R E P O R T   R E S U M E S

ED 016 534   PS 000 384
AN OBJECTIVE MEASURE OF STRUCTURAL COMPLEXITY IN CHILDREN'S
WRITING.
BY- O'DONNELL, ROY C.
                                                    PUB DATE   3 JAN 68
EDRS PRICE MF-$0.25 HC-$0.40   8P.

DESCRIPTORS- #MEASUREMENT INSTRUMENTS, #MATURITY TESTS,
#LANGUAGE DEVELOPMENT, #WRITTEN LANGUAGE, #SYNTAX, SENTENCE
STRUCTURE, WRITING SKILLS, LANGUAGE ABILITY, LANGUAGE TESTS,
OBJECTIVE TESTS, AERA

        AN INSTRUMENT FOR MEASURING LANGUAGE MATURITY SHOULD BE
EASILY ADMINISTRABLE AND SCOREABLE AND SHOULD BE BASED ON
VALID INDICES OF LANGUAGE MATURITY. A RECENT STUDY BY KELLOG
W. HUNT SUGGESTED THAT THESE INDICES BE BASED ON A NEW
SYNTACTIC UNIT, THE T-UNIT, CONSISTING OF 1 MAIN CLAUSE
TOGETHER WITH ANY CLAUSES SUBORDINATED TO IT. COORDINATED
MAIN CLAUSES, WHICH ARE USUALLY A SIGN OF IMMATURITY, WILL
THUS HAVE NO EFFECT ON THE INDEX. HUNT'S STUDIES, BASED ON
1000-WORD SAMPLES, INDICATED THAT T-UNIT LENGTH, CLAUSE
LENGTH, AND NUMBER OF CLAUSES PER T-UNIT INCREASED FROM
GRADES 4 TO 12. TO SEE IF THE INDICES COULD BE RELIABLY
COMPUTED FROM SHORTER SAMPLES, 80 CHILDREN IN GRADES 4, 6,
AND 8 WERE ASKED TO REWRITE A NARRATIVE COMPOSED OF SIMPLE
DECLARATIVE SENTENCES, COMBINING THESE INTO LONGER UNITS.
ALSO, 69 CHILDREN IN GRADES 4, 8, AND 12 WERE ASKED TO
REWRITE A STRUCTURALLY SIMPLIFIED EXPOSITORY ESSAY USING
LONGER SENTENCES. OBSERVED INCREASES WITH GRADE IN THE CLAUSE
LENGTH, T-UNIT LENGTH, AND NUMBER OF CLAUSES PER T-UNIT WERE
STATISTICALLY SIGNIFICANT. BECAUSE OF THE INTERRELATION OF
THE INDICES, T-UNIT LENGTH APPEARS TO BE THE MOST USEFUL
INDEX. FURTHER REFINEMENT OF THE NARRATIVE AND THE ESSAY IS
NECESSARY, TOGETHER WITH STUDIES TO DETERMINE THEIR GENERAL
VALIDITY AND RELIABILITY. THE PAPER WAS PRESENTED AT THE
ANNUAL MEETING OF AMERICAN EDUCATIONAL RESEARCH ASSOCIATION,
FEBRUARY 9, 1968. (DR)

AN OBJECTIVE MEASURE OF STRUCTURAL COMPLEXITY IN CHILDREN'S WRITING

In a study reported thirty years ago, John E. Anderson (1) evaluated by statistical methods three indexes of language development. He reported data on sentence length and pronoun index, but his study was primarily concerned with the subordination index which had been worked out a short time earlier by Lou L. LaBrant (3).

Anderson was interested in the generality of LaBrant's index and in the possibility of developing an easily applied and uniform measuring implement. Because of such variables as composition length, subject matter, and situations in which language is used, he drew pessimistic conclusions about generalized application of the subordination index. He recognized, however, the existence of "...a fertile field for the development of indices based on common subject matter, well-categorized scales, and adequate samples from the standpoint of length of passage."

Subsequent investigations of children's language have relied heavily on mean sentence length and the subordination index; and, until recently, not much progress seems to have been made toward an easily applied and uniform measurement of language development. Recently developed techniques of linguistic analysis, however, offer new hope that such an instrument can be devised.

To be practical, an instrument for measuring language development would possess the following characteristics: 1) It would be based on a reliable and valid index of language maturity; 2) It would be easily administered; 3) It would be easily scored. In addition, it should reduce to a minimum the effects of variables

such as length of composition, subject matter, and conditions under
which the writing samples are produced. This report deals with an
investigation of the feasibility of devising a useful instrument
based on recently developed indexes of language development.

In his study of children's grammatical structures written at
three grade levels, Kellogg W. Hunt (2) gives a critical evaluation
of the traditional indexes of language development and proposes
some new indexes. Pointing out the fact that inadequate punctuation
and indiscriminate use of and makes sentence length an unreliable
index, Hunt proposes a syntactic unit consisting of one main clause
and any subordinate clauses attached to the main clause. Such a unit
is grammatically capable of being considered a sentence; and since it
is a terminable unit, he proposes that it be called a T-unit. This
unit can be identified objectively, and it is not affected by poor
punctuation. It has the added advantage of preserving all the sub-
ordination achieved by the student and all of his coordination of
words, phrases, and subordinate clauses. It does not preserve the
student's coordination of main clauses; but, as Hunt points out,
excessive coordination of main clauses is a sign of immaturity rather
than maturity. Thus, the elimination of coordinated main clauses
from the syntactic unit to be used in language development studies is
a gain rather than a loss.

Hunt calls attention to the fact that LaBrant counted coordinated
verbs as separate clauses, and he insists that a clause is a structure
with a subject and a finite verb. Then, having identified a syntactic

unit more useful than the sentence and having clarified the definition
of the clause, Hunt proposes a revision of the subordination index.
Instead of dividing the number of subordinate clauses by the total
number of clauses, he proposes that the total number of clauses be
divided by the number of main clauses. This procedure would give the
mean number of clauses per T-unit, which can be converted into the
ratio of subordinate clauses to main clauses.

Computing these proposed indexes for the materials analyzed in
his study, Hunt found statistically significant increases in T-unit
length, clause length, and number of clauses per T-unit from grade four
to grade eight to grade twelve. It should be pointed out that the
three measures under consideration are interrelated. T-units can be
lengthened by either or both of two means: lengthening clauses or
increasing the number of subordinate clauses. Analysis based on tech-
niques of transformational-generative grammar reveals that syntactic
complexity of children's language is reflected in both clause length
and number of clauses per T-unit. T-unit length can be computed by
multiplying words per clause by number of clauses per T-unit; and since
T-unit length incorporates the other two indexes, it is evident that
T-unit length is a useful index of structural complexity.

Hunt's conclusions were based on his analysis of 1,000 words
written by each student participating in his study. Such lengthy
samples are desirable for statistical reasons, but they are too long
for easy analysis. In order to see whether the indexes Hunt proposed
could be computed from smaller samples, this investigator devised an
experimental instrument for collecting comparable writing samples from

children at various grade levels.

Writing samples were gathered from 80 pupils enrolled in the public schools of High Point, North Carolina, in the spring of 1966. There were 24 pupils in grade four, 28 pupils in grade six, and 28 pupils in grade eight. In an extension of the study, writing samples were gathered from 69 pupils (26 in grade four, 20 in grade eight, 23 in grade twelve) enrolled in the University School (FSU), Tallahassee, Florida.

In order to obtain comparable samples of children's writing, the investigator asked the High Point school children to rewrite a narrative composed of simple declarative sentences. They were asked to rewrite the story, putting the short sentences together to make longer sentences. In this way they could demonstrate their ability to combine kernel structures to form sentences of greater structural complexity. Each child's writing was segmented into T-units, and the following computations were made for each composition: 1) mean number of words per clause; 2) mean number of clauses per T-unit; and 3) mean number of words per T-unit. Statistical analysis of variance was executed to test for significance at the .01 level the mean differences at various grade levels. Subsequently, a structurally simplified expository essay was presented to University School students, with instructions to rewrite the essay using longer and better sentences. The resulting compositions were analyzed in the same manner as the narrative compositions.

In the narrative compositions, mean words per clause increased from 6.41 for grade four to 7.11 for grade six, and to 8.17 for grade eight

(F ratio, 22.78). Mean clauses per T-unit increased from 1.14 to 1.30
to 1.45 for the three grades (F ratio, 23.24). Mean words per T-unit
increased from 7.37 to 9.27 to 12.01 (F ratio, 39.20). Increments on
each of the three indices are significant at the .01 level. In the
expository compositions, mean T-unit length increased from 6.10 to
8.48 to 10.98 from grade four to grade eight to grade twelve (F ratio,
59.11). Mean clause length increased from 5.76 to 6.53 to 8.45 for
the three grades (F ratio, 56.39). Mean number of clauses per T-unit
increased from 1.05 to 1.29 to 1.30 (F ratio, 24.65). These increments
are also significant at the .01 level.

It was pointed out earlier that both clause length and number of
clauses per T-unit reflect syntactic complexity. Data resulting from
this study indicate that both clause length and number of clauses per
T-unit increase together at the lower grade levels, but that clause
length accounts for most of the growth at higher grade levels. Since
T-unit length is accounted for by a combination of these two factors
(mean clause length x mean clauses per T-unit = mean T-unit length),
it appears that T-unit length is a useful index of growth in structural
complexity over a wide age-range.

It should be noted at this point that syntactic complexity is
only one factor in language maturity. A complex unit is not necessarily
a well-constructed unit, and there is no doubt that stylistic maturity
is manifested in a writer's ability to avoid complexity as well as in
his ability to produce it. However, there is sufficient evidence to
support the assertion that syntactic complexity is a major factor in
language maturity, and a preliminary study recently completed indicates

that there is a significant correlation between complexity scores and overall rating scores on compositions written by ninth grade students.

The basic difference between the present study and Hunt's study lies in the method of obtaining writing samples. Hunt's analysis was based on large samples of writing collected over an extended period of time; this study was based on single samples taken (with the exception of a few fourth grade expository essays) in a single class period. Since the mean values of the three measures investigated in this study are comparable to those obtained by Hunt, the findings of this study lend support to Hunt's findings and suggest the feasibility of designing standardized instruments to measure language development.

When evaluated in the light of the criteria suggested earlier, the experimental instruments used in this study obviously need further refinement. The narrative instrument employed in this study was found to be unsuitable for use beyond the eighth grade, and the expository instrument was too lengthy for some fourth graders to complete in a single class period. Although it is evident that clause length, T-unit length, and number of clauses per T-unit increase with advance in grade, there are no data to show how consistently these indexes measure the structural complexity of an individual student's writing in various situations.

In a recently completed study, Professor Hunt has shown that a shortened form of the expository essay described above is useful for collecting writing samples from pupils in grades four through twelve. His analysis of writing samples collected by means of this instrument shows essentially the same rate of increment on the three indexes of

(7)

complexity as he found in his earlier study. Thus, it seems that con-
siderable progress has been made toward devising an instrument that is
easily administered and easily scored. Further study of validity and
reliability of instruments such as these is urgently needed.

REFERENCES

1. Anderson, John E. 1937. "An Evaluation of Various Indices
of Linguistic Development," Child Development, 8:1
(March), 62-68.

2. Hunt, Kellogg W. 1965. Grammatical Structures Written at
Three Grade Levels. National Council of Teachers of
English, Research Report No. 3. Champaign, Ill.:
National Council of Teachers of English.

3. LaBrant, Lou. 1933. "A Study of Certain Language Develop-
ments of Children in Grades 4-12 Inclusive," Genetic
Psychology Monographs, 14:4 (November), 387-491.

AERA 1968 Annual Meeting
Submitted by Roy C. O'Donnell
            Associate Professor
            English Education
            The Florida State University
            Tallahassee, Florida
            January 3, 1968