NOTING THAT RECENT RESEARCH POINTS TO LINGUISTIC CODE DIFFERENCES AS AN IMPORTANT FACET OF CULTURAL DEPRIVATION AND THAT THE MAJORITY OF EDUCABLE MENTALLY RETARDED (EMR) CHILDREN COME FROM LOW SOCIOECONOMIC LEVELS, IT WAS HYPOTHESIZED THAT ONE OF THE KEYS TO LEARNING FOR EMR CHILDREN IS THE COMPLEXITY OF THE TEACHER'S LANGUAGE IN RELATION TO THAT OF THE CHILDREN AND THAT THE LINGUISTIC CODE MAY BE RELATED TO THE FAILURE OF SPECIAL CLASSES TO SHOW ANY SIGNIFICANT BENEFIT FOR THE STUDENT. A SYSTEM OF ANALYZING AND CODING TEACHER-PUPIL VERBAL INTERACTION WAS DEVELOPED BY USING TAPESCRPTS FROM TWO EMR CLASSES WITH A TOTAL OF 19 CHILDREN (MEAN IQ = 117 MONTHS). TEACHERS OF BOTH CLASSES HAD BEEN SPECIAL EDUCATION MAJORS IN COLLEGE AND HAD 7 YEARS EXPERIENCE BETWEEN THEM WITH THE EMR. DIFFERENT METHODS OF WORD ANALYSIS WERE ASSESSED. THE SYSTEM DEVELOPED HAS THREE PARTS--THE LENGTH AND COMPLEXITY INDEX (LCI), THE SENTENCE CLASSIFICATION SYSTEM, AND A CONTENT ANALYSIS. THE LCI IS A SCORING SYSTEM WHICH SIMULTANEOUSLY WEIGHTS THE LENGTH OF A SENTENCE AND THE COMPLEXITY OF ITS NOUN AND VERB PHRASES. SENTENCES ARE CLASSIFIED INTO ONE OF FIVE CLASSES--SIMPLE, SIMPLE WITH A PHRASE, ELABORATED SIMPLE, COMPOUND AND COMPLEX, AND ELABORATED COMPOUND AND COMPLEX SENTENCES. THESE FIVE CLASSES FORM A LOGICAL HIERARCHY OF COMPLEXITY AND ARE DIFFERENTIALLY WEIGHTED. A SENTENCE COMPLEXITY SCORE (SCS) IS OBTAINED BY ADDING THE LCI AND THE WEIGHTED VALUE FOR THE CLASS OF SENTENCE. THE CONTENT OF A SENTENCE IS CODED ACCORDING TO THE PURPOSE OF THE COMMUNICATION INTO EIGHT CATEGORIES GROUPED INTO THREE MAIN CATEGORIES (STRUCTURING, RESPONSES REQUESTED, AND REACTING) USED AS PARAMETERS FOR THE INTERPRETATION OF THE SCS. ONE FIGURE AND A 26-ITEM BIBLIOGRAPHY ARE INCLUDED. APPENDIXES PROVIDE INSTRUCTIONS FOR THE LINGUISTIC CODING SYSTEM AND A STUDY OF THE VERBAL OUTPUT IN A SPECIAL CLASS FOR RETARDED CHILDREN. ACCOMPANYING THE APPENDIXES ARE A CODING SHEET, A TABLE, AND THREE REFERENCES. (AA/JD)
FINAL REPORT
Project No. 7-0130
Grant No. GEC-3-7-070130-1609
PA-440

LINGUISTIC ANALYSIS OF VERBAL INTERACTION IN SPECIAL CLASSES FOR THE MENTALLY RETARDED

June 1967

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Office of Education
Bureau of Research
Linguistic Analysis of Verbal Interaction in Special Classes for the Mentally Retarded

Project No. 7-0130
Grant No. OEG-3-7-070130-1609

Oliver L. Hurley

June 1967

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, represent official Office of Education position or policy.

Institute for Research on Exceptional Children
University of Illinois
Urbana, Illinois
ACKNOWLEDGEMENTS

The author wishes to express thanks to Dr. Esther Minskoff, Department of Special Education, Ferkauf Graduate School, Yeshiva University, for her graciousness in allowing the use of her tapescripts for this project. Gratitude is also expressed to Dr. Thomas Shriner, Children's Research Center and Department of Speech, University of Illinois, for the hours he spent discussing this project and the problems involved. Mr. Jerome Albert also contributed substantially to the project as research assistant.
BACKGROUND OF THE PROJECT

The impetus for this project stemmed from a dissatisfaction with the results of efficacy studies. These studies, as a group, reveal that educable mentally retarded (EMR) children in special classes neither achieve better nor adjust better than similar children remaining in regular classes. Theoretically, special class placement is beneficial to EMR children academically, socially, and emotionally. Yet efficacy studies do not clearly show any superiority of such special placement over regular class placement for the EMR. These studies are well known and will not be reviewed here. The reader is referred to Johnson's article (14) devoted to this topic or to Kirk's review of educational research (16). It is sufficient to note that there is a confusing array of research results which cannot be tolerated if the theory and philosophy underlying the special class movement are correct.

Why aren't special classes more effective? Why do not EMR children achieve closer to their "potentials?" These were the questions which motivated the undertaking of this project. The answer must lie in the classroom, in the "educational treatment" received by the EMR. The answer proposed is that one of the keys to learning for EMR children is the complexity of the teacher's language in relation to that of the children.

Preliminary to a formal investigation of the linguistic nature of teacher-pupil verbal interaction, it was necessary to develop a system of analysis. This report concerns itself primarily with the analytic system developed and with the rules for coding data. In order to put the coding system in focus, a brief review of pertinent literature and implications for the EMR will be presented. This will be followed by a description and explanation of the coding system developed.

Language Studies

This review will not be exhaustive; only some of the more representative studies will be summarized. They will be looked at from the socio-economic point-of-view. It is not surprising that evidence of linguistic differences between socio-economic levels have existed for over thirty years. It is surprising, however, that educators and school personnel have taken so little notice of the fact until recently.

McCarthy (17) studied the language of 140 pre-school children from six occupational levels. She found significant differences in (a) length of responses, (b) adapted information (naming, remarks
about or associated with immediate stimulus situations), (c) number of questions asked, (d) percentage of simple sentences, simple sentences, simple sentences with a phrase, complex and compound and elaborated sentences - all in favor of the upper occupational levels. The lower levels exhibited more incomprehensible speech and functionally complete but structurally incomplete sentences. When the children were grouped according to MA, the differences were still evident even though less striking.

Smith (24) studied children from three occupational levels roughly equivalent to professional, skilled and unskilled. Most marked differences were found in sentence length at all age levels. The high classes also asked more questions.

Davis (8) analysed the language of 166 twins, 97 only, and 173 singleton children at ages 5 1/2, 6 1/2 and 9 1/2. She summarized, "Children from the upper occupational groups are definitely superior to children from lower occupational groups in every phase of language ability." The elements she looked at were the same as those of McCarthy (above).

Templin (26) examined the articulation, sound discriminations, sentence development and vocabulary of 422 children from six occupational levels in eight discrete age groups from 3 to 8. She found consistent differences in favor of the upper socio-economic groups. The differences were significant in only 29%, 68 of the 230 comparisons. However, the lower groups exceeded the upper in only 13 instances. The lower occupational level children were significantly inferior in the articulation of vowels, grammatical complexity and in the vocabulary of recognition at the older age levels. None of these variables were significant at all age levels.

Pringle and Tanner (21) compared the speech development of 18 pairs of 4 year old children matched on sex, age, IQ and home background. One member of each pair lived in a residential institution, equated with early deprivation. They found qualitative differences between the groups. On the first five words of the WISC vocabulary, the nursery children gave definitions which tended to be in terms of active participation whereas the residential children used a "more passive mode of description." Even though there was no difference in types of words used in a free play situation, those verbs used exclusively by the nursery children tended to be active and aggressive (dance, push), whereas those used only the residential children "suggest a rather docile, helpful attitude" (mend, worry). They also found that there was more communication between the nursery school children, accompanied by 20% fewer remarks directed to adults and to "no one." The nursery children showed significantly more imaginative interpretation and pure make-believe than did the residential children.
On the basis of these and similar studies, Bernstein (2,3,4) proposes two types of language codes, elaborate and restricted. He regards these codes to be functions of differing social relationships. These social relationships correspond roughly to differences between the upper and lower social classes. He presents evidence of significant differences in language between the upper and working classes. Lower class children learn a linguistic code characterized by its paucity of content, specificity, and explicitness. Some characteristics are (a) short grammatically simple sentences, (b) little use of subordinate clauses, (c) rigid and limited use of adjectives and adverbs, and (d) frequent use of statements where the reason and conclusion are telescoped into one categoric statement (2, p. 297). The elaborated code of middle class children is characterized by the varied use of subordinate clauses, of adjectives and adverbs, and of sentence types.

Bernstein sees the different language codes as resulting in qualitatively different structuring of the individual's universe. Of a restricted language code, he says, "this radically narrows the extent and type of his object relationships. Thus, the middle class child and the lower working class child, are oriented to different orders of learning as a result of the implications of the forms of their language use." The elaborated code allows and provides for the handling and learning of the interrelationship of ideas and human contingencies since it does contain the variety and specificness needed to linguistically characterize the uniqueness of a particular event or relationship. (This ability to individuate an event or relationship is called the "individual qualification" by Bernstein"). In addition, "the working class child has to translate and thus, mediate middle class language structure through the logically simpler language structure of his own class to make it personally meaningful. Where he cannot make his translation, he fails to understand and is left puzzled." (2, p. 293)

Hess and Shipman (13) applied Bernstein's theory of social learning to the analysis of mother-child interaction in an attempt to obtain clues to "early blocks to children's learning". After recording and analysing the language used by mothers to instruct their children on certain structured tasks, they conclude

The picture that is beginning to emerge is that the meaning of deprivation is a deprivation of meaning -- a cognitive environment in which behavior is controlled by status rules (result of the lack of "individual qualification") rather than by attention to the individual characteristics of a specific situation and one in which behavior is not mediated by verbal cues or by teaching which relates events one to another and the present to the future. The environment produces a child who relates to authority rather than to
rationale, who although often compliant is not reflective in his behavior; and for whom the consequences of an act are largely considered in terms of immediate punishment and reward rather than future effects and long range goals (p. 194).

Support for their interpretation is found in Pringle and Tanner (21) study cited above.

Implications for the EMR

Since the majority of the EMH come from low socio-economic levels (20, p. 4 and 8), the above cited research applies to them. It should follow therefore, that the majority of the EMR are users of a restricted language code. Since the lower classes must translate the middle class communications of their teachers into their simpler language code in order to understand, it follows that language and language training is an important variable in the education of the EMH child.

If mentally retarded children use a restricted language and the teacher uses primarily an elaborated one, the result would be a breakdown in communication and cognitive growth. Cognitive growth requires the development of ideas and concepts. Conceptual development, in turn, requires a language with which relationships and contingencies can be precisely explicated. In addition, the restricted language code relies heavily on the authority or the status of the speaker rather than on reasons for justification of rules and behavior. Thus, the children are not encouraged to reason nor to plan or follow through on self-initiated activities; questioning is almost non-existent.

On the other hand, mix these children without specific language training with a teacher who uses primarily a restricted code and the result is likewise negative. The children are not directly helped to learn to see those relationships and behavioral contingencies within their ability to learn, but expressable only by means of an elaborated language code.

This, then was the frame of reference within which the linguistic analytic system was developed. The rest of this report will be devoted to an explanation of the system.
METHOD

The coding system was developed using two tapescripts of EMR classes. The classes had been audiotaped while an observer recorded board work and other pertinent non-verbal actions. These were then combined into one tapescript for each class. Only the academic subjects taught by the EMR teacher herself were recorded. This section will describe the EMR classes, discuss the criteria used in the development of the linguistic coding system, and illustrate the procedure used.

The Classes

The two classes were considered typical of 17 such classes for which tapescripts were available. Class A contained nine children with an average CA of 130 months, average MA of 86 months, and an average IQ of 67. The teacher had two years teaching experience, all of it with the EMR.

Class B contained ten children with an average CA of 107 months, average MA of 65 months, and average IQ of 61. The teacher had five years teaching experience, all of it with the EMR.

In summary, there was a total of 19 children in both classes. Mean IQ was 63. Mean CA was 117 months. Both teachers had been special education majors in college and had seven years experience with the EMR between them.

Criteria Used in the Development of the Linguistic Coding System

1. The system should reflect the complexity and structure of both phrases and sentences. Since the differences, as stated earlier, between elaborated and restricted codes resides in grammatical complexity and sentence structure, this criterion is very important. It was found that the simple analysis of noun and verb phases (6) or a combination of length, noun and verb phases (22) did not adequately reflect the overall structure of the sentence (15), although there is a significant correlation between the latter and the first two. It was found that both simple and complex sentences could receive the same weights. This highlighted the necessity for this criteria.

2. The system should be applicable to both teacher and pupil language. Since the ultimate purpose of this system is to facilitate comparisons of language code, the relevance of this criterion is evident. Most research has compared or described the language of children of the same age or of adults. Seldom has one attempted to compare children with adults. Thus, there are methods of looking at child language profitably which become useless with adults.
One such method is "sentence length", which does not discriminate among adults although it is a good index of language development up to about nine years of age.

3. The system should be uncomplicated enough to be used by a teacher herself. The research program of which this project was a part aimed at discovering clues to curriculum and instructional changes for the EMR. Because of this, it was felt important that teachers be able to easily understand any data resulting from the use of the system of analysis. Additionally, it was thought that this system or one like it could be used as a means of measuring the effects of language teaching in EMR classes. As such, simplicity is paramount if the average teacher, with her premium on time, is to consider using the system.

These three criteria served to channel the direction of the development of the linguistic coding system. They were considered sufficient for the purpose of defining and comparing elaborated and restricted linguistic codes. Since this type of developmental project depends a great deal on the nature of the raw data (i.e., the tapescripts), freedom to follow directions indicated by the data had to be maintained. Additional criteria would have limited this freedom.

**Procedures**

The basic procedure used was one of trying-out different methods of analysis and applying the criteria to them.

To illustrate, mean length of response, mean of five longest responses, number of one-word responses, and structural complexity classification of sentences (15, pp. 167-173) were tried out. Even though these measures yielded interesting information (See Appendix B), they failed to meet the criteria.

Mean length of response and mean of five longest responses do not reflect complexity of sentence and lose their meaning in the analysis of teacher productions. In the classroom situation, these measures and the number of one word responses lose their discriminative ability because of the vast preponderance of one word responses (Appendix B), situationally prescribed, which tends to produce spuriously high or low measures. The structural complexity classification was at first inadequate also because of the exceedingly large number of elliptical utterance; i.e., functionally complete but structurally incomplete responses. This method has also been heavily criticized for the arbitrary weights assigned to the different classes (7, 19, 22).
Shriner's method (22) was found to suffer from some of the same problems. Nevertheless, it seemed to be the single most sensitive measure. It was, therefore, used as a base for the development of the system of analysis. Modifications dictated by the classroom tapescript data were made and other methods integrated with it.
The system of linguistic analysis is a composite of Shriner's (22) system, and measure of sentence structure (15, p. 170) and Hess and Shipman's (12) content analysis. Directions for coding and a sample coding sheet are contained in Appendix A.

This coding system yields a numerical value called the Sentence Complexity Score (SCS) for each sentence analysed. The SCS is composed of four part-scores which may be treated separately. These parts are a) noun phrase points, b) verb phrase points, c) points for additional words in a sentence, and d) points for the type of sentence. The sum of the first three (a + b + c) yields another score called the Length and Complexity Index (LCI). Common statistical techniques may be applied to these scores. The content analysis does not yield a numerical score. It is intended to provide a framework within which the SCS can be interpreted.

After a discussion of the various parts of the coding system, special problems in the use of the system will be discussed.

Length and Complexity Index (LCI)

This is a modification by Shriner (22) of Cazden's (6) earlier work with noun and verb phrase complexity. The LCI is a combination of this phrase complexity with the length of response measure. As the name indicates it is a single number, the sum of the weights given to parts of a sentence.

In the adaptation used here, copulatives and negatives are counted and a weight of one given to additional words which are important to that sentence. Whereas Cazden did not count and Shriner didn't bother with infinitives, these are counted and the infinitive marker to is considered part of the verb and not counted. Cazden avoided the problem of prepositions by counting only the noun phrases themselves. Shriner treated them arbitrarily.* Assigning a weight of 1 to prepositions tends to lower the LCI. However, not counting prepositions, the marker of subordinate phrases and clauses, seemed indefensible in the light of the hypotheses of the study. An arbitrary decision was made, therefore, to count and score each preposition except the infinitive marker and the elliptical ending a sentence. Other modifications made were to count any adjectives or pronouns serving nominative functions as nouns phrases, to count all negatives, and to add certain noun phrase and verb phrase constructions found in the data.

* Personal communication
Shriner's (22) study revealed that the LCI is significantly and highly correlated with measures of mean length of response and structural complexity and these three variables were significant predictors of language development (as rated by independent judges) for the children from 2-6 to 12 years of age. Kazden's noun phrase index is based on "emergence" (5, 18) of the structures in the developmental period and the verb phrase index is based solely on complexity. Thus the LCI can be considered to be a good reflection of complexity of language.

In coding a sentence, noun phrases, verb phrases and additional words are coded separately. The scores are then added to produce the LCI value. If one desires, therefore, one can study the noun phrase or verb phrase level of development as well.

Interjudge reliability for the LCI was briefly investigated using a total of 62 complete sentences from tapscript A. Each sentence was scored for total LCI points, number of noun phrases, noun phrase points, number of verb phrases, and verb phrase points. This produced a total of 310 numerical values. Before discussion there was 92 per cent agreement. After discussion there was 99 per cent agreement, the judges continuing to disagree on 2 items.

To give an example of the utility of the LCI, figure 1 is presented here. The LCI proved to be sensitive to teacher-pupil differences. The children's complete sentences (N=141) from both tapescripts are compared to an equal number of randomly selected teacher sentences. The differences between the means of the pupils (5.05) and teachers (7.85) are significant at the .005 level. The medians (4.86 vs. 7.19) also differ significantly. It is obvious that the variances (4.19 vs. 16.07) differ. These same teacher-pupil differences were observed in each tapescript individually.

In spite of this sensitivity to complexity it became apparent that the LCI was not reflecting the linguistic relationship between the verb and noun phrases; i.e., the sentence structure.

For example, consider the two sentences.

a. "That is the way we do it."
b. "Did you buy the new one?"

According to Shriner's system sentence a receives a score of 7 while sentence b receives a higher score of 8. However, sentence b is a simple sentence whereas a is a complex sentence. To overcome this lack, the measure of sentence structure (15, p. 170) was added.
Figure 1
Comparison of Length and Complexity Index of Teachers and Pupils
Type of Sentence

This is the measure of language structure of Johnson et al. (15) somewhat modified in terms of the weighting of the different classes of sentences. The classification is in the nature of a logical hierarchy of complexity. The sentence (classes) types, in order of complexity, are a) simple, b) simple with phrase, c) elaborated simple, d) compound and complex, and e) elaborated compound and complex sentences. These types are defined in "Instructions for Coding", Appendix A. The compound and complex sentences are in the same class for two reasons. First, the number of either one or the other type is relatively small (11 out of 104 teacher sentences in a Social Studies class). Secondly, using the LCI as the independent variable, the means of compound and complex sentences do not differ even when differentially weighted.

As indicated in Appendix A, the scores 0 (simple), 5 (simple with phrase), 10 (elaborated simple sentence), 15 (compound and complex), 20 (elaborated compound and complex) are added to the LCI and provide a total value for each sentence. These weights are arbitrary. They were chosen because when added to the LCI, the mean total sentence complexity score (SCS) increases noticeably from class to class and greatly reduces the overlap between classes. For example, for the five classes of sentences in a Social Studies lesson from tape script B, the mean SCS's are 4,93, 12,50, 18,90, 24,86, 40,50. The SCS ranges are as follows: Simple, 2 to 9; Simple with phrase, 10 to 14; elaborated simple, 16 to 19 plus one score of 32; compound and complex, 22 to 29; elaborated compound and complex, 38 to 44. If supported by future research, such a distribution of scores would allow direct interpretation of the SCS values.

The same Social Studies lesson provides a further example of the importance of classifying sentence types independently. The mean LCI for the teacher sentences to which the children verbally responded was 6,11 which was not much different from the 6,44 LCI mean for all of the teacher's sentences during the lesson. Intuitively this didn't seem to reflect the teachers linguistic code when she was requesting a response. Classifying the sentences according to type revealed that of 43 sentences only 2 were not any of the three categories of simple sentences, 34 were merely simple sentences. The resultant mean SCS was 8,21 as compared with a 10,09 SCS mean for all teacher sentences in the lesson. Referring to the distribution of scores above, certain inferences can be now made; a) the teacher tends to reserve her more complex sentence structives for "structuring", b) she tends to use relatively uncomplicated sentences throughout her teaching but even more so when questioning the pupils. One wonders how social studies concepts involved in a unit on "Self" can be effectively taught using simple sentences.
In summary, the additional coding and additive weighting of sentence types serves to more adequately reflect the level of complexity of the linguistic code being used by the speaker.

Content

The content analysis used here is a very simple one. It follows the model provided by Hess and Shipman (12). Unlike the previous categories discussed, these are not weighted. The teachers seemed to use sentences of differing complexity depending upon the purpose of the communication. This classification resulted from an investigation of these differences.

The categories seem to fit the usual teacher-pupil verbal interplay in the classroom. The main teacher categories are Structuring, Response Requested, Reacting. These are further subdivided into the categories shown in Appendix A. The student categories follow the same model.

The intention is to use these categories as parameters to aid in the interpretation of the SCS. For example, as discussed in the preceding section, it was found that the teacher tended to use much more complex sentences when structuring than when requesting a response. A careful qualitative analysis of the tapescript revealed that most often the children need not attend or understand the "structuring" sentences in order to answer the questions posed. In addition it was noted that the "reacting" utterances of the teacher tended to be very short elliptical expressions. This revealed that very little teacher "expansion" or correct "modeling" occurred (5).

Used in this qualitative way, it is the author's opinion that this content analysis coupled with the SCS yields useful and pertinent information about the linguistic code of teachers and pupils in a classroom situation.

Application of the Coding System to Tapescript

Using the coding system is fairly easy. One first identifies each sentence and elliptical expression. These are numbered consecutively beginning with the number 1 on each page. Next expand all ellipses. Then one underlines the noun and verb phrases in each sentence. One then simply follows the directions for coding in Appendix A.

There are still certain problems which arise in the use of this coding system. The most pressing of these problems is that of how to handle elliptical expressions. Elliptical expressions are utterances which are functionally complete but grammatically incomplete. Ellipses are best explained through an illustration. Consider the following conversation.
Teacher: How many legs are in the picture
Pupil: 2
Teacher: Yes. Very good....

Here the pupil's response, "2", is elliptical. The child means "There are 2 legs in the picture." The meaning of his response resides in the teacher's question and his meaning is perfectly understood by the teacher and anyone hearing this snatch of conversation. Likewise the teacher's response "Yes. Very good!" is elliptical. Both can be expanded into complete sentences which would take their meaning and form from the situation and context in which they were uttered. The common "Yes" or "no" responses are ellipses.

Gunter (11) identifies two types of ellipses, telegraphic and contextual. Telegraphic ellipsis occurs when one omits certain function words in writing or speaking; e.g., "Arrive 14:30 Thursday" which means "I will arrive at 2:30 P.M. Thursday." The reader will recognize the form as that used in writing telegrams; hence, its name. Contextual ellipses would be undecipherable without knowing something about the antecedent events, the context in which they occur. This is demonstrated in the example above. Omit the teacher's question and the child's response is rendered meaningless.

Contextual ellipsis is used by everyone in daily speech with no problem. However, because 80 per cent of the children's classroom verbal productions were elliptical, they presented a special problem. Unfortunately, the ellipsis has been neglected by linguists and there is no good way of analysing the ellipsis as such (11). Since Gunter reports almost 100% agreement in the expansion of ellipses by different judges and since a reliability check by the author yielded 85% agreement, it was decided to expand the ellipses and code them as complete sentences. It should be noted that the 15% disagreement was of the kind where one judge used the word "that" and the other judge used "it". Differences of this sort do not affect the SCS value at all.

The expansion and coding of children elliptical expressions into the shortest possible complete sentences serves as a good reflection of the language used by the teacher to elicit their response. Even though this remains to be substantiated some preliminary data revealed, among other differences, significant differences in both means and variances between pupil complete sentences and the expansions of their ellipses. The expansion of the teacher's ellipses, following the same rule of parsimony, did not show mean differences when compared with the pupils' complete sentences. Even though these data are not conclusive, they certainly warrant further consideration.
The point is that ellipses have to be handled in some manner. This coding system handles them in a way which seems to illuminate the nature of the linguistic code used by teachers and pupils in the classroom.

Another problem is a mechanical one resulting from how the audiotapes were transcribed. In reading the tapescripts, there are certain characteristics which seem to stand out but for which no analysis is available. The most striking of these characteristics is the frequency with which teachers re-word their questions. Since pauses and duration of pauses are not indicated on the tapescripts used, there is no way of knowing whether or not the restating is due to the absence of student response. This becomes important because the subjective impression is that the successive rewordings become simpler in structure. This cannot be verified until pauses and their duration are known. It is recommended, therefore, that in transcribing audiotape data, due note be taken of pauses and their duration.

Summary

The parts of the coding system have been discussed. These elements—the length and complexity index, the sentence classification scheme, and context analysis—seem to satisfy the criteria discussed earlier. It is believed that this system is sensitive to differences in linguistic code and would yield information of value to teachers in planning developmental and remedial language programs. The steps in coding and problems of ellipsis and transcription were discussed.
SUMMARY AND DISCUSSION

This report is primarily concerned with a system of analysis developed to aid in the investigation of the linguistic codes of classroom teacher-pupil verbal interaction. In order to put the coding system into perspective, research substantiating two points were reviewed. These points were a) research dating back 30 years reveals significant language differences between low and middle socio-economic levels, and b) recent research points to linguistic code differences as an important facet of cultural deprivation. Noting that the majority of EMR children come from low socio-economic levels, it was hypothesized that one of the keys to learning for EMR children is the complexity of the teacher's language in relation to that of the children; that linguistic code may be related to the unflattering picture of special classes drawn by the efficacy studies.

Before a formal investigation of the linguistic nature of teacher-pupil verbal interaction, it was necessary to develop a system of analysis. This report is an explanation of the coding system.

Method

Two tapescripts of EMR classrooms provided the raw data for the development of the coding system. There was a total of 19 children in both classes. Mean IQ was 63. Mean CA was 117 months. Both teachers were special education majors in college. One had five years experience with the mentally retarded, the other two years.

The basic procedure used was one of trying-out different methods of analysis and applying three criteria to them. These criteria were a) the system should reflect the complexity and structure of phrases and sentences, b) the system should be applicable to both teacher and pupil language, and c) the system should be uncomplicated enough to be used by a teacher herself.

The Linguistic Coding System

The system developed has three parts, the length and complexity index, the sentence classification scheme, and a content analysis. Directions for coding are in Appendix A and these three parts are discussed at length in the body of this report.

The length and complexity index (LCI) is a scoring system which simultaneously weights the length of a sentence and the complexity of its noun and verb phrases. Sentences are classified
into one of five classes—simple sentences, simple sentences with a phrase, elaborated simple sentences, compound and complex sentences, elaborated compound and complex sentences. These five classes form a logical hierarchy of complexity and are differentially weighted. A sentence complexity score (SCS) is obtained by adding the LCI and the weighted value for class of sentence.

The content of a sentence is coded according to the purpose of the communication into eight categories which are grouped into three main categories—Structuring, Responses Requested, and Reacting. Children's responses are also coded as to whether the response was correct and/or relevant.

Discussion

It is fairly evident that the linguistic coding system reported here needs to be tested in a cross-section of EMR classes and the resulting data critically analyzed and evaluated. One method of validating this instrument would be to compare its results with data obtained via stimuli pictures or in a free-play situation. Another method would be to use what Shriner (22) calls "psychological scale values"; i.e., judges rate language samples on a scale ranging from least development of language to most development. Reliability studies also need to be carefully carried out.

The problem of elliptical expressions is an important one. Children's utterances were 80 per cent elliptical in the two EMR classes used. These data cannot be ignored. Expanding and coding the expansions is artificial at best. Some way of handling ellipses as such is sorely needed.

Some of the recent work in linguistics in connection with "pivot" and "open" class words and the development of morphology of language appears promising for the modification or alteration of the coding system proposed in this report. One can only watch the developments in that area.

Lastly, in spite of the reservations noted above, the linguistic coding system appears to be a viable instrument for research not only into classroom interaction but also into curriculum and the art of teaching. It seems evident to this investigator and others (1,9,10,23,25) that language or aspects of it have to be taught to "language deprived" normal children. How much more true is this for children who are mentally retarded?
REFERENCES


8. Davis, Edith A. The Development of Linguistic Skill in Twins, Singletons with Siblings, and Only Children from Age Five to Ten Years. Minneapolis: University of Minnesota Press. 1937.


APPENDIX A

INSTRUCTIONS FOR THE LINGUISTIC CODING SYSTEM

Introduction

This coding system yields a numerical value called the sentence complexity score (SCS) for each sentence analyzed. The SCS is composed of four part-scores which may be treated separately if the user so desires. These parts are a) noun phrase points, b) verb phrase points, c) points for additional words in the sentence, and d) points for the type of sentence. The sum of the first three (a,b,c) comprise another score called the length and complexity index (LCI).

Common statistical techniques may be applied to the SCS, LCI, or other scores to aid in the interpretation of the data.

The content analysis is used as a framework within which the SCS can be interpreted. The content analysis does not yield a numerical score.

Preparation for Coding

The following steps are recommended in preparing the tape-script for coding:

1. Mark off each incomplete or complete sentence with hash marks (/). Omit all direct reading.

2. Number each sentence consecutively beginning with the number "1" on each page.

3. Write in the expansions of the elliptical or incomplete sentences.

4. In each sentence underline the noun phrases with a single line and the verb phrases with a double line.

5. Record the page number and other data at the top of the Coding Sheet.

Coding Rules and Notations

These directions will treat each column in the order in which they appear on the Coding Sheet immediately following these instructions. (p. A-11)
I. T/S

This column is used to indicate which of the utterances are student utterances. Only student utterances are indicated in this column, except as otherwise stated below. The notations used in this column are as follows:

A. SC

When the student's sentence is structurally complete, this notation is used. When the student's sentence is complete, it is coded according to the directions below.

B. SCE( ) - student contextual ellipsis

This notation is used to identify a contextual ellipsis. In the parentheses is put the number of the antecedent referent sentence from which the ellipsis derives its meaning.

Ellipses are expanded into complete sentences. These expansions are then coded according to the directions below.

C. STE( ) - student telegraphic ellipsis

In the parentheses is put the number of the antecedent referent sentence if there is one.

These are also expanded and scored.

Teacher utterances are coded in this column only when they are elliptical or incomplete. The following notations should be used. The directions for these are the same as for SCE and STE.

D. TCE( ) - teacher contextual ellipsis

E. TTE( ) - teacher telegraphic ellipsis

II. The Length and Complexity Index (LCI)

All words are counted according to the following rules:

1. Contractions of the subject and predicate like "it's" and "we're" are counted as two words.

A-2
2. Contractions of the verb and the negative like "didn't" are counted as two words.

3. Hyphenated words and compound nouns, particularly proper nouns designating a single object, are counted as single words. Example: Mary Lou, Miss X.

4. Starters such as Oh, hey, and, cause, then, well, Miss X, are eliminated and not scored.

5. The word "now" is not eliminated when it serves a sequencing function.

6. All prepositions are counted except in the following situations. (a) When it is considered part of the infinitive construction, (b) When it is the last word in a sentence and is elliptical.

7. Proper names in apposition are eliminated; e.g., Joseph, what are you doing?

A. \#NP
   In this column goes the number of Noun Phrases within the sentence.

B. Pts.NP
   In this column goes the total number of points for Noun Phrases in the sentence:

   The rules for calculating the noun phrase points are:

   1. Adjectives which are functioning as nouns are counted as in a noun phrase. For example, "I want yellow".

   2. Pronouns serving a nominative function are counted as noun phrases. For example, "I don't know what to do." In this sentence the word what functions in the same manner as it does in the sentence, "What is that?" Thus, the word what is counted as a noun phrase and given a score of 1.

   3. Following are examples of NP for each score from 1 through 6. Symbols used are N(oun), M (any modifier), A (article), MA (any modifier other than an article), pl (plural inflection s or -es), and poss (possessive inflections).

   A-3
Score of 1  
N     Dog, I, he, this, that  
N + N  candy cane

Score of 2  
M + N  my mother, lunch menu  
N + pl  cars, legs

Score of 3  
A + MA + N  a funny way, the other one  
M + N + pl  my lips, several things,  
           baked beans  
N + poss + N  Bruce's pencil, John's birthday

Score of 4  
MA + MA + N  that little sentence  
A + MA + N + pl  the long legs  
M + N + poss + N  the teacher's desk

Score of 5  
MA + MA + N + pl  these two sentences  
A + MA + MA + N  a funny little elf  
N + poss + N + pl  teacher's crayons  
MA + A + MA + N  such a funny menu

Score of 6  
A + MA + MA + MA + N

C. #VP

In this column goes the number of Verb Phrases in the sentence.

D. Pts. VP

In this column goes the total number of points for Verb Phrases in the sentence.

The rules for calculating the verb phrase points are:

1. An unmarked verb is a verb which does not have an ending, a suffix. The normal suffixes are -ed, -s, or -ing. Other verbs, even though they may be in the past tense, are considered unmarked verbs. The exception to this is the word said. Said is considered to be a regular past. The rationale for this is that the word is in effect the same as sayed and thus is considered to possess an ending.

2. In infinitive constructions, the word to is considered to be part of the verb and not a preposition. Thus the word to, in this case, is not scored.

A-4
In the sentence, "I don't want to" the word to is an elliptical expression standing for an infinitive. Since it is elliptical, it is not scored.

3. Only lexical verbs and copulatives are counted. This was decided on in order to escape from the problems of deciding when particular prepositions would be considered parts of verbs and when they would not, particularly in cases other than the infinitive. For example, in the sentence, "Pick out mine" the question of whether the verb is pick or pick out would depend on such factors as intonation, normal usage of the expression by the child and other considerations which are not obtainable through a tapescript.

4. Following are examples of VP for each score from 1 through 5. Symbols used are V(erb), PrP (present participle), PP (past participle), Aux (auxiliary), and ? (question).

Score of 1
Unmarked V: I go, you put, we had, done (irregular pasts), is

Score of 2
PrP alone: he going out
3rd person singular present: it looks like
regular past (-ed): dropped, said
Aux + unmarked V I'll make, could run

Score of 3
Aux + PrP he is going out
Aux + V + ? May I have a pencil?
Aux + marked V I am finished

Score of 4
Aux + PrP + V am going to get
Aux + Aux + V would have begun
Aux + PrP + ? Are you buying your lunch?

Score of 5
Aux + Aux + V + V as objective predicate could have been eating
Aux + Aux + V + V as objective predicate I am going to try to fix that.

E. LCI

In this column goes the sum of

Pts.NP + Pts. VP + additional points according to the following rules
1. All words which are not included under noun phrases or verb phrases are given a score of 1.

2. Negatives are given a score of 1, except the elliptical or appositive No.

III. Type

A. In this column the following codes will be used for the type of sentence indicated:

1. Simple Sentences - A -
   Examples: Look up here.
   Who can read this?
   What's that?

2. Simple Sentence with a Phrase - B -
   A compound subject, object or predicate can be substituted for the phrase.
   Examples: How many legs are in the picture?
   Does anyone see this word over here?
   Where is the picture of the legs?

3. Elaborated Simple Sentence - C -
   This is a simple sentence with two or more phrases or with a compound subject or predicate and a phrase.
   Examples: Who can take the other card with the word legs and put it over the picture of the legs?
   Come up and point to the word legs.
   Come up here and attach the word legs.
   Would you like to count to ten for us?

4. Complex and Compound Sentences - D -
   These are sentences that have either one independent clause plus one dependent clause or two independent clauses.
   Examples: You don't have to make it the way we made it. Draw a picture of it and then you can write in eyes, nose and mouth.
   I said over, but I meant above.

A-6
5. Elaborated Complex and Compound Sentences - E -

These are complex sentences with more than one subordinate clauses or with a phrase or phrases. These are compound sentence with more than two independent clauses or with subordinate clauses or phrases.

Examples: If you really like the word monster perhaps some of you can make your own on a piece of paper.
We are going to learn a new word today, but before we learn our new word, we are going to work on this.
Who can find the sentence that tells me that my legs are on my body?

B. Scoring: Sentence types are scored in the following manner

- Type A - score of 0
- Type B - score of 5
- Type C - score of 10
- Type D - score of 15
- Type E - score of 20

IV. SCS

In this column goes the Sentence Complexity Score which is the sum of

\[ \text{LCI} + \text{points for sentence type} \]

V. Content

The content of the sentence or utterance is coded according to the purpose of the communication.

A. TEACHER CATEGORIES

Teacher messages are categorized according to whether they are structuring communications, seeking responses, or reacting communications.

1. Structuring (S)

Structuring utterances are those utterances which attempt to either motivate, orient or inform the child or
children. They are generally declarative sentences. They generally attempt to set up a situation to facilitate the learning of the children. There are three types of structuring.

a. **Structuring - Motivating (Sm)**

Messages are coded as motivating (Sm) are attempts to elicit the child's interest and cooperation.

1. Suggesting that the task itself would be a rewarding experience. "We are going to play a game."

2. Promising external rewards. "Donna didn't have a turn before but she has been waiting very patiently." "Don't you want Miss X to see how smart you are?"

3. Relying on the authority of the teacher. "Now I want you to work quietly." "Are you going to be a good boy?"

4. Exhortatory sentences. "Let John do the next one." Even though strictly a command, it is interpreted as eliciting another child's cooperation in a sharing procedure.

b. **Structuring - Orienting (So)**

Orienting statements are those statements used to develop a set in the child's mind for a task activity to follow.

1. A declarative sentence telling the child what he will do next or with whom he will work.
   Examples: I want you to break it into tens and ones.
   John and Marianne will work together and Beverly and Donna will work together.
   What are you going to do now, Karen?

2. Questions such as "Now what are you going to do?" when followed by the teacher telling what to do.

A-8
Questions which attempt to focus the child's attention on a particular aspect of the task although possibly calling for physical or verbal feats or responses, are coded as orienting messages. 

Examples: "Who can read the whole story?"

In this question, the orienting is couched in the structure of the question itself. For example, if the teacher had stated "Read the whole story," this would be a message seeking verbal response. If the teacher had said, "Who will read the whole story?", this would also be a message seeking verbal response. Thus, the use of the word "can" here requires a verbal response of yes or no and not the task itself. Another example would be "Can you count that?"

c. Structuring - Informing (Si)

Teacher messages are considered to be informing when she is lecturing or imparting specific information about a task. Usually these are declarative sentences telling the child something about the task. 

Examples: "For this you have these sticks."
"The assembly is going to be much later on in the day."
"This is for Donna and Beverly."

2. Responses requested (X)

Teacher messages are coded as requiring responses when they are very clearly requests for either action on the part of the children or verbal production.

a. Response requested - action (Xa)

A message is coded as seeking action whenever the teacher asks the child to produce a physical action. 

Examples: "Touch your legs."
"Put this over here."
b. **Response requested - verbal (Xv)**

The teacher is seeking a verbal response whenever she asks the child to identify attributes of stimuli or to explain or to read whatever is being presented to him. There will be messages which seek verbal responses but which will be coded under either orienting or motivating. For example, when the teacher asks, "Don't you want Miss X to see how smart you are?", she is usually seeking a **yes** answer, but this is essentially a structuring - motivating message.

If the teacher actively seeks a production which is part of the task, if the teacher seeks an evaluation of the child of his ability, if the teacher asks a rhetorical question, it is coded as structuring - motivating or structuring - orienting.

3. **Reacting (R)**

After the student has responded to the teacher's structuring or request for response, then the teacher reacts to the student's performance. These reactions are essentially either positively reinforcing, negatively reinforcing or neutral.

a. **Reacting - positive reinforcement (Rp)**

Messages coded as positively reinforcing are verbal confirming statements which immediately follow a correct response by the child.

Examples: "That's right."
"Very good."
or a repetition of what the child has said in a manner such as to indicate that it was a correct response.

b. **Reacting - negative reinforcement (Rg)**

Messages coded as negatively reinforcing are verbal negations which immediately follow incorrect responses. This category includes the correction of a wrong answer on the part of the child.

For example: "These is." (child) 
"are." (Teacher) 
"are." (child)
Other examples: "No."
or asking another child whether the answer is correct.

c. Reacting - neutral reinforcement (Rn)

Messages are coded in this category when they have become used so routinely that they lose their meaning either positively or negatively. Examples of this would be a teacher's use of the words "all right", "O.K.", "very good" (in some situations this is a positive reinforcement)

B. STUDENT CATEGORIES

In the classroom the student's production are usually responses to a teacher's request for responses or to other sentences. Because of this, the student productions will be categorized according to the teacher category to which the student is responding.

1. A student category is classified according to the teacher category responded to.

In some instances, the student's sentence may fit into one of the teacher's categories as when the child requests action or verbal response from the teacher not in response to a previous teacher utterance; or when the child is structuring a situation for the teacher or another child. In this case, use a teacher category.

The coding for the student responses would be essentially the same as the abbreviations for the teacher categories. The difference is that, for example, if the child is responding to a structuring - motivating utterance of the teacher, the child's response would be coded (RSm). Thus, we have the following codes. RS, RSm, RSo, RSi, RX, R.a, RXv, RR, RRp, RRg, RRn.

2. Student responses are also coded a + or 0 depending on whether they are relevant to the question asked and whether they are correct or incorrect. Student responses are coded:

++ if relevant and correct
+0 if relevant and incorrect
0+ if irrelevant but correct
00 if both irrelevant and incorrect

A-11
<table>
<thead>
<tr>
<th>T/S</th>
<th>Length and Complexity Index</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#NP</td>
<td>Pts.</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX B

A Study of the Verbal Output in a Special Class for Retarded Children

These data were gathered during the early stages of the project and are based solely on the length of response concept; i.e., counting the number of words, sentences, and communications and computing percentages.

The tapescript was of a class of 9 children. The children had a mean CA of 130 months, mean MA of 86 months, and a mean IQ of 67. The teacher had been a special education major in college. She had two years of teaching experience, all of it with the retarded. The tapescript represents 1 hour, 50 minutes of class time.

The results of the analysis appear in the Table 1. A communication is defined as the complete speech production of a child or teacher before another person speaks. Words were counted according to the rules of Templin (1957, p. 160). In the Table "sentences" (c) refers to structurally complete sentences (d) plus those utterances which were functionally complete but structurally incomplete. The percentages listed for a, b, c, d, and e represent the percentage of the total of the tapescript.

There are four pertinent statements which can be made from Table 1. They are:

1. Teachers do about 85% of the talking in terms of the number of words uttered. (b)
2. EMH children average less than one complete sentence per utterance while the teacher produces more than two. (j, k)
3. About 80% of a teacher's sentences are structurally complete whereas only about 20% of the children's sentences are complete. (l)
4. The children are very seldom given opportunity to use structurally complete sentences (nor are they demanded). (d)

Discussion

These observations speak for themselves. It has long been known that teachers talk too much. But it also seems that these teachers...
Table 1

Length of Response of Teacher and Pupil

<table>
<thead>
<tr>
<th></th>
<th>Teacher</th>
<th>%</th>
<th>Child</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Communications</td>
<td>484</td>
<td>51</td>
<td>455</td>
<td>49</td>
</tr>
<tr>
<td>b. Words</td>
<td>6927</td>
<td>86</td>
<td>1067</td>
<td>14</td>
</tr>
<tr>
<td>c. Sentences</td>
<td>1191</td>
<td>75</td>
<td>397</td>
<td>25</td>
</tr>
<tr>
<td>d. Complete Sentences</td>
<td>930</td>
<td>93</td>
<td>62</td>
<td>7</td>
</tr>
<tr>
<td>e. Words in Complete Sentences</td>
<td>6351</td>
<td>95</td>
<td>294</td>
<td>5</td>
</tr>
<tr>
<td>f. Mean No. Words per sentence (b/c)</td>
<td>5.81</td>
<td></td>
<td>2.68</td>
<td></td>
</tr>
<tr>
<td>g. Mean No. of Words per complete sentences (e/d)</td>
<td>6.82</td>
<td></td>
<td>4.74</td>
<td></td>
</tr>
<tr>
<td>h. Mean No. Words per incomplete sentence (b - e) (c - d)</td>
<td>2.20</td>
<td>2.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i. Mean No. Words per communication (b/a)</td>
<td>14.31</td>
<td>2.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>j. Mean No. Sentences per communication (c/a)</td>
<td>2.27</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k. Mean No. Complete Sentences per communication (d/a)</td>
<td>1.91</td>
<td>.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>l. Percentage of Complete Sentences (d/c)</td>
<td></td>
<td>78.08</td>
<td>16.62</td>
<td></td>
</tr>
</tbody>
</table>
neither demand nor provide much opportunity for the children to use structurally complete sentences. The implication is not only that very few situations are structured so that the children must produce such sentences, but also that very few situations are planned which require from the children the kind of specificity of language found in other than simple sentences.

These observations startlingly point up a need for further investigation of language as it is used to convey strange content to cognitively limited children. The elements of language comprise the very building blocks of the content of instruction. Without these elements learning would be hindered. Without even the opportunity to practice using elements of "elaborated" language, it is questionable whether EMH children will learn them to the extent necessary to render them functional. It is relatively clear that the children, at least in the classroom, are fairly restricted in variety and complexity of language.

These observations lead to questions of the following type. Are these observed patterns widespread enough to be of concern? Can we systematically increase the variety and complexity of the language of EMR children? Since this is not the traditional view of language arts, what changes in curriculum and methodology are necessary?

There are suggestions in the literature which may be applicable to EMR children. Stearns (1967) presents a promising set of guidelines which governed the development of specific lessons in his language project. These guidelines focus on three areas—response elaboration, verbal definition and verbal feedback. The Bereiter and Englemann (1966) direct teaching program contains some useful ideas.
References


