected when we consider the consequences of verified left hemisphere lesions in children, which are notable for not producing persistent

dysphasia (Guttman, 1942; Alajouanine & Lhermitte, 1965; Lenneberg, 1967; Hecaen, 1976). Even in the extreme case of removal of the entire left hemisphere, language will not be permanently affected if the removal occurs early enough (Krynauw, 1950; Basser, 1962). Some authors have claimed that early left hemisphere removal or damage can result in subtle linguistic impairments (Woods & Teuber, 1973; Dennis & Kohn, 1975; Dennis & Whitaker, 1976) but these do not resemble the deficits found in developmental language disorders, and there is reason in any case to be sceptical of some of these claims.[1]

Since early left hemisphere damage does not appear to be a sufficient cause for the development of language disorders in children, alternatives must be considered. Several possibilities suggest themselves:

1. Gordon (1964), Taylor (1964) and Benton (1964) have proposed that subcortical systems may be malfunctioning. One suggestion is that efferent pathways in the reticular formation, responsible for the inhibition of auditory input, may be disrupted. Such an account could explain the variable audiometric responses obtained with many children with receptive problems (Reichstein, 1964), although other explanations for this phenomenon are possible (Mark & Hardy, 1958). Gordon (1964) counters the objection that a lesion in the reticular formation would be bound to have multiple consequences by suggesting that it is a failure of connections to develop, rather than a gross structural lesion, which is responsible.

[1]Dennis & Kohn (1975) base their conclusions on tiny samples with no normal control group. The linguistic deficiencies of their left hemidecorticated patients consisted of inability to perform above chance level on reversible passive sentences in the affirmative and negative, and this was shown clearly in only 2 patients. Performance as poor as this was found in over one quarter of normal children aged between 8 and 13 years on TROG, and it is suspected that this is within the normal range of performance for adults.

2. An alternative is that bilateral cerebral damage is involved. If homotopic areas of both hemispheres are affected, then take-over of function cannot occur. It is known that acquired childhood aphasia associated with bilateral lesions has a poor prognosis (Van Dongen & Loonen, 1977), and it will be remembered that Landau et al's (1960) case had bilateral pathology. A single case of expressive disorder reported by Sato & Dreifuss (1973) was found to have a focus of paroxysmal electrical activity in the right temporal lobe, and a mirror focus in the left. The difficulty with any theory proposing that bilateral cortical damage is involved is that, as noted by Sugar (1952), such lesions would be expected to result in more neurological signs and greater intellectual impairment than is the case. However, this argument need not trouble those who propose aplasia or agenesis rather than gross lesions (Worster-Drought, 1943).

3. It could be that specifically left hemisphere damage or dysfunction is implicated, but that the circumstances are such that the right hemisphere does not take over language functions. It is known (Milner, 1974) that early left hemisphere damage will not result in right hemisphere language representation unless the primary speech areas are involved, and it is possible that with only mild involvement, such transfer does not occur. A similar view had been proposed by Kinsbourne (1974):

> "In left-hemispherectomized patients, the right hemisphere can compensate because it is totally free from competitive inhibition from the other side. But if the early left-sided damage is mild, it will suffice to impair the left hemisphere's language potential, but it will not release the right hemisphere for compensatory purposes."
>
> p. 265

4. Orton (1937) proposed that developmental language disorder is

associated with failure to develop normal cerebral dominance. However, the assumption behind this notion, that lack of lateralization results in poor language, does not appear to be justified. Milner, Branch & Rasmussen (1966) described several patients with bilateral speech representation, as demonstrated by the Wada technique, but no mention is made of their having language problems. Studies assessing cerebral dominance in language disordered populations have had to rely on less direct methods, such as dichotic stimulation. Witelson & Rabinovitch (1972) studied a population described as suffering from "auditory-linguistic deficits", who seemed characterised more by reading difficulty than by general language problems, although all did poorly on a test of auditory discrimination. Results from a dichotic test were inconclusive. Differences in accuracy between ears did not reach significance for either experimental or control children, although there was a significant interaction between Ear and Groups, i.e. normal children reported more right ear than left ear digits, and experimental children the converse. On the basis of a post hoc analysis where children were divided up according to whether more left or right ear digits were reported (a procedure of debatable validity; see Satz, 1977), Witelson & Rabinovitch concluded that there may be two types of children with auditory-linguistic deficits - one with left hemisphere speech and one with right hemisphere speech. The cynical observer might be tempted to conclude that Witelson and Rabinovitch have rather shown that neither normal nor abnormal children have established cerebral dominance! Sommers & Taylor (1972) tested a sample of children who had been diagnosed 2 years previously as having serious speech delay, and compared them with a control group on tasks involving the recall of dichotic words and digits. It should be noted that most of their

experimental subjects had greatly improved language functioning at the time of the study, and did not differ from the control group on either the Receptive or the Expressive section of NSST. Results on the digit task were quite remarkable - 9 out of 10 experimental subjects had a left ear advantage, whereas all ten control children had the usual right ear advantage. However, with word stimuli, all children except for one experimental subject had a right ear advantage. There is no simple explanation for these puzzling results, and Sommers & Taylor offer no theoretical account. Rather than enabling us to draw any conclusions about the cerebral lateralization of this sample, these results demonstrate the unreliability of ear advantage scores across different tasks. Springer & Eisenson (1977) studied 10 children aged 8-7 to 12-9 with reading disorders, five of whom also had expressive or receptive language disorders. Dichotic consonant-vowel pairs were presented and children required to identify the stimuli they had heard from a multiple choice card. It seems strange that a response mode requiring the subject to read should be used with children with reading problems, but Springer & Eisenson claim that all children could associate the written symbols with the sounds in a pretest. It was found that the experimental group did more poorly overall than the controls, but that both groups demonstrated right ear advantages. A secondary post hoc analysis in which Springer & Eisenson claim to show that size of ear asymmetry is related to severity of linguistic impairment is not at all convincing, since the ear asymmetries of the exerimental group are well within the range of normal controls.

Linguistic aspects

A common question to encounter in this field is "is language

Table 13

Delay vs deviance

Interpretation A: is pattern of abilities normal?

e.g.

i) is language in keeping with nonlinguistic development?

ii) is expressive speech in keeping with comprehension?

iii) are syntax, semantics and phonology at the same level of development?

iv) within syntactic, semantic and phonological systems, is there a normal pattern of development?

Interpretation B: is behaviour atypical for any level of development?

e.g.

i) is phonology bizarre rather than immature?

ii) is semantics " " " "

iii) is syntax " " " "

delayed or deviant?" (Leonard, 1972; Menyuk, 1973). Attempts to answer this question have resulted in some confusion, probably because the distinction has been interpreted in a variety of ways. At least two types of interpretation are possible, as shown in table 13. Interpretation A defines deviance and delay in terms of patterns of behaviour. Here deviance refers to abnormal patterning of abilities, while delay is diagnosed when all abilities are in keeping with one another. Which label is applied to a child's language will further depend on the level of comparison used, as shown in the table. At a general level (Ai) we may ask "is language in keeping with nonlinguistic development?" By this criterion, children with developmental language disorders are by definition deviant. A ii to A iv are examples of increasingly specific comparisons, and it can be seen that deviance at one level does not imply deviance at another. Some examples may clarify this point. Lenneberg, Nichols & Rosenberger (1964) studied the linguistic abilities of 61 noninstitutionalized subjects with Down's syndrome aged from 3 to 22 years. Subjects' language performance was divided into four categories: mostly babble, mostly words, primitive phrases, and sentences. It was found that language abilities correlated significantly with motor development, suggesting that by the criterion of A i, these children have delayed rather than deviant language, with verbal and nonverbal skills in keeping with one another. This result is complicated by Lenneberg et al's finding that there was not a significant correlation between IQ and language development, a result which led them to conclude:

> "advance in language is more closely controlled by maturational factors than by intellectual ability as expressed by traditional IQ."
>
> p. 136

However, this analysis seems to have been misguided, since IQs were compared across a wide age range. A 20-year-old with an IQ of 30 will function at a higher level than a 4-year-old with an IQ of 60; the appropriate measure to use when comparing different aged subjects would have been "mental age". Godan & Panagos (1976) using a task in which children with Down's syndrome were required to repeat sentences varying in transformational complexity did find a relation between mental age and linguistic ability, and Karlin & Strazzulla (1952) found that when age at which words and sentences were first used was the dependent variable, IQ did relate to language development in a group of mentally retarded children (etiology unspecified). Taken together, it would seem that in Down's syndrome, language development is related to nonlinguistic ability.[1] Lenneberg et al also provide data relevant to question A ii, where they report that presence or absence of babbling is related to understanding, and:

> "apparently all of these aspects develop with certain synchrony in mental retardation."
>
> p. 125

i.e. delay by definition A ii. In terms of A iii, however, language is deviant, with phonology lagging behind expressive vocabulary.

Interpretation B offers a different concept of the distinction between deviance and delay. An aspect of language is termed deviant if it is never observed in the course of normal development (Crystal, Fletcher & Garman, 1976; pp 28-29). Here again, we may have different levels of comparison: a child's phonology might be quite normal, but he may construct neologisms using that phonology. Or he may have a normal vocabulary, but use it in a syntactically bizarre

[1] In fact the answer to question A i will depend on the definition of "nonlinguistic development": as Mittler (1972b) and Spradlin (1974) have pointed out, "mental age" will vary with different tests.

way. Returning to the study by Lenneberg et al (1964) we find that by this criterion, those authors' impression is that subnormal children are delayed rather than deviant:

> "the intellectual limitation does not produce bizarre behavious forms; it merely results in arrest at primitive, but "normal" stages of development."
>
> p. 134

In contrast, the language of autistic children, characterized by echolalia and pronoun reversal, is often held to be deviant in the sense of definition B (Kanner, 1946; Bartak, Rutter & Cox, 1975). However, Clark (in press) has produced evidence that pronoun reversal does occur in very young normal children before they master correct usage of these terms, so this might not be deviance so much as delay. However, Shapiro, Roberts & Fish (1970) compared autistic children with younger normal subjects and demonstrated that imitated speech of the former group differed qualitatively from the normal children.

So long as we define our terms carefully, the distinction between deviance and delay is useful, both theoretically and in remedial work. It may also prove to be of value in prognosis. An example comes from Rutter & Yule's (1973) work on reading disorders. These authors found that specific reading retardates, whose reading was poorer than anticipated from their general intelligence (c.f. A i: deviance) had a poorer prognosis than backward readers whose reading was commensurate with their intelligence (c.f. A i: delay).

Work on the language of children with developmental language disorders will be discussed in terms of the framework outlined in table 13.

Relation between expressive and receptive language

Research on this question has tended to concentrate on the

phonological level, asking whether children with abnormal articulation have normal phonemic perception. Most studies have been correlational, and results have been conflicting. Weiner (1967) showed that a positive relationship between auditory discrimination and articulation is most evident in studies which include children with four or more misarticulations. If the range of articulatory efficiency in the population studied is small, then it is hardly surprising that low correlations between discrimination and articulation result. This may be a better explanation for the diversity of results than Stewart's (1968) suggestion that a relationship between articulatory errors and sound discrimination holds only for a subset of children with articulatory disorders. However, there are other studies suggesting that it is useful to discriminate different types of articulation disorder (Crookes & Greene, 1963). Renfrew & Geary (1973) found that inability to discriminate similar sounding words was one test that picked out those children whose articulatory difficulties failed to improve after 6 months at school. Whitacre, Luper & Pollio (1970) also found that even for 6 to 7-year-old children simply diagnosed as having defective articulation (as opposed to developmental language disorder or dysphasia) other linguistic disabilities could be demonstrated. In particular these children did very poorly in a task where they were required to choose between acceptable and unacceptable phoneme clusters. Tallal & Stark (1976) studied a group of children with developmental language disorders and found that those children who had greatest difficulty in perceiving stop consonants did worst on all items in a speech production test.

Lerea (1958) devised comprehension and expression tests of

vocabulary and grammatical structure which he administered to 16 brain injured children with poor language and 16 children diagnosed as "aphasoid". He computed the ratio of expression to comprehension scores and found that this did not differ between the aphasoid group and a normal control group, but was reduced for the brain-injured group who did more poorly on the expressive scale of grammatical structure than on the corresponding comprehension scale.

Relation between semantics/syntax/phonology

No study to date has comprehensively examined the relationships between all three aspects of language in language-disordered children, but several studies have considered two systems.

phonology and syntax

Shriner, Holloway & Daniloff (1969) studied 30 children aged from 5-11 to 8-11 diagnosed as having severe articulation problems. Analysis of speech samples from these children showed that their use of syntactic structures was underdeveloped compared to a control group (matched on age, sex, race, socio-economic status and mental age). Shriner et al consider several hypotheses that might account for this result. It may be that such children deliberately shorten their utterances to reduce the probability of articulatory errors. However, anecdotal evidence against this view comes from Benton (1964) who described cases where syntactic abnormalities persisted after articulation had been normalized, or where syntactic errors were evident in the written output of such children. However, as Shriner et al observe, there could be a vicious circle in operation so that the child who misarticulates speaks less often and thus fails to experiment with new aspects of syntactic complexity, his language

suffering accordingly. Shriner et al consider alternatives however. Theories of generative grammar regard syntactic structures as input to the phonological system. Shriner et al argue that a defective syntactic system might therefore be the basis of the phonological disorder. They also consider the possibility that the disruption might be in the feedback mechanisms regulating syntactic and phonological performance, though they do not elaborate this point. A study by Menyuk & Looney (1972b) supports the notion that the deficit might be primarily a syntactic one. These workers gave their subjects repetition tasks involving real or nonsense words, and found a high correlation between syntactic errors and phonological errors in repetition of real words, but the number of phonological errors on real words was not significantly related to phonological errors in repetition of nonsense syllables; i.e. phonological processing of meaningful and meaningless sequences was quite different. Marquardt & Saxman (1972) considered receptive language in a group of children described as "articulation deficient". 30 experimental children and 30 control children were given Carrow's syntactic test and the Wepman auditory discrimination test. The experimental group did poorly on both, and interestingly, within the experimental group only, there was a positive correlation between performance on the two tests.

The study by Whitacre et al (ibid) on a similar population found significant correlations between ability to discriminate acceptable from unacceptable phoneme clusters, and tests designed to test knowledge of word form class and of sentence structure.

A contrasting result was obtained by Foley & Locke (1971) who found a nonsignificant correlation between performance on a task of

judging permissible and nonpermissible syllable sequences and a morphology test similar to that used by Berko (1958). However, these workers used a sample of only 20, presumably normal, children, and the low correlation probably reflects the limited range of scores obtained on the phonology test.

In sum, there are several pieces of evidence supporting the idea that the syntactic and phonological difficulties experienced by language disordered children are related to one another.

semantics and syntax

Two studies have matched children on MLU, which may be regarded as a crude index of syntactic development, before comparing their use of semantic relations.[1] These studies, by Freedman & Carpenter (1976) and Leonard, Bolders & Miller (1976) obtained similar results, namely that there was virtually no difference between language disordered children and younger MLU-matched controls in the semantic relations expressed. Freedman and Carpenter concluded:

> "that the language impaired group expressed the same semantic relations with as much variability and flexibility as the normal group suggests that the language impaired group seemed not only to possess those basic cognitive structures necessary for the development of corresponding linguistic structures but they have organized and coded their prelinguistic experiences as have the normal children."
>
> p. 792

An experiment by Liles, Shulman & Bartlett (1977) using a different technique showed that in some respects the semantic development of language disordered children is in advance of their syntactic development. These investigators required children to

[1] An extremely clear account of what is meant by "semantic relations" is given by Brown (1973; pp 7-11), who also lists the major types of semantic relation. A more detailed treatment is given by Fillmore (1968).

judge the acceptability of sentences which were normal, semantically anomalous or syntactically anomalous. The latter sentences included those with deviant word order or other errors such as failure of noun-verb agreement. 15 language disordered children and 15 controls aged from 5 to 7 years, all with PPVT scores within 1 s.d. of the mean, were asked to judge the correctness of sentences, and to remedy those they found incorrect. Language disordered children proved to be as competent as controls in detecting semantic errors, but did significantly more poorly on syntactic errors. Fewer experimental children could satisfactorily correct sentences they judged to be wrong.

Patterns of performance within linguistic categories

phonology

A number of authors (e.g. Weber, 1970; McReynolds & Huston, 1971; Leonard, 1973) have argued that the articulation abnormalities of some children cannot be explained in terms of developmental delay and do not seem determined by peripheral motor constraints. Oller (1973), for example, has written substitution rules for speech samples of five language disordered children and demonstrated that their substitutions are regular but do not invariably involve producing a less complex phoneme than the target. Indeed some children, while failing to produce a phoneme correctly in its proper context, articulate it perfectly in other contexts. McReynolds & Huston (1971) showed that distinctive feature analysis (Chomsky & Halle, 1968) produced a more parsimonious description of the misarticulations of some children than analysis in terms of phonemes, and that several children appeared to be using different rules rather than producing simplifications. This approach may be simply summarized in the terms suggested by Pollack and Rees (1972) who distinguish between phonetic

and phonemic articulatory errors. Phonetic errors may result from a variety of peripheral factors, e.g. problems in auditory processing, structural anomalies of the speech apparatus, muscular and neuromuscular disorders. Such a child may have adult phonological competence, but be unable to demonstrate it because of performance constraints. Children who make phonemic errors, on the other hand, correspond to those traditionally labelled as having functional articulatory disturbances. These are the children whose misarticulations are best described in terms of distinctive feature theory, and who seem to lack adult competence and use a deviant set of rules.

A fascinating study by Kornfeld (1976) asked not whether language disordered children were poor at perceiving phonemes, but rather whether such children perceived and produced distinctions imperceptible to normals. Kornfeld noted that both normal preschool children and children with language disorders frequently do not accept adult imitations of their own speech as genuine imitations. Such children appear to give the same response ("kaeb") when asked to say "crab" and "cab", but will not accept an adult's saying "kaeb" as a correct imitation of their output. Kornfeld claimed that the spectrographs of pairs of child's utterances that sound identical to adults are different, and show acoustic regularities. Thus, she argued that these children might be perceiving, representing and producing distinctions that are imperceptible to adult listeners. Unfortunately it is not clear from Kornfeld's account whether the spectographic analyses were performed blind. She also cited a study by Goehl & Golden (1974) who tested a group of children whose production of /r/ and /l/ sounded like /w/. They found that these children

discriminated r/l distinctions produced by adults and had significantly greater success in recognizing their own productions of /r/ than did control groups. This would suggest that their phonological systems are deviant in terms of definition B. However, this does not seem to apply to all children. Panagos & King (1975) tested whether children with deviant speech understood their own deviant productions (imperative sentences elicited in a repetition task) and found that, like normal children, they understood normal speech better than their own output.

semantics

Leonard et al (ibid) found no difference in semantic relations expressed by language deficient children and MLU-matched controls, although when compared with age-matched controls some differences emerged, namely the experimental group was less likely to use the Dative case, and more likely to use Essive + designator and verb + Agentive constructions[1]. Freedman & Carpenter (ibid) found only one difference between language disordered children and MLU-matched controls, namely that language disordered children used introducer + Essive constructions with greater diversity. In sum, the semantic relations expressed by language disordered children do not appear to be deviant in the sense of definition B.

Research is lacking at the lexical level. It would be interesting to find whether some types of word give greater difficulty than others: e.g. do abstract words pose relatively more difficulty for language disordered children than for controls? Are semantically related words (e.g. lock and key) more frequently confused by language disordered children than by controls at a similar linguistic level?

[1]These cases are explained by Fillmore (1968)

syntax

Menyuk (1964; 1969) investigated language samples from ten children aged from 3-0 to 5-10 diagnosed as having infantile speech, and ten control children, matched on age, IQ and social class). She found that many structures were used relatively infrequently by the experimental group, e.g. conjunction was found in the samples of all normal children but in only four of the language disordered children. The language disordered samples also contained many omissions, and some sentences violated class and sequence rules; e.g. "big the dog" or "any more not pictures?" Such sentences were very rare in the control samples. Comparison of the speech of a 3-year-old language disordered child and a normal 2-year-old revealed more dissimilarities than similarities. Menyuk (1969) proposed that language disordered children appear to use a very basic grammatical structure, which they fail to elaborate. Expansions of noun phrases do occur, but restrictions on co-occurrence are frequently not observed and articles are usually omitted. Verb phrases are almost never expanded. Conjunctions are rare, and when they do occur, conjoined phrases do not agree with one another in number, tense, etc. Menyuk also stated:

> "In some of the sentences that the children in the deviant speaking population produce there is also some evidence that they may be using rules or forming hypotheses about the structure of language which are different from those of the normal-speaking population."
>
> p. 134

i.e. Menyuk is claiming that the syntax of these children is deviant in the sense of definition B.

Lee (1966) developed an assessment procedure, based on a developmental interpretation of Chomsky's (1957) transformational

grammar, which she used to compare the language of a normal 3-year-old with a language delayed 4-year-old. She found many differences between the children, some of which might qualify as instances of definition B deviance; e.g. the language delayed child used "be" as a pivot in structures such as "be finished", "be here", and omitted some types of structure, such as two-word designatives, which were common in the speech of the normal child (e.g. "that horse", "there ducks").

Leonard (1972) compared elicited speech samples from 9 children with defective language to 9 controls matched on age, PPVT and class, using Lee and Canter's (1971) method of analysis. He noted that while the number of children using various syntactic structures did not differentiate groups, the frequency with which each structure was used did. Leonard stresses that no qualitative differences between groups were apparent, but the relative frequency of usage of various structures did differentiate groups (deviance: Aiv; delay: B).

Morehead & Ingram (1973) compared 15 linguistically deviant children with 15 younger controls. Selection of subjects was on the basis of MLU, using the five stages of linguistic development described by Brown (1973), three children in each group being selected for each stage. It should be noted that the control group was not matched on other variables, and was likely to be of different socio-economic status from the experimental group since the former came from a University nursery school. Morehead and Ingram used an adaption of Chomsky's (1965) transformational grammar to describe the linguistic output of their subjects. Groups were compared in their use of i) phrase structure rules; ii) transformations; iii) "construction types" - i.e. surface realization of major syntactic

categories; iv) inflectional morphology; v) particular lexical items. No difference was found between groups except in two respects. Comparison v) revealed that the experimental group used more pronouns, but did not differ on usage of other lexical items studied. Comparison iii) showed that the deviant group did not use the major linguistic categories in as many different contexts or syntactic frames as did the normal group, i.e. the former did not have bizarre linguistic systems, but rather used normal systems in a rather limited way. Morehead & Ingram conclude that the prime disorder of children with developmental language disorders is prelinguistic rather than linguistic, i.e. they lack the ability to formulate the representational ideas behind language. However, it is possible that their result is an artifact of social class differences. It is known that there are class differences in language usage, and these are likely to affect content as well as structure.

Johnston & Shery (1976) looked specifically at the use of grammatical morphemes by language disordered children. Language samples from 287 children aged from 3 to 16 were analysed by trained teachers. All children were in classes for children with severe language disorders, and had IQs within 2 s.d.s of the mean. MLU was used to classify children into five stages of development (Brown, 1973). Johnston & Shery showed that order of difficulty of morphemes was the same as for normal children studied by other investigators. They concluded that the order of difficulty of acquisition of morphemes was the same for normal and language disordered children, but that the latter acquired morphemes one or two stages later than normal children. This latter conclusion may be viewed with some

scepticism, since the comparison group was of different social class, and further, the selection criteria for the experimental group were sufficiently lenient to allow the inclusion of children of educationally subnormal intelligence.

Apart from analyses of speech samples, the most widely used technique for studying the syntactic development of language disordered children has been sentence repetition. There is considerable evidence that comprehension as well as expression is involved in such tasks (Menyuk, 1963; Rodd & Braine, 1971; Slobin & Welsh, 1973; but c.f. Fraser, Bellugi & Brown, 1963).

In an uncontrolled study, Berry-Luterman & Bar (1971) presented four language disordered subjects with sentences that were either grammatically correct or contained grammatical errors. They noted that two of their subjects did as poorly on sentences with correct word-order as on reversed sentences, and concluded that these subjects were not helped by grammatical structure.

A better designed study was conducted by Menyuk & Looney (1972b) who compared 13 language disordered children with 13 controls (roughly equivalent on PPVT). Two sets of sentences were used. Set A included various grammatical transformations. Set B sentences were so designed that all consonants appeared at least once in initial, middle and final positions. The experimental group made many more errors than did the control group and were especially hampered by transformational complexity. Their errors consisted predominantly of omission of articles.

Other recent studies examined the types of errors made by language disordered children in sentence repetition tasks. Weiner (1969a), investigating mildly impaired children, found that a sentence

repetition test differentiated between experimental and control children better than any other test. The pattern of errors was similar for both groups, with omissions and substitutions comprising the bulk of errors, although the experimental group did make more errors involving morphology and syntax. Dukes and Panagos (1973) noted that when repeating sentences, children with expressive disorders tended to omit words, rather than adding to or substituting words in the sentences. Bliss and Peterson (1975) obtained a similar result for children repeating sentences which they understood. In none of these experiments were sequencing errors prominent.

Very few studies have looked at comprehension. Tallal (1975) gave dysphasic children the Token Test and found that they performed more poorly than a control group overall, but that the order of difficulty for the different sections of the test was the same for the two groups, namely section 4, containing lexically dense items, was more difficult than the syntactically complex section 5. Thus it would appear that her experimental group did not have a specific difficulty with syntax.

Stanton (1976) was specifically concerned with the question of whether language delayed preschoolers' comprehension of grammatical structure showed an abnormal patterning. Children were asked to act out a variety of sentences. She did not find significant differences between language delayed children and a control group, and the order of difficulty of six structures was the same for both groups. She concluded that, as far as comprehension of grammar was concerned, these children were neither delayed nor deviant. However, she presented sentences under various conditions (well-formed, telegraphic, scrambled, etc) and did find a significant interaction between groups

and conditions. The statistical analysis she presents is inadequate and does not allow for specific comparisons to be made, but inspection of her data suggests that language delayed children do almost as poorly with well-formed and telegraphic presentation as with scrambled sentences, whereas the control children obtain much lower scores on scrambled sentences (as poor as those obtained by the experimental group) than in other conditions. Thus it would appear that the language delayed group profit far less than the controls from grammatical structure.

Overview

Most of the research reviewed here supports the idea that language disordered children have especial difficulty with syntactic and phonological components of language, with relatively better semantic performance. One result which disagrees with this view is that of Tallal (1975) who found that syntactically simple but lexically dense instructions gave more difficulty than syntactically complex ones. The other studies reviewed here have simply not considered the variable of lexical density, and it is possible that if language disordered children's output were examined in terms of this variable then deficiences would become apparent. Dukes and Panagos (1973) found that adjectives and determiners were the most likely items to be omitted in a sentence repetition task, which is in keeping with the idea that these children are limited in the number of nonredundant words they can process in a sentence.

Psychological approaches

Comparatively few psychologists concerned with developmental language disorders have regarded the disordered language itself as

the appropriate focus for attention. More commonly, research has concentrated on uncovering a primary disorder, not essentially linguistic in character, which could result in disordered language development. Such research will be considered before discussing more specifically linguistic hypotheses.

Perceptual disorder

By definition, developmental language disorders cannot be accounted for in terms of deafness. Nevertheless, several authors have commented on the inconsistent auditory thresholds obtained in several cases of receptive disorder (e.g. Reichstein, 1964; Lea, 1965) and it has been suggested by Benton (1964) that some high-level auditory perceptual deficit may be present.

> "It is not so much a question of elevation of auditory thresholds (it is evident that many of these children show no such elevation) but of deficits in the orientational, discriminative and integrative aspects of auditory cognition."
>
> p. 51

McReynolds (1966), using an operant conditioning procedure, demonstrated that language disordered children were capable of discriminating between speech sounds as well as normal children, so long as the sounds were presented in isolation. However, the experimental group did more poorly than controls when the sounds were embedded in nonsense words. McReynolds rejected the notion that language disordered children have poor auditory discrimination and suggested that they could have a problem in sequencing input, an hypothesis which will be discussed below. Weiner (1969a) found that a group of language disordered children had no difficulties in an auditory discrimination test. These results do not mean that a primary perceptual disorder is not involved in some cases. Weiner's

subjects had mild disorders. McReynolds did not give a detailed description of her subjects, merely saying that they displayed sufficient aphasic involvement to be included in the therapy programme of an Aphasia Institute. It is noteworthy that one of her subjects was excluded because of inability to form the basic association between pressing a lever and hearing a sound necessary to do the task. This child might have had a perceptual difficulty of the kind suggested by Benton. It may require particularly sensitive tests to detect some perceptual disorders. De Ajuriaguerra et al (1976) do not describe their test battery in detail, but claim that in a series of 40 children with developmental language disorders, no child had normal perception for phonemes and nonsense syllables, although all had adequate articulation and normal responses on pure tone audiometry.

Tallal & Piercy (1973a, b, 1974, 1975) have proposed that language disordered children have a specific difficulty in perceiving rapid auditory stimuli. In an initial series of experiments (Tallal & Piercy, 1973a, b) these workers demonstrated that language disordered children could be trained to make a specific response to each of two tones, A and B, but when required to make discriminative responses to tone pairs such as AB, AA, BA, BB, they did poorly at short tone durations, or if the tones were separated by 150 msec or less, but did as well as a control group when longer tone durations or intertone intervals were used. This result held both when children were required to indicate whether the two members of a pair were the same or different, and when they were required to press two panels in sequence to correspond to the tones they heard. In an analagous experiment using visual stimuli (light flashes of two shades

of green), language disordered children did as well as controls. In a second set of experiments, Tallal & Piercy (1974, 1975) presented children with synthetic vowel-vowel and consonant-vowel syllables. Consonants are characterized by rapid transitional formants, and Tallal & Piercy were interested in the question of whether it was the brevity of the discriminable components of consonants or their transitional character that made them harder to discriminate than vowels for language disordered children. They showed that language disordered children have no difficulty in processing transitional stimuli as such, so long as these are "stretched" in time, but they do have difficulty with brief stimuli, whether transitional or not. Of relevance to this hypothesis is a study by Blosser, Weidner & Dinero (1976) who studied a population of children whose scores fell below the 10th percentile on the NSST scales. These children were presented with the Token Test at normal and slowed rates of speech (effected by a Rate Changer). As might be expected, significant main effects were obtained for language disorder and rate, as well as for sex. However, although some significant interactions were obtained, these are hard to interpret. The interaction of interest, between language-disorder and rate was not significant. This result is not in keeping with Tallal & Piercy's hypothesis.

Associating sound with meaning

Another hypothesis considered by Benton (1964) is that language disordered children have difficulty in attaching meaning to sounds, although the latter are perceived adequately. A similar hypothesis has been proposed by Worster-Drought (1943) who remarks:

> "the difficulty is not in the auditory perception of the spoken sounds but in the recognition of the acoustic word-imagery."

We might expect that a child with such a disorder would also have difficulty in attaching meaning to nonverbal sounds. Children with such difficulties have been described by Benton (1964) and Worster-Drought (1943, 1968), but no systematic study has been carried out.

An experiment by Wilson, Doehring & Hirsch (1960) can be interpreted in terms of this hypothesis. Their subjects were described as primary sensory (i.e. receptive) aphasics. The purpose of the experiment was to investigate these children's ability to discriminate short and long tones and noises. They were required to associate 4 sounds with corresponding letters. The experimental group did more poorly than a control group on this task, but it was unclear whether this was owing to failure to discriminate the auditory stimuli, failure to discriminate the visual stimuli, or failure to make the necessary association between auditory and visual stimuli. Observation of a single subject a year later suggested that the latter could be the case, since this boy remained unable to associate auditory and visual stimuli despite demonstrating adequate discrimination between auditory stimuli. Wilson et al commented on the extreme variability shown by their subjects - 8 of their 14 experimental subjects did almost as well as controls, whereas the remainder could not learn at all.

It should be noted that failure to derive meaning from sound could lead to apparent perceptual problems. Mark & Hardy (1958) suggested that if a class of stimuli has no meaning for the child his responses to them will habituate, and he will appear not to perceive them, just as normal children appear not to respond to background noises, such as traffic.

Sequencing

Orton (1937) first proposed that language disordered children's difficulty is not so much in perceiving elements, as in sorting out the order in which they occur. Others have taken up this idea, but there is disagreement as to the modalities and units affected. For example, Monsees (1968) has suggested a disorder restricted to the auditory processing of phonemes, whereas Stark (1966) proposes a more general deficit affecting visual as well as auditory stimuli. What is the evidence for such theories?

An experiment by Furth (1964) contrasted the performance of deaf and normal children with a group suffering from receptive language disorders. Deaf children are an important control group because they allow us to see whether any deficit observed in a language disordered group might arise as a secondary consequence of a more basic perceptual disorder. Furth trained his subjects to make discriminative responses to visual stimuli comprised of pairs of nonsense patterns in a particular order. No differences were found between the three groups of subjects in this test. Furth & Pufall (1966) extended this design to include stimuli consisting of three elements in a particular order in both visual and auditory modalities. They tested 2 groups each of deaf and aphasic children aged 6 to 7 years and 10 to 11 years. No differences were found between deaf and aphasic children except for 6- to 7-year-olds on auditory stimuli. However, there is some evidence that the aphasic children may have had more difficulty in discriminating auditory stimuli than the deaf group. All aphasic children had hearing losses of 50 to 100 db, and two aphasic subjects were excluded because of inability to discriminate the auditory test stimuli. Possibly other subjects included in the

aphasic group had sufficient difficulty in discriminating stimuli to disrupt their performance in the sequencing test.

Lowe & Campbell (1965) tested a group of children termed "aphasoid" aged between 7 and 14 years, on two tasks. In the first, the children had to judge whether two sounds they heard were successive or simultaneous, whereas in the second, they had to report the order of two stimuli. Only scanty details of experimental procedure are given, so it is not clear whether or not these children were required to make a verbal response. No difference was found between aphasoid and control children on the detection of successivity, but the aphasoid children were worse at ordering stimuli.

Stark (1966, 1967) examined the performance of children with predominantly receptive language disorders on the ITPA, and found that they did poorly on all tasks involving sequencing, in both visual and auditory modalities. However, Stark relied on the ITPA normative group for control data, an unsatisfactory procedure in view of the restricted range of intelligence of the normative group. Olson (1961) using an early version of the ITPA found that language disordered children did more poorly than deaf children on tests corresponding to verbal expression, auditory reception, grammatical closure and visual association, but the sequencing tests did not differentiate groups.

Stark, Poppen & May (1967) compared 8 children aged 7-5 to 9-5, diagnosed as aphasic, with normal children aged from 4-2 to 7-11. This control group does not appear to have been matched with the experimental group on any variable, and included many siblings of the aphasic children. Subjects heard sequences of 1 to 3 words and were required to press keys corresponding to those words in the same order. Three aphasic children who did very well on this task were omitted

from later analyses. For the remainder of the experimental group, errors tended to occur for the first item in the sequence, unless this was stressed. There are many difficulties with this experiment. The control group is inadequate, and the post hoc exclusion of three experimental subjects unjustified. Even if we accept Stark et al's results, there is no evidence that it is sequencing that is impaired in the aphasic children. The errors of these children tended to be omissions, rather than ordering errors. Poppen et al (1969) gave receptive aphasic children a variety of visual and auditory sequencing tasks, and found that they did more poorly than a control group, but once again, the errors made by the experimental group do not appear to have been order errors. Further, a brief report by Withrow (1964) indicates that although language disordered children do more poorly than a control group on the recall of sequential visual stimuli, deaf children do no better, suggesting that the difficulty that aphasic children have with this type of task may be a consequence rather than a cause of their language disorder.

Monsees (1968) studied a group of more mildly affected children, aged 6 to 13 years, whose expressive speech was of low intelligibility. Children were asked to repeat phonemes after the experimenter. The poor performance of the experimental group led Monsees to conclude that they had a difficulty in auditory temporal sequential processing, but yet again, this conclusion is not supported by the type of errors made by these children. Substitutions of different phonemes rather than rearrangement of phonemes tended to occur. Similarly, Weiner (1969a) reported that in tests of repetition of vowels and sentences language disordered children do not tend to make errors of sequence, although their performance is poor.

Rees (1973) points out that those advocating defective sequencing as the basis for developmental language disorders invariably quote the work of Efron (1963) showing that adult aphasics may require ten times the normal interval between brief auditory stimuli in order to discriminate their temporal order. She notes that elsewhere Efron (1967) states that this defect is not related to comprehension of spoken language, so cannot be invoked to account for receptive deficits. Rees questions the theoretical as well as the practical rationale of sequencing theories. In common with Warren & Ackroff (1976), she stresses that normal perception of speech is not achieved by separate analysis of each sound segment in sequence. Nor is analysis of grammatical structure performed sequentially. Those suggesting that language disordered children have difficulty sequencing either phonemes or words are out of touch with modern psycholinguistics.

In sum, there is much evidence that language-disordered children do poorly on sequencing tasks. However, the pattern of their errors does not suggest that it is sequencing that is the crucial variable. Relevant here also is the work on sentence repetition reviewed on pp. 120 to 121, which shows that language disordered children are poor at repeating sentences, but their errors are not typically ones of word order.

Memory

Masland & Case (1968) noticed that three children with delayed language development had difficulty in repeating unstructured strings of digits, and suggested that a short-term memory deficit might be responsible for their language disorder. Smith (1967) found that children with mild to severe functional articulation problems did more poorly than IQ-matched controls in several digit repetition tasks, while not differing in tests of visual short-term memory. The

problem with both these studies is that the child must speak in order to perform the memory tasks - when children do poorly it cannot be told whether it is their memory or their poor expressive ability which is responsible.

Menyuk (1969) noted the poor sentence repetition of a group of language disordered children, and also the high correlation between sentence length and ability to repeat which was found for language delayed but not for normal children. She suggested:

> "...as memory capacity or storage capacity increase, normal-speaking children expand the rules in their grammar, observe more contextual constraints, and, in some instances reorganize rules to achieve greater definition and economy. These rules are stored in long-term memory for regeneration of various structures. If the 3 year old child could not keep in short-term memory storage more than a two-to-three morpheme length utterance it is difficult to see how this development of a deepening analysis could be carried out, since only a very limited amount of data could be handled at one time. Further, although short-term memory capacity increases over time in this population, it seems that the hypotheses or rules derived from this limited data and to be used in the comprehension or generation of utterances are also limited."
>
> p. 142

However, the evidence for this hypothesis consisted of repetition of grammatically correct sentences. It could be argued that failure to make use of the syntactic structure of the sentences was the primary disturbance which led children to treat them as they would unstructured strings.

In later work, Menyuk seems to reject the notion of a primary disorder of short-term memory. Menyuk & Looney (1972a) found that sentence structure was more important than length in determining whether or not repetition would be accurate, and concluded:

"Apparently the factor creating the greatest difficulty in repetition for the experimental group is not difficulty in remembering all the items in the sentences but rather difficulty in understanding the transformational structure of those sentences."

p. 276

This is not to say that language deficiences cannot result from memory limitations: the work of Graham (1968) for example suggests that this might be the case for educationally subnormal children. However, the evidence reviewed suggests that for children with developmental language disorders this is not such a likely hypothesis.

Symbolic function

Piaget (1970) distinguishes between operative and figurative aspects of cognitive functions. Operative activities attempt to transform reality, whereas figurative activities represent reality as it appears. The semiotic or symbolic function overlaps with, but is not identical to, figurative activity, and involves the ability to represent an object which is absent by means of symbols or signs, which are not identical to their referents. Language belongs to the symbolic function, as do imitation, mental imagery and drawing.

Workers in the Piagetian framework (de Ajuriaguerra, 1966; de Ajuriaguerra et al, 1976; Inhelder, 1976) have proposed that language disordered children are deficient in symbolic functioning, and have demonstrated that such children do poorly on tests designed to assess such activities. However, it is not clear whether they regard this deficit as the cause, consequence or concomitant of language disorder.

Affolter, Brubaker & Bischofberger (1974) demonstrated that a group of language disordered children were very poor at drawing, and were deficient in imitative behaviour, including nonverbal aspects

such as gesture and facial expression. However, their sample does not appear to have been a very pure one. Many of their subjects had hearing losses of a greater or lesser degree, and performance on several nonverbal tasks was low. The authors argued that this latter result was due to the children not understanding task requirements, rather than inability to solve the problems, but nowhere do they present satisfactory evidence that their sample was not generally mentally retarded.

Lovell, Hoyle & Siddall (1968) found very little difference in the amount of symbolic play engaged in by language disordered and normal nursery school children, although the former group demonstrated less social play.

Griffiths (1972) has argued on the basis of her own observations that children with developmental language disorders can generally use and understand quite complex gestures, and may show considerable ingenuity in producing detailed and expressive drawings.

Hughes (1974/5) attempted to teach four children with severe receptive disorders the nonverbal "language" designed by Premack (1971), "Premackese". This language was originally designed to be taught to a chimpanzee. Words are represented by plastic shapes which can be arranged on a board to make sentences. Hughes found that the functions "word", "sentence" and "class concept" were acquired by all four subjects with little difficulty - i.e. they learned to associate shapes with objects, to comprehend simple strings of shapes representing sentences such as "Jenny give doll", and to learn the meaning of shapes that stood for general concepts such as "colour". Performance was less good on the concept of negation

(i.e. a shape was used to represent "not") and various question forms, although one child did master all of these. This work shows that language disordered children do possess at least a rudimentary symbolic ability. Hughes nevertheless draws attention to the limitations of "Premakese" and points out that it lacks the full status of language.

Primary linguistic disorder

Crystal, Fletcher & Garman (1976) feel that in the study of language disordered children, the primary focus of attention should be the disordered language itself, rather than supposed underlying factors. This case has been put strongly by Cromer (1977) who argues in particular that theories stressing sequencing impairments are misguided. Like Rees (1973), he points out that since Chomsky's (1957) demonstration of the inadequacies of Markov process models of language, linguists have come to accept that grammar has an hierarchical rather than a sequential structure. Cromer suggests that the deficits of language disordered children are not related to sequencing problems but rather to difficulties in abstracting the hierarchical structure of language. He points out that research demonstrating that language disordered children have difficulties in detecting and reproducing rhythmic patterns (de Ajuriaguerra, 1966; Griffiths, 1972; Kracke, 1975) is in keeping with this hypothesis, since rhythms are better described in terms of hierarchical rather than sequential structure (Martin, 1972). Cromer compared the written language of ten children with receptive language disorders with that of deaf children. Subjects were shown a puppet show designed to elicit a variety of grammatical structures, and asked to describe what they had seen. Superficially there were few

differences between deaf children and those language disordered children who produced analysable output (six out of ten subjects). However, detailed grammatical analysis revealed a number of dissimilarities. Language disordered children used significantly less phrase structure categories (e.g. noun phrase, verb phrase, etc) per sentence, and were more restricted in the sentence types they used. No language disordered child attempted to use future or perfect tenses, negatives, questions or complements. Conjoined and embedded sentences occurred only in the output of the deaf. These results are similar in many respects to those found by Menyuk (1964, 1969) whose studies on the expressive speech of language disordered children have already been discussed. She found that the output of these children was characterized by a simple basic structure which was largely unelaborated, most sentences being of the form:

$$S \rightarrow NP + VP$$

$$\text{or} \quad S \rightarrow NP + V + NP$$

where the noun phrase (NP) was rarely expanded into determiner + noun, and the verb phrase (VP) was almost never expanded. Like Cromer, she found that conjunction and embedding were very rare. Kerschensteiner & Huber (1975) reported a single case of developmental language disorder in a German-speaking man who had persisting language problems at the age of 23 years. The most notable feature of both comprehension and expressive speech was difficulty in handling grammatical structure. Like Menyuk's subjects, this patient observed only very general forms of grammatical transformation and neglected many grammatical constraints.

Cromer's hypothesis has the interesting implication that, far

from being the result of disordered sequencing ability, language disordered children's language might result from their reliance on an inefficient sequential approach in the absence of ability to abstract hierarchical structure. The type of language described by Cromer, Menyuk, and Kerschensteiner & Huber would seem to be what one would expect on the basis of Cromer's hypothesis. Without some grasp of the total structure of a sentence, elaboration could not occur, and failure of different parts of the sentence to agree might be expected if the child did not grasp the structural relations between parts.

Evidence against the notion of a specifically grammatical impairment comes from Tallal (1975) who showed that, like normal children, language disordered subjects do no worse on the grammatically complex part 5 of the Token Test than on the less complex part 4.

Discussion of psychological theories

Two important general comments apply to the research described above.

First, it must be noted that out of necessity experiments in this area are correlational. It would be neither possible nor ethical to induce language disorders in children and observe the consequences of such manipulation. Thus if a psychological deficit is discovered in a language disordered group, it cannot be assumed that the deficit is responsible for the disordered language. The deficit may be a result of the language disorder, or both may be caused by a third factor.

To illustrate this point, let us consider the theories proposing

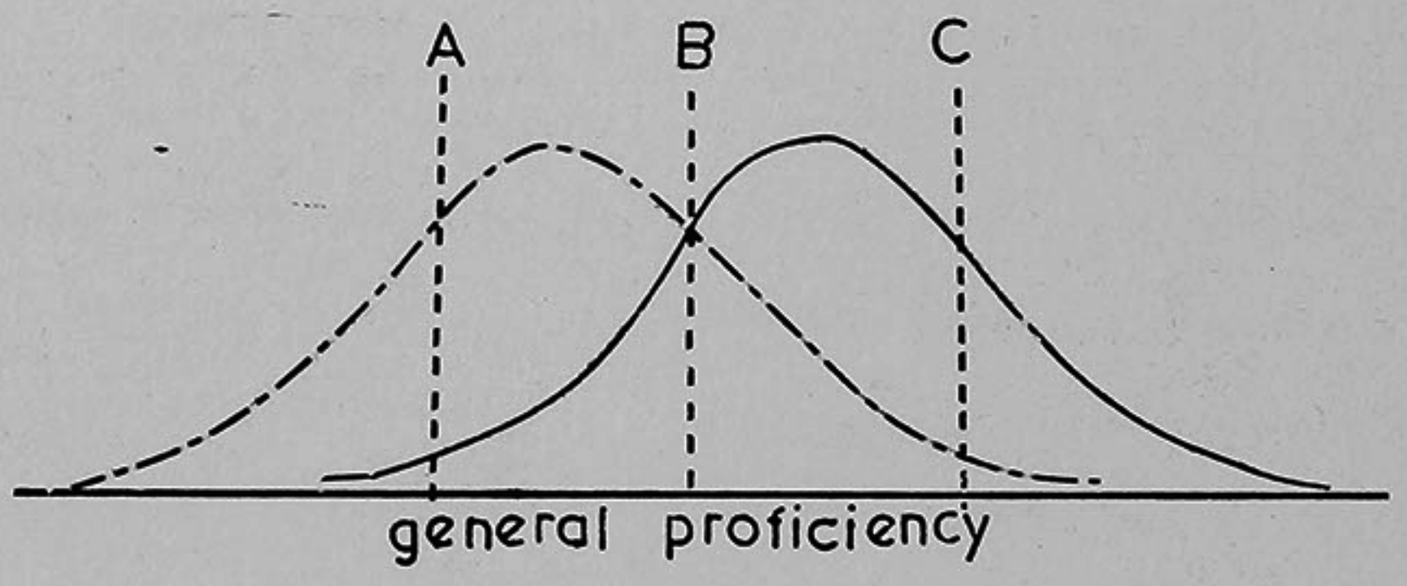
defective sequencing of auditory stimuli. It may be that normal children use implicit verbal coding of stimuli in tasks designed to assess auditory sequencing, and that language disordered children cannot apply this strategy. In this case, the language disorder would be responsible for the observed sequencing deficit, rather than the converse. Alternatively, memory limitations might be responsible for both the poor sequencing ability and the disordered language. This is but one example: similar arguments can be applied to other theories discussed above.

Some workers use a deaf control group. If psychological deficits occurring in language disordered children are also found in the deaf, then they are very likely secondary consequences of disorder rather than causal factors. However, the converse is not true. If deficits are found in language disordered children which are not observed in the deaf, then it is not valid to conclude that the deficits are the primary basis of the language disorder. Language disordered and deaf children differ in many respects, not only in symptomatology but also in the type of education that they receive, and the differential deficits could be a consequence of these discrepancies.

In order to test between theories it is necessary to derive refutable predictions from them. The field of potential theories is narrowed particularly by considering language-disordered children's abilities rather than their disabilities. Consider for example the notion that there is a deficit in figurative thought. From such a hypothesis we would predict that not only verbal language but also sign language should be disturbed in these children. Demonstration of impairment in sign language would not confirm this hypothesis

Fig. 22

Discriminating power of tests requiring different levels of general proficiency



----- group X

——— group Y

since such a disturbance could result from other disorders, but demonstration of a good capacity for learning sign language would refute the hypothesis.

A second difficulty concerns the problem of discriminating between general and specific impairments. A specific impairment is generally demonstrated by contrasting performance on a control task with that on a task relevant to the specific factor in question. Thus workers interested in sequencing compare performance on tasks using single stimuli with those using sequences of stimuli, and demonstrate that language-disordered children are poor on the latter task only. However, such selective deficits may be artifacts due to differences in the discriminating power of control and sequencing tasks. This important point has been made in detail by Chapman & Chapman (1973)[1]. The argument is essentially as follows. Consider two groups of subjects, X and Y, neither of whom has a specific disability, who differ only in respect of what we will term "general proficiency", i.e. nonspecific variables affecting task performance such as level of motivation, alertness, ability to follow instructions, general intelligence, etc. Distributions of general proficiency for these theoretical populations are given in figure 22. It would be possible to devise a pair of analagous visual and auditory tasks such that group X would appear to have a specific visual deficit, by using tasks of different difficulty. Suppose we used an auditory task of difficulty level A, such that subjects whose general proficiency fell to the left of the cutoff A failed the task, the remainder passing. We can see that about 70% of group X and

[1]Many of the points made by Chapman & Chapman are anticipated in an article by Baumeister (1967).

95% of group Y will pass this test. Consider then a visual task of difficulty level B, such that subjects whose general proficiency falls to the left of the cutoff line B will fail, and the remainder pass. Such a task will be passed by about 30% of group X and 70% of group Y. The difference between groups is much bigger for the visual than for the auditory task, and we might be tempted to conclude that group X have a specific visual difficulty. However, by reversing the difficulty level[1] of the visual and auditory tasks we could obtain the reverse effect. It is suspected that this argument, which is an extension of the familiar concept of floor and ceiling effects, may apply to much of the work reviewed here. Consider for example the experiments on sequencing, most of which used rather easy control tasks, which are probably less discriminating than the experimental tasks. Inspection of Tallal & Piercy's (1973a) data suggests that the same principle might account for their results: at long intervals normal children obtained at or around 100% correct, whereas aphasics obtained around 95% correct: obviously the two groups did not differ significantly. At shorter intervals, normal children obtained around 92% correct whereas aphasics' performance dropped to around 60% correct, this difference between groups being significant. It could be argued that when long intervals are used the task is so easy that only a low level of general proficiency is required to perform correctly, so the task does not discriminate between groups, whereas when shorter intervals are used, more general proficiency is required. Indeed, Tallal & Piercy's (1973b) own data support this interpretation: when the task is made more difficult by increasing the length of tone sequence, then the aphasic group does significantly

[1]Difficulty level is not the only factor affecting discriminating power of tests, but is the easiest to illustrate here. For a more detailed account see Chapman & Chapman (1973).

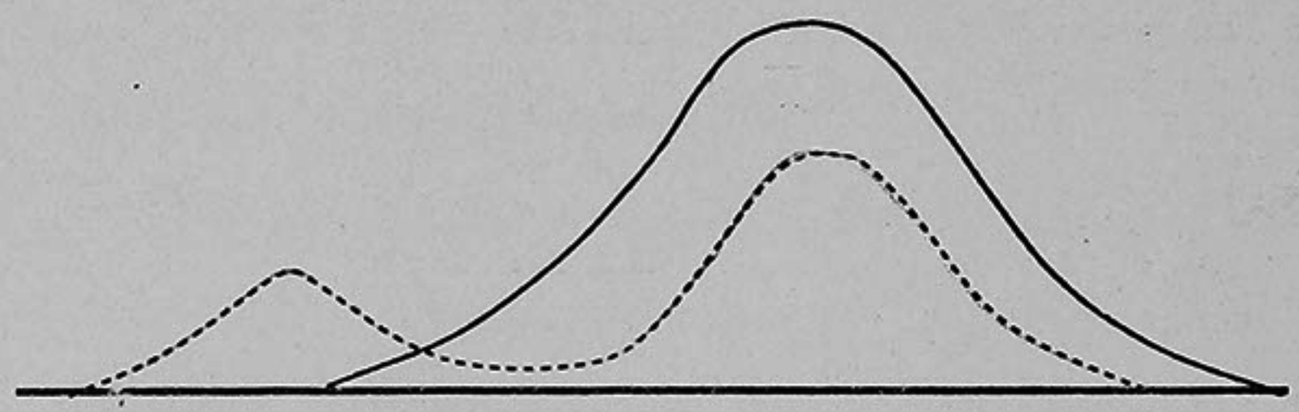
more poorly than the control group at long as well as at short tone durations. However, the picture is complicated by Tallal's (1976) demonstration that her aphasic group do better than younger children ($4\frac{1}{2}$- and $5\frac{1}{2}$-year-olds) at long interstimulus intervals, but worse than this group at short intervals. This suggests that Tallal & Piercy's results cannot be totally explained away as an artifact of task difficulty.

Surprisingly few answers have come out of the research on developmental language disorders. The weakest hypotheses are those proposing memory deficits or general symbolic disability as the primary disorder, since there is evidence that language disordered children have adequate abilities in both of these areas. There is insufficient evidence to enable us to decide between the other hypotheses discussed above. It remains uncertain whether the primary disorder is perceptual or cognitive, auditory or multimodal.

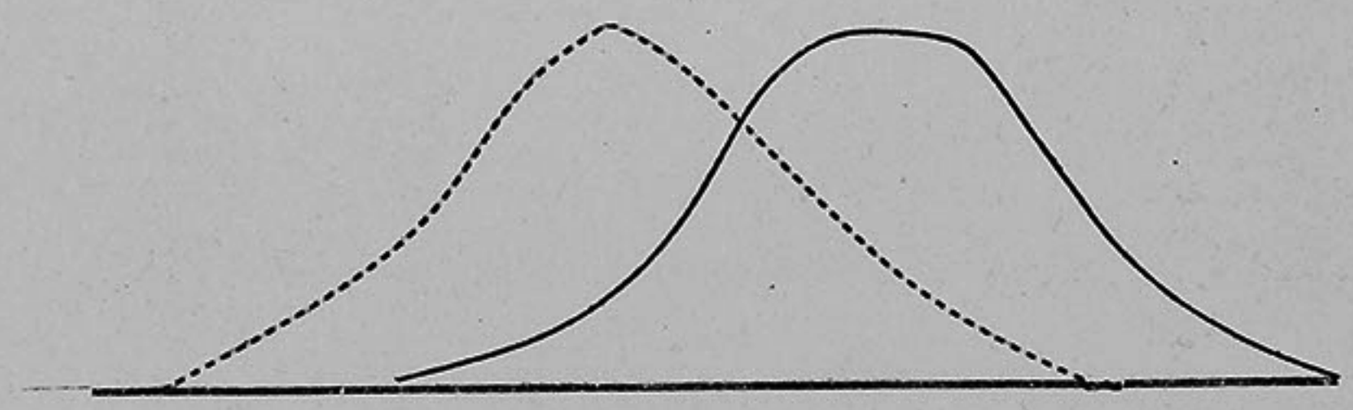
Fig. 23

Expected distributions of comprehension scores

A. QUALITATIVE MODEL



B. QUANTITATIVE MODEL



—— normal

-----language disordered

Chapter 6: Experimental study: language disordered children

Introduction

Language comprehension in language disordered children was investigated using TROG, with a particular interest in the following questions:

1. Does the distinction between receptive and expressive disorders reflect a qualitative or quantitative difference?

As discussed in chapter 5, the conventional distinction made between expressive and receptive disorders implies these are two qualitatively distinct categories, but some writers suggest that the distinction is a matter of degree. Eisenson (1968) makes the stronger claim that most so-called expressive disorders have their basis in comprehension problems. This claim can be tested by examining performance of a mixed group of language disordered children on comprehension tests. If Eisenson is correct, we would expect nearly all children to have comprehension difficulties of a greater or lesser degree, and on a standard scale of language comprehension, we would predict that there would be a unimodal distribution of scores, shifted to the left of the normal distribution. If there is a true qualitative difference between expressive and receptive disorders, we would predict that the distribution of standard comprehension scores would be bimodal, with children with expressive disorders falling in the normal range and those with receptive disorders falling below the normal range (see figure 23). Questions about the classification of developmental language disorders are not just of theoretical interest. They have important implications for our understanding of etiology. If we demonstrate that there is only a quantitative difference

between expressive and receptive disorders then it does not necessarily follow that there is a unitary etiology, but on the other hand, if we can demonstrate discontinuity between these disorders at the behavioural level, then it is likely that they have distinct etiologies.

2. <u>What is the relationship between expressive and receptive language?</u>

It is known that comprehension is not dependent on expressive speech (Lenneberg, 1962; Fourcin, 1975) and it is assumed by many workers that comprehension is normal in children with expressive language disorders (e.g. McGinnis, 1963; Morley, 1972). However, Eisenson (1968) predicts not only that most language disordered children have comprehension difficulties but also that expressive and receptive language correlate highly in these children. If we demonstrated that a large number of language disordered children had expressive difficulties that were not mirrored by comprehension problems, then Eisenson's hypothesis would be disconfirmed. Once again, this question is not only of theoretical interest. A child whose problem is purely one of output is likely to require different forms of therapy from one who has a difficulty with both comprehension and production.

3. <u>Within the context of receptive language, are grammatical and vocabulary development in keeping with one another?</u>

This is but one aspect of the delay versus deviance question discussed in chapter 5. The answer will have important implications for psychological theories of developmental language disorders. Theorists such as Cromer (1977) who regard developmental language disorders as disorders of hierarchical structuring would predict that

the grammatical aspects of language should pose greater difficulty than vocabulary, whereas those proposing more general perceptual disturbances would not predict such specific impairments.

4. Do some grammatical structures give disproportionate difficulty?

This is another aspect of the delay versus deviance question. Here we are interested in whether language disordered children's comprehension of grammar resembles that of normal children at some particular level of development, or whether it has an abnormal pattern not found in normal children at any stage of development. According to Cromer's hypothesis we would predict that structures involving complex transformations should give disproportionate difficulty. On the other hand, if sentence length proved more important than grammatical complexity in determining performance, then short-term memory might seem to be implicated.

These questions were investigated by comparing the performance of a group of language disordered children with the normal group described in chapter 4.

Materials

The three tests used in the second test session with the normal group, TROG, PPVT and Picture Completion (see pp. 42-50) were given to all subjects except for 6 subjects who had been given PPVT in the previous 6 months for whom this test was not readministered.

Subjects

The normal sample from the second test session were used as a comparison control group (see p. 52).

The language disordered group consisted of all available pupils (n = 76) from two residential schools for children with developmental language disorders. Children are admitted to the schools on the basis of their having a severe language disorder not attributable to deafness, mental subnormality or emotional disturbance. Both schools use similar teaching techniques, with an intensive structured attack on language. The Paget-Gorman Sign System (Paget, Gorman & Paget, 1972) is used in conjunction with, rather than as an alternative to, spoken language. Available background information on these children was of variable quality. The schools do not have the facilities to administer detailed neurological, psychological and audiometric examinations, and rely on referring sources to supply such data. Speech therapists on the staff however do give comprehensive language assessments. Both schools utilized the distinction between expressive and receptive disorders to allocate children to classes, without necessarily assuming these are clearcut categories. Such allocation was made on the basis of all available information on the child. Despite the stringent selection criteria of the schools, three children had been found to have hearing losses of 80 db or more. These children were tested, but their data were excluded from the analyses discussed here.

Method

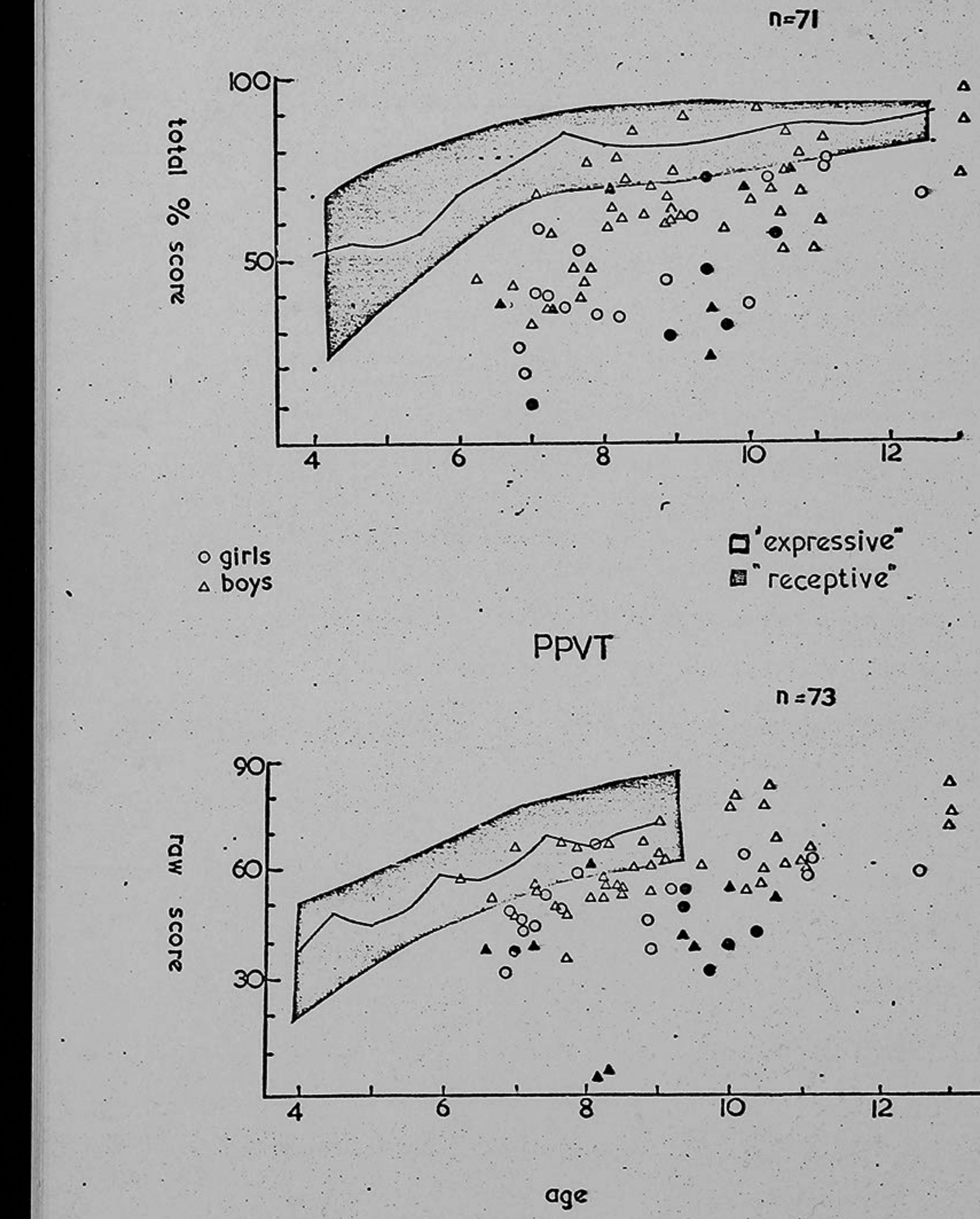
This was essentially the same as for the normal children (see pp. 52 to 54).

Results and discussion

Two children made so many errors on the vocabulary pretest that calculation of %-scores on TROG would have been invalid. For the

Fig. 24

Language comprehension: experimental group in relation to controls



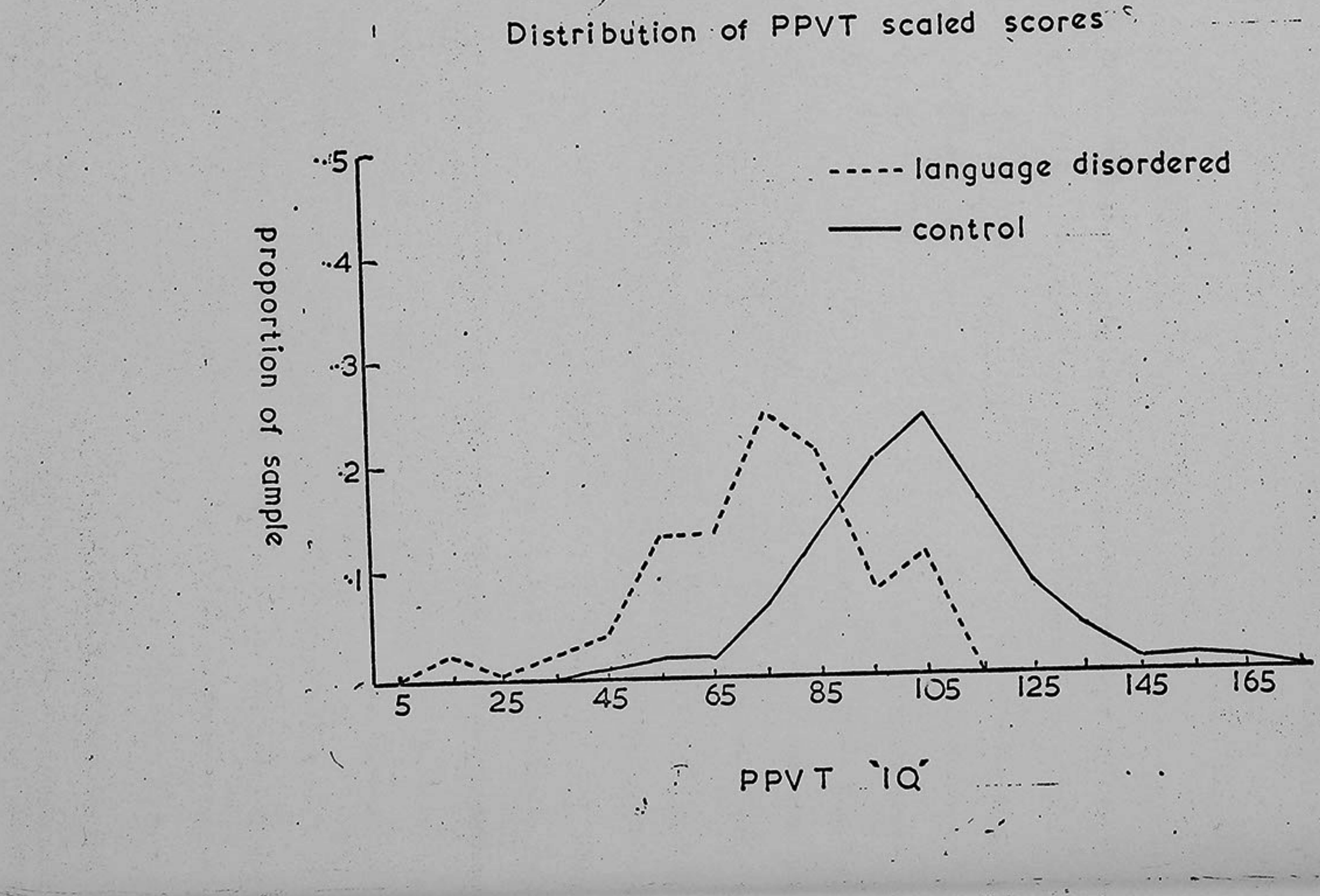
remainder of the sample, %-scores were calculated for each grammatical scale and for the test as a whole, by the procedure described in appendix 3. Results will be presented and discussed with reference to the questions outlined above.

1. Receptive/expressive distinction

Figure 24 shows the total %-scores of language disordered children plotted in relation to the control group. Median scores of the controls are shown as a bold line surrounded by a shaded area denoting the smoothed 10th and 90th percentile limits. The experimental group are plotted as points. It can be seen that there is considerable overlap between children classified by the schools as having receptive disorders (henceforth "receptives") and those with expressive disorders ("expressives"), and that most "expressives" perform below age level. However there is a significant difference between "expressives" and "receptives" after controlling for age and nonverbal (picture completion) score (t-test on matched pairs: $t = 2.73$; $d.f = 12$; $p<.01$). Figure 24 also shows performance on the PPVT relative to the control group. It can be seen that receptive vocabulary separates the two groups better than does TROG (t-test on matched pairs: $t = 4.62$; $d.f = 14$; $p<.001$), although the "expressives" still perform below age level.

This sample is too small to draw more than tentative conclusions from the shape of the distribution of scores for the language disordered sample. It did not seem valid to calculate z-scores on TROG from the mean and standard deviation values obtained with the normative group, as the number of children in each age group was small, and such estimates would have been very unreliable (see also footnote on p. 82). However, it did seem feasible to use scaled

Fig. 25


Distribution of PPVT scaled scores
----- language disordered
—— control
proportion of sample
.5
.4
.3
.2
.1
5
25
45
65
85
105
125
145
165
PPVT 'IQ'

("IQ") scores from the PPVT to examine the distribution of vocabulary comprehension scores in normal and language-disordered groups. These distributions are shown in figure 25 and may be compared with the theoretical distributions of figure 23. The most obvious feature of figure 25 is that the language-disordered distribution is very similar to the normal one but shifted to the left, closely resembling that predicted on the basis of a quantitative model. The significance of the small hump at the low range of PPVT IQ's for the language disordered group is uncertain. This hump is due to two boys with receptive problems who resemble the classical case of congenital auditary imperception described by Worster-Drought & Allen (1929). These were the two children for whom TROG scores could not be calculated because of limited vocabulary. One child demonstrated a fair degree of grammatical comprehension when TROG was readministered using the Paget Gorman Sign System, whereas the other did poorly with signed instructions. It is possible that there is a true qualitative difference between these rare cases of congenital auditory imperception and other types of developmental language disorder, but firm conclusions cannot be reached on the basis of two cases. With this possible exception, the data clearly support a quantitative model of developmental language disorders which holds that the traditional distinction between expressive and receptive types is a matter of degree. Further, the data support Einsenson's contention that the majority of children with expressive disorders do have comprehension problems[1] (although it remains to be demonstrated that the receptive disorder is primary).

[1]Language disordered children also performed significantly more poorly than controls on the nonverbal Picture Completion task, suggesting that their poor comprehension scores might reflect low general ability rather than a specific linguistic problem. However, it will be demonstrated below that the difference still holds when language disordered children are compared with normal children matched on Picture Completion scores.

Table 14

Comprehension scores: sex differences in distribution for language disordered children

TROG

overall score	♂	♀	
⩽ 50%	11	12	$\chi^2 = 7.14$
> 50%	38	10	$P < .01$

PPVT

raw score	♂	♀	
⩽ 50	9	13	$\chi^2 = 12.16$
> 50	41	9	$P < .01$

However, the picture is complicated by sex differences. In figure 24 males are denoted by triangles and females by circles. Ignoring the division into "receptive" and "expressive" groups, we may reclassify children according to TROG performance, setting an arbitrary cutoff at 50% correct. The proportions falling above and below cutoff are different for the two sexes. The same applies to the PPVT, where a cutoff of raw score 50 produces an almost identical picture (see table 14). These results are not artifacts of differences in age or nonverbal ability which do not differ significantly for the two sexes (see table 15). The sex difference is also evident for the PPVT if standard ("IQ") rather than raw scores are used. This confirms objectively the observation that for expressive disorders boys outnumber girls, whereas for receptive disorders the sex ratio is equal (Rutter, Bartak & Newman, 1971) and conforms to the pattern described by Taylor & Ounsted (1972):

> "The apparent paradox that the sex least frequently affected by a disorder is the one relatively more severely affected....is an unexplained rule common throughout medicine."
>
> p. 220

The qualitative model has no difficulty with this result - if we regard expressive and receptive forms as different disorders then we are not surprised to find different sex ratios. However, although more difficult to explain on the basis of a quantitative model, this result is similar to that found in several other conditions which would appear to be quantitative in character (see Taylor & Ounsted, 1972). Taylor & Ounsted do propose an hypothesis to account for this type of sex difference. They draw an analogy with height, and point out that a female who is 6 ft 2 in. tall is more likely to be suffering from some syndrome of gigantism than is a male of that

Table 15

Language disordered children: mean age and Picture Completion scores*

	boys n = 51	girls n = 22
age	9-1 (1-8)	8-10 (1-8)
Picture Completion raw score	25.32 (6.15)	24.07 (6.73)

*standard deviations in brackets

height, although males of that height are more numerous. Further, a male of that height who does suffer from gigantism would be less severely affected than a gigantic female of the same height. However, the force of this example depends on the different means for males and females in the distribution of height. There is no such difference for language comprehension. Taylor & Ounsted attempt to remedy such objections by proposing that males develop more slowly than females, so that for a given characteristic (in this case language development) there is greater variance in the distribution for boys than for girls, so relatively more males are represented at the extreme. The simplest interpretation of this account is shown in figure 26. (Taylor & Ounsted point out that the underlying distributions need not be normal and need not share the same mean). Beyond a certain threshold, (t) individuals may be regarded as being affected with the disorder in question, and:

> "although boys will be more numerous than girls, the girls will have required a relatively greater degree of divergence from their norm and hence will be the 'worse' affected."

Taylor & Ounsted (1972) p. 232

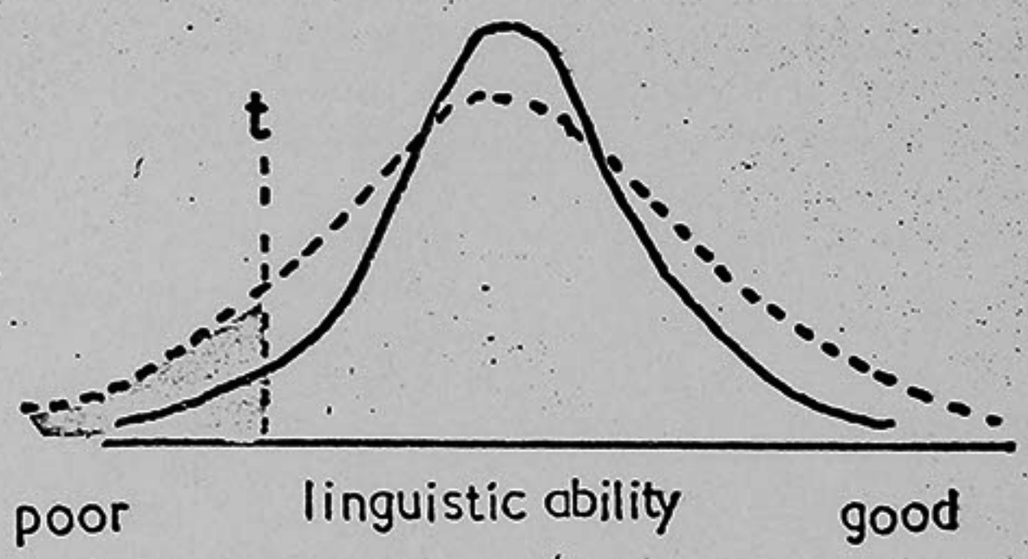
i.e. if one calculated standard scores separately for boys and girls, a girl falling below threshold will have a more extreme standard score than a boy, even though her raw score may be less extreme with respect to the total population. However, this account cannot accommodate the result found here where girls had more severe comprehension problems relative not only to other girls, but to all children.

A model which handles this result better has been proposed by Reich, Cloninger & Guze (1975) and Cloninger, Reich & Guze (1975) to

Fig 26

Models of sex differences in incidence of disorders

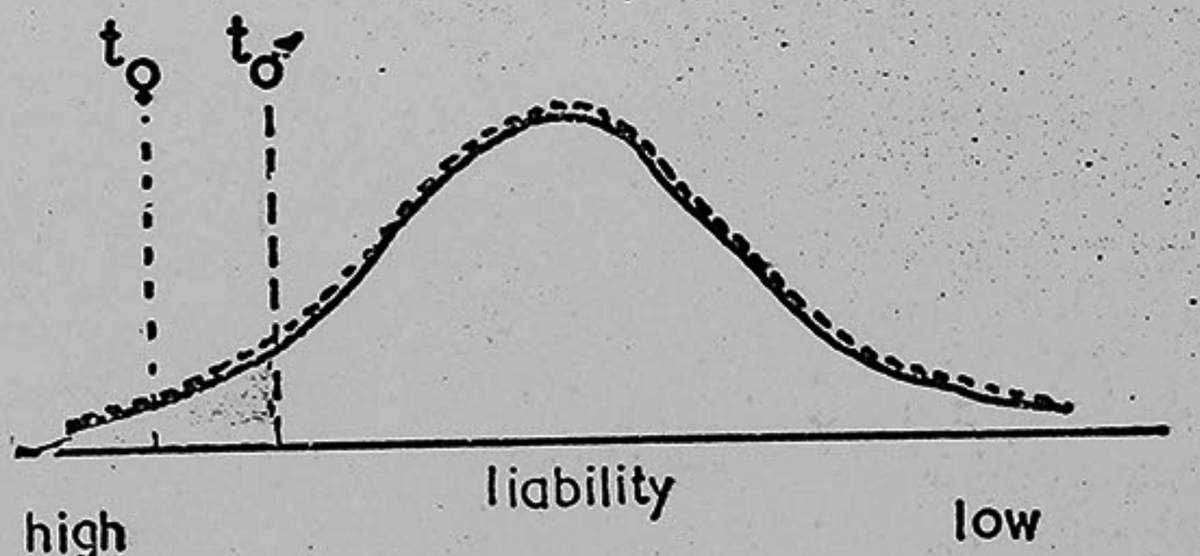
TAYLOR & OUNSTED



t: threshold

affected persons

REICH, CLONINGER & GUZE



___girls

----boys

account for sex differences in incidence and severity of sociopathy. This model may be applied to other disorders which can be viewed as semi-continuous variables, i.e. disorders where affected individuals can be graded, but unaffected individuals cannot. It is proposed that all individuals in a population can be regarded as falling on a continuum of "liability to develop the disorder". This liability cannot be directly measured, and is assumed to be the sum of genetic and environmental influences affecting the disorder. When liability exceeds a certain threshold, the disorder is manifest. It is suggested that in disorders where the incidence and severity differ between sexes, it can be assumed that the underlying liability distributions for the two sexes are congruent, but that the thresholds differ (see figure 26). The more extreme threshold of females means that affected females will be less frequent but more deviant than affected males. This model offers more than a parsimonious description of the observed situation - it makes clearcut preductions about the incidence of the disorder in relatives of affected individuals. In particular, the incidence of the disorder in the relatives of affected females should be higher than the incidence in the relatives of affected males.

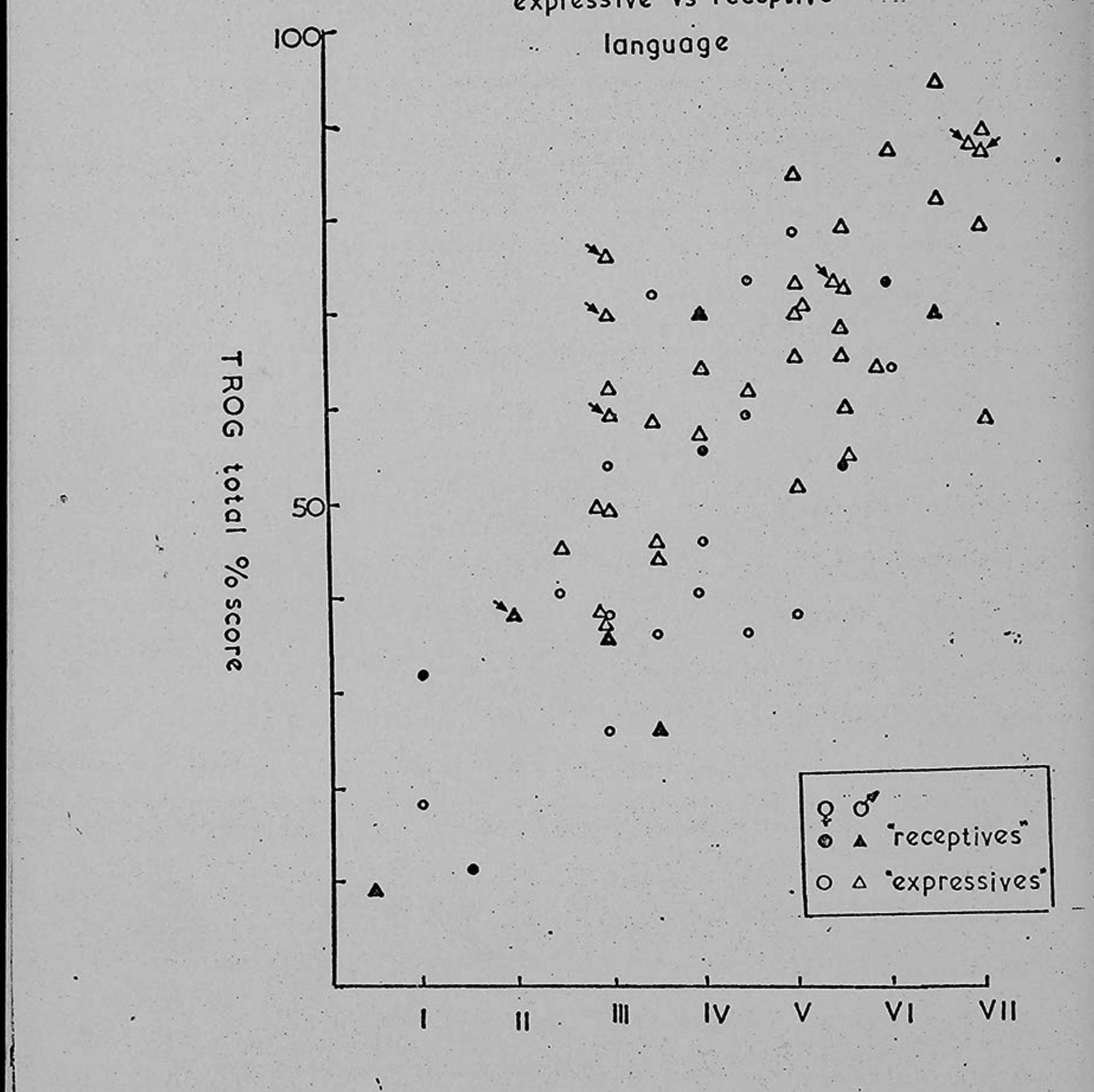
2. Relationship between expressive and receptive language

Estimates of grammatical complexity of expressive speech obtained by Crystal, Fletcher & Garman's (1976) Language Assessment, Remediation and Screening Procedure (LARSP) were available for 63 language disordered children[1]. LARSP estimates are obtained from both

[1]These data were kindly provided by the school staff. They had been collected by experienced speech therapists trained in the use of LARSP. TROG scores were not available at the time of LARSP analysis, and LARSP estimates were not known to the author at the time of TROG administration

Fig. 27

GRAMMATICAL DEVELOPMENT

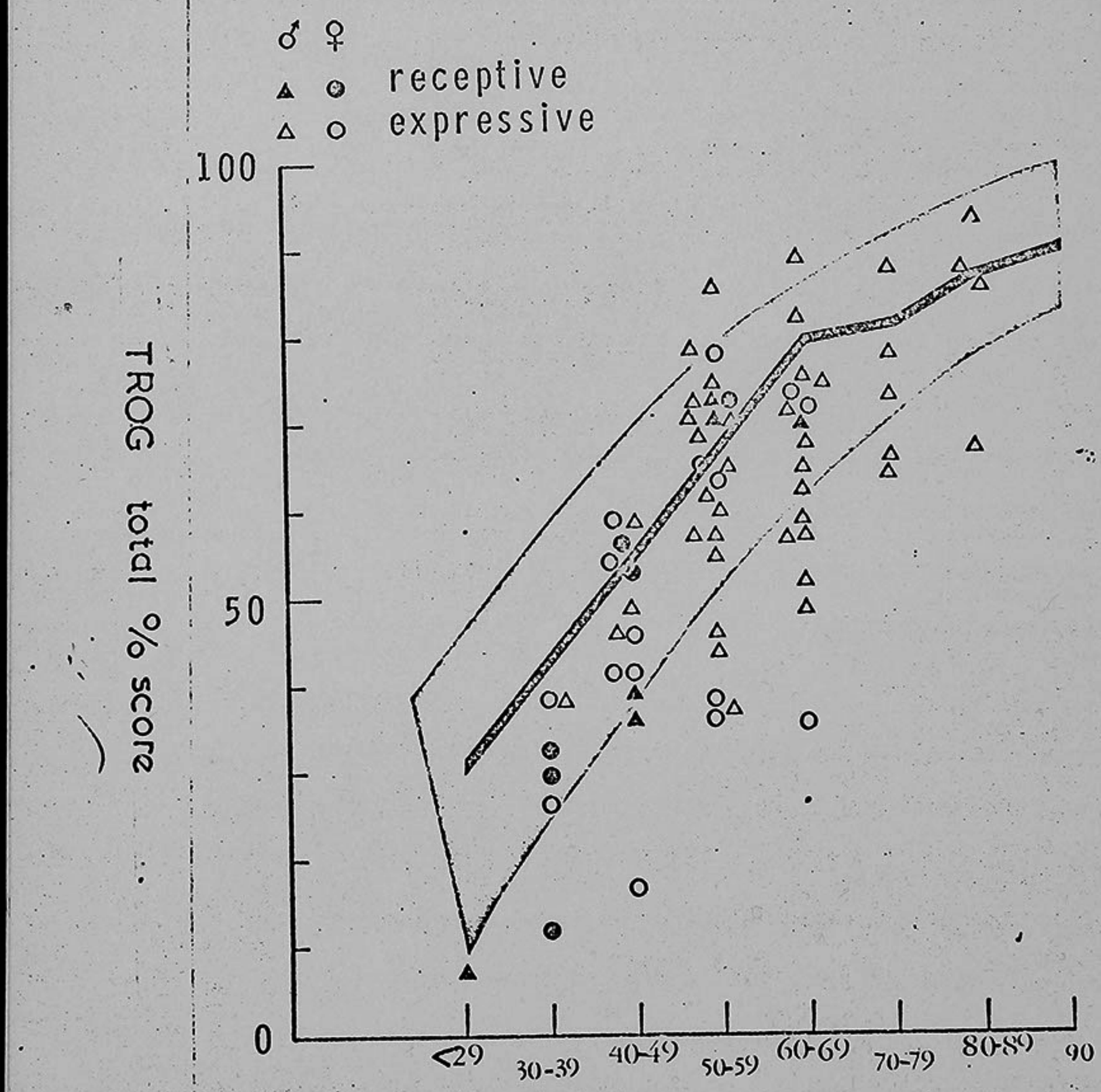


spontaneous speech samples and elicited imitations, on the basis of which expressive speech is allocated into one of seven developmental stages. In practice, children were often allocated to intermediate stages, so that LARSP estimates fell on an ordinate scale with 12 points. Where spontaneous and imitated samples gave different estimates, the higher estimate was used here. The relationship between LARSP stage and TROG total %-scores is shown in figure 27. Spearman rank order correlation between these measures was .73 which is highly significant ($p<.001$) and supports the hypothesis that a strong relationship exists between grammatical complexity of expressive and receptive language. Seven children who were reported to have motor or structural abnormalities of the speech apparatus are indicated by arrows in figure 27. These children would be predicted to be most likely to have good comprehension relative to their expressive speech. In fact, although inspection of figure 27 suggests a trend in this direction, TROG scores of these children were not significantly different from other language disordered children who were individually matched with them on sex and LARSP stage (selected at random from the subset of suitable children). (matched pairs t-test; $t = 1.44$; $d.f. = 6$).

In sum, these results seem in general to support Eisenson's hypothesis. Not only do "expressive" children have comprehension problems, but the degree of grammatical complexity in their expressive speech is related to ability to understand grammatical structures. It would be interesting to go on to test motor function of the speech apparatus in these children using objective tests, to see whether there was a tendency for those with poor expressive speech in relation to their comprehension to be dyspraxic.

Fig. 28

TROG scores in relation to PPVT



3. Relation between receptive grammar and vocabulary

In order to answer this question, normative data on TROG were replotted with respect to raw PPVT score rather than age. These data are shown in figure 28 where once again, median scores of control children are plotted as a bold line surrounded by a shaded area representing 10th and 90th percentile limits, and the experimental data are plotted as points. 70% of the experimental sample have TROG scores falling within the range expected on the basis of their PPVT score. Virtually all of the remaining children have grammatical scores below that predicted on the basis of vocabulary. The converse is very rare. Thus the answer to the question posed is equivocal. For the majority of children comprehension of grammar and vocabulary are at a similar level, but a minority of children appear to have a more specific problem with grammatical comprehension.

4. Pattern of performance within TROG

It was noted above that the experimental sample did more poorly than the control group on the nonverbal Picture Completion test as well as on the linguistic measures. Before proceeding with more detailed comparisons it was desirable to control for this factor. It will be remembered that Picture Completion and PPVT data were not available for the control children aged over $9\frac{1}{2}$ years. Therefore the following procedure was adopted. The experimental group was divided into those aged from 6-0 to 9-5 years and those aged over 9-5. For each of these groups, two new matched samples were selected from the main control group, control 1 being matched on the basis of age, and control 2 on the basis of PPVT raw score (see table 16 for further details). Two-way analysis of variance (with groups

Table 16

Criteria for selection of control groups

	young group	old group
experimental	language-disordered children aged 6-0 to 9-5, for whom both types of control could be selected, excluding those with hearing losses of 30 db or more in best ear (n = 30)	all language-disordered children aged 9-6 and over except those with Picture Completion scaled scores of 6 or less, and those with hearing losses of 30 db or more in best ear (n = 20)
control 1	individually matched with experimental group on age (within 6 months), sex, and Picture Completion raw score (within 2 points)	individually matched with experimental group on age (within 6 months and sex
control 2	individually matched with experimental group on PPVT raw score (within 4 points)	individually matched with experimental group on PPVT raw score (within 4 points)

Table 17

Analysis of variance: summary tables

A. Children aged 6-0 to 9-5

source*	d.f.	s.s	m.s.	f
A	2	32861	16430	19.51
between subjects	88	74102	842	
B	3	123605	41201	182.13
A x B	6	58881	9813	43.38
within groups	264	59724	226	

B. Children aged over 9-5

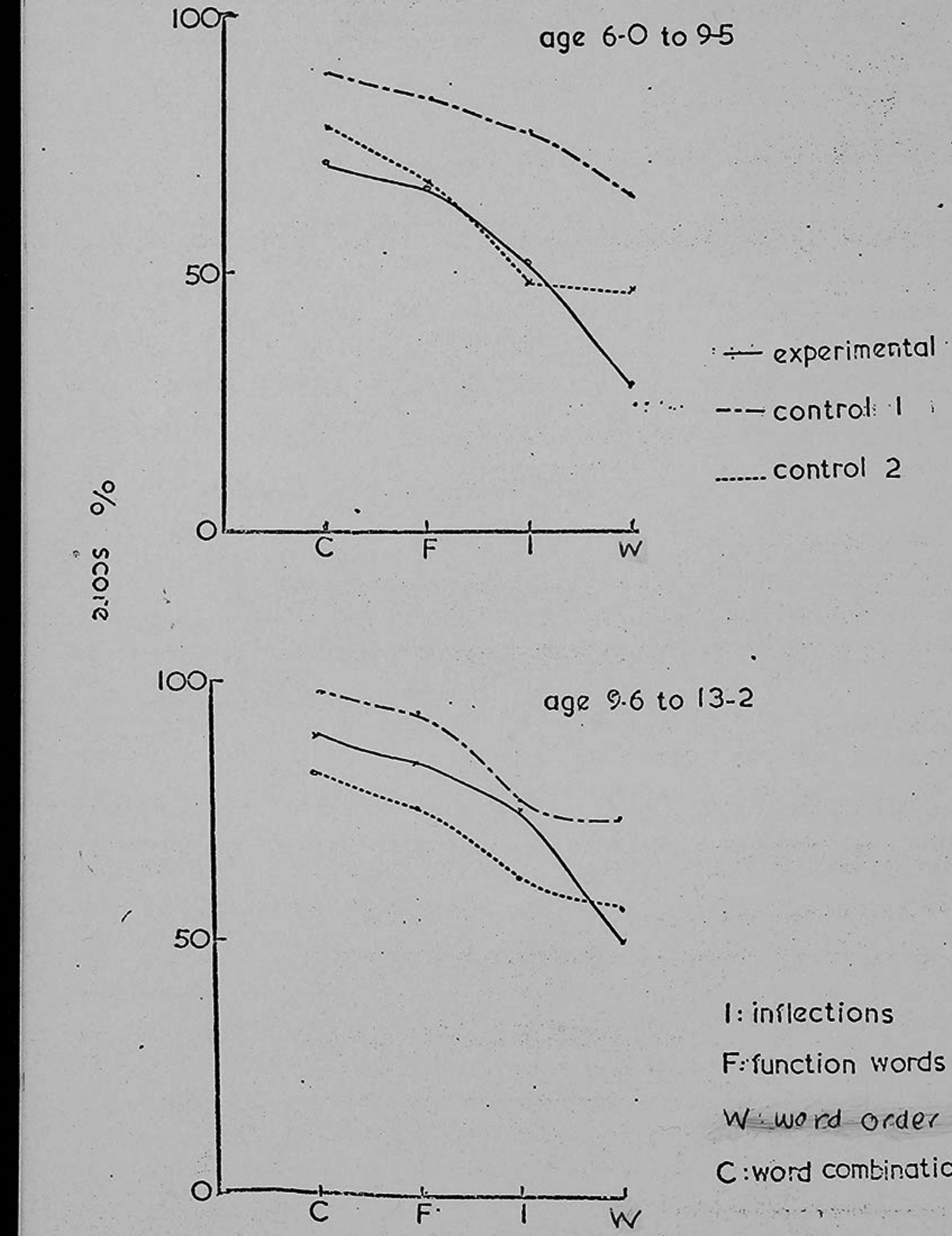
source	d.f	s.s	m.s.	f
A	2	12221	6111	13.50
between subjects	57	25796	452	
B	3	86753	28918	150.88
A x B	6	48072	8012	41.80
within groups	171	32774	192	

* A: language disordered vs control 1 vs control 2

B: Inflections vs Function Words vs Word Combinations vs Word Order

Fig. 29

GRAMMATICAL SCALES



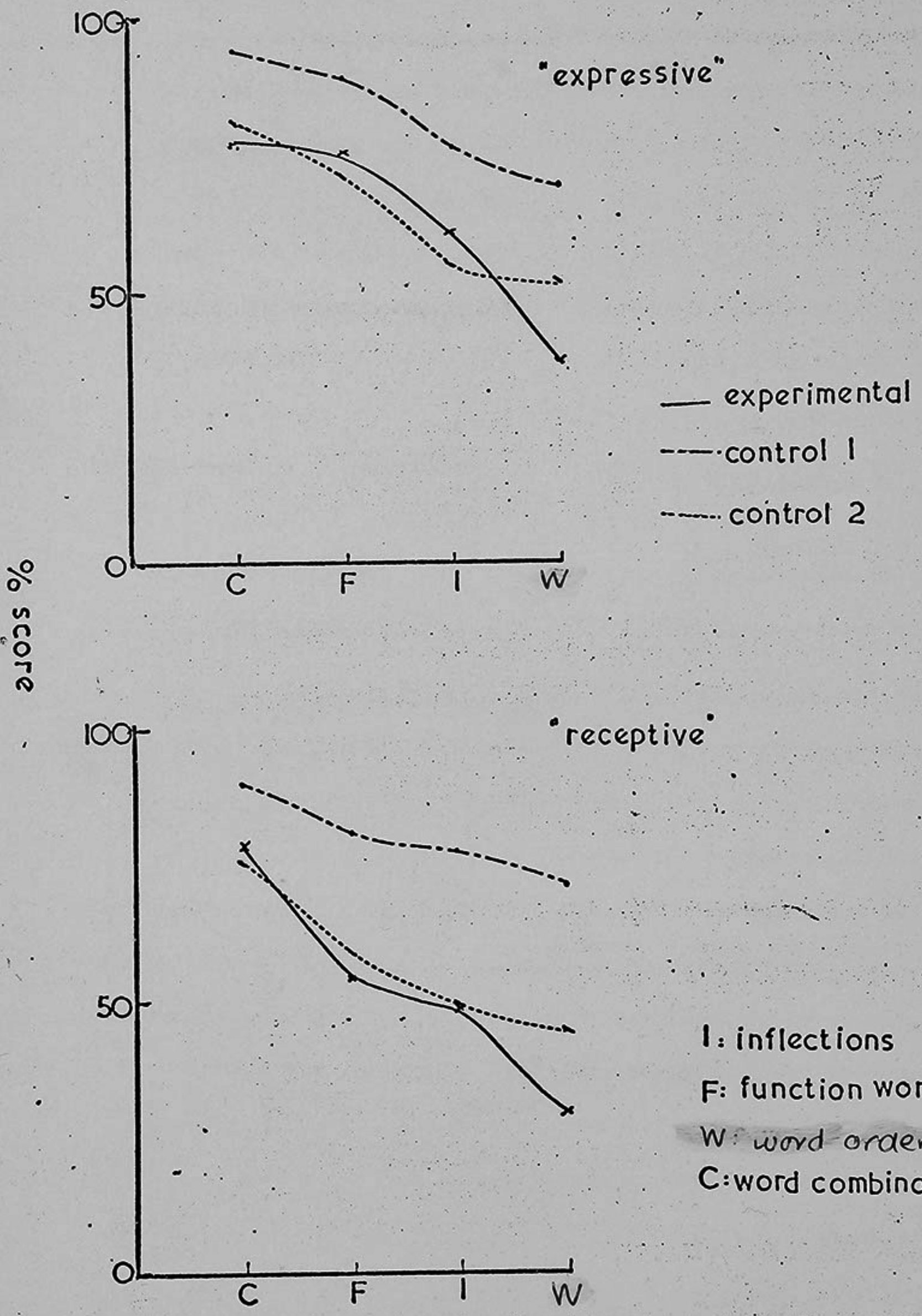
as between subjects main factor and grammatical scores as within subjects main factor) was carried out for the old and young groups separately. Significant effects (.01 level) were found for both main factors and for the interaction between them (see table 17 and figure 29). A Tukey test of specific comparisons showed that for the younger group the language disordered did more poorly than control 1 (age-matched) on all subtests, and more poorly than control 2 (vocabulary-matched) on Word Order only. The older experimental group were not significantly different from either control group except on Reversible Structures where they did significantly more poorly than control 1. Alternative analyses of variance where "receptives" and "expressives" were considered separately (summing across age groups) showed an essentially similar pattern for these two groups (see figure 30).

The particular difficulty with Word Order could be accommodated by theories stressing impairment in sequential processing or by those arguing for a more specifically linguistic disorder in hierarchical structuring. If sequencing of words in general were affected, we would expect difficulty with all types of Word Order structure. If the problem lay in assigning the correct deep structure to sentences, we might predict that grammatically simple sentences and phrases, such as prepositions + word order and reversible actives, would give no particular problem, whereas passive and embedded sentences, which have a more complex transformational history, and where surface structure does not correspond to deep structure in a simple way, should be especially difficult.

Since in this comparison we were more interested in the pattern of performance than the general level of performance, it was decided

Fig. 30

"Expressive" and "receptive" groups



to select a control group who were individually matched with the language disordered group on total %-score on TROG (i.e. the controls were younger than the experimental group). It was possible to select such a control (matched within 4 points) for all but five experimental subjects. Children were scored as passing or failing contrasts according to the same criteria as were used in the first analysis with normal children (see p. 56), i.e. a pass was awarded if all items testing a given contrast were correct, a fail otherwise. Thus for each pair of subjects and for each grammatical contrast four outcomes were possible:

1. Control child fails, experimental child fails
2. Control child passes, experimental child passes
3. Control child fails, experimental child passes
4. Control child passes, experimental child fails

Our null hypothesis is that the different contrasts should be of the same degree of difficulty for experimental and control groups, i.e. for each structure, outcomes 3 and 4 should be equally likely to occur. Observed frequencies of all 4 outcomes are shown for all contrasts in table 18. It can be seen that very easy contrasts (e.g. comparative/superlative) and very difficult ones (e.g. embedded sentences) do not discriminate well between groups, outcomes 3 and 4 being very rare. However, interesting results emerge from other contrasts where these outcomes are more common.

Testing the significance of the observed results is hampered by the fact that the results for different structures cannot be regarded as independent. It can be argued that since experimental and control members of a pair are matched on total TROG %-scores, poor performance on one structure is likely to be balanced by good

Table 18

Outcomes for matched pairs on grammatical contrasts

	number of pairs with a given outcome			
outcome*	1	2	3	4
comparative/superlative	0	60	3	2
negative	1	58	5	6
singular/plural nouns	16	30	14	5
pronouns	30	12	20	3
prepositions (standard)	46	4	8	17
prepositions (+ word order)	23	11	15	16
reversible actives	22	24	6	13
reversible passives	46	3	0	16
tense	53	1	11	0
embedded	57	2	3	3
singular/plural verbs	58	2	3	2
2-word combinations	2	46	4	13
3-word combinations (active)	6	46	6	7
3-word combinations (passive)	20	17	10	13

* 1: control fail; experimental fail

2: control pass; experimental pass

3: control fail; experimental pass

4: control pass; experimental fail

performance on another. The dependence is not total however, since total %-scores and number of structures passed need not be highly correlated. Consider for example 2 subjects who make a total of 14 errors on TROG. At one extreme we may find a subject who makes one error on each grammatical category, so is scored as passing none. At the other extreme, we may have a subject whose errors are concentrated exclusively on embedded sentences, reversible passives and tense items, so is credited with passing 11 categories.[1]

In view of these considerations, it was decided to use the Binomial Theorem to calculate the probability (2-tailed) of frequencies as extreme as those observed for outcomes 3 and 4, adopting a significance level of .001, to guard against type 1 errors.

By this criterion, differences in frequency of outcomes 3 and 4 differ significantly from those predicted on the null hypothesis for pronouns ($p = .0004$), tense ($p = .001$) and reversible passives ($p = .00004$). Language disordered children are more likely than controls to pass pronouns and tense items, and more likely to fail reversible passives.

Do these results enable us to choose between theories? The result on tense is not predicted by any theory. It is interesting to note that the result on pronouns is consistent with one study on the expressive speech of language disordered children, who were found to produce more pronouns than younger controls at a similar linguistic level (Morehead & Ingram, 1973).

The result on reversible passives, while by no means conclusive, does lend more support to Cromer's (1977) hypothesis than to a

[1] In fact, for the language disordered sample, Spearman correlation coefficient for number of structures passed and total %-score was .392 ($n = 71$).

sequencing theory. It will be remembered that we predicted on the basis of the sequencing theory that all word order contrasts should be difficult for language disordered children, whereas Cromer's hypothesis predicts particular difficulty with grammatically complex reversible passives and embedded sentences. Embedded sentences did not discriminate between groups in this case, but this was because virtually all children failed these. However, prepositions + word order, reversible actives and reversible passives were more discriminating, yet only on reversible passives did the experimental group do much worse than the controls.

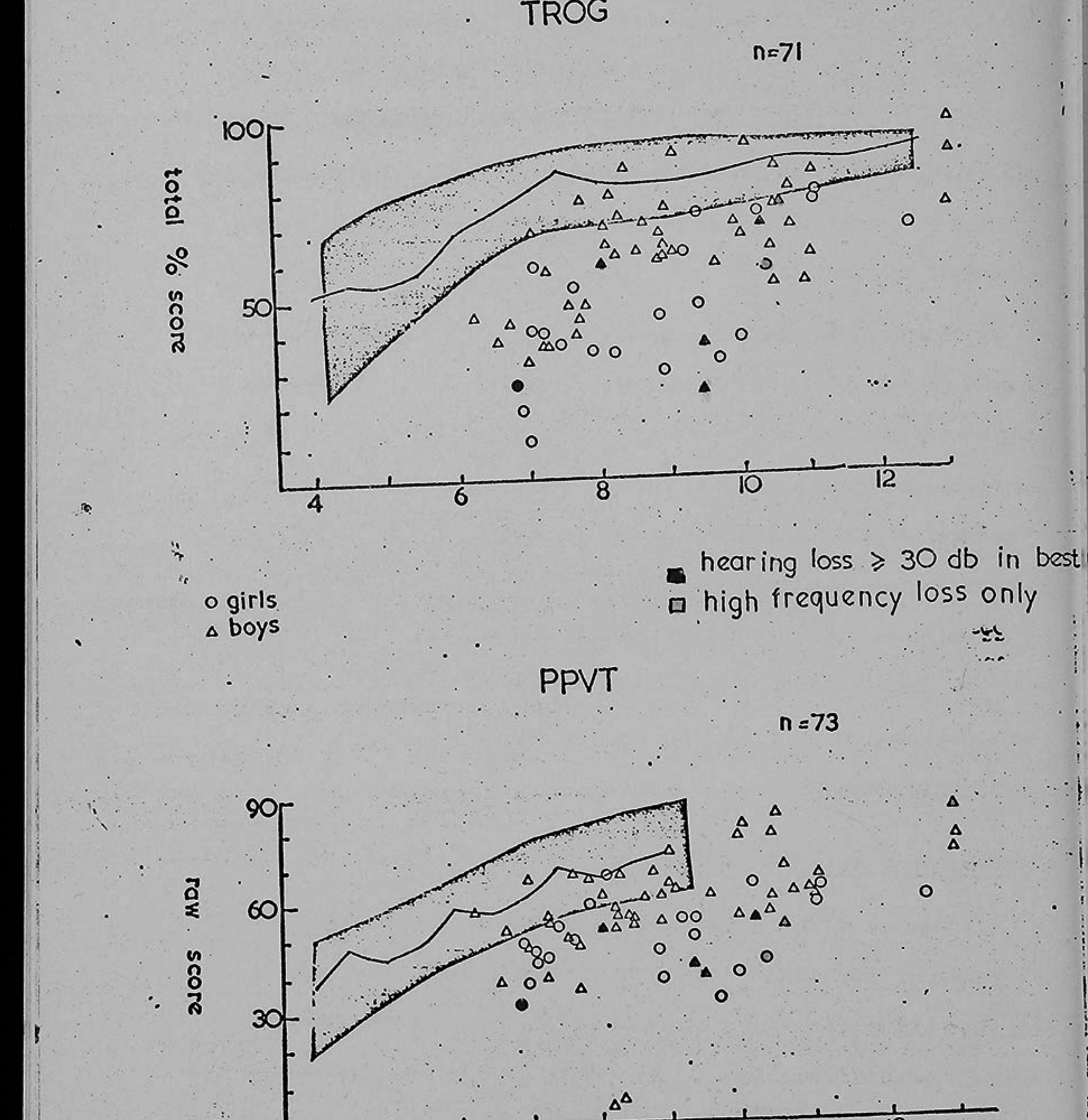
However, in interpreting these results, a caution is necessary. It must be remembered that language disordered children were undergoing intensive language therapy. One school uses a system based on that outlined by Conn (1971), which tackles syntax in a highly structured way, avoiding complicated structures in the early stages. The difficulty that these children had with passives, and their relatively good performance on pronouns could simply reflect the fact that the former are not taught until children have acquired a reasonable degree of linguistic proficiency whilst the latter are taught from the start.

Comprehension performance and background variables

As mentioned above, the background information on these children was gathered from school records containing reports from various sources, and therefore was of variable quality. Nevertheless it was felt that it would be of interest to consider comprehension scores in relation to some background variables, so long as it is remembered that such analysis can only raise tentative suggestions

Fig. 31

Relation of hearing loss to comprehension scores



for further research rather than being in any way conclusive.

hearing loss

Figure 31 shows the incidence of recorded hearing loss in the sample. Only five children had losses of moderate severity (greater than 30 db) and all could hear at least some frequencies at 50 db. Two children with hearing losses greater than 30 db and two with smaller losses used hearing aids. The five children with hearing losses of 30 db or more were individually matched with five language disordered children according to age (within 5 months) and Picture Completion raw score (within 2 points), the controls being selected at random from the pool of suitable children. T-tests on matched pairs indicated that the group with hearing loss did significantly more poorly than those with normal hearing on PPVT ($t = 2.77$; $d.f = 4$; $p<.05$) but not on TROG ($t = 0.92$; $d.f = 4$; n.s.).

positive family history

Figure 32 shows those children recorded as having relatives with difficulties in speech (including stammering) or reading. Data from two adoptees and one fostered child are omitted. 21.4% of boys and 22% of girls had mention of a positive family history in the records, a result which does not lend support to application of Cloninger, Reich & Guze's (1975) model of sex differences to language disorders (see p. 149). Inspection of figure 32 does not suggest that there is a significant association between positive family history and comprehension test performance, and this impression was verified by t-tests contrasting those with and without such histories.

late onset

Three girls and one boy were reported as developing language disorders after a period of normal language development. In two of these children, onset was thought to be associated with routine

Fig. 32

Positive family history in relation to comprehension scores

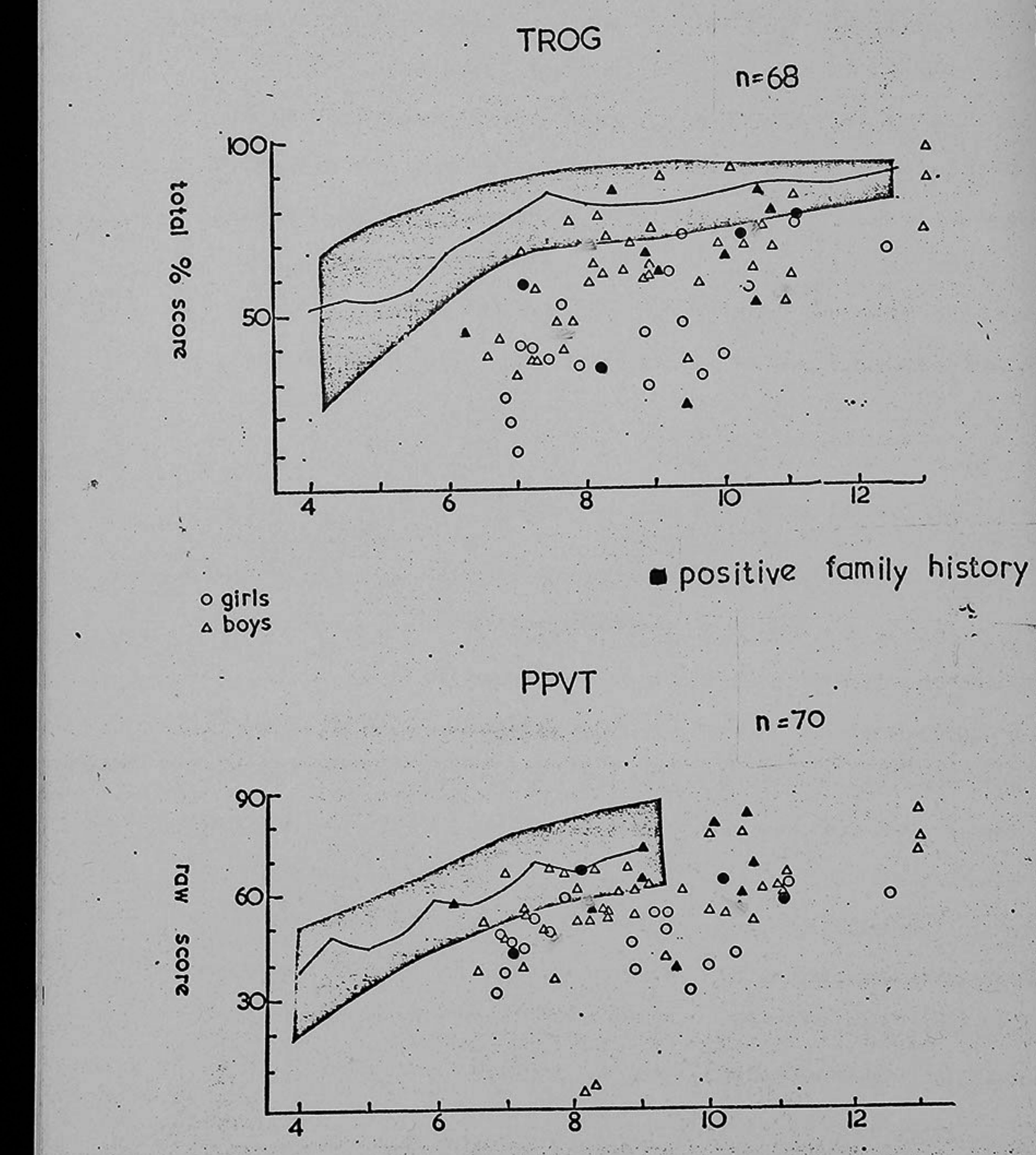


Fig 33

Children with late onset: comprehension scores

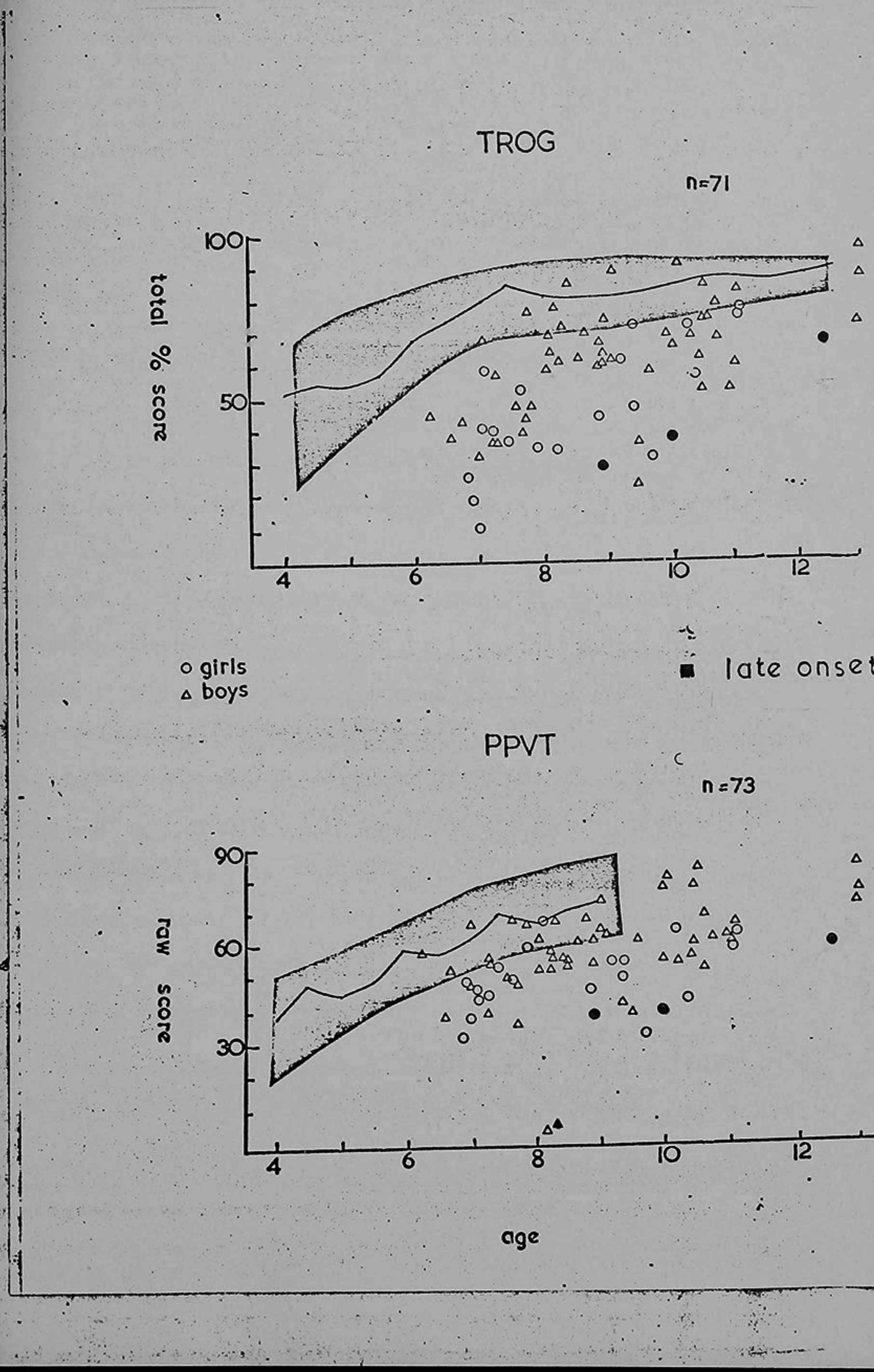


Table 19

Neurological abnormality: allocation of children to categories

category A

neurological signs: n = 4

1. cerebellar ataxia
2. suprabulbar palsy and right arm paresis
3. minimal left hemiplegia; hypotonic tongue
4. slight spasticity both legs; squint; poor control of tongue

category B

other abnormalities: n = 6

1. abnormal EEG. Described as brain-damaged at birth
2. convulsions until age 4 following measles at 9 months
3. intracranial haemmorhage as baby
4. bilateral EEG abnormality
5. abnormal EEG
6. abnormal EEG

category C

suggestion of abnormality: n = 9

1. hypoglycaemia and convulsions (neonatal)
2. "blue baby"; 3 months premature
3. severe toxaemic pregnancy; afterbirth preceded baby
4. toxaemic pregnancy; breech birth; jaundiced
5. difficult birth; anoxia; "brain damage"
6. anaemic birth; neonatal fits
7. ptosis; gross motor difficulties
8. premature baby; caesarian after haemmorhage; stopped breathing; squint
9. slow to breathe on birth; went black; lost consciousness aged 3 weeks

category D

neurologically normal: n = 54

no mention of anything abnormal: n = 41

abnormalities of uncertain significance (includes premature; threatened miscarriage; "difficult birth"; toxaemic pregnancy): n = 13

vaccination. In the others, speech was described as fading away. In none of these cases were fever or convulsions at time of onset specifically mentioned, nor were any abnormal EEGs reported.

In chapter 5 it was noted that children with late onset tended to have receptive problems. Tentative support for this view comes from this study, as can be seen from figure 33 although the numbers are too small for meaningful statistical analysis. (One of the children with late onset was untestable on TROG because of limited vocabulary.)

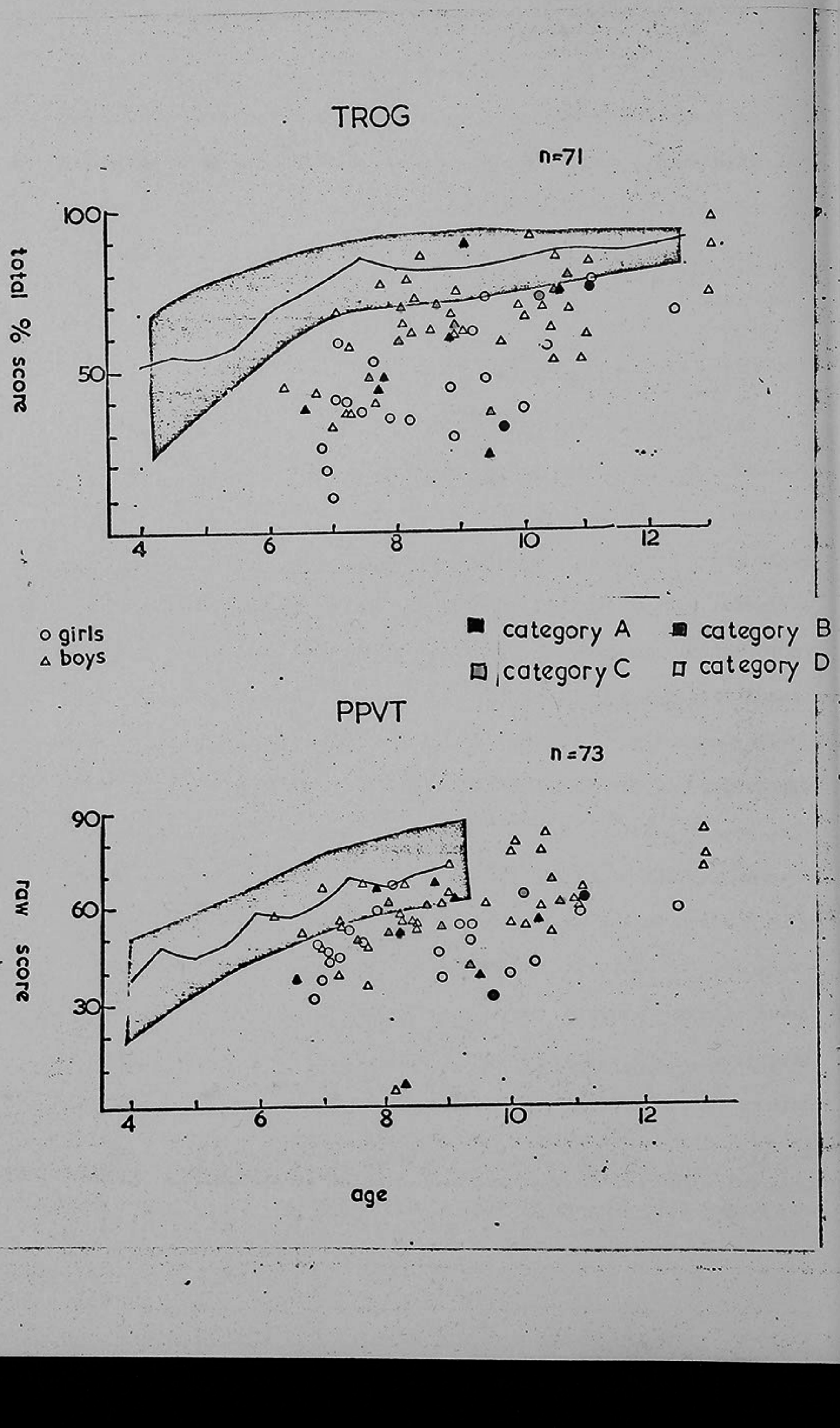
neurological abnormality

Information on the neurological status of the sample was very varied. In many cases information was inadequate to make a firm statement of neurological abnormality, while suggesting that the child was not normal. In other cases, factors in the pre or perinatal history were reported which raised the probability of neurological abnormality, but could not be regarded as definite causal factors, since many children with similar histories would develop normally. Within the limitations of the data available, it was decided to divide the children into four categories: A. Those with neurological signs; B. Those with a record of definite EEG abnormality but without a record of neurological signs; C. Those with factors in the history suggestive of possible neurological disorder, but no further evidence of neurological dysfunction; D. Those with no evidence, or only very weak suggestion of neurological disorder. Table 19 gives further details of children in these categories[1]. Figure 34 shows the comprehension scores of these

[1] I would like to thank Dr Gillian Sleigh for advising me on the allocation of children to categories, while stressing that she is not responsible for any anomalies that may exist.

Fig. 34

Neurological abnormality in relation to comprehension scores



children. Matched pairs t-tests contrasting children in categories A and B with others of the same age and sex in cateogry D confirmed that neurologically abnormal children do no worse on these comprehension tests than those without such abnormality (for TROG, t = 1.65; for PPVT, t = 0.53; d.f. = 8). More detailed standardized objective neurological description of these children would be valuable.

Summary & Conclusions

Although at an empirical level we have arrived at several answers to the questions raised in the introduction, the theoretical significance of these findings is by no means certain. The empirical results may be summarized as follows:

1. Most children with developmental language disorders have comprehension deficits.

2. The comprehension scores of children with receptive and expressive types of disorder do not clearly cluster into separate categories, but form a unimodal distribution.

3. Indices of receptive and expressive grammatical development correlate significantly in these children.

4. Although fewer girls than boys have developmental language disorders, girls tend to have more severe comprehension problems.

5. In the majority of cases, comprehension of grammar and vocabulary are in keeping with one another, but a not insubstantial minority does more poorly on grammatical than vocabulary comprehension.

6. Within the grammatical test, children with developmental language disorders have an atypical pattern of performance, with Word Order giving a disproportionate amount of difficulty.

7. Language disordered children have relatively better comprehension of pronouns and tense inflections and relatively worse comprehension of reversible passives than controls of the same general level of comprehension ability.

As far as classification is concerned, these results on the whole lend support to a quantitative model which holds that differences between expressive and receptive types of disorder are a matter of degree. It is possible that cases resembling classical "congenital auditory agnosia" should be regarded as a qualitatively distinct group. The finding of different sex ratios for mild and severe comprehension deficits is not inconsistent with either model.

We may consider the results on the comprehension tests in terms of the deviance versus delay argument outlined in chapter 5. It can be seen that different answers are appropriate at different levels. Insofar as expressive and receptive grammatical scores tend not to be discrepant, we may talk of delay. The finding of similar levels of comprehension in grammar and vocabulary may also be described as an instance of delay. The patterning of performance within the grammatical comprehension task, however, is deviant.

These results do not allow us to decide between the psychological theories outlined in chapter 5, except to reject the notion that language disordered children have a simple memory impairment. It was found that the Word Combinations scale, which included some of the longest items, was for both normal and language-disordered children the easiest scale in TROG. This would not be predicted if memory was a problem. Evidence for and against other hypotheses is conflicting. The finding that in general comprehension of grammar and vocabulary

are in keeping with one another is consistent with all theories except those advocating a specific grammatical deficit (Cromer, 1977) or a problem in sequencing words. However, the disproportionately poor performance on the word order scale in TROG is difficult to account for except in terms of the latter two theories. Detailed analysis suggests Cromer's hypothesis can best account for different patterns of performance on grammatical structures for language disordered and normal children.

At the risk of appearing unparsimonious, we must say that the comprehension of language disordered children cannot easily be described in terms of a single deficit. In addition to a general impairment in comprehension, affecting vocabulary and grammar alike, a specific difficulty in handling complex grammatical structures is apparent. Of course, these two aspects may be related, although it would be difficult to demonstrate the direction of causality; e.g. it could be argued that a general difficulty in perceiving language will result in a severe grammatical deficit. Support for this view comes from the finding that children with peripheral deafness are known to have difficulty in handling grammatical structure (Tervoort, 1970; Cromer, 1977). Furthermore, as de Ajuriaguerra (1966) has argued, it may be that those morphemes which are important in the analysis of grammatical structure are just those which are most difficult to perceive, because they are variable, and often barely audible.

Conversely, it might be argued that inability to analyse language hierarchically makes the individual unaware of the various semantic roles that words can play in a sentence and generally less practised in language skills. Consequently, vocabulary will suffer.

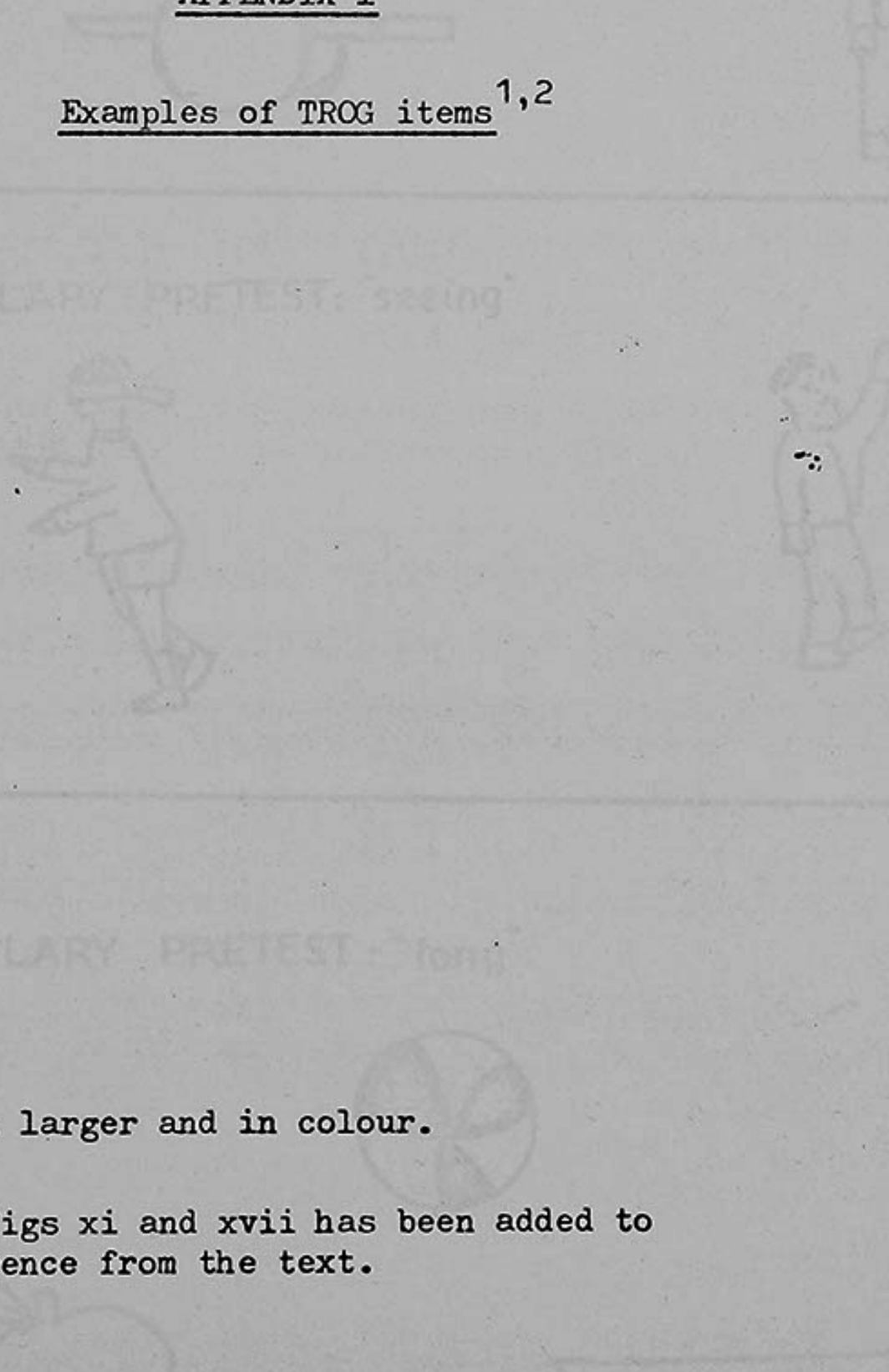
One way of testing more directly between some of the theories outlined in chapter 5 would be to examine language-disordered children's comprehension of sign language. Vocabulary tests are not very suitable for this purpose because the vocabularies of most sign languages are very limited. The author found attempts to translate the PPVT into the Paget-Gorman system fruitless, because only 44 of the first 75 words had Paget equivalents, and many of these could have been passed by children understanding only part of the sign.[1] TROG, however, could be easily translated into Paget-Gorman, particularly since this sign system maps directly onto spoken language, and has signs for inflections and function words as well as content words.

All theories stressing that language disordered children have a specifically auditory deficit (e.g. Tallal & Piercy, 1973b; Benton, 1964) would predict that these children should do much better with signed than spoken language, whereas those stressing disorders of symbolic thought, hierarchical structuring or sequencing, which are not regarded as modality specific, would predict that signed language should be as difficult as spoken.

[1] The Paget-Gorman sign system uses a set of basic signs to denote major semantic categories. These are inflected to give more precise meanings. E.g. a sign is used for "animal" which is given a long neck, a tail, whiskers, etc. to make it mean giraffe, rabbit, cat, respectively. A child tested on the word "cat" could get this right by just understanding the sign for animal, if the distractors were inanimate.

APPENDIX 1

Examples of TROG items[1,2]



[1]Originals are much larger and in colour.

[2]The lettering in figs xi and xvii has been added to facilitate reference from the text.

fig i)

VOCABULARY PRETEST : "ball"

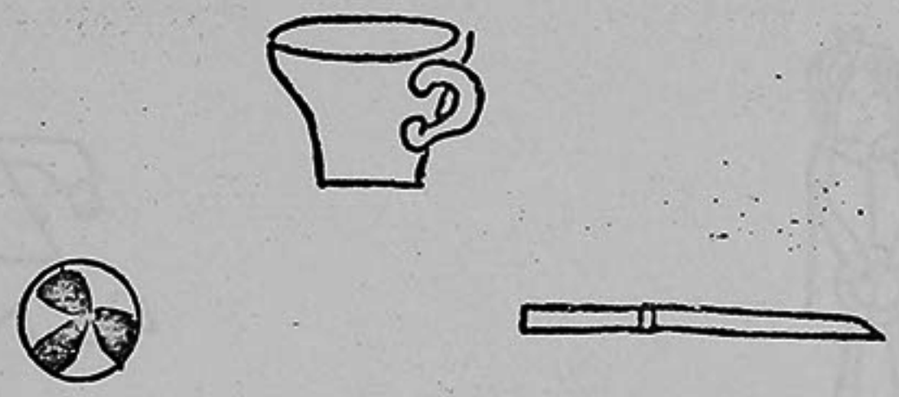


fig ii)

VOCABULARY PRETEST: "cutting"

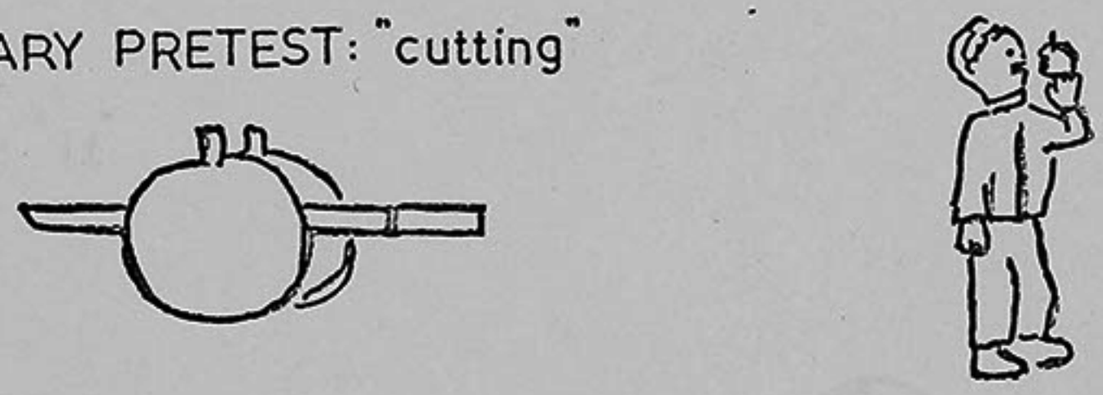


fig iii)

VOCABULARY PRETEST: "seeing"



fig iv)

VOCABULARY PRETEST : "long"

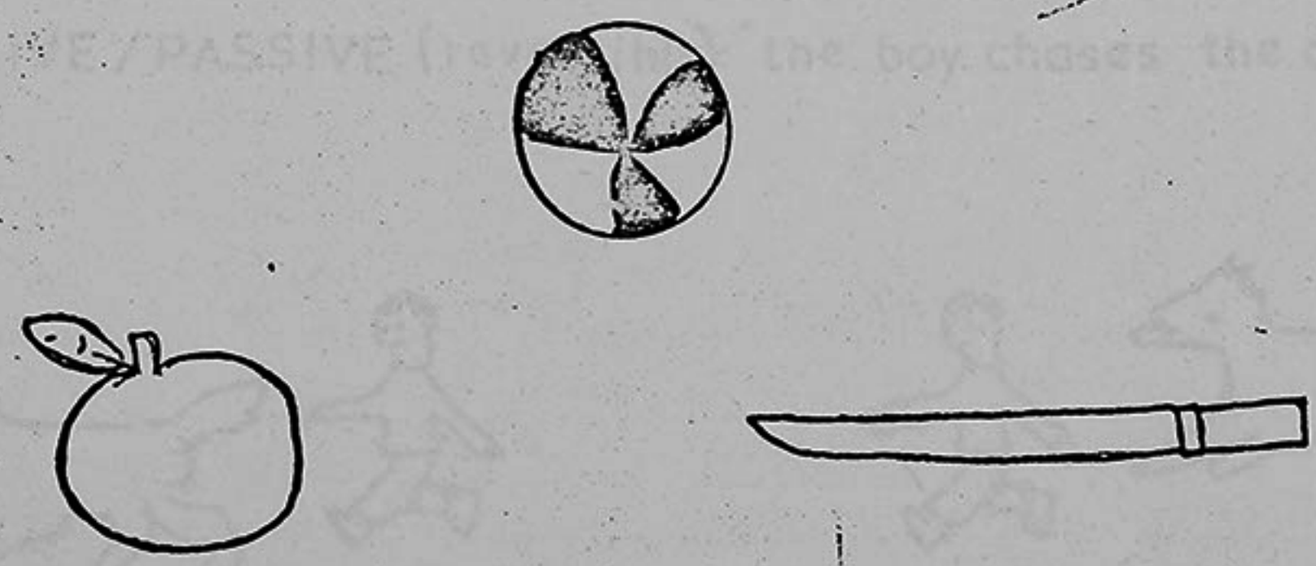


fig v)

NEGATIVE: "not running"

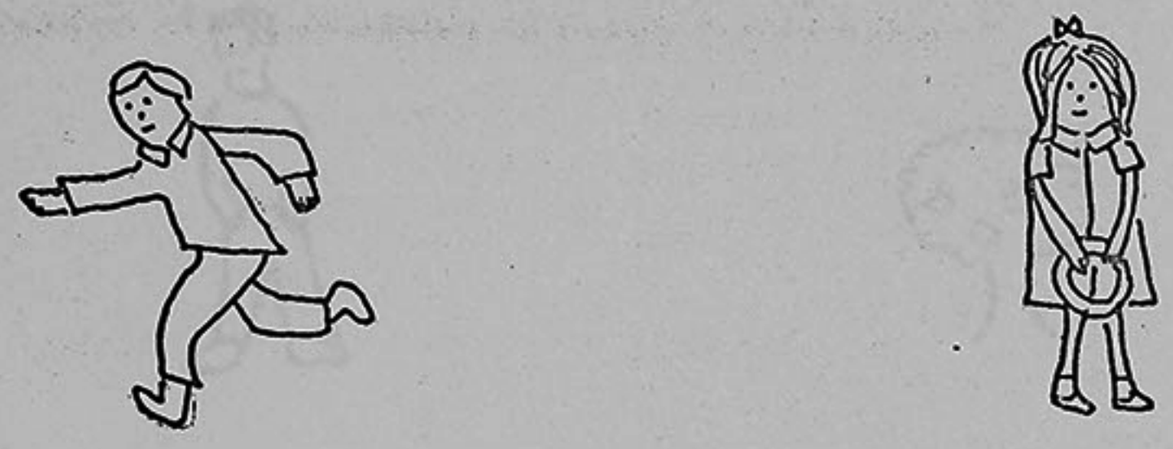


fig vi)

PRONOUN: "it"

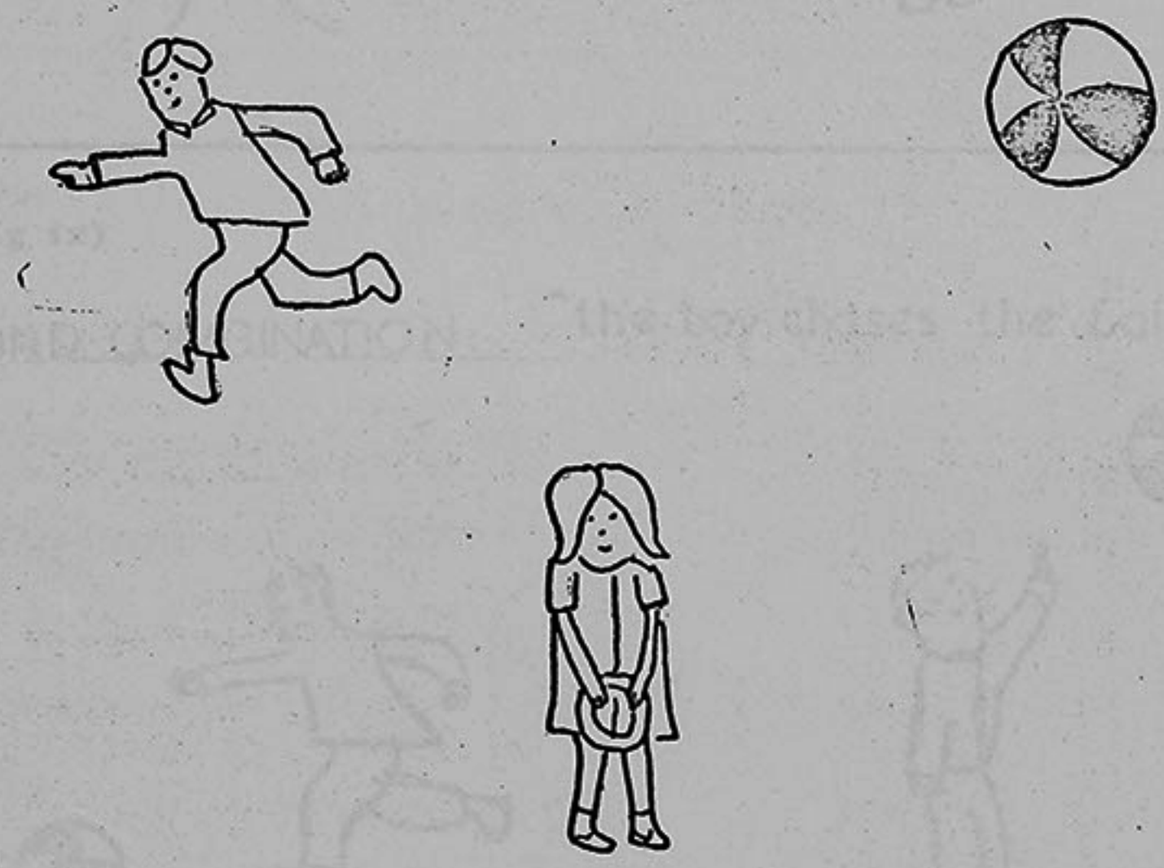


fig vii)

ACTIVE / PASSIVE (reversible): "the boy chases the dog"

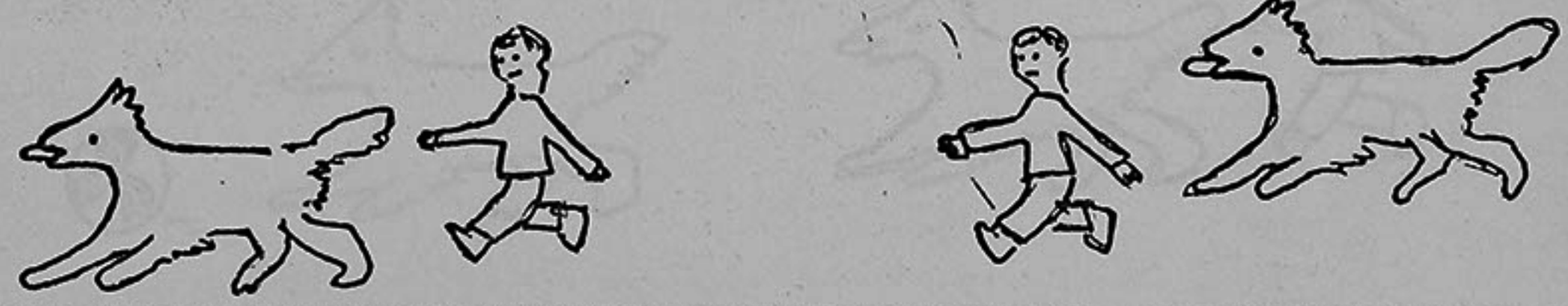


fig viii)

2-WORD COMBINATION: "big tree"

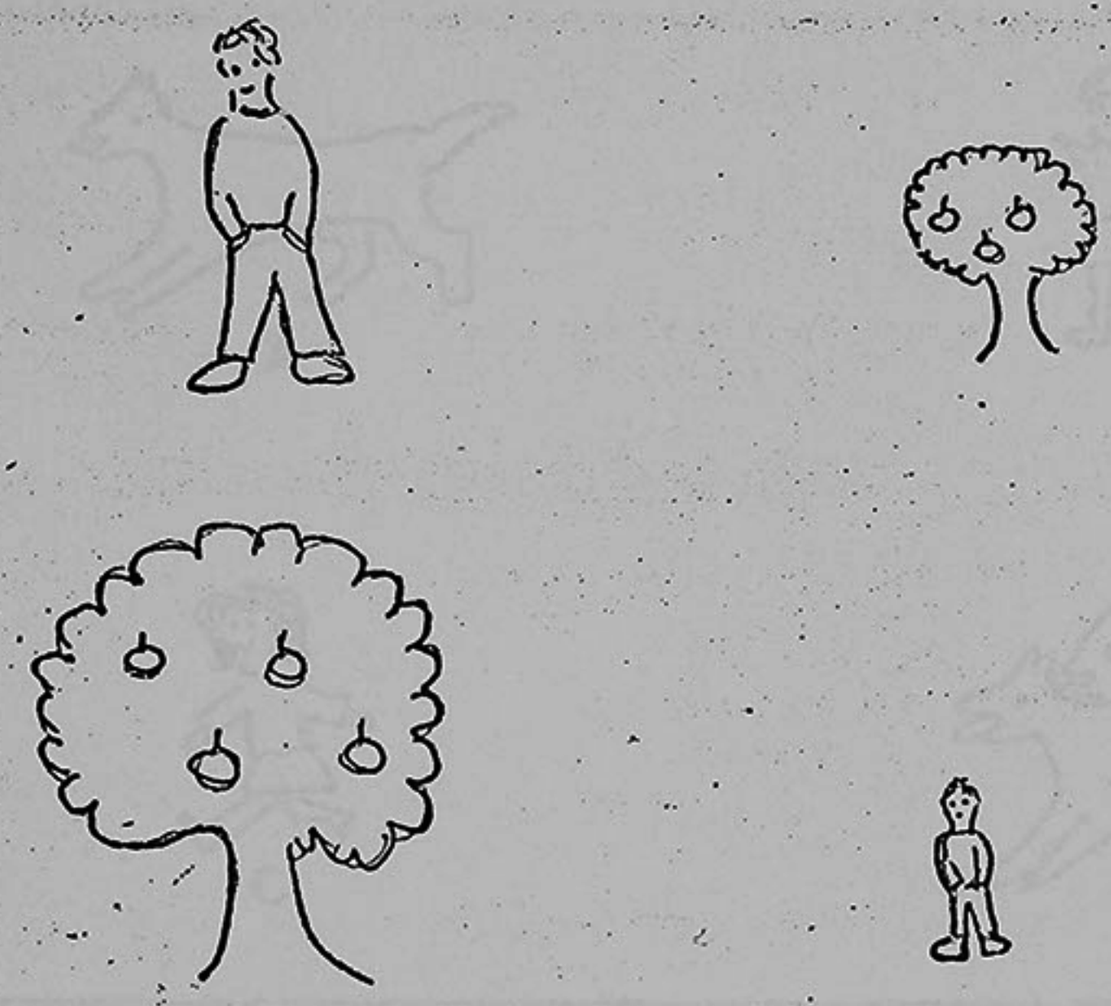


fig ix)

3-WORD COMBINATION: "the boy chases the ball"

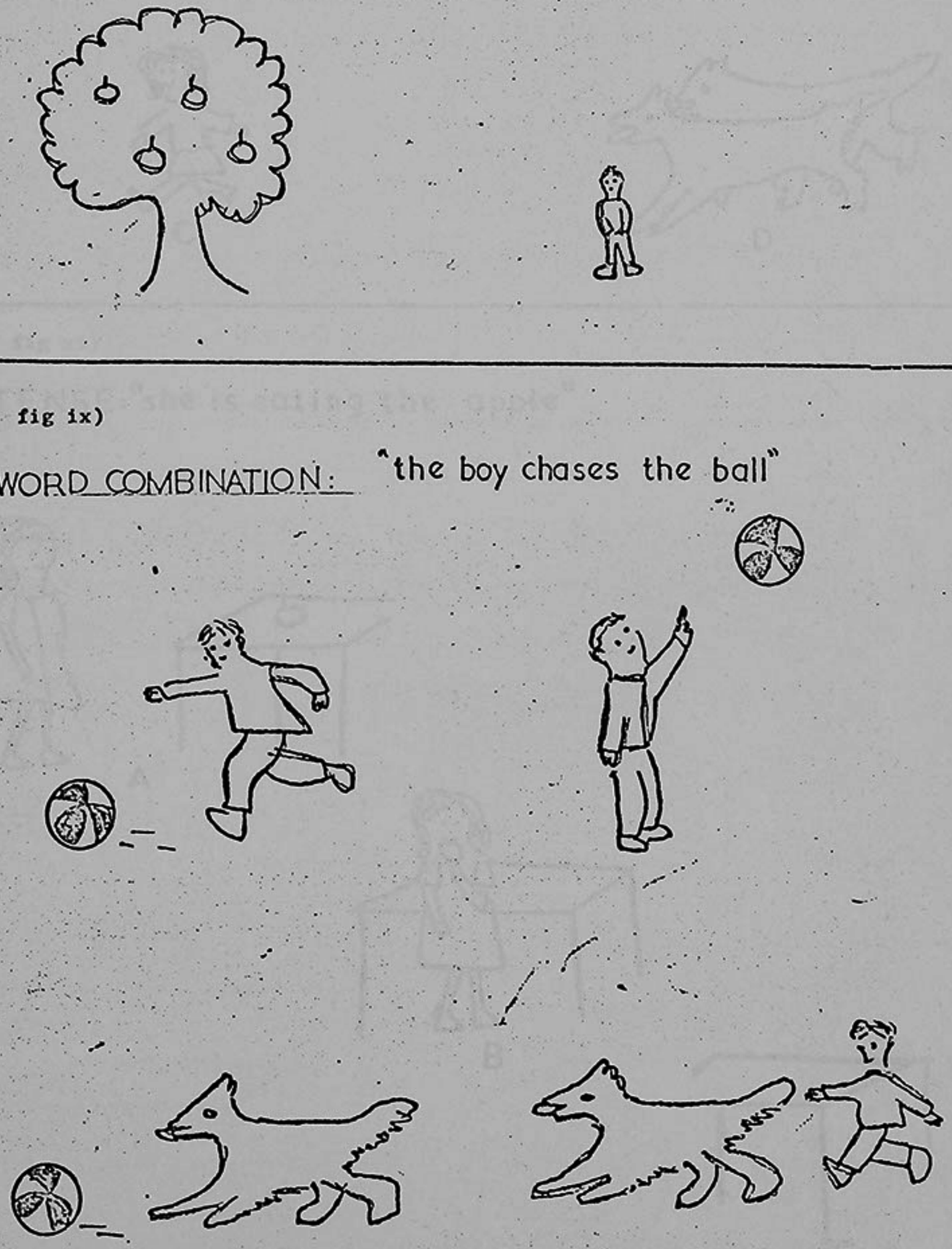


fig x)

NOUN. SINGULAR/PLURAL: "the dogs"



A



B



C

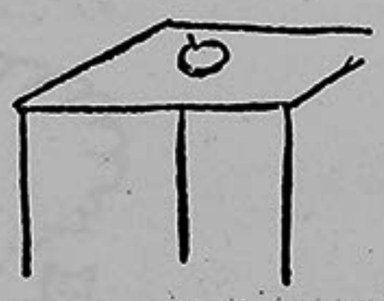


D

fig xi)

TENSE: "she is eating the apple"

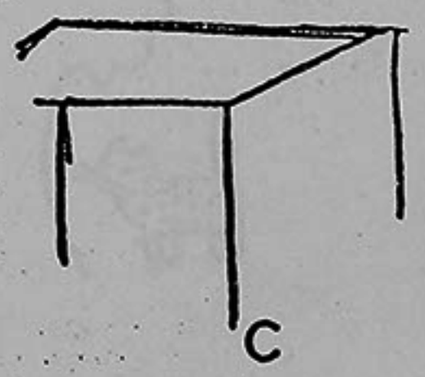




A



B



C



fig xii)

TENSE: "the boy will run"







fig xiii)

WORD COMBINATION: "the spotty cat sits under the tree"





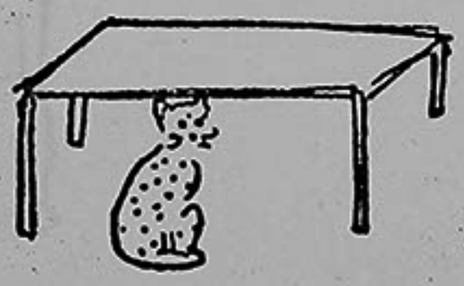




fig xiv)

EMBEDDED: "the ball the boy chases is spotty"

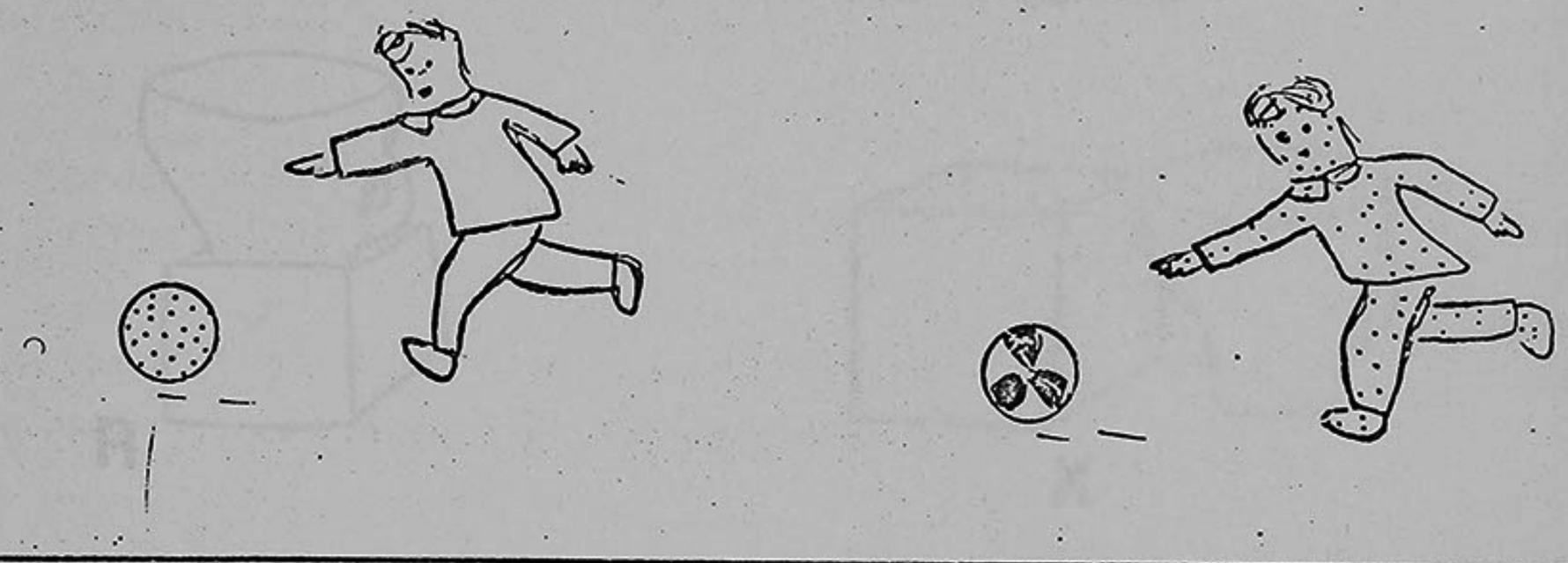


fig xv)

VERB SINGULAR/PLURAL: "the sheep chase the cat"



fig xvi)

PRONOUN: "their ball"

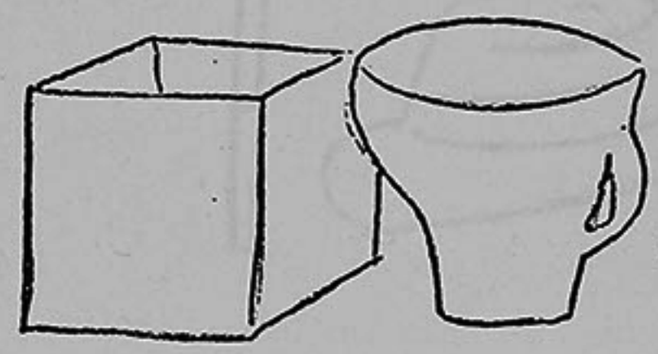


fig xvii)

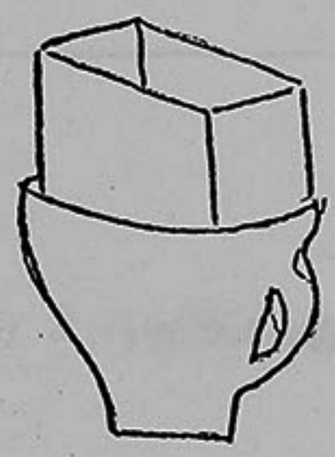
PREPOSITION (session 1) "the box in the cup"



R



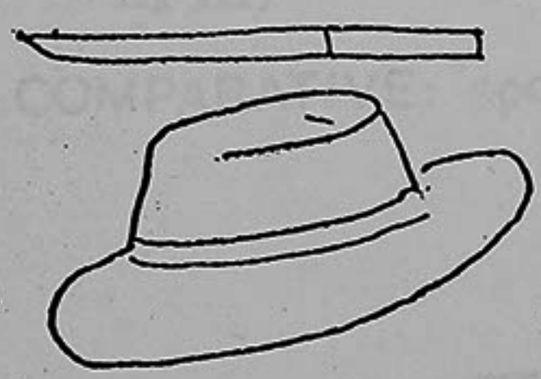
X



C

fig xviii)

PREPOSITION (session 2; + word order) "the knife under the hat"



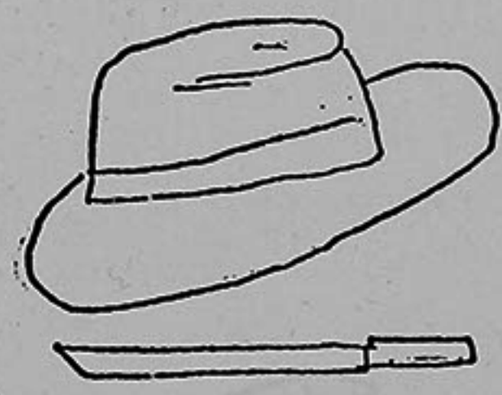


fig xix)

PREPOSITION (session 2; standard): "the hat under the knife"

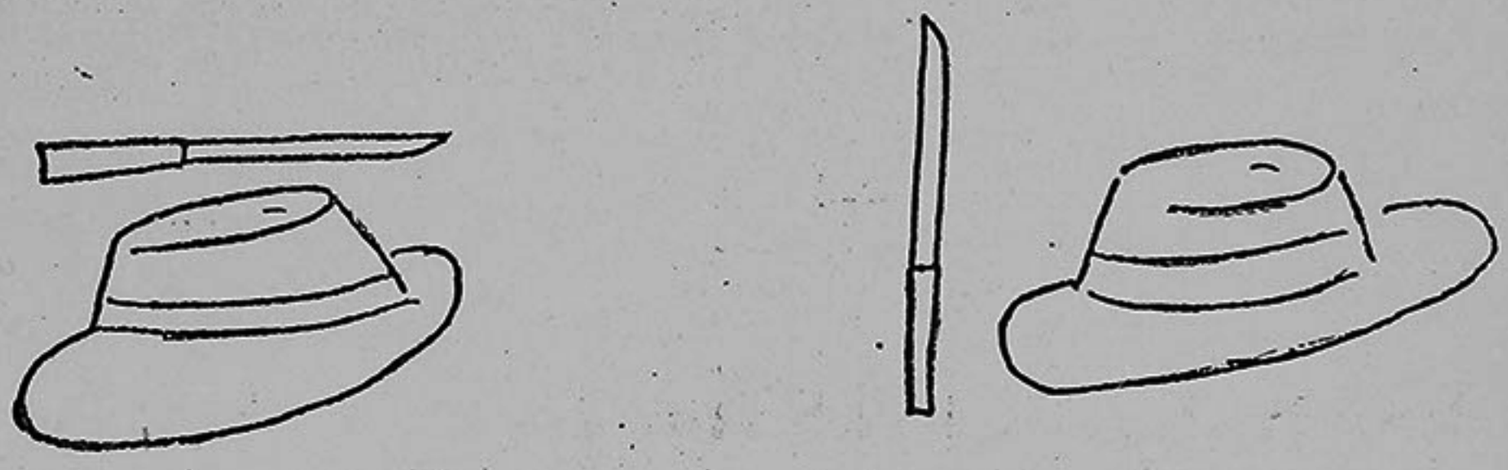


fig xx)

PREPOSITION (session 2; + word order): "the ball on the cup"



fig xxi)

COMPARATIVE: "spottier"

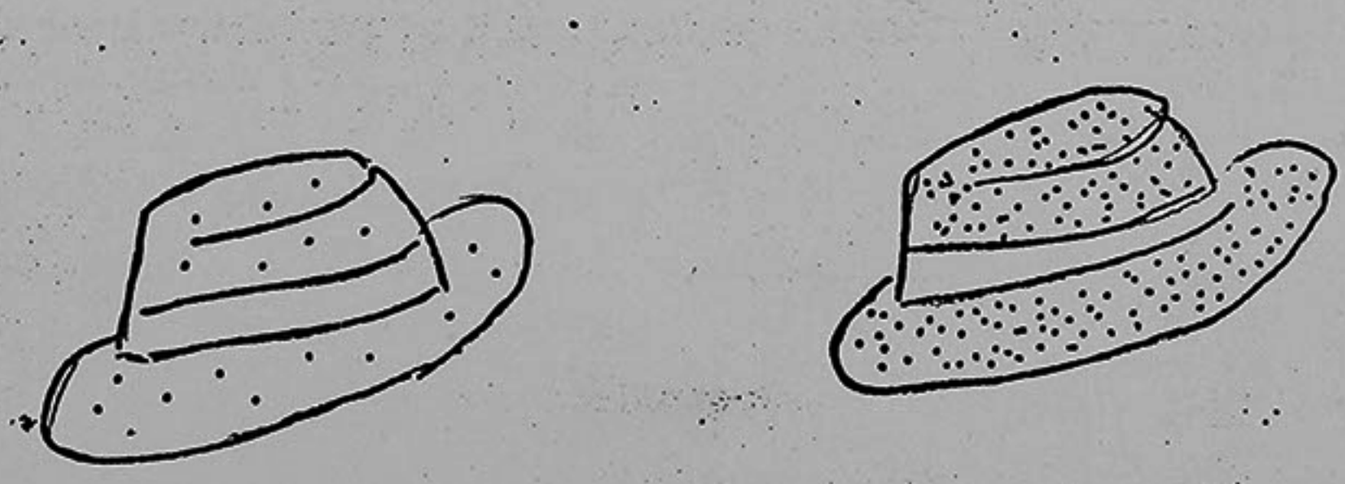


Table 20

WISC-R IQ at age 8½

	completed WPPSI	refused 1-3 subtests	refused 4-5 subtests	refused 6 or more subtests
n	96	7	8	9
mean	98.7	96.6	86.5	84.4
s.d.	13.1	8.1	15.4	16.0

Table 21

IQ and preschool experience

	with preschool experience	without preschool experience
WPPSI		
n	55	34
mean IQ	100.2	97.7
WISC-R		
n	68	43
mean IQ	96.8	93.5

APPENDIX 2

Possible biases in samples of preschool children

On p. 52 two possible sources of bias in preschool samples are discussed. First, only those children who will co-operate with testing are included, and second, in this study, only nursery school pupils were tested. Data on whether degree of co-operation and nursery school attendance are associated with systematic differences in intelligence are available from another project in which the author is involved. All children registered with a rural general practice in Oxfordshire who were $4\frac{1}{2}$ years old between April 1972 and April 1974 were tested on the WPPSI[1]. Retesting on the WISC-R at age $8\frac{1}{2}$ is being carried out by the author[2]. At the time of writing, 119 children have been seen both at $4\frac{1}{2}$ and $8\frac{1}{2}$. 24 children refused one or more subtests of the WPPSI. None so far have refused the WISC-R. IQ data for refusers and nonrefusers are presented in table 20. Although the differences in IQ between refusers and nonrefusers are not statistically significant, they are large enough for us to be worried about assuming the groups are comparable.

Detailed background information on most of the children is also available[3], making it possible to divide the sample into those who attended some kind of preschool nursery or playgroup, and those who did not. IQ data at ages $4\frac{1}{2}$ and $8\frac{1}{2}$ for these groups are presented in table 21. Differences between groups are neither large nor statistically significant.

[1]Testing was carried out by Dr G. Butterworth. Results are reported in Fairweather & Butterworth (1977).

[2]Testing is due to be completed in April 1978.

[3]The author would like to thank Dr Anthony Steel who has provided the facilities for this project, and who gathered background information from the mothers.

APPENDIX 3

Calculation of %-scores for TROG

Since different items used different numbers of distractors, Guilford's (1936) correction for guessing was applied to make scales comparable. The formula is as follows:

$$c = \frac{kp - 1}{k - 1}$$

where c = corrected proportion of correct answers
p = obtained proportion of correct answers
k = number of choices for each item

%-scores are obtained by multiplying c by 100.

All items testing Word order are 2-choice, and all items testing Word-Combinations are 4-choice, so calculation of c for these scales simply involves using the formula above, making k = 2 and 4 respectively. Inflections and Function Words include both 2- and 3-choice items[1]. Therefore, for each of these scales, c was obtained separately for the subsets of 2- and 3-choice items (c_2 and c_3 respectively). Then for these scales,

$$\%\text{-score} = \frac{c_2 \cdot n_2 + c_3 \cdot n_3}{n_2 + n_3} \times 100$$

where n_2 = number of 2-choice items
n_3 = number of 3-choice items

Total %-scores were calculated by the following formula:

$$\text{total \%-score} = \frac{C_2 \cdot N_2 + C_3 \cdot N_3 + C_4 \cdot N_4}{N_2 + N_3 + N_4} \times 100$$

where C_2 = value of c for all 2-choice items and N_2 = no. of 2-choice items
C_3 = value of c for all 3-choice items and N_3 = no. of 3-choice items
C_4 = value of c for all 4-choice items and N_4 = no. of 4-choice items

[1] As mentioned in the text, noun singular/plural items were treated as 2-choice items.

Table 22

a) incidence of missing data

	normative group (n = 204)	language disordered group (n = 73)
% of children with no missing data	89.2	91.8
% of children with 1 or 2 items missing	8.8	8.2
% of children with 3 or more items missing	1.9	0

b) incidence of V-errors

	normative group	language disordered group
% of children with no V-errors	81.4	69.9
% of children with 1 or 2 V-errors	13.7	19.2
% of children with 3 or 4 V-errors	3.9	5.5
% of children with 5 or more V-errors	0.9	5.4[1]

[1] 2 children (2.7%) made so many V-errors that it was deemed invalid to calculate %-scores.

ge disordered group
(n = 73)

91.8

8.2

0

ge disordered group

69.9

19.2

5.5

5.4[1]

s deemed invalid

procedure for missing data*

This procedure was adopted in the following situations:

1. Child points to more than one picture
2. Child refuses to point to any picture
3. Child will not complete test

In these cases, items are treated as if the child had guessed at random. Missing items are designated as 2-, 3- or 4-choice, then the following are counted as correct:

every other 2-choice item

every third 3-choice item

every fourth 4-choice item

TROG %-scores are then calculated in the normal way.

vocabulary-errors*

Errors that arise because the child does not understand the vocabulary necessary to perform correctly will be called vocabulary-errors (V-errors), and are omitted from consideration in computing %-scores. To detect V-errors, words which the child did not understand are listed, and all errors on items including these words are marked. These are not all necessarily V-errors, e.g. if a child does not understand "black", and then selects a black cow for the item "black cat", this is not a V-error, because the error cannot be attributed to failure to understand "black". However, if the same child picks a white cat for the item "black cat", this is a V-error. Thus each case was checked to see if the particular failure made could indeed have arisen from failure to understand a word from the pretest, and was only considered a V-error if this was the case.

If a child did make V-errors, %-scores were calculated by the formula described above, omitting all V-error items from consideration. If a child made more than 4 V-errors on a scale, %-scores for that scale were not calculated.

*Table 22 shows the percentage of cases in which these adjustments to scores were necessary.

REFERENCES

AFFOLTER, F., BRUBAKER, R. & BISCHOFBERGER, W. (1974) Comparative studies between normal and language disturbed children. Acta Oto-Laryngologica, supplement 323.

ALAJOUANINE, T. & LHERMITTE, F. (1965) Acquired aphasia in children. Brain, 88, 653-662.

ANYAN, W.R. & QUILLIAN, W.W. (1971) The naming of primary colours by children. Child Development, 42, 1629-1632.

ARAM, D.M. & NATION, J.E. (1975) Patterns of language behaviour in children with developmental language disorders. Journal of Speech and Hearing Research, 18, 229-241.

ARNDT, W.B. (1977) A psychometric evaluation of the Northwestern Syntax Screening Test. Journal of Speech and Hearing Disorders, 42, 316-319.

AUERBACH, V.H., DIGEORGE, A.M., BALDRIDGE, R.C., TOURTELLOTTE, C.D. & BRIGHAM, M.P. (1962) Histidinaemia. Journal of Pediatrics, 60, 487-497.

BAIRD, R. (1972) On the role of chance in imitation-comprehension-production test results. Journal of Verbal Learning and Verbal Behaviour, 11, 474-477.

BALDIE, B.J. (1976) The acquisition of the passive voice. Journal of Child Language, 3, 331-348.

BARTAK, L., RUTTER, M. & COX, A. (1975) A comparative study of infantile autism and specific developmental receptive language disorder. 1. The children. British Journal of Psychiatry, 126, 127-145.

BASSER, L.S. (1962) Hemiplegia of early onset and the faculty of speech with special reference to the effects of hemispherectomy. Brain, 85, 427-460.

BAUMEISTER, A.A. (1967) Problems in comparative studies of mental retardates and normals. American Journal of Mental Deficiency, 71, 869-875.

BECKWITH, L. & THOMPSON, S.K. (1976) Recognition of verbal labels of pictured objects and events by 17- to 30-month old infants. Journal of Speech and Hearing Research, 19, 690-699.

BENEDICT, H. (1976) Language comprehension in 9 to 15 month old infants. Paper presented at the Psychology of Language conference, University of Stirling, June, 1976.

BENTON, A.L. (1964) Developmental aphasia and brain damage. Cortex, 1, 40-52.

BERKO, J. (1958) The child's learning of English morphology. Word, 14, 150-177.

BERRY-LUTERMAN, L. & BAR, A. (1971) The diagnostic significance of sentence repetition for language-impaired children. Journal of Speech and Hearing Disorders, 36, 29-39.

BEVER, T.G. (1970) The cognitive basis for linguistic structures. In J.R. Hayes (ed) Cognition and the Development of Language. New York: John Wiley & Sons.

BLISS, L.S. & PETERSON, D.M. (1975) Performance of aphasic and nonaphasic children on a sentence repetition task. Journal of Communication Disorders, 8, 207-212.

BLOOM, L. (1970) Language Development: Form and Function in Emerging Grammars. Cambridge, Mass.: MIT Press.

BLOOM, L.M. (1974) Talking, understanding and thinking. In R.L. Schiefelbusch & L.K. Lloyd (eds) Language Perspectives - Acquisition, Retardation, and Intervention. Baltimore: University Park Press.

BLOSSER, J.L., WEIDNER, W.E. & DINERO, T. (1976) The effect of rate-controlled speech on the auditory receptive scores of children with normal and disordered language abilities. Journal of Special Education, 10, 291-298.

BRIMER, M.A. & DUNN, L.H. (1962) English Picture Vocabulary Test. Bristol: Educational Evaluation Enterprises.

BROWN, G.W., SKLAIR, F., HARRIS, T.O. & BIRLEY, J.L.T. (1973) Life-events and psychiatric disorders. Part 1: some methodological issues. Psychological Medicine, 3, 74-87.

BROWN, H.D. (1971) Children's comprehension of relativized English sentences. Child Development, 42, 1923-1936.

BROWN, R. (1957) Linguistic determinism and the part of speech. Journal of Abnormal and Social Psychology, 55, 1-5.

BROWN, R. (1973) A First Language: the Early Stages. London: George Allen & Unwin Ltd.

BUTLER, N.R., PECKHAM, C. & SHERIDAN, M. (1973) Speech defects in children aged 7 years: A national study. British Medical Journal, 1, 253-257.

BYERS, R.K. & MCLEAN, W.T. (1962) Etiology and course of certain hemiplegias with aphasia in childhood. Pediatrics, 29, 376-383.

BYRNE, M.C. (1977) A clinician looks at the Northwestern Syntax Screening Test. Journal of Speech and Hearing Disorders, 42, 320-322.

CALNAN, M. & RICHARDSON, K. (1976) Speech problems in a national survey: assessments and prevalences. Child:Care, Health and Development, 2, 181-202.

CAMBON, J. & SINCLAIR, H. (1974) Relations between syntax and semantics: are they "easy to see"? British Journal of Psychology, 65, 133-140.

CAREY, S. (1976) "Less" never means "more". Paper presented at the Psychology of Language conference, University of Stirling, June, 1976.

CARR, D.L., BROWN, L.F. & RICE, J.A. (1967) The PPVT in the assessment of language deficits. American Journal of Mental Deficiency, 71, 937-939.

CARRELL, J. & BANGS, J. (1957) Disorders of speech comprehension associated with idiopathic language retardation. The Nervous Child, 9, 64-76.

CARROW, E. (1975) Test for Auditory Comprehension of Language. Texas: Learning Concepts.

CARROW, M.A. (1968) The development of auditory comprehension of language structure in children. Journal of Speech and Hearing Disorders, 33, 99-111.

CARTWRIGHT, L.R. & LASS, N.J. (1974) A comparative study of children's performance on the Token Test, Northwestern Syntax Screening Test, and Peabody Picture Vocabulary Test. Acta Symbolica, 5, 19-29.

CHAPMAN, L.J. & CHAPMAN, J.P. (1973) Problems in the measurement of cognitive deficit. Psychological Bulletin, 79, 380-385.

CHAPMAN, R.S. & MILLER, J.F. (1975) Word order in early two and three word utterances: does production precede comprehension? Journal of Speech and Hearing Research, 18, 355-371.

CHOMSKY, C. (1969) The Acquisition of Syntax in Children from 5 to 10. Cambridge, Mass: MIT press.

CHOMSKY, N. (1957) Syntactic Structures. The Hague: Mouton.

CHOMSKY, N. (1965) Aspects of the Theory of Syntax. Cambridge, Mass: MIT press.

CHOMSKY, N. & HALLE, M. (1968) The Sound Pattern of English. New York: Harper & Row.

CICCHETTI, D.V. (1972) Extensions of multiple-range tests to interaction tables in the analysis of variance: a rapid approximate solution. Psychological Bulletin, 77, 405-408.

CLARK, E.V. (1973) Nonlinguistic strategies and the acquisition of word meanings. Cognition, 2, 161-182.

CLARK, E.V. (in press) From gesture to word: On the natural history of deixis in language acquisition. In J.S. Bruner & A. Garton (eds) Human Growth and Development: Wolfson College Lectures, 1976.

CLARK, H.H. (1970) The primitive nature of children's relational concepts. In J.R. Hayes (ed) Cognition and the Development of Language. New York: John Wiley & Sons.

CLONINGER, C.R., REICH, T. & GUZE, S.B. (1975) The multifactorial model of disease transmission: III Sex differences in the familial transmission of sociopathy (antisocial personality). British Journal of Psychiatry, 127, 11-22.

COHEN, D.J., CAPARULO, B. & SHAYWITZ, B. (1976) Primary childhood aphasia and childhood autism. Journal of the American Academy of Child Psychiatry, 15, 604-645.

COMPTON, A.J. (1970) Generative studies of children's phonological disorders. Journal of Speech and Hearing Disorders, 35, 315-339.

CONN, P. (1971) Language Therapy 1 - remedial syntax. London: Invalid Children's Aid Association.

COOK, V.J. (1975) Strategies in the comprehension of relative clauses. Language and Speech, 18, 204-212.

CRITCHLEY, M. (1967) Aphasiological nomenclature and definitions. Cortex, 3, 3-25.

CROMER, R.F. (1970) "Children are nice to understand": surface structure clues for the recovery of a deep structure. British Journal of Psychology, 61, 397-408.

CROMER, R.F. (1972) The learning of surface structure clues to deep structure by a puppet show technique. Quarterly Journal of Experimental Psychology, 24, 66-76.

CROMER, R.F. (1977) Hierarchical disability in the written syntax of aphasic children. Paper presented at the International Society for the Study of Behavioural Development biennial conference, "Biosocial Aspects of Development", University of Pavia, Italy, September, 1977.

CROOKES, T.G. & GREENE, M.C.L. (1963) Some characteristics of children with two types of speech disorder. British Journal of Educational Psychology, 33, 31-40.

CRYSTAL, D., FLETCHER, P. & GARMAN, M. (1976) The Grammatical Analysis of Language Disability. London: Edward Arnold.

DE AJURIAGUERRA, J. (1966) Speech disorders in childhood. In E.C. Carterette (ed) Brain Function, vol. 3. Berkeley, L.A.: University of California Press.

DE AJURIAGUERRA, J., JAEGGI, A., GUIGNARD, F., KOCHER, F., MAQUARD, M., ROTH, S. & SCHMID, E. (1976) The development and prognosis of dysphasia in children. In D.M. Morehead and A.E. Morehead (eds) Normal and Deficient Child Language. Baltimore: University Park Press.

DENNIS, M. & KOHN, B. (1975) Comprehension of syntax in infantile hemiplegics after cerebral hemidecortication: left-hemisphere superiority. Brain and Language, 2, 472-482.

DENNIS, M. & WHITAKER, H.A. (1976) Language acquisition following hemidecortication: linguistic superiority of the left over the right hemisphere. Brain and Language, 3, 404-433.

DE RENZI, E. & VIGNOLO, L.A. (1962) The token test: a sensitive test to detect receptive disturbances in aphasia. Brain, 85, 665-678.

CROMER, R.F (1976) The cognitive hypothesis of language acquisition & its implications for child language deficiency. In D.M. Morehead & A.E. Morehead (eds.) Normal & Deficient Child Language. Baltimore: University Park Press.

DE VILLIERS, J.G. & DE VILLIERS, P.A. (1973) Development of the use of word order in comprehension. Journal of Psycholinguistic Research, 2, 331-341.

DE VILLIERS, P.A. & DE VILLIERS, J.G. (1972) Early judgements of semantic and syntactic acceptability by children. Journal of Psycholinguistic Research, 1, 299-310.

DONALDSON, M. & BALFOUR, G. (1968) Less is more: a study of language comprehension in children. British Journal of Psychology, 59, 461-471.

DONALDSON, M. & WALES, R. (1970) On the acquisition of some relational terms. In J.R. Hayes (ed) Cognition and the Development of Language. New York: John Wiley & Sons.

DOZIER, P. (1937) Etiological and clinical types of so-called "nerve deafness": congenital word-deafness. Laryngoscope, 47, 516-519.

DUCHAN, J.F. & ERICKSON, J.G. (1976) Normal and retarded children's understanding of semantic relations in different verbal contexts. Journal of Speech and Hearing Research, 19, 767-776.

DUKES, P. & PANAGOS, J.M. (1973) Repetition of comprehensible sentences by children with deviant speech. British Journal of Disorders of Communication, 8, 139-145.

DUNN, L.M. (1965) Expanded manual for the Peabody Picture Vocabulary Test. Minnesota: American Guidance Service, Inc.

EDWARDS, M. (1973) Developmental verbal dyspraxia. British Journal of Disorders of Communication, 8, 64-70.

EFRON, R. (1963) Temporal perception, aphasia and deja vu. Brain, 86, 403-423.

EFRON, R. (1967) Comments in response to Hirsch. In C.H. Millikan & F.L. Darley (eds) Brain Mechanisms underlying Speech and Language. New York: Grune & Stratton.

EISENSON, J. (1968) Developmental aphasia (dyslogia): a postulation of a unitary concept of the disorder. Cortex, 4, 184-200.

EVERITT, B.S. (1975) Multivariate analysis: the need for data, and other problems. British Journal of Psychiatry, 126, 237-240.

FAIRWEATHER, H. & BUTTERWORTH, G. (1977) The WPPSI at four years: a sex difference in verbal performance discrepancies. British Journal of Educational Psychology, 47, 85-90.

FERNALD, C.D. (1972) Control of grammar in imitation, comprehension and production: problems of replication. Journal of Verbal Learning and Verbal Behaviour, 11, 606-613.

FILLMORE, C. (1968) The case for case. In E. Bach & R. Harms (eds) Universals in Linguistic Theory. New York: Holt, Rinehart & Wilson.

FOLEY, H.M. & LOCKE, J.L. (1971) Young children's knowledge of morphological and phonological rules. Journal of Communication Disorders, 4, 259-262.

FORREST, T., EISENSON, J. & STARK, J. (1967) EEG findings in 113 non-verbal children. Electroencephalography and Clinical Neurology, 22, 291.

FOURCIN, A.J. (1975) Language development in the absence of expressive speech. In E.H. Lenneberg & E. Lenneberg (eds) Foundations of language development: a multidisciplinary approach, vol. 2. New York: Academic Press.

FRASER, C., BELLUGI, U. & BROWN, R. (1963) Control of grammar in imitation, comprehension and production. Journal of Verbal Learning and Verbal Behaviour, 2, 121-135.

FREEDMAN, P.P. & CARPENTER, R.L. (1976) Semantic relations used by normal and language-impaired children at stage 1. Journal of Speech and Hearing Research, 19, 784-795.

FRIEDLANDER, B.Z. (1970) Receptive language development in infancy. Merril-Palmer Quarterly, 16, 7-51.

FURTH, H.G. (1964) Sequence learning in aphasic and deaf children. Journal of Speech and Hearing Disorders, 29, 171-177.

FURTH, H.G. & PUFALL, P.B. (1966) Visual and auditory sequence learning in hearing impaired children. Journal of Speech and Hearing Research, 9, 441-449.

FUSILIER, F.M. & LASS, N.J. (unpublished) A comparative study of children's performance on the Illinois Test of Psycholinguisitc Abilities and the Token Test.

GAER, E.P. (1969) Children's understanding and production of sentences. Journal of Verbal Learning and Verbal Behaviour, 8, 289-294.

GARVEY, M. & GORDON, N. (1973) A follow-up study of children with disorders of speech development. British Journal of Disorders of Communication, 8, 17-28.

GASCON, G., VICTOR, D., LOMBROSO, C. & GOODGLASS, H. (1973) Language disorder, convulsive disorder and electroencephalographic abnormalities. Archives of Neurology, 28, 156-162.

GHADIMI, H., PARTINGTON, M.W. & HUNTER, A. (1961) A familial disturbance of histidine metabolism. New England Journal of Medicine, 265, 221-224.

GHADIMI, H., PARTINGTON, M.W. & HUNTER, A. (1962) Inborn error of histidine metabolism. Pediatrics, 29, 714-728.

GLEITMAN, L.R., GLEITMAN, H. & SHIPLEY, E.F. (1972) The emergence of the child as grammarian. Cognition, 1, 137-164.

GODAN, W.L. & PANAGOS, J.M. (1976) Developmental transformational capacity of children with Down's syndrome. Perceptual and Motor Skills, 43, 967-973.

GOEHL, H. & GOLDEN, S. (1974) A psycholinguistic account of why children do not detect their own errors. Unpublished MS. Department of Speech, Temple University, Philadelphia.

GOLDSTEIN, R., LANDAU, W.M. & KLEFFNER, F.R. (1958) Neurologic assessment of some deaf and aphasic children. Transactions of the American Otological Society, 46, 122-136.

GORDON, N. (1964) The concept of central deafness. Clinics in Developmental Medicine, 13, 62-64.

GORDON, N. (1970) Delayed speech and histidinaemia. Developmental Medicine and Child Neurology, 12, 104-106.

GRAHAM, N.C. (1968) Short term memory and syntactic structure in educationally subnormal children. Language and Speech, 11, 209-219.

GRIEVE, R., HOOGENRAAD, R. & MURRAY, D. (1977) On the young child's use of lexis and syntax in understanding locative instructions. Cognition, 5, 235-250.

GRIFFITHS, P. (1972) Developmental aphasia: an introduction. London: Invalid Children's Aid Association.

GUILFORD, J.P. (1936) The determination of item difficulty when chance success is a factor. Psychometrika, 1, 259-264.

GUILFORD, J.P. & FRUCHTER, B. (1973) Fundamental Statistics in Psychology and Education. Tokyo: McGraw-Hill Kogakusha Ltd.

GUTTMAN, E. (1942) Aphasia in children. Brain, 65, 205-219.

HARNER, L. (1976) Children's understanding of linguistic reference to past and future. Journal of Psycholinguistic Research, 5, 65-84.

HECAEN, H. (1976) Acquired aphasia in children and the ontogenesis of hemispheric functional specialization. Brain and Language, 3, 114-134.

HERRIOT, P. (1969) The comprehension of tense by young children. Child Development, 40, 103-110.

HOBSBAUM, A. & MITTLER, P. (1971) Sentence Comprehension Test: experimental edition. Unpublished.

HOLLAND, V.M. & PALERMO, D.S. (1975) On learning "less": language and cognitive development. Child Development, 46, 437-443.

HUGHES, J. (1974/5) Acquisition of a non-vocal "language" by aphasic children. Cognition, 3, 41-55.

HUSKISSON, J.A. (1973) Acquired receptive language difficulties in children. British Journal of Disorders of Communication, 8, 54-63.

HUTTENLOCHER, J., EISENBERG, K. & STRAUSS, S. (1968) Comprehension: relation between perceived actor and logical subject. Journal of Verbal Learning and Verbal Behaviour, 7, 527-530.

HUTTENLOCHER, J. & STRAUSS, S. (1968) Comprehension and a statement's relation to the situation it describes. Journal of Verbal Learning and Verbal Behaviour, 7, 300-304.

HUXLEY, R. (1970) The development of the correct use of subject personal pronouns in two children. In G.B. Flores D'Arcais & W.J.M. Levelt (eds) Advances in Psycholinguistics. Amsterdam: North Holland Publishing Co.

INGRAM, D. (1974) The relationship between comprehension and production. In R.L. Schiefelbusch & L.K. Lloyd (eds) Language Perspectives - Acquisition, Retardation, and Intervention. Baltimore: University Park Press.

INGRAM, T.T.S. (1959) Specific developmental disorders of speech in childhood. Brain, 82, 450-467.

INGRAM, T.T.S. (1963) Report of the Dysphasia Subcommittee of the Scottish Pediatric Society. Unpublished.

INGRAM, T.T.S. (1969) Developmental disorders of speech. In P.J. Vinken & G.W. Bruyn (eds) Handbook of Clinical Neurology, vol. 4. New York: John Wiley & Sons.

INGRAM, T.T.S. (1972) The classification of speech and language disorders in young children. Clinics in Developmental Medicine, 43, 13-32.

INHELDER, B. (1976) Observations on the operational and figurative aspects of thought in dysphasic children. In D.M. Morehead & A.E. Morehead (eds) Normal and Deficient Child Language. Baltimore: University Park Press.

JOHNSON, D.J. & MYKLEBUST, H.R. (1967) Learning Disabilities: Educational Principles and Practices. New York: Grune & Stratton.

JOHNSTON, J.R. & SCHERY, T.K. (1976) The use of grammatical morphemes by children with communication disorders. In D.M. Morehead & A.E. Morehead (eds) Normal and Deficient Child Language. Baltimore: University Park Press.

KANNER, L. (1946) Irrelevant and metaphorical language in early infantile autism. American Journal of Psychiatry, 103, 242-246.

KARLIN, I.W. & STRAZZULLA, M. (1952) Speech and language problems of mentally deficient children. Journal of Speech and Hearing Disorders, 17, 286-294.

KARMILOFF-SMITH, A. (1976) The interplay between syntax, semantics and phonology in language acquisition processes. Paper presented at the Psychology of Language conference, University of Stirling, June, 1976.

KEENEY, T.J. & SMITH, N.D. (1971) Young children's imitation and comprehension of sentential singularity and plurality. Language and Speech, 14, 373-382.

KEENEY, T.J. & WOLFE, J. (1972) The acquisition of agreement in English. Journal of Verbal Learning and Verbal Behaviour, 11, 698-705.

KERSCHENSTEINER, M. & HUBER, W. (1975) Grammatical impairment in developmental aphasia. Cortex, 11, 264-282.

KESSEL, F.S. (1970) The role of syntax in children's comprehension from ages 6 to 12. Monographs of the Society for Research in Child Development, 35, no. 6.

KINSBOURNE, M. (1974) Mechanisms of hemispheric interaction in man. In M. Kinsbourne & W. Lynn Smith (eds) Hemispheric Disconnection and Cerebral Function. Springfield, Ill: C.C. Thomas.

KIRK, S.A., MCCARTHY, J.J. & KIRK, W.D. (1968) Illinois Test of Psycholinguistic Abilities, revised edition. Urbana, Ill: University of Illinois Press.

KORNFELD, J.R. (1976) Implications of studying reduced consonant clusters in normal and abnormal child speech. Paper presented at the Psychology of Language conference, University of Stirling, June, 1976.

KRACKE, I. (1975) Perception of rhythmic sequences by receptive aphasic and deaf children. British Journal of Disorders of Communication, 10, 43-51.

KRYNAUW, R. (1950) Infantile hemiplegia treated by removing one hemisphere. Journal of Neurology, Neurosurgery and Psychiatry, 13, 243-267.

LA DU, D.M., HOWEL, R.R., JOCOBY, G.A., SEEGMILLER, J.E., SOBER, E.K., ZONNONI, V.G., CANBY, J.P. & ZIEGLER, L.K. (1963) Clinical and biochemical studies in two cases of histidinaemia. Pediatrics, 32, 216-227.

LAHEY, M. (1974) Use of prosody and syntactic markers in children's comprehension of spoken sentences. Journal of Speech and Hearing Research, 17, 656-668.

LANDAU, W.M., GOLDSTEIN, R. & KLEFFNER, F.R. (1960) Congenital aphasia: A clinicopathologic study. Neurology, 10, 915-921.

LANDAU, W. & KLEFFNER, F. (1957) Syndrome of acquired aphasia with convulsive disorder in children. Neurology, 7, 523-530.

LARSON, G.W. & SUMMERS, P.A. (1976) Response patterns of pre-school-age children to the Northwestern Syntax Screening Test. Journal of Speech and Hearing Disorders, 41, 486-497.

LASS, N.J., DEPAOLO, A.M., SIMCOE, J.C. & SAMUEL, S.M. (1975) A normative study of children's performance on the short form of the token test. Journal of Communication Disorders, 8, 193-198.

LEA, J. (1965) A language scheme for children suffering from receptive aphasia. Speech Pathology and Therapy, 8, 58-68.

LEE, L.L. (1966) Developmental sentence types: a method for comparing normal and deviant syntactic development. Journal of Speech and Hearing Disorders, 31, 311-330.

LEE, L.L. (1970) A screening test for syntax development. Journal of Speech and Hearing Disorders, 35, 103-112.

LEE, L.L. (1971) Northwestern Syntax Screening Test (NSST). Evanston, Ill: Northwestern University Press.

LEE, L.L. (1977) Reply to Arndt and Byrne. Journal of Speech and Hearing Disorders, 42, 323-327.

LEE, L.L. & CANTER, S.M. (1971) Developmental sentence scoring: a clinical procedure for estimating syntactic development in children's spontaneous speech. Journal of Speech and Hearing Disorders, 31, 311-330.

LENNEBERG, E.H. (1962) Understanding language without ability to speak: case report. Journal of Abnormal and Social Psychology, 65, 419-425.

LENNEBERG, E.H. (1967) Biological foundations of language. New York: John Wiley & Sons.

LENNEBERG, E.H., NICHOLAS, I.A. & ROSENBERGER, E.F. (1964) Primitive stages of language development in mongolism. In D.McK. Rioch & E.A. Weinstein (eds) Research Publications: Association for Research in Nervous and Mental Diseases, 42, 119-137.

LEONARD, L.B. (1972) What is deviant language? Journal of Speech and Hearing Disorders, 37, 427-446.

LEONARD, L.B. (1973) The nature of deviant articulation. Journal of Speech and Hearing Disorders, 38, 156-161.

LEONARD, L.B., BOLDERS, J.G. & MILLER, J.A. (1976) An examination of the semantic relations reflected in the language usage of normal and language disordered children. Journal of Speech and Hearing Research, 19, 371-392.

LEREA, L. (1958) Assessing language development. Journal of Speech and Hearing Research, 1, 75-85.

LIEBER, C.W. & SPITZ, H.H. (1976) Inference of word-meaning from syntax structure by normal children and retarded adolescents. Journal of Psychology, 93, 3-12.

LILES, B.Z., SHULMAN, M.D. & BARTLETT, S. (1977) Judgements of grammaticality by normal and language-disordered children. Journal of Speech and Hearing Disorders, 42, 199-209.

LOVELL, K. & DIXON, E.M. (1967) The growth of the control of grammar in imitation, comprehension and production. Journal of Child Psychology and Psychiatry, 8, 31-39.

LOVELL, K., HOYLE, H.W. & SIDDALL, M.Q. (1968) A study of some aspects of the play and language of young children with delayed speech. Journal of Child Psychology and Psychiatry, 9, 41-50.

LOWE, A.D. & CAMPBELL, R.A. (1965) Temporal discrimination in aphasoid and normal children. Journal of Speech and Hearing Research, 8, 313-314.

MARATSOS, M.P. (1974) Children who get worse at understanding the passive: a replication of Bever. Journal of Psycholinguistic Research, 3, 65-74.

MARK, H.J. & HARDY, W.G. (1958) Orienting reflex disturbances in central auditory or language handicapped children. Journal of Speech and Hearing Disorders, 23, 237-242.

MARQUARDT, T.P. & SAXMAN, J.H. (1972) Language comprehension and auditory discrimination in articulation deficient kindergarten children. Journal of Speech and Hearing Research, 15, 382-389.

MARTIN, J.G. (1972) Rhythmic (hierarchical) versus serial structure in speech and other behaviour. Psychological Review, 79, 487-509.

MASLAND, M.W. & CASE, L.W. (1968) Limitation of auditory memory as a factor in delayed language development. British Journal of Disorders of Communication, 3, 139-142.

MCGINNIS, M.A. (1963) Aphasic Children. Washington: Alexander Graham Bell Association for the Deaf, Inc.

MCNEILL, D. (1971) The capacity for the ontogenesis of grammar. In D.I. Slobin (ed) The Ontogenesis of Grammar. New York: Academic Press.

MCREYNOLDS, L.V. (1966) Operant conditioning for invesitgating speech sound discrimination in aphasic children. Journal of Speech and Hearing Research, 9, 519-528.

MCREYNOLDS, L.V. & HUSTON, K. (1971) A distinctive feature analysis of children's misarticulations. Journal of Speech and Hearing Disorders, 36, 155-166.

MENYUK, P. (1963) A preliminary evaluation of grammatical capacity in children. Journal of Verbal Learning and Verbal Behaviour, 2, 429-439.

MENYUK, P. (1964) Comparison of grammar of children with functionally deviant and normal speech. Journal of Speech and Hearing Research, 7, 109-121.

MENYUK, P. (1969) Sentences Children Use. Cambridge, Mass: MIT press.

MENYUK, P. (1973) The developmental implications of deviant language acquisition. In J.I. Nurnberger (ed) Biological and Environmental Determinants of Early Development. Research Publication of the Association for Research in Nervous and Mental Disease, 51. Baltimore: Williams & Wilkins.

MENYUK, P. (1974) Early development of receptive language: from babbling to words. In R.L. Schiefelbusch & L.K. Lloyd (eds) Language Perspectives - Acquisition, Retardation, and Intervention. Baltimore: University Park Press.

MENYUK, P. & LOONEY, P.L. (1972a) A problem of language disorder: length versus structure. Journal of Speech and Hearing Research, 15, 264-279.

MENYUK, P. & LOONEY, P.L. (1972b) Relationships among components of the grammar in language disorder. Journal of Speech and Hearing Research, 15, 395-406.

MILNER, B. (1974) Hemispheric specialization: scope and limits. In F.O. Schmitt & F.G. Worden (eds) The Neurosciences: Third Study Program. Cambridge, Mass: MIT press.

MILNER, B., BRANCH, C. & RASMUSSEN, T. (1966) Evidence for bilateral speech representation in some non-right handers. Transactions of the American Neurological Association, 91, 306-308.

MITTLER, P. (1972a) Psychological assessment of language abilities. Clinics in Developmental Medicine, 43, 106-119.

MITTLER, P. (1972b) Language development and mental handicaps. Clinics in Developmental Medicine, 43, 136-146.

MITTLER, P., JEFFREE, D., WHELDALL, K. & BERRY, P. (1974) Assessment and remediation of language comprehension and production in severely subnormal children. Unpublished MS.

MITTLER, P. & WARD, J. (1970) The use of the ITPA on British 4-year-old children: a normative and factorial study. British Journal of Educational Psychology, 40, 43-54.

MONSEES, E.K. (1968) Temporal sequence and expressive language disorders. Exceptional Children, 35, 141-147.

MOREHEAD, D.M. & INGRAM, D. (1973) The development of base syntax in normal and linguistically deviant children. Journal of Speech and Hearing Research, 16, 330-352.

MORLEY, M. (1972) Development and Disorders of Speech in Childhood. Edinburgh: Churchill Livingstone.

NELSON, K. & BENEDICT, H. (1974) The comprehension of relative, absolute and contrastive adjectives by young children. Journal of Psycholinguistic Research, 3, 333-342.

NEVILLE, B.G.R., BENTOVIM, A., CLAYTON, B.E. & SHEPHERD, J. (1972) Histidinaemia: study of relation between clinical and biological findings in 7 subjects. Archives of Disease in Childhood, 47, 190-200.

NOLL, J.D. & LASS, N.J. (1972) Use of the token test with children: two contrasting socio-economic groups. Paper presented at the American Speech and Hearing Association Annual Convention, San Francisco, November, 1972.

NURSS, J.R. & DAY, D.E. (1971) Imitation, comprehension and production of grammatical structures. Journal of Verbal Learning and Verbal Behaviour, 10, 68-74.

OLLER, D.K. (1973) Regularities in abnormal child phonology. Journal of Speech and Hearing Disorders, 38, 36-47.

OLSON, J.L. (1961) Deaf and sensory aphasic children. Exceptional Children, 27, 422-424.

ONIONS, C.T. (1956) The Shorter Oxford English Dictionary. London: Oxford University Press.

ORTON, S.T. (1937) Reading, Writing and Speech Problems in Children. New York: W.W. Norton.

OSGOOD, C.E. (1957a) A behaviouristic analysis of perception and language as cognitive phenomena. In J.S. Bruner, E. Brunswick, L. Festinger, F. Heider, K.F. Muenzinger, C.E. Osgood & D. Rapaport (eds) Contemporary Approaches to Cognition. Cambridge, Mass: Harvard University Press.

OSGOOD, C.E. (1957b) Motivational dynamics of language behaviour. In M.R. Jones (ed) Nebraska Symposium on Motivation. Lincoln, Neb: University of Nebraska Press.

PAGET, R., GORMAN, P. & PAGET, G. (1972) A Systematic Sign Language. London: Association for Experiment in Deaf Education, Ltd.

PALERMO, D.S. (1973) More about less: a study of language comprehension. Journal of Verbal Learning and Verbal Behaviour, 12, 211-221.

PANAGOS, J.M. & KING, R.R. (1975) Self and mutual speech comprehension by deviant- and normal-speaking children. Journal of Speech and Hearing Research, 18, 653-662.

PARASKEVOPOULOS, J.N. & KIRK, S.A. (1969) The Development and Psychometric Characteristics of the revised Illinois Test of Psycholinguistic Abilities. Urbana: University of Illinois Press.

PARISI, D. (1971) Development of syntactic comprehension in preschool children as a function of socioeconomic level. Developmental Psychology, 5, 186-189.

PARISI, D. & ANTINUCCI, F. (1970) Lexical competence. In G.B. Flores D'Arcais & W.J.M. Levelt (eds) Advances in Psycholinguistics. Amsterdam: North Holland Publishing Co.

PETRETIC, P.A. & TWENEY, R.D. (1977) Does comprehension precede production? The development of children's responses to telegraphic sentences of varying grammatical adequacy. Journal of Child Language, 4, 201-209.

PIAGET, J. (1970) Piaget's theory. In P.H. Mussen (ed) Carmichael's Manual of Child Psychology, vol. 1. New York: John Wiley.

POLLACK, E. & REES, N.S. (1972) Disorders of articulation: some clinical applications of distinctive feature theory. Journal of Speech and Hearing Disorders, 37, 451-461.

POPPEN, R., STARK, J., EISENSON, J., FORREST, T.& WERTHEIM, G. (1969) Visual sequencing performance of aphasic children. Journal of Speech and Hearing Research, 12, 288-300.

PREMACK, D. (1971) Language in chimpanzee? Science, 172, 808-822.

RATUSNIK, D.L. & KOENIGSKNECHT, R.A. (1975) Internal consistency of the Northwestern Syntax Screening Test. Journal of Speech and Hearing Disorders, 40, 59-68.

REES, N.S. (1973) Auditory processing factors in language disorders: the view from Procrustes' bed. Journal of Speech and Hearing Disorders, 38, 304-315.

REICH, T., CLONINGER, C.R. & GUZE, S.B. (1975) The multifactorial model of disease transmission: 1. Description of the model and its use in psychiatry. British Journal of Psychiatry, 127, 1-10.

REICHSTEIN, J. (1964) Auditory threshold consistency in differential diagnosis of aphasia in children. Journal of Speech and Hearing Disorders, 29, 147-155.

RENFREW, C.E. (1975) Action Picture Test. Oxford: C.E. Renfrew.

RENFREW, C.E. & GEARY, L. (1973) Prediction of persisting speech defect. British Journal of Disorders of Communication, 8, 37-41.

REYNELL, J.K. (1969a) Reynell Developmental Language Scales. Windsor: NFER Publishing Co.

REYNELL, J.K. (1969b) A developmental approach to language disorders. British Journal of Disorders of Communication, 4, 33-40.

REYNELL, J.K. & HUNTLEY, R.M.C. (1971) New scales for the assessment of language development in young children. Journal of Learning Disabilities, 4, 549-557.

RODD, L.J. & BRAINE, M.D. (1971) Children's imitations of syntactic constructions as a measure of linguistic competence. Journal of Verbal Learning and Verbal Behaviour, 10, 430-443.

RONDAL, J.A. (1976) Tests and nontests in psychology in general and in language assessment in particular: some reflections. Scientia Paedagogica Experimentalis, 13, 68-85.

ROSE, F.C. (1972) Speech and language disorders in children. Transactions of the Medical Society (London), 88, 72-76.

RUEDA, R. & PEROZZI, J.A. (1977) A comparison of two Spanish tests of receptive language. Journal of Speech and Hearing Disorders, 42, 210-215.

RUTTER, M., BARTAK, L. & NEWMAN, S. (1971) Autism - a central disorder of cognition and language? In M. Rutter (ed) Infantile Autism: Concepts, Characteristics and Treatment. Edinburgh: Churchill Livingstone.

RUTTER, M. & YULE, W. (1973) Specific reading retardation. In L. Mann & D. Sabatino (eds) The First Review of Special Education. New York and London: Grune & Stratton.

SATO, S. & DREIFUSS, F.E. (1973) Electroencephalographic findings in a patient with developmental expressive aphasia. Neurology, 23, 181-185.

SATZ, P. (1977) Laterality tests: an inferential problem. Neuropsychologia, 13, 208-212.

SCHOLNICK, E.K. & ADAMS, M.J. (1973) Relationships between language and cognitive skills: passive-voice comprehension, backward repetition and matrix permutation. Child Development, 44, 741-746.

SCOTT, C.I. (1976) Achondroplastic and hypochondroplastic dwarfism. Clinical Orthopaedics and Related Research, 114, 18-30.

SECKEL, H.P.G. (1960) Birdheaded Dwarfs: Studies in Developmental Anthropology Including Human Proportions. Springfield, Ill: C. Thomas.

SHAPIRO, T., ROBERTS, A. & FISH, B. (1970) Imitation and echoing in young schizophrenic children. Journal of the American Academy of Child Psychiatry, 9, 548-567.

SHAW, H.J., MATHEWS, C.G. & KLOVE, H. (1966) The equivalence of WISC and PPVT IQs. American Journal of Mental Deficiency, 70, 601-606.

SHERIDAN, M.D. (1973) Children of seven years with marked speech defects. British Journal of Disorders of Communication, 8, 9-16.

SHIPLEY, E.F., SMITH, C.S. & GLEITMAN, L.R. (1969) A study in the acquisition of language: free responses to commands. Language, 45, 322-342.

SHOUMAKER, R.D., BENNETT, D.R., BRAY, P.F. & CURLESS, R.G. (1974) Clinical and EEG manifestations of an unusual aphasic syndrome in children. Neurology, 24, 10-16.

SHRINER, T.H., HOLLOWAY, M.S. & DANILOFF, R.G. (1969) The relationship between articulatory deficits and syntax in speech defective children. Journal of Speech and Hearing Research, 12, 319-325.

SINCLAIR, A., SINCLAIR, H. & DE MARCELLUS, O. (1971) Young children's comprehension and production of passive sentences. Archives de Psychologie, 41, 1-22.

SLEIGH, G. (1970) A study of some symbolic processes in young children. British Journal of Disorders of Communication, 5, 163-175.

SLOBIN, D.I. (1973) Cognitive prerequisites for the development of grammar. In C.A. Ferguson & D.I. Slobin (eds) Studies of Child Language Development. New York: Holt, Rinehart & Wilson, Inc.

SLOBIN, D.I. & WELSH, C.A. (1973) Elicited imitation as a research tool in developmental psycholinguistics. In C.A. Ferguson & D.I. Slobin (eds) Studies of Child Language Development. New York: Holt, Rinehart & Wilson, Inc.

SMITH, C.R. (1967) Articulation problems and ability to store and process stimuli. Journal of Speech and Hearing Research, 10, 348-353.

SOMMERS, R.K. & TAYLOR, M.L. (1972) Cerebral speech dominance in language-disordered and normal children. Cortex, 8, 224-232.

SPELLACY, F.J. & SPREEN, O. (1969) A short form of the Token Test. Cortex, 5, 390-397.

SPRADLIN, J.E. (1974) Discussion summary - development of receptive language. In R.L. Schiefelbusch & L.K. Lloyd (eds) Language Perspectives - Acquisition, Retardation & Intervention. Baltimore: University Park Press.

SPRINGER, S.P. & EISENSON, J. (1977) Hemispheric specialization for speech in language-disordered children. Neuropsychologia, 15, 287-293.

STANTON, A.M. (1976) A qualitative assessment of comprehension and imitation in language delayed preschool children in comparison with normal children of the same age. British Journal of Disorders of Communication, 11, 63-71.

STARK, J. (1966) Performance of aphasic children on the ITPA. Exceptional Children, 33, 153-158.

STARK, J. (1967) A comparison of the performance of aphasic children on three sequencing tests. Journal of Communication Disorders, 1, 31-34.

STARK, J., POPPEN, R. & MAY, M.Z. (1967) Effects of alterations of prosodic features on the sequencing performance of aphasic children. Journal of Speech and Hearing Research, 10, 849-853.

STARR, S. (1974) Discrimination of syntactical errors in children under two and one half years. Developmental Psychology, 10, 381-386.

STEWART, A. (1968) Disorders of articulation in children. British Journal of Disorders and Communication, 3, 121-129.

STROHNER, H. & NELSON, K. (1974) The young child's development of sentence comprehension: influence of event probability, nonverbal context, syntactic form and strategies. Child Development, 45, 567-576.

SUGAR, O. (1952) Congenital aphasia: an anatomical and physiological approach. Journal of Speech and Hearing Disorders, 17, 301-304.

TALLAL, P. (1975) Perceptual and linguistic factors in the language impairment of developmental dysphasia: an experimental investigation with the Token Test. Cortex, 11, 196-205.

TALLAL, P. (1976) Rapid auditory processing in normal and disordered language development. Journal of Speech and Hearing Research, 19, 561-571.

TALLAL, P. & PIERCY, M. (1973a) Defects of non-verbal auditory perception in children with developmental aphasia. Nature, 241, 468-469.

TALLAL, P. & PIERCY, M. (1973b) Developmental aphasia: impaired rate of non-verbal processing as a function of sensory modality. Neuropsychologia, 11, 389-398.

TALLAL, P. & PIERCY, M. (1974) Developmental aphasia: rate of auditory processing and selective impairment of consonant perception. Neuropsychologia, 12, 83-94.

TALLAL, P. & PIERCY, M. (1975) Developmental aphasia: perception of brief vowels and extended stop consonants. Neuropsychologia, 13, 69-74.

TALLAL, P. & STARK, R.E. (1976) Relation between speech perception and speech production impairment in children with developmental dysphasia. Brain and Language, 3, 305-317.

TANZ, C. (1977) Learning how "it" works. Journal of Child Language, 4, 224-235.

TAYLOR, D.C. & OUNSTED, C. (1972) The nature of gender differences explored through ontogenetic analyses of sex ratios in disease. In C. Ounsted & D.C. Taylor (eds) Gender Differences: their Ontogeny and Significance. Edinburgh & London: Churchill Livingstone.

TAYLOR, I.G. (1964) The Neurological Mechanisms of Hearing and Speech in Children. Manchester: Manchester University Press.

TAYLOR, L.J. (1975) The Peabody Picture Vocabulary Test: what does it measure? Perceptual and Motor Skills, 41, 777-778.

TERVOORT, B. (1970) The understanding of passive sentences by deaf children. In G.B. Flores D'Arcais & W.J.M. Levelt (eds) Advances in Psycholingustics. Amsterdam: North Holland Publishing Co.

TURNER, E.A. & ROMMETVEIT, R. (1967) The acquisition of sentence voice and reversibility. Child Development, 38, 649-660.

VAN DONGEN, H.R. & LOONEN, M.C.B. (1977) Factors related to prognosis of acquired aphasia in children. Cortex, 13, 131-136.

WARREN, R.M. & ACKROFF, J.M. (1976) Two types of auditory sequence perception. Perception and Psychophysics, 20, 387-394.

WARYAS, C. & RUDER, K. (1974) On the limitations of language comprehension procedures and an alternative. Journal of Speech and Hearing Disorders, 39, 44-52.

WEBER, J.L. (1970) Patterning of deviant articulation behaviour. Journal of Speech and Hearing Disorders, 35, 135-141.

WECHSLER, D. (1967) Manual: Wechsler Preschool and Primary Scale of Intelligence. New York: Psychological Corporation.

WECHSLER, D. (1974) Manual: Wechsler Intelligence Scale for Children - Revised. New York: Psychological Corporation.

WEENER, P., BARRITT, L.S. & SEMMEL, M.I. (1967) A critical evaluation of the ITPA. Exceptional Children, 33, 373-380.

WEINER, P.S. (1967) Auditory discrimination and articulation. Journal of Speech and Hearing Disorders, 32, 19-29.

WEINER, P.S. (1969a) The perceptual level functioning of dysphasic children. Cortex, 5, 440-457.

WEINER, P.S. (1969b) The cognitive functioning of language deficient children. Journal of Speech and Hearing Research, 12, 53-64.

WETSTONE, H.S. & FRIEDLANDER, B.Z. (1973) The effect of word order on young children's responses to simple questions and commands. Child Development, 44, 734-740.

WHITACRE, J.D., LUPER, H.L. & POLLIO, H.R. (1970) General language deficits in children with articulation problems. Language and Speech, 13, 231-239.

WHITAKER, H.A. & NOLL, J.D. (1972) Some linguistic parameters of the Token Test. Neuropsychologia, 10, 395-404.

WILCOX, S. & PALERMO, D.S. (1974) "In", "on" and "under" revisited. Cognition, 3, 245-254.

WILSON, L.F., DOEHRING, D.G. & HIRSCH, I.J. (1960) Auditory discrimination learning by aphasic and nonasphasic children. Journal of Speech and Hearing Research, 3, 130-137.

WITELSON, S.F. & RABINOVITCH, M.S. (1972) Hemisphere speech lateralization in children with auditory-linguistic deficits. Cortex, 8, 412-426.

WITHROW, F.B. (1964) Immediate recall by aphasic, deaf, and normal children for visual forms presented simultaneously or sequentially. Asha, 6, 386.

WOODS, B.T. & TEUBER, H.L. (1973) Early onset of complementary specialization of cerebral hemispheres in man. Transactions of the American Neurological Association, 98, 113-115.

WORSTER-DROUGHT, C. (1943) Congenital auditory imperception (congenital word-deafness) and its relation to idioglossia and allied speech defects. Medical Press and Circular, 210, 411-417.

WORSTER-DROUGHT, C. (1968) Speech disorders in children. Developmental Medicine and Child Neurology, 10, 427-440.

WORSTER-DROUGHT, C. (1971) An unusual form of acquired aphasia in children. Developmental Medicine and Child Neurology, 13, 563-571.

WORSTER-DROUGHT, C. & ALLEN, I.M. (1929) Congenital auditory imperception (congenital word-deafness): with report of a case. Journal of Neurology and Psychopathology, 9, 193-208.