CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (r. 82/08)

Canada

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Division
Division des thèses canadiennes

Ottawa, Canada
K1A 0N4

53984

0-315-06064-6

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

THOMAS B. MACNEIL

Date of Birth — Date de naissance

DECEMBER 25, 1942

Country of Birth — Lieu de naissance

CANADA

Permanent Address — Résidence fixe

68 REGENT CRESCENT, BRANDON, MANITOBA R7B 2W9

Title of Thesis — Titre de la thèse

The Effect of Sentence-Combining Practice on the Listening, Reading Comprehension and Written Syntactic Skills of Ninth Grade Students

University — Université

ALBERTA

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

PhD

Year this degree conferred — Année d'obtention de ce grade

1981

Name of Supervisor — Nom du directeur de thèse

JOHN E. OSTER

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date

September 28, 1981

Signature

Thomas B. MacNeil

NL-91 (4/77)

THE UNIVERSITY OF ALBERTA

THE EFFECT OF SENTENCE-COMBINING PRACTICE ON THE DEVELOPMENT OF READING COMPREHENSION AND THE WRITTEN SYNTACTIC SKILLS OF NINTH GRADE STUDENTS

by

© THOMAS B. MACNEILL

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF SECONDARY EDUCATION

EDMONTON, ALBERTA

FALL, 1981

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Thomas B. MacNeill

TITLE OF THESIS The Effect of Sentence-Combining Practice on the Development of Reading Comprehension and the Written Syntactic Skills of Ninth Grade Students

DEGREE FOR WHICH THESIS WAS PRESENTED Doctor of Philosophy

YEAR THIS DEGREE GRANTED 1981

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(Signed) T.B. MacNeill

PERMANENT ADDRESS:

Dept. of Curriculum & Instruction
Faculty of Education
Brandon University
Brandon, Manitoba, Canada

Dated Sept. 28, 1981

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled The Effect of Sentence-Combining Practice on the Development of Reading Comprehension and the Written Syntactic Skills of Ninth Grade Students submitted by Thomas B. MacNeill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Secondary Education.

John E. Oster
Supervisor

J. B. Parsons

V. R. Nyberg

[illegible]

[illegible]

[illegible]
External Examiner

Date September 28, 1981

DEDICATION

To My Two Girls

Abstract

The purpose of this study was to examine the effect of the O'Hare Sentencecraft (1975) program on the written syntactic skill of average-ability, ninth-grade students, using written argumentative compositions as the data source. To clarify the impact of mode on written syntax, the syntactic frequencies achieved by students in this study were compared with those found by Hunt (1965) and Loban (1976). A second purpose was to determine the effect of the Sentencecraft program on the reading comprehension "Level" and "Speed" skills of the same students as measured by the Davis Reading Test, Series 2 (Davis and Davis, 1962). An experimental group/control group--pre-test/post-test/delayed post-test design was used. Students were divided into three ability groups of two classes each on the basis of a school-wide English Department exam for ninth graders. Classes at each ability level were randomly assigned to treatment and were taught by the same teacher. The experimental group (n=75) completed the Sentencecraft program over a nine-week period. The control group (n=68) did their regular classroom work excluding grammar study and reading skill development work. All students wrote one essay per week in English class. Mode for essays varied except for the six testing occasions when one-sixth of the students wrote on each argumentative topic. The six occasions were the pre-test, three bi-weekly tests, the post-test, and the delayed post-test which followed eight weeks later. Reading tests were administered on the three principal occasions with one-third of the students writing each test form on each occasion. Selected syntactic counts were made from the

compositions of the complete samples: the number of words, T-units, clauses, words per T-unit, words per clause, and clauses per T-unit. Within-clause syntactic structures were also counted using compositions by representative subsamples of thirty students from each group on the three principal occasions. Structures counted were those found by Hunt (1965) to be most indicative of syntactic maturity. The results were as follows: 1) two-way analyses of variance with repeated measures using the six major writing indices revealed no significant differences between the groups' means which were attributable to the experimental treatment; 2) two-way analyses of variance with repeated measures revealed no significant differences between the subsamples' means on the sixteen syntactic counts; 3) a two-way analysis of variance with repeated measures using comprehension "Level" scores revealed no significant difference between the groups which were attributable to the sentence-combining treatment; 4) a two-way analysis of variance with repeated measures did reveal a significant increase in comprehension "Speed" scores beyond the 0.01 level; and 5) the comparison with previous studies revealed that these grade nine students, on the pre-test occasion, used certain within-clause syntactic structures at a level near or above the mean use exhibited by the Hunt and Loban grade twelve samples, except for genitives. The findings raise questions about the efficacy of the Sentencecraft program for inducing syntactic maturity in average grade nine students while re-affirming the importance of mode of writing in setting syntactic norms. The results also clarify the influence of sentence-combining activity on reading comprehension skills.

Acknowledgements

Appreciative thanks are expressed to all members of my Committee for their erudite comments which added significantly to the study and this report. Their cooperation in reading and commenting on the preliminary drafts displayed their true professionalism.

Special thanks are extended to my advisor, Dr. John Oster, whose interest in and work on this document went far beyond the call of duty. Also contributing in a special way was Dr. Glenn Martin who provided the stimulus for the study.

To my wife, typist, and friend, I express my deepest appreciation for support throughout my graduate work. And finally, my thanks to Shandra for willingly entertaining herself when I ignored her.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	THE PROBLEM	1
	Purpose of the Study	3
	Statement of the Problem	3
	Need for the Study	4
	Hypotheses and Questions	8
	Limitations	10
	Significance of the Study	10
	Definition of Terms	11
II.	REVIEW OF THE LITERATURE	13
	Writing and Syntactic Development	14
	Tracing Syntactic Development in Writing: Descriptive Studies	14
	Improving Written Syntax: Experimental Studies	31
	Reading Comprehension and the Influence of Syntax	46
	Reading Comprehension and Syntax: Descriptive Studies	48
	Improving Comprehension by Improving Knowledge of Syntax or Grammar: Experimental Studies	55
	Reading Comprehension and Written Syntactic Ability: The Relationship	58
	The Relationship Between Reading Comprehension and Syntax: Descriptive Studies	58
	Improving Reading Comprehension By Improving Written Syntax: Experimental Studies	63
	Conclusions from the Review of Literature	71
	Writing and Syntactic Development	71
	Reading Comprehension and the Influence of Syntax	72

TABLE OF CONTENTS (CONTINUED)

CHAPTER		PAGE
	Reading Ability and Writing Ability: the Relationship	73
III.	DESIGN OF THE STUDY	74
	Overview	74
	Sample	75
	The Treatment	77
	Instrumentation	81
	Data Collection Procedures	87
	Scoring Procedures	88
	Conclusion	91
IV.	ANALYSIS OF DATA	92
	IQ: Experimental vs Control	94
	Writing Analysis: Experimental vs Control Samples	95
	Writing Analysis: Experimental vs Control Subsamples	105
	Reading Analysis	123
V.	DISCUSSION OF RESULTS	131
	Writing: Experimental vs Control	131
	Writing: Pre-test Means and Comparison with Previous Studies on Major Indices	131
	Writing: Pre-test Means Compared with Previous Studies on Minor Indices	138
	Reading: Experimental vs Control	145
VI.	SUMMARY, CONCLUSIONS, AND IMPLICATIONS	147
	Summary	147
	Conclusions	148
	Implications	149

TABLE OF CONTENTS (CONTINUED)

***

	PAGE
BIBLIOGRAPHY	152
APPENDIX A. STUDY PROPOSAL	160
APPENDIX B. PILOT STUDY NOTES	168
APPENDIX C. ARGUMENTATIVE WRITING HANDOUT	169
APPENDIX D. STUDY SCHEDULE	173
APPENDIX E. ANALYSED COMPOSITION	174
APPENDIX F. ADDITIONAL TABLES	177
APPENDIX G. ADDITIONAL PERMISSION CORRESPONDENCE. . . .	191
VITA	196

LIST OF TABLES

Table	Description	Page
1	Major Indices of Syntactic Development	18
2	Additional Indices of Syntactic Development	19
3	Results from Three Studies Using Sentence-combining: Experimental Groups Only	39
4	Syntactic Structures Which Increase Sentence Difficulty	79
5	Experimental and Control Group Comparison: IQ	94
6	One-way Analysis of Variance: IQ	95
7	Comparison of Topics: Pre-Test	96
8	Mean Number of Words, Repeated Measures	97
9	Treatment Effect on Number of Words Written	97
10	Mean Number of T-units Across Occasions	99
11	Treatment Effect on Number of T-units	99
12	Mean Number of Dependent Clauses	100
13	Treatment Effect on Number of Dependent Clauses	100
14	Mean Number of Words per T-unit	101
15	Treatment Effect on Words per T-unit	102
16	Mean Number of Words per Clause	102
17	Treatment Effect on Words per Clause	103
18	Mean Number of Dependent Clauses per T-unit	104
19	Treatment Effect on Dependent Clauses per T-unit	104
20	Chi-Square: Subsamples versus Remainder of Subsamples	106-7
21	Mean Number of Noun Clauses per T-unit	109
22	Mean Number of Adjective Clauses per T-unit	110

LIST OF TABLES (CONTINUED)

Table	Description	Page
23	Mean Number of Adverb Clauses per T-unit	110
24	Mean Percentage of Short T-units	111
25	Mean Percentage of Medium T-units	111
26	Mean Percentage of Long T-units	112
27	Mean Number of Gerunds per Clause	113
28	Mean Number of Factive Infinitives per Clause	113
29	Mean Number of Nominalized Verbals per Clause	114
30	Mean Number of Modal Auxiliaries per Clause	115
31	Mean Number of Perfect Forms per Clause	115
32	Mean Number of Passives per Clause	116
33	Mean Number of Verb Forms per Clause	116
34	Mean Number of Prepositional Phrases as Noun Modifiers per Clause	117
35	Mean Number of Infinitives as Noun Modifiers per Clause	118
36	Mean Number of Present Participles as Noun Modifiers per Clause	118
37	Mean Number of Past Participles as Noun Modifiers per Clause	119
38	Mean Number of Verb Forms as Noun Modifiers per Clause	120
39	Mean Number of Inflected Genitives per Clause	120
40	Mean Number of Phrasal Genitives per Clause	121
41	Mean Number of Genitives per Clause	122
42	Mean Number of Adjectives per Clause	122

LIST OF TABLES (CONTINUED)

Table	Descriptive	Page
43	Davis Reading Test Level Scores Compared by Form	124
44	Davis Reading Test Speed Scores Compared by Form	125
45	Mean Reading Level Across Occasions	126
46	Reading Level: Analysis of Variance Results	126
47	Mean Reading Speed Across Occasions	128
48	Reading Speed: Analysis of Variance Results	128
49	Reading Speed: Test of Simple Treatment Effects	129
50	Reading Speed: Newman-Keuls Comparison—Experimental Group	129
51	Mean Number of Words, T-units, and Clauses Written: Pre-test	134
52	Comparison of Results of Various Studies: W/TU, W/CL, CL/TU	135-6
53	Comparison of Hunt (1970) and the Present Study on Percentage of Short, Medium, and Long T-units	140
54	Comparison of Results of Hunt (1970) and the Present Study: Noun, Adjective, and Adverb Clauses	141
55	Mean Number of Syntactic Structures per Clause	144

LIST OF FIGURES

FIGURE	PAGE
1. Words per Essay	132
2. T-units per Essay	132
3. Clauses per Essay	132
4. Words per T-unit	132
5. Words per Clause	133
6. Clauses per T-unit	133

CHAPTER I

THE PROBLEM

Two of the most significant developments in man's communicative history are his acquisitions of the ability to read and to write his language. In an effort to improve facility in communication, numerous experiments and studies have been conducted to determine how the acquisition of language can be most efficiently fostered. Most have met with limited success and English teachers have experienced continuing frustration in their attempts to improve significantly their students' performance in language skills. In 1972, Roy O. Freedle and J.B. Carroll wrote:

> There is evidence that competence develops only as the individual is exposed over long periods of time to increasingly difficult materials; somehow the individual is able to use this material as a basis for increasing not only his general knowledge and understanding, but also his basic linguistic competence. To a degree, the material itself can teach him new and complex syntactical constructions. In stating these generalizations, however, we realize that actually we do not know enough about how the individual's linguistic competence is developed and how it can be fostered. (p. 360)

Until quite recently this basic lack of knowledge left teachers with only the traditional grammar approaches to use in their attempt to improve student writing. Despite Braddock, Lloyd-Jones, and Schoer's (1963) conclusion that such a program had no positive effect on student writing ability, no alternative program was available. Bateman and Zidonis (1966) pointed out the problems facing the composition teacher:

> The composition teacher, not having been provided with an adequate theory of language or grammar, is forced to develop or secure curricular materials that will stimulate and challenge his students to write--hopefully, to write better. (p. 6)

In fact, the scarcity of research data supporting the effectiveness of traditional grammar instruction in improving writing led the same authors to state that "the composition teacher seems to be incidental to whatever process it is that transforms a writer of fragments or poor sentences into a writer of acceptable prose." (p. 6) Only since the advent of transformational grammar have researchers had an intuitively plausible theory of sentence production to test. And it is only since the advent of sentence-combining programs that researchers (O'Hare, 1971; Combs, 1973; Pedersen, 1977) have been able, consistently, to induce growth in the syntactic aspects of student composition. In such programs, the student is given practice in joining short, simple sentences to produce longer, more syntactically complex sentences. No grammar study is involved in the most successful programs. In terms of transformational grammar theory, sentence combining requires a student to take short "kernel" sentences and "transform" them into more syntactically mature "surface" sentence structures while maintaining the "deep structure" meaning.

There are two types of sentence combining problems. The "signalled" combining problems require the student to follow the signals indicating the transformations to be applied. Thus the student is forced to practice in his writing what he intuitively understands about the operation of the language. He is aided in this by the signals. The "open" combining problems require the student to apply the rules of the language, which he intuitively understands, to produce a

grammatically correct, more syntactically mature surface structure.

Because sentence-combining practice is becoming an integral part of, or at least a supplement to school writing programs, researchers must determine the effects of such programs on student language skills.

Purpose of the Study

The purpose of the present study was to measure the effects of a sentence-combining program on the language-processing and production skills of grade nine students. The main points of interest were (1) whether this program results in syntactic growth which can be called true syntactic development, that is, development which is not forced or artificial; and (2) whether this program results in an increase in the levels of reading comprehension or speed of comprehension of the students involved.

Statement of the Problem

The problems with which the present study was concerned can thus be stated as:

1) To determine whether practice in sentence combining results in writing which is significantly more syntactically mature at the grade nine level.

2) To determine whether practice in sentence combining results in a significant improvement in reading comprehension skills at the grade nine level.

Need for the Study

Authorities on language acquisition (McNeill, 1970) and linguistics (Lefevre, 1970) agree that improvement in a student's ability to process and produce the various syntactic patterns is highly desirable. McNeill, in The Acquisition of Language (1970), writes:

> Virtually everything that occurs in language acquisition depends on prior knowledge of the basic aspects of sentence structure . . . . The concept of a sentence is the main guiding principle in a child's attempt to organize and interpret the linguistic evidence that fluent speakers make available to him. (p. 2)
>
> Between sound and meaning stands syntax. The relation between sound and meaning is therefore understood to the degree that the syntax of a language is understood. (p. 155)

And Carl Lefevre in Linguistics, English, and the Language Arts (1970) states:

> Greater attention should be given to developing sentence sense in reading and writing . . . . It is probable that given a mastery of basic sentence structure, vocabulary would largely take care of itself, because basic sentence patterns of English can be filled with an almost unlimited number of words. But no number of individual words can themselves combine into a single structural and meaning-bearing language pattern. (p. 23)
>
> An understanding of basic sentence patterns and their variations is crucial to the development of sentence sense, which in turn is fundamental to successful writing and to reading comprehension. (p. 63)

Despite the assertiveness of these quotations, there is some question about the degree of relationship between the abilities required in the receptive and expressive modes with which we are concerned here. McNeill (1970) points this out:

> There are persons who believe that children acquire two grammars, one for production and one for comprehension, and that acquisition is faster for the comprehension grammar. . . . While it is possible that comprehension and production are separate to some degree, we do not know why one should be possible before the other. (p. 101)

If it is true that students do acquire two grammars, there is a need for studies which shed some light on the relationship between them to determine if an improvement in the productive capability will indeed increase the comprehension capability. Results of earlier studies (Fisher, 1973; Hughes, 1975; Combs, 1975; Smart and Ollila, 1978; Evans, 1979; Sullivan, 1979) have been inconclusive and will be discussed later. The point being made here is that a program which resulted in significant development in the student's ability to manipulate syntax and also improved the student's comprehension ability would be a welcome addition to a language arts program.

It is important at this point to insist on the necessity for "true" language development. There have been two major approaches evolving in the study of language development. One group has carried out descriptively-oriented research and attempted to define what happens in writers' syntax as they mature. The prime movers in this area were A.F. Watts (1944), Walter Loban (1966, 1967, 1976), Kellogg Hunt (1965, 1966, 1967), and Roy O'Donnell et al. (1967). A second group carried out experimentally-oriented research and attempted to improve ability to manipulate syntactical patterns. The pioneers in this area were Bateman and Zidonis (1966), John Mellon (1969), Joel Gajadharsingh (1970), Frank O'Hare (1973), Warren Combs (1975), Elray Pedersen (1977), and Maureen Sullivan (1979).

Apart from the preliminary study by Donald Bateman and Frank Zidonis (1966) researchers have not, in any detailed manner, investigated this question: Is the writing growth attributed to the practice in syntactic manipulation as suggested by O'Hare (1973) the same as the normal development in syntactic growth which has been

identified by Hunt (1965) and Loban (1976), or is it a kind of artificial growth much of which dissipates in a relatively short period of time as Combs (1975) found? On a very superficial level, this possibility was suggested by the fact that sixty-five percent of the increase in the O'Hare experimental group's words per T-unit was due to the relatively unsophisticated addition of subordinate clauses. And sixty-four percent of the increase in the Combs experimental group's words per T-unit was due to the same syntactical manipulation.

There was, then, a great need, because of the increasing popularity of sentence-combining programs and the increasing number of studies involving sentence-combining practice, to determine whether the resulting syntactic growth was true developmental growth. There was also a need for carefully designed research that attempted to determine the degree of transference between the development of reading comprehension skills and growth in written syntax induced by sentence combining. The answers to both of these questions would have important curricular implications, as was suggested earlier.

Apart from the question of whether or not sentence combining induces true growth in language skills, two additional considerations suggested the necessity for studies similar in design to the present one. The first was that the majority of sentence-combining studies (Mellon, 1969; O'Hare, 1973; Hughes, 1975; Combs, 1975; Pedersen, 1977) had been done with grade seven students in the hope of producing a significant increase in such writing indices as words per T-unit. Bateman and Zidonis (1966) and Fisher (1973) each worked with a group of grade nine students. However, the Bateman and Zidonis study was a two year transformational grammar study and not the more simple and

seemingly more fruitful O'Hare sentence-combining practice. Fisher was working with a self-made program and a very small experimental group of fourteen students. There was need then for a study which used the O'Hare program and a large number of upper junior high students. Such a study would help to determine whether the significant growth obtained with grade sevens was also attainable with grade nine students.

The third consideration had to do with the modes of discourse which had been used in the various studies mentioned so far. Researchers, most notably Hunt (1965) and Loban (1976), analyzed samples of student writing as one method of determining syntactic maturity. These writing samples were mainly descriptive and narrative (O'Hare, 1973), expository (Hunt, 1970), or a mixture (Loban, 1976). The writing samples would be acceptable if the researcher wanted to identify syntactic ability as exemplified by the "usual writing" of students in a particular grade. However, if the researcher wants to measure the upper limit of students' syntactic ability in writing, he cannot ignore the influence of mode. Perron (1976) showed that, even as early as grade five, student writing in the argumentative mode was significantly more mature (see "Definition of Terms", p. 11), as measured by words per T-unit, words per clause, and clauses per T-unit, than writing in any of the other three traditional modes. He had fifty-one grade five students write in each of the four modes. The order for modes from simple to complex was descriptive, narrative, expository and argumentative. All differences were significant beyond the .001 level. In general, Crowhurst (1977, p. 74) confirmed Perron's findings in her study involving students at the grade six and grade

ten level. It is necessary to begin establishing norms using the words per T-unit, words per clause, and clauses per T-unit indices while at the same time controlling for mode.

The need for this study was predicated upon the following:

1. the need to determine if the growth in writing achieved through sentence-combining was true language development;

2. the need to determine whether the sentence-combining practice which induces this growth in writing also resulted in improvement in reading comprehension skills;

3. the need to determine whether the growth achieved with grade seven students was achievable with grade nine students; and

4. the need to attempt to measure the upper limit of students' syntactic ability as indicated by argumentative writing.

Hypotheses and Questions

The hypotheses and questions were divided into two main areas of interest: those concerned with the effects of sentence-combining practice on written syntactic development and those concerned with the effect of sentence-combining practice on reading comprehension.

Writing. 1. There will be no significant difference between the mean scores of classes of ninth grade students who have and who have not received the sentence-combining treatment as measured by argumentative compositions written on six occasions using the following indices: number of words, number of T-units, number of dependent clauses, words per T-unit, words per clause, and dependent clauses per T-unit.

2. There will be no significant difference between the mean scores of samples of ninth grade students who have and who have not received the sentence-combining treatment as measured by argumentative compositions, written on pre-test, post-test, and delayed post-test occasions, using the following indices:

a) the mean number of noun, adjective, and adverb clauses per T-unit;

b) the percentage of short (8 words or less), medium (9-20 words), and long (21 or more words) T-units;

c) the mean number of nominalized verbals (gerunds and factive infinitives) per clause;

d) the mean number of modal auxiliaries, perfect forms, and passives per clause;

e) the mean number of prepositional phrases as noun modifiers per clause;

f) the mean number of verb forms (infinitives, present participles, and past participles) as noun modifiers per clause;

g) the mean number of genitives (inflected and phrasal) per clause; and

h) the mean number of adjectives per clause;

Reading. 1. There will be no significant difference between the mean scores of classes of ninth grade students who have and who have not received the sentence-combining treatment as measured by the "Level of Comprehension" subtest of the Davis Reading Test on three different occasions: pre-test, post-test, and delayed post-test.

2. There will be no significant difference between the mean scores of classes of ninth grade students who have and who have not received the sentence-combining treatment as measured by the "Speed of Comprehension" subtest of the Davis Reading Test on three different occasions: pre-test, post-test, and delayed post-test.

Additional Question. How do the means computed for the various syntactic indices compare with the Hunt and Loban results?

Because of the influence of mode on student writing, these comparisons should provide important, new information about the influence of mode of discourse on the indices selected as well as providing some benchmarks for future researchers.

Limitations

This study did not include an assessment of writing quality. Rather, it relied on past research in sentence combining as support for the assumption made here, that sentence combining would not have a deleterious effect on the quality of compositions written by ninth graders.

It was not possible in the study to isolate ability level and teacher effect variables, so any interpretation of the results is limited to random groupings rather than stratified groupings of students. This is due to the uncontrolled and possibly interactive effect of the teacher variable with student ability level. It is, therefore, not justifiable to generalize to high, middle, and low ability groups although these were involved in the study.

Significance of the Study

The present study was an attempt to throw some light on the question of whether basic linguistic performance could be significantly developed in a relatively short period of time. It attempted to determine whether such a development, if possible, was reflected in the students' writing and reading comprehension.

The problem under study in writing could best be summed up by the following statement: It is incumbent upon researchers to determine whether T-unit growth resulting from sentence-combining is growth toward maturity. It is also important in terms of curricula to determine whether sentence-combining practice can be useful in improving

reading comprehension. A program which improved both these language skills would represent a highly significant development for the teaching of English.

Walter Loban in his 1976 summary of the findings from his massive and complex language development study concluded:

> If a little knowledge is a dangerous thing, no-one is at present out of danger in the study of language. There is need for many more interested researchers observing children's language in varied situations and making syntactic records of that language. (Complex truth is always an aggregate; each of us offers only part of an evolving mosaic.) (p. 90)

The experiment represents one attempt to add an important piece to that emerging picture.

Definition of Terms

Argumentation: "using language that, in the main, argues a point of view, defends a position, expresses an emotional inclination, or tries to persuade". (Perron, 1976, p. 6)

Clause: "a structure containing a subject (or coordinated subjects) and a finite verb (or coordinated finite verbs) . . . . " (Hunt, 1965, p. 15)

Comprehension: refers to the "derivation of the literal and/or inferred meaning of a sentence or passage" (Rutland, 1975, p. 2)

a/ Level of Comprehension - "indicates the depth of comprehension displayed by the student in reading the kinds of material he is ordinarily required to read in high school and college" (Davis, 1962, p. 5).

b/ Speed of Comprehension - "indicates the rapidity and

accuracy with which he understands the same material." (Davis, 1962, p. 5).

Conjoined adjectives: adjectives joined by coordinate conjunctions

Elaboration of language: the use of various strategies of syntax through which a T-unit is expanded beyond a simple subject and predicate. It deals with all strategies of expansion.

Factive Infinitive: an infinitive verb form used as a noun.

Fluency: the ability to write ideas in an easy and smoothly flowing, flexible, and logical way.

Maturity of syntax: "the observed characteristics of the sentence structure of older writers." (Hunt, 1965, p. 4)

Nominal: "any structure that functions as a noun, and its modifiers" (Hunt, 1965, p. 99). (Note: This does not include predicate adjectives).

Subordination ratio: $\frac{\text{Number of subordinate clauses}}{\text{Number of principal clauses}}$

Syntax: that part of the grammar that consists of rules for combining words and constituents into sentences.

T-unit or Communication unit (C-unit): "an independent clause with its modifiers" (Hunt, 1970, p. 14).

Two-Count Structures: passives, paired conjunctions, dependent clauses, comparatives, participles, infinitives as subjects, appositives, and conjunctive adverbs. (Evanechko et al. 1974, p. 317)

Chapter II

REVIEW OF THE LITERATURE

Psychologists, linguists, and educators have been studying language development for years but it was with the advent of the transformational grammarians and their psycholinguistic interpretations of syntactic complexity (Chomsky, 1957, 1965) that researchers were provided with an intuitively plausible theory to test. Since that time there have been numerous studies dealing with the various aspects of expressive and receptive language development. Some studies have concentrated on tracing this development using various indices and establishing norms for developmental levels. Other studies have attempted to improve students' syntactic skills in reading and writing by carrying out experimental treatments. This review of the literature first discusses the descriptive and experimental research concerned with written syntactic development, to provide the background information related to the analysis of writing carried out in the present experiment. Secondly, this review discusses the research which attempted to find a relationship between reading comprehension and the ability to process various syntactical forms, in order to establish the rationale for the present study's attempt to determine whether written syntactic practice might increase comprehension skills. The third section of the review is concerned with those studies which relate more directly to the present one, namely those that attempted to determine the relationship between syntactic

ability in writing and ability in reading comprehension. Selected descriptive studies and experimental studies have been reviewed.

Writing and Syntactic Development

There is little doubt that the ability to manipulate syntactic structures to convey meaning accurately and precisely is a mark of maturity in writing. Numerous studies have been carried out to determine what is involved in syntactic maturation and how it can be enhanced. Because of the growing number of studies listing syntactic frequencies for various grades and various writing indices this review concentrates on: 1) those studies involving grades seven to twelve; and 2) those indices which have shown promise for indicating syntactic development beyond grade seven.

Tracing syntactic development in writing -- descriptive studies.

In 1933, Lou Labrant conducted a pioneering study of written language development. She examined the writing of children in grades four to twelve as well as the writing of a group of published psychologists. Even at this early stage in the analysis of syntactic development, Labrant noted what Hunt (1965), Loban (1976), and others, have also found, that the increase in adjective clauses was one of the most significant indices of maturation in writing. Labrant did not have the benefit of the theory of transformational grammar to provide her with an analytic framework and thus concentrated on subordinate clauses and word counts. She did, however, find one additional item of interest which was elaborated on later by other researchers:

"While dependent clauses increased in frequency with increasing maturity of the writers, they also increased in complexity and clarity of thought." (p. 462)

Eleven years later, A.F. Watts (1944, p. 123) in The Language and Mental Development of Children reported the percentage of dependent clauses used by pupils at various age levels. Use of dependent clauses increased from sixteen percent in grade two to forty-three percent in grade six. Watts' study, which confirmed Labrant's findings, helped lay the foundation for the frequency count studies which followed upon the theories of the transformational grammarians.

In 1953, Walter Loban began what was to result in a very significant contribution to the study of language development. He undertook a study of the language skills of three hundred and thirty-eight students in Oakland, California. At the end of the thirteen year study he released his comprehensive analysis called Language Development: Kindergarten Through Grade Twelve (1976). This and his two interim reports Language Ability: Grades Seven, Eight, and Nine (1966), and Language Ability: Grades Ten, Eleven, and Twelve (1967) are important to the present study. Loban, however, was only one of three important researchers working in the area at this time. In 1965, when Loban's students had reached grade twelve, Kellogg W. Hunt published Grammatical Structures Written at Three Grade Levels in which he carried out a comprehensive syntactical analysis of student writing in grades four, eight, and twelve. He was searching for two things: 1) a method for quantitative study of syntactic structures, and 2) developmental trends in the frequencies of various syntactic structures. His results have often been referred to as representing

average ability at the three grade levels. In 1966, Hunt took his study one step further and analyzed the writing of eighteen average and eighteen superior students in each of grades four, eight and twelve, as well as the writing of nine authors in each of Harper's Magazine and Atlantic whom he referred to as Superior Adults. In 1970, he investigated the differences in the sentence structures written by schoolchildren at various stages of chronological maturity and the writing of certain adults with "all writers saying virtually the same thing" (Hunt, 1970, p. 10). To achieve this 'sameness', he used Roy O'Donnell's "Aluminum" passage which was composed of thirty-two kernel sentences which the subjects were asked to rewrite "in a better way". Although these results showed some of the same trends as did his earlier studies, this study had the following limitations: 1) the exercise was rather artificial as a writing activity; 2) the resulting compositions appeared to be unreliably short; and 3) the limitations placed on good writers almost certainly curtailed their overall fluency. Thus the majority of his data with import for the present research lies in his two earlier studies.

The third major contributor to this field of study was Roy C. O'Donnell. Along with William J. Griffin, and Raymond C. Norris, he carried out a study in 1967 reported as "A Transformational Analysis of Oral and Written Grammatical Structures in Grades Three, Five, and Seven". Like Hunt and Loban, they hypothesized that there were "significant differences among the grammatical structures of children at various grade levels, with structures increasing in syntactic complexity with advance in grade level" (p. 35).

Tables 1 and 2 (pp.18-19) contain the major findings of Hunt, Loban, and O'Donnell et al. as they relate to written syntactic development. Also included in Table 1 are the results of Perron's (1976) study in which he analyzed grade five compositions written in three modes. His data for his fifty-one grade five students' argumentative writing are also included to indicate the major effect caused by mode on the frequency of written syntactic structures.

The various syntactic indices were tested statistically by Hunt (1965) to determine which were the best indicators of a student's grade level. Only for those indices significant for grade at least at the .05 level was a contingency coefficient calculated. (Hunt, 1965, p. 23) Therefore, only they have been reported in Tables 1 and 2.

The most significant measure of both chronological and mental maturity in writing during the public school years according to Hunt (1966) is T-unit length. This measure eliminates the problem of improper punctuation and capitalization and overuse of sentence coordination by young writers. T-unit length was related significantly to grade at the 0.01 level and had a contingency coefficient of 0.69. Hunt concluded that T-unit length was "a more valid measure than sentence length and should replace it in all serious studies" (p. 45). Loban's (1976) longitudinal study confirmed Hunt's claim. Hunt (1965) found an average yearly increase of 0.73 words per T-unit. Loban (1976) found an average yearly increase of 0.86 words per T-unit between grades seven and twelve. A brief look at Table 1 shows that this growth is not even. Loban's data indicate a dramatic spurt between grades nine and ten, followed by a drop from grades ten to

Table 1

Major Indices of Syntactic Development (As reported by: Hunt, 1965 (N=18/grade); Loban, 1976 (n=211); O'Donnel et al., 1967 (N=30); and Perron, 1976 (N=51). (c) means that figures are calculated from data.

Index	Researcher	Sig. Level	Grade 5	7	8	9	10	11	12	Superior Adults
Words/T-unit	Hunt	.01			11.5				14.4	20.3
	Loban			8.94	10.37	10.05	11.79	10.69	13.27	
	Perron		13.06							
Words/clause	Hunt	.01			8.1				8.6	11.5
	Loban(c)			6.98	6.91	6.84	7.76	7.37	8.29	
	Perron		7.72							
Dep.Cl./T-unit	Hunt	.01			.42				.68	.74
	Loban(c)			.28	.50	.47	.52	.45	.60	
	Perron(c)		.72							
Noun Cl./T-unit	Hunt(c)	.05			.16				.29	.23
	Loban(c)			.08	.19	.15	.16	.15	.18	
	O'Donnel			.18						
Adj.Cl./T-unit	Hunt(c)	.01			.09				.17	.25
	Loban(c)			.05	.11	.13	.16	.18	.20	
	O'Donnel			.05						
Adv.Cl./T-unit	Hunt(c)	NS			.13				.20	
	Loban(c)			.15	.21	.19	.18	.12	.22	
	O'Donnel			.18						
% T-unit: short	Hunt	.01			39				23	19
medium	Hunt	.01			54				61	43
long	Hunt	.01			7				16	38

Table 2

Additional Indices of Syntactic Development

Index	Researcher	Sig. Level	Grade 8	Grade 12
Nom. Verbals/Clause:				
Gerunds	Hunt(c)	.01	.03	.07
Factive Infinitives	Hunt(c)	.01	.04	.07
Prep. Phrases as				
Noun modifiers/clause	Hunt(c)	.01	.10	.15
Verb Forms as				
Noun Modifiers/clause	Hunt(c)	.01	.07	.09
Genitives: Inflected &				
Phrasal/clause	Hunt(c)	.01	.29	.38
Adjectives/clause	Hunt(c)	.05	.39	.43

(c) means that figures are calculated from the data supplied.

eleven, and another dramatic rise in grade twelve. Nevertheless, Loban concluded that "the average number of words per T-unit has proven to be one of the most critical measures of language fluency developed during the course of this investigation" (p. 26).

Directly related to this index is one which was mentioned by Hunt (1965) but which has not received much attention. He noted that "the amount of writing done in 'short' T-units (1-8) words is a highly significant [negative] index of maturity. It has a contingency coefficient of 0.70, even slightly higher than that for T-unit length" (p. 52). The relationship between words per T-unit and this index is obvious, but it does provide additional information on T-unit growth.

The index of syntactic maturity which Hunt (1965) found to be the second major indicator of syntactic maturity was words per clause, which was significant for grade at the 0.01 level of confidence. He noted that "if we take the maturity of the skilled adult as the goal, the results show that words per clause is as good an index as words per T-unit" (p. 6). He had already noted that "clause length alone accounts for seventy-five percent of superior adults' superiority over grade twelve students in sentence length and T-unit length" (p. 141). A look at Table 1 reveals the very slow growth in words per clause which averages 0.26 when calculated from Loban's data and 0.13 in Hunt's data. Owing to the small sample size employed in the Hunt study, Loban's data are the more reliable. But again, the growth in words per clause is uneven with reductions from grades seven to nine, a spurt in grade nine, a consolidating plateau in grade eleven, and another spurt in grade twelve. In reviewing Loban's (1976) study,

O'Donnell (1977) wrote that the T-unit and elaboration index "show that growth in children's language does in fact follow a predictable sequence, but that the velocity or relative yearly growth rates are uneven" (p. 50). He suggested that negative data should be interpreted as plateaus in development. Despite these apparent plateaus in words per clause at various grade levels, Loban's data did indicate significant overall growth. Hunt's (1965) study had indicated that superior adults average about three and a half more words per clause than twelfth graders. He suggested that "clause length is especially worth studying as a place where substantial growth can occur" (p. 142).

One must keep in mind that the two previous indices, words per T-unit and words per clause, are interrelated with the index which Hunt (1965) found to be the third best indicator of syntactic maturity, the number of clauses per T-unit, which showed significant increases between grades four, eight, and twelve at the .01 level. This index along with words per clause represent the only possible ways of lengthening T-units. That is, the writer can add more dependent clauses or he can lengthen his clauses by adding more words or phrases, often through reductions of clauses to structures which are less than a clause. Loban (1976) observed a rather uneven growth from grades eight to twelve, where Hunt (1965, p. 20) found such significant growth in dependent clauses per T-unit. Loban postulated that the more nature students preferred tighter ways to coil their thoughts than dependent clauses permitted. Nevertheless, the number of dependent clauses per T-unit is still a significant indicator of syntactic growth at the junior and senior high school level.

As a result, when doing frequency counts to determine whether

syntactic growth has taken place, it is not sufficient to use words per T-unit alone. The growth that one finds may be due to over-reliance on the use of dependent clauses and the more significant growth in words per clause may be negligible. Because these three indices are interrelated, the researcher must determine where any measured growth is coming from.

Another major set of counts has been the number of noun, adjective, and adverb clauses used in subordination. Hunt (1965) found that the writer in grade twelve adds a dependent noun clause thirteen more times per hundred T-units than does his counterpart in grade eight. However, by grade twelve Loban's (1976) group still had not reached the frequencies found in Hunt's writing sample. In fact, they were below the frequency which they themselves attained in grade eight. Hunt's (1966) results indicated that his earlier findings were somewhat spurious. He made an important observation about the possible determinants of noun clause frequencies. The writing sample he analyzed was one thousand words of "whatever the students were writing". He noted that:

> While reading the themes of older students and tabulating the noun clauses, the investigators repeatedly noticed that one theme on one kind of subject would be full of noun clauses, while another theme by the same student on a different topic had almost none . . . . The two older, different groups write about the same number of noun clauses per given number of words as the two groups of fourth graders. (pp. 56-7)

Thus we can conclude that an increase in noun clauses is not indicative of increasing syntactic maturity from grade eight onward. Nevertheless, these structures were counted in the present study in order to gather further data about this sentence structure.

Increased use of adjective clauses was found to be indicative of increasing syntactic maturity by all researchers listed in Table 1. O'Donnell et al. (1967, p. 36) found a significant increase between grades three and five and Hunt between grades eight and twelve. Loban (1976), similarly, found a significant increase from grade three to grade twelve. The overall increase in adjective clauses, as Hunt (1965) observed, is "the most important developmental trend" (p. 80) in written syntax, well into maturity. He also concluded that "adjective clauses are the most significant factor in the increase of T-unit length" (Hunt, 1965, p. 108) which itself is the best single index of maturity in writing.

The frequency of adverb clauses per T-unit, however, "tells more about the mode of discourse and subject matter than about maturity . . . . The ceiling on adverbial clauses is reached in the middle grades" (Loban, 1976, p. 56). Hunt (1966, p. 58) reached much the same conclusion. The proportion of adverb clauses in his study dropped from forty-three percent in grade four to thirty-two percent in grade eight to thirty percent in grade twelve. (p. 62) His results for his first study (Hunt, 1965) found in Table 1 were also non-significant for this index. Although Loban (1976, p. 57) did suggest that better writers used more adverb clauses of concession, the differences were not significant, and he concluded that "the topic of any writing or speaking shifts the frequency of dependent clause functions". (p. 57) An increase in adverb clauses therefore, is not an indicator of syntactic maturity.

As Hunt (1966) observed: "among the three kinds of subordinate

clauses frequently used, only adjective clauses provide a highly significant index of both chronological and mental maturity from grade four to maturity". (p. 64) Earlier Hunt (1965) had stated that "adding clausal [illegible] more effective as a means of lengthening T-units, but adding non-clausal modifiers is more effective in achieving conciseness". (p. 408) With the exception of adjective clauses, then, one must look within the clause to find more indices of syntactic maturity. Table 1 shows that clause lengthening accounts for nearly half the growth in T-unit length between grades eight and twelve in Loban's study, 1.38 words per clause compared with 2.90 words per T-unit. Increased use of within-clause structures account for the remaining increase. Hunt (1965) provided more information on the syntactic trends within high school students' clauses than did Loban. Thus the information of concern to this study was derived from Hunt's data (as Table 2 indicates) and was supplemented with information from Loban whenever possible. The figures on syntactic structures used in clause expansion (Table 2) were computed from Hunt's data and are presented as frequency per clause. If frequency per T-unit is desired, the given frequency is simply multiplied by: 1 (for the independent clause) plus the dependent clause per T-unit figure.

Hunt found that average clause length correlated significantly with several kinds of noun modifiers: prepositional phrases (.567), adjectives (.547), infinitives (.431), genitives (.396), and present participles (.278). These structures accounted for "the major lengthening of the clause" (Hunt, 1965, p. 143). And as Hunt also observed, the number of these non-clause modifiers was ten to fifteen

times as great for each grade as the number of adjective clauses. Loban (1976, p. 45) corroborated Hunt's findings. Earlier in the same study, Loban (1966, p. 42) had observed that the higher ability students invariably used more nouns amplified by modifiers, more compound nouns, more clauses, and more infinitives. However, he did not provide data for these syntactic units.

Hunt (1965) detected a significant increase in the use of nominalized verbs and clauses as well. Factive infinitives, gerunds, and genitive modifiers especially showed great increases as is observable from Table 2.

> So the chief factor which lengthens clauses appears to be the increasing of nonclause modifiers of nouns and the nominalization of clauses. This factor and the increase in adjective clauses account in the main for the increased length of T-units. (Hunt, 1965, p. 143)

Loban (1976, p. 68) also recorded a measurable increase in gerunds and infinitives as the student matured. He made a rather interesting observation.

> In the history of the English language, the use of nonfinite verbal constructions has been increasing for the last five centuries. They are a way of simplifying, and they are forceful; they help us to express and subordinate thought effectively and directly. (p. 69)

Hunt (1965) also found a significant increase (.01) in the frequency with which six modals, "be" as an auxiliary, and the "can" and "have" auxiliaries were used by students in each successive grade. Although this expansion of the verb had only a slight effect on clause length, "this tendency suggests an increased modulation of the verb" (Hunt, 1965, p. 124). Variety in verb forms represents an increasing syntactic ability and these forms were tabulated in the present study.

In the midst of this frequency counting by researchers such as Hunt, important questions arose. The first question was in regard to the validity of the writing samples. Except for Hunt's (1970) study using the "Aluminum" passage and the study by O'Donnell et al. (1967), none of the authors controlled sufficiently well for topic and mode. And as Braddock et al. (1963) pointed out in their "Suggested Methods of Research":

> Before conducting a frequency count or using the results of one, a person should determine what his purpose is and then ascertain that the compositions used are appropriately controlled or sampled according to topic, mode of discourse, and characteristics of the writer. (p. 17)

Perron's (1976) findings also indicated that mode must be controlled in order to properly compare the syntactic ability of groups across grades. He found that mode, and therefore topic, produced significant differences in syntactic complexity as indicated by fifth grade writing when words per T-unit, words per clause, and clauses per T-unit were counted. The order for modes from simple to complex was descriptive, narrative, expository, and argumentative. All differences were significant beyond the .001 level. Perron's results thus amply demonstrated what Braddock suggested. If a researcher is attempting to determine what a student's written syntactic capability is, he should provide the student with the opportunity to write in the mode that elicits the most complex syntax. Crowhurst (1977, 1980) and Crowhurst and Piche (1979) confirm that the argumentative mode elicits the most complex syntax in student writing. Thus Hunt's frequencies of syntactic structures taken from "whatever the student wrote on" (Hunt, 1965, p. 1) or "whatever they were writing" (Hunt, 1966, p. 2)

and Loban's frequencies derived from "thirty sequential C-units obtained in the spring of each school year, grades three through twelve" (Loban, 1976, p. 8) cannot be considered as indicative of the student's syntactic capability at these grade levels. Their results are confounded by what Hunt (1965) observed as "a shift away from narratives" (p. 133) to expository writing. Henceforth, in attempting to establish grade norms for various frequency counts, researchers should control and report topic and mode.

The second question which arose during the frequency count research was also very important. The researchers who were interested in quantitative analysis of writing could not ignore the arguments of Moffett(1968) and Christensen (1968, p. 572) that the elaboration of clauses makes for "cursed hard reading". If they were right, the teaching of more complex and varied syntactic structures would have been a waste of time. With the recent increase in the number of sentence-combining studies, other writers have joined in to urge caution. Kinneavy (1979) has added to the earlier questioning of the sentence-combining goal.

> The readability people are generally concerned with lowering the number of words per sentence to a readable level, given the assumption that many of the sentences of modern prose are entirely too complex and too long. Yet increasing length and complexity of sentence structure is an avowed goal of many sentence-combiners. . . . At some point, syntactic maturity may move into syntactic senility. (p. 70)

Stewart (1979) re-iterates Kinneavy's concern: "The primary objective of sentence-combining curricula should not be that of simply securing gross increases in syntactic indices" (p. 11). However, accumulating research results suggest that elaborated syntax is at least not

detrimental to quality in writing. O'Hare (1973), Combs (1975), Golub and Kidder (1974), Loban (1976), Pedersen (1977), and Crowhurst (1980), all found that those compositions rated highest by teachers were the ones that used the most complex syntax. O'Hare and Combs used words per T-unit, words per clause, and clauses per T-unit to show that their experimental treatment groups wrote more complex syntax than did their control groups. The compositions in each study were marked holistically as well. Both of these researchers found that the experimental group's papers were judged significantly superior to the control group's papers by experienced English teachers who were unaware of the purpose of the experiments. Combs also found that the post-test compositions of his experimental group were rated significantly superior to their pre-test compositions, thus indicating that increased syntactic maturity does indeed enable students to write better. This was borne out by the analysis of Golub and Kidder (1974) who

> identified twelve variables which correlated significantly with teachers' judgements of written language samples. That is, if the occurrence of these twelve variables was high in a child's written language sample, the chances were also good that a teacher would put a high score on a child's paper . (p. 1129)

Their "Syntactic Density Score" included words per T-unit, words per clause, and clauses per T-unit plus others which dealt with elaboration of syntax within the clause. These were indices such as the number of prepositional phrases, the number of modals, and the number of "be" and "have" auxiliary forms.

Loban (1976) in his comments on teacher grading of compositions

throughout his thirteen year study wrote:

> In measure after measure, the subjects whose language power impressed numerous teachers are the ones who show empirically: longer C-units; greater elaboration of subject and predicate; and more embeddings, especially multi-base deletion transformations. (p. 85)

Pedersen (1977) in his replication of the O'Hare and Combs experiment found that at both post-test and delayed post-test occasions his experimental subjects "wrote more syntactically fluent compositions [W/TU; W/Cl] than control subjects wrote" (p. 67). He had "six professional teachers of composition . . . without any knowledge of the student or treatment" (p. 67) perform the two quality assessments: (a) overall quality assessment of matched compositions; and (b) semantic differential assessment using fourteen indices. "Subjects engaging in sentence-combining practice achieved, at the .001 level, better judged overall quality of writing, confirming earlier research results." (Pedersen, 1977, p. 2) More importantly, however, Pedersen also found that

> The comparison of group means suggests that the writing of experimental students who engage in extended sentence-combining is perceived as measurably different from comparable control students' writing in idea development, in maturity, and in concreteness of content. (p. 58)

Such results should help to allay skeptics' fears that "professional teachers of composition" are rating compositions high on overall quality on the basis of syntactic complexity alone.

While Pedersen used narrative writing for the assessments, Crowhurst (1980) in her study with grades six, ten, and twelve students used the argumentative mode. Four experienced teachers at each level rated the compositions. Syntactic counts were made for words

per T-unit; a holistic rating was given; and a composition instrument was used to rate the compositions on seven qualities. One of these qualities was sentence-structure. The results were interesting and important to the present study which also used the argumentative mode:

> Grade 10: High complexity arguments scored significantly higher than low complexity arguments on all three quality measures.
>
> Grade 12: High complexity arguments scored significantly higher than low complexity arguments on sentence-structure and composition quality scale. On the holistic scoring, high complexity arguments scored higher than low complexity arguments, but the difference between the two was not significant. (Crowhurst, 1980, p. 229)

It seems safe to agree with the conclusion reached by Kerek, Daiker, and Morenberg (1979) in their very careful study involving the relationship between sentence structure and writing quality in the compositions of college freshmen. They wrote:

> Undoubtedly, the major thrust of this study is in the evidence it has yielded for a significant correlation of these syntactic maturity gains with improved writing in free compositions, no matter how writing quality is measured. (p. 85)

Evidently, the ability to use complex and varied syntactic structures is an asset to a writer. The structures reported in Tables 1 and 2 are the best indices of syntactic maturity which research can provide at this time. In 1965, Hunt had concluded his study by suggesting the implications his results might have for curricula:

> This study suggests a kind of sentence building program that probably has never been produced or at least not systematically and fully. The aim would be to widen the student's span of grammatical attention and concern. The method would be for him to reduce independent clauses to subordinate clauses and non-clauses, consolidating them with adjoining clauses and T-units. He could work up to structures of considerable depth and complexity comparable to

> those exhibited by twelfth graders and superior adults. (p. 157)

Whether Hunt was aware of the attempt being made by Bateman and Zidonis, published in 1966, to produce such a sentence-building program, is here only a matter of speculation. But in 1966, they released the results of their two year experiment. This was followed by a series of studies involving sentence building.

Improving written syntax -- experimental studies. When Donald R. Bateman and Frank J. Zidonis published their results in 1966, they concluded as had Braddock et al. (1963), that traditional grammatical analysis would not result in improved student writing. They also concluded that corrective, functional, and formal grammars, which they refer to as "pseudo-grammars",

> do not account for the ways in which sentences are produced; in fact, they can offer no grammatical explanation for the process of sentence formation . ... . (p. 1)

In their study, they theorized that a grammar which gave students practice in generating various sentence patterns would result in increased student ability to use these sentence structures. They selected the ninth-grade class at Ohio State University School and randomly assigned twenty-five students to the experimental treatment, a study of transformational grammar. The other twenty-five acted as the control group. The treatment lasted for two years with writing samples collected during the first three months of ninth grade and the last three months of the tenth grade. The results were somewhat inconclusive. The experimental-group scores were superior to the

control group scores on "grammatical operations per sentence" but most of their increase was attributable to four students. The experimentals wrote proportionately more well-formed sentences but only at the .06 level of significance. Students in the experimental group also wrote with fewer mistakes measured by an "error change score" and this change was significant at the .01 level. Bateman and Zidonis concluded that "the study of a systematic grammar which is a theoretical model of the process of sentence production is the logical way to modify the process itself" (p. 37). The Bateman and Zidonis results pointed out the possibility that such programs could produce significant results in writing and suggested the need for more research on this question.

John Mellon (1967) argued that all Bateman and Zidonis did was restate the claim of traditional grammar in generative terms, that " the learning of grammatical formulations can result in their being consciously applied in the production of mature sentence structure" (p. 20). Mellon worked with two hundred and forty-seven seventh grade students who were assigned to three groups: transformational sentence combining, conventional parsing, and placebo. His aim was

> to direct a maximum of the students' attention to the way that content initially expressed in collections of separately represented kernel sentences could be collapsed into single statements . . . the student was given a set of kernel sentences plus directions for combining these sentences into a single complex statement, which he was then required to write out. (p. 32)

One other group studied traditional grammar and the placebo read extra literature and received direct method instruction in techniques for varying sentence structure when writing. The experimental

treatment lasted five months with the four previous months devoted to learning the transformational rules. The results were quite impressive with the experimental group experiencing significant growth on twelve indices of syntactic maturity. Growth on words per T-unit was from 9.98 to 11.25 and was significant at the .01 level. Growth for the traditional grammar group was non-significant (9.94-10.21). Because the students were simply directed to "write as well as they could", the results for the experimental group were very impressive. What was not so impressive was the finding that the experimental compositions were judged inferior to the control compositions for overall quality and indistinguishable from those of subjects who had studied no grammar. (p. 104) Thus, Mellon could not show that transformational sentence-combining practice involving the learning of transformational rules would result in increased quality of student writing.

Francis Christensen's Notes Toward A New Rhetoric was also published in 1967. In it, Christensen criticized current composition teaching practices saying that "we do not really teach our captive charges to write better--we merely expect them to" (p. 3). In 1968, in an article entitled "The Problem of Defining a Mature Style" he questioned Hunt's emphasis on nominalizations and relative embeddings:

> The sentences of all but the most immature or inept writing are made long, in part, by a class of constructions far different in rhetorical effect from nominalizations and relative embeddings. This class is the so-called sentence modifier or free modifier. (p. 576)

Christensen pointed out that there is not much scope for adding free modifiers within a clause where mostly word modifiers or bound modifiers are used. He claimed that these latter modifiers restricted

the writer's rhetorical freedom.

In 1970, in a "cosmopolitan" area of Edmonton, Joel L. Gajadharsingh carried out a well-designed but little-reviewed study in sentence-combining based on Christensen's "rhetoric of the sentence". He wanted to know whether or not students who received such instruction would write "more stylistically mature prose than students who received instruction in the principles of writing through traditional grammar" (p. 73). His results showed that the experimental group achieved significantly higher adjusted mean scores on the post-test (three paragraphs, eighteen T-units) than students in the control group. The experimental group also used a significantly higher number of free modifiers, a significantly greater adjusted mean number of words per T-unit, and a significantly greater adjusted mean number of intra-T-unit coordination. (p. 249) He also found no significant difference in the performance of students in the experimental group on a delayed post-test, given three months later, when compared to their post-test scores. He concluded with the remarks that the study "indicated that under the conditions of the experiment, instruction in the rhetoric of the sentence has a beneficial effect on the written composition of grades seven, eight, and nine students" (p. 252).

However, despite such seemingly impressive results, this researcher like Moffett (1968) tends to be somewhat skeptical of such cumulative sentence techniques. Moffett questioned Christensen for making the "descending clause-modification structure a doctrinaire kind of absolute good" (p. 43). Moffett thus pointed out the absurdity of advocating one syntactical pattern to the exclusion of the others. He

also noted that "Syntactic complexity is no virtue in itself. The point is to be able, not obliged to complicate one's sentences" (p. 171). Research now suggests that acquiring facility with syntactic patterns may be achievable through sentence-combining practice.

Frank O'Hare (1971) designed his experiment to provide students with practice in the use of the same syntactical patterns that Mellon had used. He hypothesized that such sentence-combining practice, not dependent on any formal knowledge of grammar, would help seventh graders to write syntactically better compositions than students of similar ability who did not receive such practice. O'Hare wondered if Mellon's rules were too complicated for seventh graders. Presumably, by increasing the ability of students to manipulate syntactic units by patterned practice, uninhibited by grammatical terminology, the teacher would be expanding the students' number of syntactic options as Moffett(1968, pp. 176-177) recommended. O'Hare's experimental group was composed of forty-one grade seven students from the Florida State University School. His control groups contained forty-two students. They were randomly assigned to treatments which lasted for eight months and involved one and a half hours per week. The experimental groups practiced the same sentence-combining patterns as Mellon's group. The control group did a language study unit consisting of teacher-made study sheets and exercises on vocabulary study, dictionary skills, punctuation, capitalization, and usage. (O'Hare, 1973, p. 34) Both groups wrote the same number of compositions, with a minimum of two and a maximum of four pages per week. O'Hare analyzed the first ten T-units in each of five pre-test and post-test

compositions which ranged over the narrative, descriptive, and expository modes.

O'Hare's results caused a stir in the composition world. Mellon (1969) had shown that up to two years growth was possible when measured by words per T-unit and eleven other similar indices. Using six of these indices: words/T-unit, clauses/T-unit, words/clause, noun clauses/T-unit, adverb clauses/T-unit, and adjective clauses/T-unit, O'Hare found his experimental group of seventh graders achieved at a level similar to, and on four occasions superior to, Hunt's (1965) norms for twelfth graders. Only on noun clauses were they below the norms reported by Hunt. (O'Hare, 1973, p. 55) The experimental groups established a highly significant superiority at the .001 level over the control group on all six factors. But even more impressive was that, after "normalizing the papers of fifteen controls and fifteen experimental students who were paired on the basis of sex and IQ, eight experienced English teachers, with no knowledge of the experiment, judged the papers of the experimental group to be significantly better at the .001 level when rated for 'general impression'" (O'Hare, 1973, pp. 62-69). As O'Hare stated,

> The results raise questions about the belief that growth in writing ability is necessarily a slow and difficult process. In showing that significant qualitative and syntactic gains can be achieved in approximately eight months, the study suggests that, at least for seventh graders, a part of the composing process is definitely amenable to alteration. (p. 73)

O'Hare pointed out other positive aspects of his program such as the increase in student confidence, the absence of a need to learn terminology, the program's simplicity, and the high student interest it generated. As well, he found no interaction between treatment and

teacher, sex, or ability of students. O'Hare's study lent credence to the idea that a writer's ability to manipulate syntax is related to quality in writing and that complexity achieved by subordination does not necessarily result in writing that is "cursed hard reading" as Christensen and others had suggested.

The studies on sentence combining that followed closely upon O'Hare's did not really add much information to the writing improvement picture. Fisher (1973) conducted a five week (twenty-five minutes per day) treatment with small classes of grade fives, sevens, and nines. O'Donnell's "Aluminum" passage and a "Cotton" passage developed by the author were used as both pre- and post-tests. The results showed that all experimental groups wrote more maturely as measured by T-unit length, clause length, and clauses per T-unit at the .05 level. However, one must be skeptical about the use of a measure requiring students to combine kernel sentences into more complex sentences. Such measurement is biased in favor of the group receiving the sentence-combining treatment. Perhaps such increases are due to the interaction of the treatment and the instrumentation.

Hughes (1975) carried out a similar study with small groups of twelve grade-seven students. She also used the "Aluminum" passage both as a pre- and post-test. After thirteen weeks and an average of thirty-one hours in manipulating adjective clauses, Hughes concluded that "writing fluency made large gains according to teacher observation" (p. 54). However, the same problems were encountered in the Hughes study as in Fisher's study. The gains may be attributable to the basic task similarities of the "Aluminum" passage and the sentence-combining activities. Comparing the two groups on increased

writing ability using such instruments is not a fair test. It is questionable whether Fisher or Hughes would have found the same syntactic growth in their subjects' free-writing. The major concern of composition teachers must remain with syntax in free-writing, not with students' ability to combine a given set of sentences after practicing such combinations.

Thus, not until Warren Combs' (1975) study was important information added to that which O'Hare had accumulated. Combs was interested in the effect of sentence-combining practice on the writing of seventh graders and included a delayed post-test in his design. He used two intact experimental classes (forty-nine students) and two intact control classes (fifty-one students) in a suburban Minneapolis junior high school. He took care to construct his sentence-combining problems in the Mellon and O'Hare tradition. As a result, his treatment differed from O'Hare's "only in the number of sentence-combining problems and the complexity of the base sentence of the sentence-combining exercises" (Combs, 1975, p. 45). The amount of time spent was roughly equivalent to the number of hours of sentence-combining treatment which O'Hare's experimental group received. However, Combs' treatment lasted only ten weeks and O'Hare's lasted eight months. Combs had matched-compositions (experimental/control; pre-test/post-test) evaluated on the basis of words per T-unit and words per clause. A comparison of the results on the syntactic indices which Mellon, O'Hare, and Combs found for their experimental groups is presented in Table 3. All three used free-writing samples for their analyses. The results from Combs' (1975) experiment

Table 3

Results From Three Studies Using Sentence Combining:

Experimental Groups Only

Index	Author	Pre-test	Post-test	Delayed Post-test
W/TU	Mellon	9.98	11.25	N/A
	O'Hare	9.63	15.75***	N/A
	Combs	9.48	11.65***	10.99**
W/CL	Mellon	N/A	N/A	N/A
	O'Hare	7.06	8.55***	N/A
	Combs	7.03	7.74**	7.57*
CL/TU	Mellon	N/A	N/A	N/A
	O'Hare	1.36	1.84***	N/A
	Combs	1.35	1.51	1.45

* Asterisks indicate the level of significant differences between pre-test and post-test, and between post-test and delayed post-test: *$p < .05$; **$p < .01$; and ***$p < .001$.

confirm O'Hare's conclusion that sentence-combining practice results in increased syntactic ability as measured by the three indices in Table 3. However, the increase attained by Combs' experimental group was closer to the increase achieved by Mellon's (1967) group than to the increase achieved by O'Hare's group. Pedersen (1977) replicated the studies by O'Hare and Combs with grade seven students. He carried out two of the main frequency counts (W/TU and W/CL) that the other two researchers used. The control group scores did not change significantly on either index. The experimental group did not significantly increase their mean words per clause; however, they achieved an increase in words per T-unit which was significant beyond the .01 level. The comparison with the control group showed the experimental group to have significantly outperformed the control group consistently on both the post-test and delayed-post-test occasions.

Robinson (1979) conducted a study using O'Hare's Sentencecraft program with forty grade seven students but no control group for comparison. He used the "Aluminum" passage for both the pre-test and post-test. The increase in T-unit length following completion of the program was found to be significant. Robinson noted that "all students in this selected sample seemed able to handle more complex embeddings after they had experienced the Sentencecraft program" (p. 4)

Sullivan (1979) used the Getting It Together program which was constructed under the direction of Charles R. Cooper. Her study involved grade nine and eleven students. The grade nine results indicated that "the growth in the two syntactic factors, mean T-unit length and frequency of the use of final free modifiers were

significant at the .01 level and the .001 level". (p. 85) These results demonstrate the efficacy of sentence-combining in increasing the syntactic complexity in student writing. Other studies have been carried out at university level, the most impressive being the one by Kerek, Daiker, and Morenberg, (1979). The trends reported in the lower levels seem to carry on. These three researchers conclude that

> The results of this experiment strongly confirm the claims advanced for SC in recent years: SC clearly helps accelerate syntactic growth even among young adults, and it is significantly more effective than the conventional essay-analysis approach in increasing the overall writing skills of college freshmen . . . . The subordination ratio, clauses per T-unit, post-test differences between the two groups remained not significant, as they were on the pretest. But in the mean clause length, words per clause, which Hunt claims is the syntactic factor that best differentiates adolescents from professional writers, the experimental group gained .89, nearly one whole word whereas the control group dropped .13. This difference is statistically significant at and beyond the .001 level of confidence. Note that in fifteen weeks the treatment group increased its clause by almost half of the growth experienced, according to Hunt, in the previous eight years of normal development. The same group showed a nearly as impressive .74 word gain in T-unit size [words per T-unit] in contrast to the .05 word drop by the control group. The post-test difference between the groups, 1.1 word, is again significant at better than .001. (p. 170)

Not only is the design of the study impressive, the results are impressive also. However, a few words of caution are in order at this point for two reasons. First, as Manning (1979) points out very clearly, not all sentence-combining studies have used identical treatments and this affects the generalizability of the studies mentioned above. He writes:

> The SC treatment is acting as a fixed variable in the statistical model. As a fixed variable it represents all the treatments to which inferences can be made in the experiment. . . . The studies are limited in claiming success to the particular treatment used in the study. They cannot claim success for the population called sentence-combining

> . . . . Replication is the solution to the problem of limited generalizability. (pp. 3-4)

The second caution concerns the question of syntactic growth as measured by the words per T-unit type of index. In the studies discussed above, the experimental groups achieved significantly greater increases on such indices than did the control groups. But as Hunt (1965) warned

> forced growth is not always firm growth. Perhaps older students' proficiency comes only as a result of years of psychological and experiential maturing. It may come only with the development of all thought processes. In that event, attempts to force the growth will be futile. . . . More than one child has been debilitated by excessive self-consciousness. (p. 158)

The question raised here is extremely important. If the syntactic increases resulting from sentence combining are artificial or forced, then the proponents of this technique have a major problem.

Combs' (1975) results suggest that at least part of this growth is forced. His experimental group gained 2.17 words per T-unit during the experiment but quickly lost 0.66 words per T-unit (roughly one year's growth), as measured by a delayed post-test eight weeks later. Perhaps the same thing happened with Combs' seventh graders as Ney (1974) found with his fourth graders. Ney noted that conjoined adjectives were extremely rare in children's writing at this grade level. Nevertheless, he attempted to teach his class to join two sentences to form a third with coordinated adjectives. After two or more training sessions of one-half hour each the error rate was still over sixty percent. Yet after the same amount of time the error-rate for the "WH-be deletion" transformation was four percent. Ney also noted that although he could train his fourth graders to write

sentences using present participial phrases as noun modifiers, they did not use sentences with these in their free compositions. (Ney, 1974, p. 167) He attributed his students' poor improvement on the coordinated adjectives to their being developmentally unready. Possibly grade seven is too early to expect major growth and retention in free writing of the more complex sentence structures which Mellon, Combs, and O'Hare had their students practice. For example, Mellon (1967) concluded that grade seven was too early to try to enhance development in the use of logical conjunctions such as "if, although, etc." (p. 70) Forcing the student to practice some of these syntactic forms before he is developmentally ready appears to be a useless exercise if free composition is used as the measure.

The post-test decrease in words per T-unit in the Combs' study may also have another explanation. It may have been due to the cessation of concerted practice in writing. Combs' groups completed only one major writing assignment in Language Arts in the eight week delay period. The fact that his students had also ceased the sentence-combining treatment, especially the rewriting involved, might easily have resulted in the "significant decrease". One composition in eight weeks would certainly put the grade seven student in an artificial situation. He would be 'coming in cold' to the delayed post-test. It is therefore questionable whether the procedure gave a fair indication of student syntactic ability.

In his study with grade seven students, Pedersen (1977) found that his experimental group sustained the significant pre to post-test growth through to the delayed post-test eight weeks later. However,

Pedersen's experimental students wrote in journals each week as well as receiving "instruction and practice in the writing of developed paragraphs related to the content of the two novels" during the post-test to delayed post-test period. (p. 46) Despite these more positive results, the question of temporary, perhaps forced, growth requires further study.

The question of whether the growth achieved through sentence-combining practice is artificial has not been explored since Mellon (1967) did transformation counts in his study. Hunt (1965) had already done his preliminary investigation into establishing mean frequency counts for the various syntactic structures used in elaborating language but Mellon counted only the larger structures such as subordinate clauses. Apart from Mellon's preliminary data, researchers in the sentence-combining area have neglected frequency counts in favor of the much simpler indices of words per T-unit, words per clause, and clauses per T-unit. However, this omission can lead to unjustified conclusions. Because there are two ways to achieve T-unit lengthening in terms of these indices, words per T-unit should not be used alone to measure syntactic growth. For example, O'Hare's (1973) results show a mean increase of 6.12 words per T-unit in the experimental group's compositions. A simple arithmetical calculation reveals that sixty-five percent of this increase was due to additional use of subordination. But as Loban (1976) observed, dependent clauses are not always the most sophisticated strategy for subordinating elements of thought. In fact, Hunt's (1965) research shows that true development, in terms of superior adult writing, requires clause

growth rather than simple increased subordination.

> Superior adults show a 47 percent gain in sentence length over twelfth graders . . . only a 6 percent gain in the number of subordinate clauses . . . but a whopping 36 percent gain in clause length. (p. 57)

From this statement, it can be concluded that a sentence-combining program which increases the amount of subordination until 84 percent of the T-units contain a subordinate clause, as did O'Hare's experimental group, is artificially inflating the words per T-unit index of growth. In his 1975 study, Combs' experimental group subordination frequency is only 0.45 as measured by the delayed post-test. For Loban's (1976) grade twelves, the figure was 0.60, and for Hunt's (1965) grade twelves it was 0.68. Despite the fact that the Combs figure is low in terms of these grade twelve norms, fifty-one percent of the growth in words per T-unit which Combs' experimental group did achieve was due to increased subordination. The more pertinent index for measuring growth, in the light of this tendency of sentence-combining practice to result in a large increase in the number of clauses per T-unit, is words per clause.

Loban (1976) found an average yearly increase of 0.26 words per clause between grades seven and twelve. The O'Hare study shows the most significant results for this index. The mean developmental increase for his experimental group using this index was 5.7 years. Combs' mean experimental group increase was 2.7 years but fell off to 2.0 years by the delayed post-test. O'Hare's results are certainly the more impressive. But his design did not include a delayed test; therefore it is not known what effect a period of delay would have had on his results. What syntactic structures contributed to this

increased clause length is also unknown as neither O'Hare nor Combs included frequency counts of the various within-clause structures. It should be noted that the increase in words per clause found in both studies could be accounted for by the simple addition of one or two adjectives per clause. A researcher would be on pretty shaky footing if he equated this with the true syntactic development traced by Hunt and Loban.

It is important then for researchers to determine whether the increases in words per T-unit and words per clause which result from a sentence-combining program are really forced and artificial growth or whether such a program fosters true syntactic development.

Reading Comprehension and the Influence of Syntax

A number of authors have concluded that syntax must have a major effect on reading comprehension but few researchers have been able to satisfactorily measure this influence. In a study entitled "Sentence Structure as a Factor Affecting Comprehension" Pavlak (1973), undertook a review of all "available, scientifically accurate studies conducted within the past twenty-four years." (p. 1) The strongest conclusion he could reach was this: "There is a relationship between children's knowledge of grammatical structure and their ability to comprehend in reading." (p. 180) The correlation between the two variables, however, was very low for each study he reviewed.

Yet the rationale for suggesting the importance of syntax appears to be very sound. Gilliland (1972) demonstrates this importance with a very simple example.

> It would be very convenient if we could work out the meaning of a sentence by adding up the meanings of the words considered separately. This is unfortunately not possible because the way in which words interact prevents us from doing this. Whereas, the sequence of words rarely affects a sentence in Latin for example, the linear sequence of words in patterns in English is extremely important. . . a "venetian blind" does not mean the same as a "blind Venetian" . . . . Our ability to extract meaning from the phrase is closely related to the grammatical structure underlying it. (p. 69)

Despite the general difficulty in establishing syntax as a major variable affecting reading comprehension, as reported by Pavlak, a number of researchers have achieved some success. Stolurow and Newman (1959) carried out a factor analysis of readability elements using adults and found that ten of their twenty-three factors or variables accounted for 92.8 percent of the total variance. These were than rotated, and two major factors emerged: "relative difficulty of words" (thirty-four percent) and "relative sentence difficulty (twenty percent). (Stolurow and Newman, 1959, p. 245) Klare (1963) in his book The Measurement of Readability refers to a study by James Brinton and Wayne Danielson (1958). They found six of their twenty factors with correlations significantly different from zero. Two emerged as being more important. They were called a "vocabulary factor" and a "sentence factor". (Klare, 1963, pp. 163-164)

Thus, as Little (1972) concluded from his review of research, there is sufficient theoretical and empirical evidence upon which to base the claim that "the syntactic structure of written language has a major effect on the reading comprehension of that language" (p. 23).

Reading comprehension and syntax--descriptive studies. There are many studies dealing with the reading problems experienced by students in the elementary grades as a result of overly complex syntax. Albanese (1972) used a reading test with limited vocabulary which was designed to measure the effects of varying structural depth and structural organization on reading comprehension. Her subjects were grades three and five students. She found that they had more trouble processing adverb and adjective clauses than adverb and adjective phrases but that all four of these structures made comprehension more difficult. Bormuth, Manning, Carr, and Pearson (1970) found that large proportions of grade four students were "unable to demonstrate a comprehension of the most basic syntactic structures" (p. 353). Rutland (1975) also used grade four students and tested them with cloze passages with controlled syntax. Although he was not concerned with "basic" syntactic structures, he did find that complex syntax posed reading difficulties for the students. His results showed that:

> 1. A sentence written with a nominalization of an active verb was significantly harder than a sentence written with a detransformed active verb.
> 2. The embedded clause structure was significantly more difficult than the right-branching clause structure.
> 3. The relative clause was more difficult than an adjective. (pp. 97-107)

Bormuth (1966) wrote:

> Unfortunately, many adults and children fail to understand what they read, not because the concepts are too difficult or because they lack basic reading skills, but simply because of the complexity of the language in which those concepts are presented. (p. 81)

In his study, involving grades four through eight, he used cloze tests from twenty passages representing a variety of prose styles and

ranging in difficulty from about grade four to grade eight. He concluded that comprehension difficulty was undoubtedly caused by the language itself, but these linguistic features constituted only a part of the causal variables. "Semantic, organizational, and context variables undoubtedly constitute other causes of difficulty". (Bormuth, 1966, p. 128)

Fagan (1971) hypothesized that the presence of a particular sentence structure could affect a child's reading comprehension of a particular passage. He used cloze tests of basal passages with grades four, five, and six students. In each passage, he inserted a simple transformation, an embedding, a conjoining, a deletion, or position shift transformation. His results indicated that

> ... . deletion and embedding transforms tend to make sentences and passages difficult for children to read. It appears that pupils have difficulty processing the information of these structures and consequently experience difficulty in understanding the sentences in which they appear (p. 170-171).

Fagan listed the most difficult structures: appositives, 'ing' -nominalizations, genitive pronouns, common elements deletion, and negatives. He also found that sentence difficulty was more dependent on difficulty of transforms than was the difficulty of a passage, and he suggested that this was accounted for by passage redundancy.

Breneman (1975) also used cloze tests as one of her instruments in assessing the relationship between linguistic awareness and reading comprehension with fourth, fifth, and eighth graders. She found that her "subjects' linguistic awareness" was positively related to

reading comprehension as measured by the MAT reading subtest. She also found that this relationship remained constant over grades four, six, and eight. Therefore, she concluded that "syntactical utilization is probably fundamental to all stages of the reading process" (p. 83).

This conclusion is in line with Smith's (1969) findings in a study using cloze tests on passages written at grade four, eight, and twelve syntactic levels based on Hunt's (1965) norms. Smith concluded that the level of syntactic complexity does affect comprehension. Grade eight students found fourth-grade passages harder to comprehend than those written on their own linguistic level. However, they found twelfth-grade passages too difficult to comprehend. Takahashi (1975) found a decreasing correlation between grade-six and grade-nine achievement on the Nelson Reading Test and her test of syntactic comprehension. She concluded that syntax plays a decreasing role in comprehension as the student matures. (p. 42) Taking her study and Smith's together, one would conclude that syntax is an important factor in reading comprehension at least through the middle school years but that its role becomes less important in later years. However, most researchers agree with Carroll (1970), quoted in Takahashi (1975). Based on a study involving third, sixth, and ninth graders Carroll concluded that:

> Acquisition of lexico-grammatic meaning was a slow process, far from complete by the ninth grade. Each succeeding grade did better than the preceding one. (Takahashi, 1975, p. 28)

This discussion of syntax and its effect on comprehension has concentrated on establishing that syntax does continue to affect

reading comprehension at least up to the senior high school level. Olson (1973) gave a plausible reason for the problems posed by complex syntax for the immature reader:

> We are coming to realize that although purely syntactic considerations in language acquisition are of great intrinsic interest, changes in the semantic structure of language and in many underlying cognitive processes are necessarily related to syntactic development. One important mental ability needed by the competent language performer would have been taxed if I had read the previous sentence aloud to you. That sentence, both because of its length and its structural complexity, would have required that you store large segments of information in immediate memory in order for you to come up with its meaning--if in fact you could. (p. 145)

This mnemonic skill needed for comprehending syntactically complex writing has been explored by a number of researchers working with subjects at the university level. Simison (1970) hypothesized that sentences which were left-branching (subject/noun phrase/complement) would be harder for undergraduates to process than right-branching (object/noun phrase/complement). She found this to be true and concluded that this was probably because the hearer, as Olson later suggested, was required to hold a large part of the information in storage before he could process the complete sentence. (p. 59)

Savin and Perchonock (1965) undertook a study to determine which kind of sentences took the most space in immediate memory. Their undergraduate subjects were given a sentence followed by a string of words. They were required to recall the sentence perfectly and then recall as many of the extra words as possible. The researchers theorized that the number of extra words recalled would indicate the amount of memory capacity remaining after the sentence had been encoded. Their results confirmed the order of difficulty as follows:

emphatic passive, passive-negative question, passive question, negative passive, emphatic, negative question, negative, passive, question, relativization, and kernel. The kernels, of course, were simplest. The authors concluded:

> Because of the nature of English grammar, grammatical complexity is partly confounded with the number of words in a sentence. It is clear, however, that length in words cannot by itself account for the present results. Neither the question transformation nor the emphatic lengthens any of the sentences in this experiment; yet both transformations increase the difficulty of remembering sentences. (p. 352)

Despite the relative crudity of this method for measuring storage space in immediate memory, it does suggest that more complex syntactic structures are harder to process and therefore harder to understand.

Gough (1965) and Slobin's (1966) findings supported those of Savin and Perchonock. But like Letson (1958), they were also concerned with the influence of complex syntactic structures on speed of comprehension. Slobin measured response times of subjects in establishing the congruity or incongruity of syntactically complex statements and pictures. His subjects were kindergarten students, grades two, four, six and summer-school university students. He found that "the main effect of grammar was statistically significant at the .001 level" (p. 223). He further concluded that his findings supported the notion that the syntactic complexity of sentences could be concretely reflected in decoding time while at the same time admitting that his measure was too crude for accurately measuring speed of comprehension. Gough (1965) also used response times. His subjects were introductory psychology students. The procedure was much the same as Slobin's.

He found that active sentences were confirmed significantly faster than passives; affirmative, faster than negative; and true, faster than false. But he also suggested that part of the reason for these findings might have been the frequency with which the various sentence patterns were encountered. He concluded that "the present results demonstrate that syntax is related to speed of understanding". (p. 111)

Letson (1958) attempted to measure reading rate in accordance with the difficulty of the material. His subjects were college freshmen. They were asked to read a selection (2500-3000 words in length), mark their place, and answer thirty multiple choice questions. Reading rate and comprehension scores were computed. He found that the correlation between rate and comprehension on easy material tended to be high, but that as the difficulty of the material increased, the coefficients decreased.

It seems reasonable to conclude from this body of research that syntax is a definite factor in determining the comprehensibility of a passage and the speed with which the passage can be understood. The concern of the present study is the degree and speed of comprehension.

Hunt (1966) pointed out an interesting phenomenon relating to student ability to comprehend and write complex syntax.

> Students in the middle grades can read with 80 percent comprehension, sentences whose clauses are as long as those they will write only when they have become about four years older. They can read syntax that is about four years beyond what they can write. (p. 82)

Power of comprehension scores should not suffer as much as speed of comprehension scores at the grade nine level when syntax of a passage

is made more difficult. By providing the student with written practice in manipulating the more complex syntactic forms, he should achieve a certain facility or speed in processing these forms. Such practice should not only result in increased maturity in written syntax but also in speed of comprehension. Ney (1974) suggested this possibility when he wrote that growth in syntactic fluency was not due to

> . . . changes in the linguistic ability of students but to . . . changes in certain skills which students utilize in the writing process. These skills include 1) mnemonic skills, 2) sentence-processing (or reprocessing) skills, and 3) skills connected with the raising to conscious control of linguistic resources which are innate to the student (p. 168).

Stotsky (1975) in a comprehensive review of the effects of sentence-combining practice advocated that more studies be undertaken and gave the following rationale for using sentence combining as a means of improving reading comprehension:

> Inasmuch as reading, speaking, listening and writing are all language-based activities, one can assume an interrelationship among all the language arts. It is theoretically plausible to maintain that growth in one area should be reflected to some extent in other areas. It is also plausible to maintain that the nature of these relationships may change as children move into higher stages of intellectual growth. As older students begin to think with and about complex language structures in their efforts to write coherently, it is conceivable that these efforts could contribute to, as well as flow from, their linguistic and intellectual development. (p. 66)

Researchers have tried to improve comprehension by improving students' knowledge of syntax or grammar but have failed. Before reviewing the attempts to improve comprehension through sentence-combining practice, a review of these grammar-oriented studies is in order.

Improving comprehension by improving knowledge of syntax or grammar--experimental studies. Roy O'Donnell made the most concerted effort to improve reading comprehension by improving subjects' knowledge of syntactical forms and their relationships. He began his series of studies in this area in 1961 with grade twelve students as his subjects and constructed his Test of Structural Relationships in English using nonsense words. He theorized that this test would enable him to control vocabulary and measure subjects' ability to recognize the various structural relationships of words in English sentences. After finding a correlation of 0.44 between these linguistics structure scores and Level scores on the Davis Reading Test (Form 2C), O'Donnell concluded:

> The correlation between awareness of str[illegible] relationships of words in sentences and ability [illegible]ding comprehension as indicated in this study is not sufficiently high to give conclusive evidence in support of the teaching of linguistic structure as a major means of developing reading comprehension. (O'Donnell, 1961, p. 90)

O'Donnell considered speed scores "irrelevant". His decision not to examine the effect of increased knowledge of syntax on speed of comprehension may have resulted in loss of valuable information. It is difficult to see how he reached this conclusion.

In 1974, O'Donnell and King carried out a study with seventh graders who were reading below grade level to see if students' reading comprehension could be improved by increasing their deep structure recovery skills. Students were given practice in de-composing complex syntactic structures into their constituent kernels and recombining them in paraphrase form. At the end of the project, correlations of 0.59 and 0.69 were obtained between Simons' Deep

Structure Recovery Test and cloze test scores, and between the DSRT and the California Test of Basic Skills Reading Comprehension scores. But the gain from pre-test to post-test was very slight and "could not be attributed to the influence of the instructional materials" (p. 337). The researchers did conclude, however, that the results "seemed to indicate a strong positive relationship between reading comprehension and certain aspects of sentence structure" (O'Donnell, 1976, p. 3) They attributed the small gains to poor morale and lack of readiness to profit from the program.

In 1975, O'Donnell and King undertook a second study involving grade nine students "to explore the possibility of increasing sensitivity to syntactic structure by direct instruction" (p. 258). They provided a twenty-four page module of programmed exercises dealing with basic sentence patterns, parts of speech, subordinated sentences, sentence analysis and synthesis. They used O'Donnell's Perception of Alternate Structures Test which used nonsense vocabulary but English inflections and function words (e.g. The birtle scared the ilbid). After four weeks of instruction, their grade nine subjects scored as high as tenth graders. They found a correlation of 0.41 between the PAST and reading comprehension scores. (O'Donnell, 1976, p. 3) O'Donnell noted that this correlation was nearly the same as for his earlier test. He then developed the Agnate Sentence Test using conventional English vocabulary. "It was designed to measure ability to perceive similarity of meaning of sentences that were structurally different." (O'Donnell, 1976, p. 3) He found a correlation of 0.65 between eighth graders' scores on this test and

reading scores, nearly the same as for the DSRT (which also used conventional vocabulary). O'Donnell (1976) concluded from this series of studies that:

> The good reader does have to sense the relationship of a word to other words in a sentence. But I do not think the structural cues are frequently sufficient for the recovery of meaning; when they are not, we fall back on syntactic cues. Thus, ability to sense relationships of words is essential, but the function of syntactic cues is that of supporting and/or clarifying cues of semantic structure. Thus it is to be expected that measures of sensitivity to syntactic cues, while correlating positively with measures of reading comprehension, would correlate less highly than semantic knowledge with such measures. (p. 7)

O'Donnell's findings were quite inconclusive. It is easy to see why an increased knowledge of grammar would not necessarily improve reading comprehension. However, his results did suggest that improvement in student facility in using various syntactic forms can result in improved reading comprehension. As Latham (1973) commented,

> Being a native speaker of a language means possessing the knowledge that certain words group together and that these groups can inter-relate in a variety of ways, and thus, produce specific underlying meanings. It does not involve why certain words can be integrated into groups or why particular relationships can obtain between such groups so that certain meanings are generated. . . . It is clear then, that to achieve the comprehension of the sentences of written language, the accomplished adult reader will possess facility in synthesizing individual words to form the grammatical word groups used in the production plan for the sentence. (p. 123)

It is with Latham's last point in mind that the literature highlighting the relationship between the two language arts of reading and writing is reviewed here.

Reading Comprehension and Written Syntactic Ability: The Relationship

In his longtitudinal study, Loban (1963, 1966, 1967) found the same results over and over again: students who read well also wrote well. In his 1976 profile of kindergarten to grade twelve language development he concluded that "In measure after measure, the subjects whose language power (written and oral) impressed numerous teachers are the ones who show empirically . . . higher scores on tests of reading ability." (p. 80) In the light of these findings, the following two sections of this review of the literature will look at the descriptive studies which confirm Loban's findings about the relationship between reading comprehension and written syntax, and the experimental studies using sentence-combining which attempted to improve reading comprehension while improving written syntactic ability.

The relationship between reading comprehension and syntax—descriptive studies. In her 1975 study with seventh graders, Mildred Kuntz made two comments which provide important insight into the relationship between a person's ability to comprehend language using certain syntactic forms and his ability to write these same forms:

> Reading achievement and written syntactic attainment are not mutually exclusive. Rather, it could be said that they thrive in a symbiotic relationship which is mutually beneficial. They share common elements, and it is suggested that one of these elements is a knowledge of language—how it works and how it is used. (p. 137)

Earlier in the same work she had noted that:

> An examination of the reader's understanding and management of sentence elements should provide valuable insights into the language that the reader brings to reading . . . . One must be careful not to confuse reading comprehension with a mastery of sentence structure; however, the ability to utilize syntactic understandings may be considered as an antecedent to comprehension. Since the sentence is regarded by linguists as the basic unit of language (Lefevre, 1964, p. VII), it would seem that an understanding of the structure and patterns of language as expressed through the sentence is such an advantage to the reader that this knowledge constitutes a preorganizing, as it were, of the materials to be read. (Kuntz, 1975, p. 3)

In her study, she used the Gates McGinitie Reading Test, Survey E to measure reading skills and the Sentence Construction Test to measure syntactic ability. This test provides the student with an example of the kind of construction he is to use and has him transform sentences into new sentences using these constructions. Based on a sample of ninety-six grade seven students she found correlations of 0.68 and 0.81 between total syntactic attainment and total reading achievement for two different schools. These correlations were significant at the .001 level. She concluded that "the knowledge of syntax that the students acquire through direct instruction or from transfer in other language areas is a valuable skill in efficient reading " (p. 134). In view of the fact that syntax is only one factor affecting reading comprehension, it is somewhat surprising to find correlations as high as 0.81.

Perron (1976), working with fifty-one fifth grade students, used samples of their free writing and calculated correlations between their words per T-unit and words per clause indices and their reading comprehension scores on the Gates-McGinitie Reading Tests. He found smaller correlations than Kuntz, 0.33 between reading

comprehension and words per T-unit, and 0.42 between reading comprehension and words per clause. Perhaps the two indices which he used were not sufficiently sensitive to true syntactic ability. Kuntz's Sentence Construction Test measured the ability of a student to produce a variety of syntactic forms, whereas the words per T-unit index used by Perron does not measure the same skill. This insensitivity may account for the wide difference in correlation coefficients.

Smith (1969) and Siedow (1973) controlled for content in writing by using O'Donnell's "Aluminum" passage, and Smith's "Bees" passage. Siedow also used her own "Ant" passage. Both authors used cloze versions of these passages, written on syntactic levels which were based on Hunt's (1970) findings, as comprehension tests. Smith's (1969) study was carried out with ninety students in each grade from four to twelve. In general, he found that students read best the material written at or near their own level of syntactic complexity. He concluded that the level of syntactic complexity of material to be read did affect comprehension and that the effect was particularly noticeable at the lower grade levels. Students in grades eight to twelve read eighth grade material better than anything harder. They also found materials written on the 'easier' syntactic level, grade four, harder to comprehend than grade eight level materials. Siedow (1973) found that the clause length variable in a subject's writing was significantly related to comprehension. However, she failed to find consistent significant positive correlations between syntactic ability and reading comprehension.

Evanechko, Ollila, and Armstrong (1974) took Smith and Siedow's

research a step further. They studied the free writing of grade six students and analyzed it according to the Botel and Granowsky Formula. They measured reading achievement with the Bond-Balow-Hoyt New Developmental Reading Test (Intermediate Level). The number of T-units and two-count structures were the best predictors on all pretests of reading achievement. Sentence patterns SVIO and SVO complement were the other two structures which contributed significantly to reading achievement. The researchers wrote:

> Of the four significant predictors of reading achievement, two were consistently first in the regression equation. These two were the number of T-units and Two-Count Structures. The former is essentially a measure of the fluency of expression in language in that it is an index of the number of ideas expressed and as such is a more accurate measure of language output than the sentence, which is more difficult of definition. The latter, Two-Count Structures (passives, paired constructions; dependent clauses; comparatives; participles; infinitives as subjects, appositives and conjunctive adverbs), is essentially a measure of the flexibility or complexity and sophistication of expressions in language in that it focuses upon means used to elaborate utterances and provides for variety in expression. Also, it appears that these same two language competencies, fluency and control of syntactic complexity, underly all measured reading behaviors. (p. 324)
>
> If fluency and control of syntactic complexity are the key language competencies underlying reading achievement, then building these two competencies may well improve reading performance. (p. 325)

Before concluding this section of the literature review it should be mentioned that not all such descriptive studies support the reading skill/writing skill relationship claims. Evans (1979) also attempted to investigate the relationship between reading comprehension and the production of written syntactic structures. (p. 129) His studies involved three groups of fifty students each in eighth and twelfth grade and senior level university. He had students do a rewrite

exercise similar to O'Donnell's "Aluminum" passage, using the "Bees" passage which was developed by Smith (1971). This instrument was used to provide three indices of syntactic ability: 1) T-unit length; 2) a subordination index; and 3) a nominalization index ("the mean number of noun clauses and phrases per sentence"). (Evans, 1979, p. 130) Reading comprehension was measured by cloze tests made from nine prose passages, three in each of the expository, narrative, and descriptive modes. Evans' results were unexpected, and contradicted much research carried out previously. He found "a significant inverse relationship between the three criteria for written syntax and the composite cloze scores" (p. 132). In short, as the cloze score means increased, the writing indices scores decreased. The cloze means showed an overall decline across grades. In fact, "eighth graders performed better on all the passages than college seniors did" (p. 133). Evans was hard pressed to explain these results. He simply concluded with the statement that "there is evidently a need for further investigation of the relationships addressed here" (p. 135).

Due to the large number of such studies devoted to studying the relationship between reading comprehension and written syntactic skill, reviewing them all is virtually impossible. Representative pieces have been included to provide insights into the procedures and results to which such studies often give rise.

Although, as Kuntz (1975) noted, reading and writing are not unitary skills, and therefore a mastery of syntax is not enough to ensure good writing and reading results (p. 4), the student needs a

facility with basic sentence patterns in all their variety. (p. 5) This syntactic skill is fundamental to successful writing and to good comprehension. A number of researchers have attempted to improve reading comprehension by improving the written syntax of students. Some have used sentence-combining programs to produce this improvement in writing and reading skills. These studies will be reviewed here.

Improving reading comprehension by improving written syntax--experimental studies. Sandra Stotsky (1975) in her comprehensive review of studies using sentence-combining techniques wrote:

> Theoretical justification for explicit syntactic manipulation in writing as a means of improving reading comprehension may derive from the possibility that such exercises clarify both the meaning and use of complex structures for children. Complex structural block-building exercises, so to speak, may help students better understand syntactic relationships within the sentence. They may also develop the capacity for synthesizing a larger number of elements within a 'total idea'. (p. 32)

> It is reasonable to assume that the child who understands how to compose a highly complex sentence also understands that sentence. (p. 59)

> Inasmuch as reading, speaking, listening, and writing are all language-based activities, one can assume an interrelationship among all the language arts. It is theoretically plausible to maintain that growth in one area should be reflected to some extent in other areas. (p. 66)

All of these points seem eminently reasonable. Researchers who attempted to improve reading through syntactic manipulation in writing all based their research on some version of these principles.

Fisher (1973) worked with students in grades five, seven, and nine. His experimental and control groups in each grade were relatively small with an average of sixteen per class. The sentence-

combining treatment was given while the control group took reading instruction for five weeks. Writing skill was measured by O'Donnell's "Aluminum" passage and a self-made "Cotton" passage. Reading comprehension was measured by the Stanford Paragraph Meaning Test and by three cloze tests written at grade four, eight, and twelve syntactic levels (Hunt, 1970) and on ninth and tenth grade vocabulary levels according to the Dale-Chall formula. Those tests were administered at pre-test and post-test occasions. Fisher found that

> the sentence-combining practice did enable the experimental groups to read the Stanford Paragraph Meaning Test and the fourth grade syntactic maturity cloze reading test better than the control groups (sig. at 0.014), but the practice did not enable the experimental groups to read the eighth and twelfth grade syntactic maturity cloze tests better than the control groups . . . . Therefore one cannot conclude that in its present form the sentence-combining course would enable students to read better after taking the course. (p. 85)

The correlation between words per T-unit on the post-tests and the Stanford Paragraph Meaning Test scores was 0.13. Such results, in light of the built-in bias due to the sentence-combining format of the writing test are probably not valid.

Shockley (1974) also investigated the effects of training in syntax on reading comprehension. She selected forty-six grade seven students of average intelligence who were experiencing comprehension problems and who scored one year or more below level. Her treatment was somewhat different from regular sentence-combining programs. Whereas Fisher's program was based on the manipulation of sentences individually, Shockley's sentence-combining program involved the re-writing of passages on a modified kernel level. Subjects rewrote the original surface structure representation of selected fables. To

test the effectiveness of her program, she used the Metropolitan Achievement Test for reading and the Shockley Syntactic Structures Test for measuring syntactic changes in writing. The latter test used a multiple-choice format employing answers involving various transformations. The student had to select the sentence with the same meaning as the key sentences. A pilot study using ten subjects yielded a reliability for the SSST of 0.80. Shockley found no significant difference between pre- and post-test scores on her SSST. (pp. 56-7) She found a significant difference (.05) between pre- and post-test comprehension scores on the MAT for both experimental and control groups. She suggested that the positive attitude toward reading resulting from work with the fables could have resulted in the success of both groups on the post-tests. She also hypothesized that her expectation of transfer from implicit usage of transformations in context (fables) to an explicit usage of transforms in isolation may have been ill-founded. (p. 57) Her explanations may well be correct. However, Shockley's study shed little light on whether reading comprehension could be improved by practice in sentence combining.

The two previous studies were only concerned with power of comprehension and not speed of comprehension. Hughes (1975) hypothesized that sentence-combining practice

> should speed up the visual process by increasing student's familiarity with the phrases and clauses resulting from the embedding process . . . . Therefore experimental students, practicing sentence-combining, should show an increase in reading speed, as compared to control students. (p. 17)

She studied twelve experimental and twelve control subjects in grade seven. The experimental subjects were given an average of thirty-one

hours of sentence-combining practice with adjective clauses and [illegible] reductions over an eleven week span. She used the Gates-McGinitie Reading Tests, Survey E to measure reading comprehension and speed and O'Donnell's "Aluminum" passage on [illegible] and post-test occasions to measure syntactic ability. Hughes concluded that

> there appears to be an extremely close link between a student's reading level and that same student's syntactic maturity level based on measures developed by Hunt and O'Donnell [i.e. the "Aluminum" passage]. (p. 53)

She also suggested that "although the Gates-McGinitie tests are widely used for measuring reading ability, the speed and accuracy section used to measure these qualities in this study seemed not to be particularly valid" (p. 58). More will be said on this point later. Hughes did not control for syntax or content in her reading tests. She administered a different form at each testing period, thus confounding her results. Hughes wrote in her "Implications" section, however, that

> the close correlations found in this study between the reading level of students and their syntactic maturity levels on the "Aluminum" instrument suggests the need for a study using a large population and examining all the measures of syntactic growth as identified by Hunt (1970). (Hughes, 1975, p. 60)

Combs' results were also published in 1975. He replicated O'Hare's (1973) sentence-combining treatment as closely as possible but limited it to ten weeks rather than eight months. The treatment involved twenty-five hours of sentence-combining work, nearly the same as O'Hare's. Like Hughes (1975), Combs used the Gates-McGinitie Reading Test to measure reading comprehension and speed. Unlike Hughes, he used pre- and post-test compositions to measure syntactic

ability instead of the "Aluminum" passage. Although he also measured syntactic ability with a delayed post-test, he did not measure reading comprehension skill on that occasion. In view of the dramatic decline on writing scores which he found, this oversight is regrettable. Combs did control for syntax and content in the reading test by assigning one-half the subjects to each form at pre-test and post-test. His results showed the significant change on the Gates-McGinitie Comprehension subtest to be negative. Combs suggested that this indicated the test was probably not sensitive to the effect the treatment had on students receiving the sentence-combining practice. (p. 80) He also found that the experimentals and controls gained significantly on the Gates-McGinitie Rate subtest indicating that the syntactic practice had no differential effect on the students' reading ability as measured by this test.

Smart and Ollila (1978) also carried out a similar study using the Gates-McGinitie Reading Test, Survey E and sentence-combining problems based on O'Hare's format. The results for their grade sevens showed non-significant differences between experimental and control group scores in both reading and writing. However, both groups showed significant pre-test to post-test reading gains. The fact that different forms of the reading test were used for each occasion suggests that these significant gains registered by both groups may have been due to some factor other than maturation. In light of the fact that reviews of this test have questioned the validity of the speed section, one wonders why it was ever used. For example, Jason Millman in Reading Tests and Reviews edited by Oscar K. Buros (1975) writes:

> The Speed and Accuracy sub-tests consist of thirty-six very short passages with the last word of the passage to be selected from among four options. The speed score is the number of passages completed; the accuracy score is the number of passages completed correctly. . .the directions for administration and the shortness of the test (four minutes) have undoubtedly lessened its reliability. The directions do not make clear what it takes to do well on the test, nor do they give the student a running start. Further, four minutes is just too short a time in which to assess reading speed and accuracy. (p. 64-5)

And Farr (1969) noted in his book, Reading: What Can Be Measured?

> Certainly the independent measurement of rate and comprehension may have diagnostic value, but the measurement of the speed at which a reader comprehends a reading selection also has value . . . . The teacher does not have to know that a reader can pass over words at 300, 800, or 1200 words per minute; what he needs to know is how long it takes the reader to comprehend the material for a given purpose. (p. 66)

And as Lennon (1962) commented,

> Our problems in the measurement of rate stem from the fact that we are never really concerned with pure speed, that is, with just the rapidity with which the subject can move over a given number of words or lines of written material. Rate is only meaningful as it defines the rapidity with which the reader covers the material at a particular level of comprehension. (p. 335)

Keeping these comments in mind, it is reasonable to suggest that the Gates-McGinitie Reading Test is not the most appropriate test to use in studying the effect of sentence combining on reading comprehension level and speed. These were the skills that the researchers (Hughes, 1975; Combs, 1975) expected the students to improve, but deficiencies in their reading test itself may have forced both researchers to conclude that reading rate was not differentially affected. (Combs, 1975, p. 86) And, as has been argued earlier, sentence-combining practice at the grade seven level, based on O'Hare's materials, is more likely to increase speed of comprehension than it is to increase level of

comprehension. It is difficult to believe that the majority of grade seven students, given a non-hurried reading situation, could not comprehend the sentences which the Mellon, O'Hare, and Combs' students practiced. It is reasonable to suggest that junior high school students already understand (level of comprehension) O'Hare's sentence-combining structures, and any increase in reading skill would be in the area of acquired fluency and facility in understanding sentences using these structures (speed of comprehension). Therefore, future studies in this area at the junior high level should use reading tests which are designed to measure speed of comprehension.

Straw (1979) supplied additional information on the possible effect of sentence-combining practice on reading comprehension. His excellent review of the literature provides some interesting insights and raises some important questions. In noting the attempts by researchers to measure only the syntactic component of reading comprehension while ignoring the "graphophonic and semantic cue systems" as outlined by Goodman (1973), Straw observed:

> Measures of reading comprehension used in studies of the effect of sentence combining should not be measures that assess only a student's syntactic competence in reading because, if that is all we measure, we cannot claim that sentence-combining instruction has an effect on reading comprehension as a whole, but only on the syntactic aspect of reading comprehension. By the same token, we should not employ measures that only measure the semantic aspect of reading comprehension with the syntactic element parcelled out, for the same reason. (p. 44)

He went on to criticize research in which the Gates-McGinitie Reading Test was used because the comprehension subtest "measures the semantic component of reading comprehension almost totally and does not attempt to measure the syntactic component" (p. 45). Straw questioned whether

the cloze tests used by Fisher (1973) and Hughes (1975) were too difficult either syntactically or semantically and therefore provided little insight into the effect of sentence-combining practice on reading comprehension. He concluded that "the 'best' test of reading comprehension has yet to be developed" (p. 48).

Warren Combs (1979) contributed to the same volume and also discussed the effect of sentence combining on reading comprehension scores. He concluded his review in a manner similar to Straw.

> As a group, the empirical explorations of the fit of SC practice and reading comprehension remain ambiguous at best. Standardized measures consistently uncover non-significant or negative differences between experimental and control groups. The results of cloze tests are varied. And the results of special-made measures are largely positive. These results do not allow substantive conclusions; quite rightly, the conclusions advanced by most researchers are tentative. (p. 55)

Combs (1979) went on to suggest a logical explanation for some of these results:

> Since reading competence precedes writing competence in syntactic complexity (Smith, 1974), one is aware that SC exercises employed in the above studies have by and large, given students practice with structures less mature than those needed to encourage gains in reading comprehension. This may mean that gains in reading comprehension may be limited to students with less syntactically mature reading skill. But whatever the case, the SC treatments and reading measures need to contain passages of greater syntactic complexity. (p. 55)

Researchers are also well advised to control as many as possible of the variables outlined by Rutland (1975), who observed that

> while many studies exist which have attempted to determine the relative difficulty of different syntactic structures, almost all of the studies suffer from one or more of the following problems: either the meaning of sentences was not held constant as the syntax changed, thus confounding syntactic difficulty with semantic difficulty; or the syntactic changes were studied in passages in which many

> difficult transformations existed, thus confounding the effect of one syntactic change alone; or the syntactic changes were studied in sentences in isolation thus reducing the generalizability of relative syntactic difficulty to longer prose passages. (p. IV)

While the present study was not concerned with determining the 'relative difficulty' of the different syntactic structures, nevertheless, it attempted to determine the effect of syntactic manipulation practice on reading comprehension and speed. Therefore, an attempt was made in the design of the study to control for the effect of syntactic and semantic variation. The study was limited to studying the effect of sentence-combining practice on reading comprehension level and speed as defined by Davis (1962) and as it relates to prose passages. It attempted to shed further light on the effects of sentence-combining practice on these two necessarily related reading skills.

Conclusions From the Review of Literature

Writing and syntactic development. In an article published in 1977, Kellogg Hunt summarized his findings as they related to syntactic development in writing:

> On the basis of my studies of free writing in the sixties, I made two broad claims. One was that as schoolchildren get older, the T-units they write tend to get longer, measuring length as the mean number of words per T-unit. This claim might be called the T-unit length hypothesis . . . . So many researchers have found that their evidence tended to confirm the claim that I suppose it is now accepted by all persons who know about such matters. (p. 93-94)

> A second broad claim which came from my study of free writing in the sixties is that as schoolchildren get older they tend to consolidate into their T-units a larger and larger number of what transformational grammarians call S-constituents.
>
> Roughly speaking an S-constituent is the abstract structure that underlies the simplest of sentences--what used to be called kernel sentences. . . .My claim was that as schoolchildren grow older, they consolidate a larger and larger number of such S-constituents into their actual T-units. We might call this the number of consolidations hypothesis. (p. 94)

Thus, all the structures listed in Tables 1 and 2 (pp. 18 and 19) contribute to T-unit length because any time a structure less than a T-unit is added, in effect the length of the T-unit to which you add them is increased. Loban (1976) confirmed Hunt's findings.

But research has not only provided the English teaching profession with normative data about these structures, it has also suggested a means for aiding the student in acquiring more mature syntax at an earlier age. Mellon (1967), O'Hare (1973), Combs (1975) and Pedersen (1977) have all shown significant experimental group increases on certain indices of syntactic maturity following sentence-combining practice (See Table 3, p. 39). What remains to be demonstrated is whether this increase is true syntactic growth as obtained by Hunt and Loban or some kind of artificial growth, and whether it is permanent or forced growth which quickly dissipates.

Reading comprehension and the influence of syntax. Numerous researchers have attempted to calculate the degree of relationship between syntactic complexity and reading comprehension difficulty. The present review of literature has demonstrated that syntax does indeed affect comprehension right up to university level. Although

many of the measuring instruments have been relatively c[illegible] demonstrate that syntax is a variable in reading compreh[illegible]. While attempts to improve reading comprehension by simply improving awareness of syntactic structures have been relative failures, O'Donnell's (1976b) results do suggest that improving student facility with various syntactic forms could improve comprehension.

Reading ability and writing ability: the relationship. Researchers have found correlations between reading ability and writing ability from 0.13 (Fisher, 1973) to 0.81 (Kuntz, 1975). The crudity of measuring instruments may be responsible for most of the variation in these correlations. However, based on the longitudinal study by Loban (1976) it is indeed reasonable to conclude that the relationship between reading and writing is "so striking as to be beyond question" (Loban, 1976, p. 82).

A number of attempts have also been made to improve reading comprehension while improving a subject's written syntax. Because of the high degree of interrelationship among the language [illegible] appears to be a logical undertaking. But again, due pa[illegible] problems encountered in attempting t[illegible]e growth in the two arts, the results of such efforts have be[illegible]clusive. More refined research is needed in this area.

CHAPTER III

DESIGN OF THE STUDY

Overview

The present study was designed to measure the effect of O'Hare's (1975) sentence-combining program, Sentencecraft, on the writing and reading comprehension skills of a group of grade nine students. The treatment lasted for nine weeks and was followed by an eight-week delay period during which no sentence combining was practiced, to test for any decline in attained increases, such as was found by Combs (1975). Writing ability was measured on six different occasions: pre-test, three occasions during the treatment, post-test, and delayed post-test. All six essays were written in the argumentative mode. Reading ability was measured by a standardized reading test on pre-test, post-test, and delayed post-test occasions.

The study used a true experimental design which was derived from Campbell and Stanley's (1963) "Pretest-Post-test Control Group Design". (p. 13) The present study additionally employed a delayed post-test and more frequent repeated measures on the writing variables. Two sets of variables were measured in the present study: reading comprehension and written syntax. The design for the writing variables was as follows:

	Pre-test			Post-Test		Delayed Post-Test
Experimental Group	O_1	O_2	O_3	O_4	O_5	O_6

	Treatment Period					Delay Period
Control Group	O_7	O_8	O_9	O_{10}	O_{11}	O_{12}

Reading variables were measured on occasions O_1, O_5, and O_6 for experimental groups and O_7, O_{11} and O_{12} for the control groups. In the diagram above, O_2, O_3, O_4 and O_8, O_9, and O_{10} were the testing occasions for writing, which occurred at intervals during the treatment.

Sample

The experiment was carried out from mid-February to mid-June, 1978. Six grade nine classes in Spruce Grove Composite High School, Spruce Grove, Alberta were used for this study. (See Appendix A) This school is the only high school in a town of 7,000 people, ten miles from the city of Edmonton, Alberta, Canada. According to the administration, the school population (grades 9-12) is drawn from a cross-section of the socioeconomic spectrum ranging from professionals to laborers. Many of the people commute to jobs in Edmonton. About seventy percent of the school population lives in the town and thirty percent are from the surrounding rural area. The school population was approximately 740. Five criteria were applied in determining the acceptability of the sample, and all were met.

1. The school population must not be atypical socio-economically.
2. The school must be on a full-year schedule rather than on a

semester system in order that the total grade nine population be involved.

3. Each teacher must teach one experimental and one control class.
4. The experimental and control classes must be deemed roughly comparable in ability by the classroom teachers, the school administration, and the researcher.
5. There must be no special, concurrent programs in reading or sentence-structure practice conducted in these classes for the duration of the experiment.

Two of the grade nine classes in the school were not included in the study because they were taught by different teachers. These two classes were used in the pilot study. (See Appendix B.) Students from the six remaining grade nine classes which were acceptable for inclusion in the experiment were streamed for Language Arts on the basis of student scores on a school-wide, staff-constructed, mid-term English exam. The students were divided into two classes each of high, middle, and low ability. No systematic bias in assignment to classes within the three levels was introduced. One teacher taught both classes at each of these levels. Classes at each level were randomly assigned to be in the experimental or control group. Eleven students were lost due to experimental mortality. Six students were from the low-ability group. Various circumstances led to the losses ranging from one student taking a trip to Europe to five students dropping out of school to take up jobs.

The Treatment

The treatment variable in the proposed experiment was the sentence-combining program constructed by Frank O'Hare, entitled Sentencecraft. This program, based on the exercises used in O'Hare's original study, uses a workbook format which gives the student practice in combining kernel sentences into more complex and mature syntactic forms. More specifically the experimental students received practice in manipulating a number of syntactic constituents as indicated by the following chapter descriptions and examples:

1. The use of free modifiers (Fred dashed into the room, lunged at Knuckles, and missed, falling in a heap in the corner.)
2. The use of adverb phrases (when, where, how) (When playing hockey, he doesn't hear the fans.)
3. Noun clauses (the fact that, that) (The fact that I haven't called you doesn't mean I've been busy.)
4. It-Inversion (it . . . that) (It is true that the world is nearly round).
5. Subordinate clauses (who, what, where, when, why, how) (I never understood what made him so popular.)
6. It-Inversion (it . . . who, it . . . what, it . . . where, it . . . when, it . . . how, it . . . why) (It isn't clear how he does that trick.)
7. WH & Infinitive as object (when to, what to, who to, whom to, how to) (She didn't know when to change gears.)
8. Ing-Nominalization ('s, of) (The audience was delighted by Joan's singing of the anthem.)

9. Genitive('s) (We are here to protest the government's failure to stop inflation.)
10. Infinitive nominal (for . . . to) (It was easy for Marie to learn French.)
11. Relativization (which, that, who, whom) (There is the car that we saw yesterday.)
12. Relativization (whose, when, where, why) (That is the man whose car we stole.)
13. WH -- Be Deletion (He is the old man.)
14. More practice with WH — Be Deletion
15. Clauses of Time, Cause, Concession (e.g. after, because, if) (He went to bed because he was tired.)
16. Nominal Expansion: Gerundive adjective phrases (ing), and prepositional phrases (with) (Bursting through, he saw the men fighting.)
17. Colon and Dash practice. (I was ten years old when another boy came — Jack, who had been an orphan all his life.)
18. Additional Practice in all previous transformations.
19. Practice in "open" sentence-combining problems (without the help of signals.)

The nine-week treatment format, as outlined in Sentencecraft (pp. 9-10), was followed in this study. It was introduced as a program designed to help students write better.

To demonstrate that the Sentencecraft program could be an effective method of improving reading comprehension, Table 4 lists the structures identified in the "Review of Literature" as being most

Table 4

Syntactic Structures Which Increase Sentence Difficulty

Author	Grade	Nominalized Verbals	Adjective Clauses	Embedded Clauses	Appositives	Genitives	Negatives	Deletions	Passives	Nominalizations: "For. . . to" "The fact that"	Prepositional Phrase Modifiers
Albanese	3,5		+								+
Rutland	4	+		+				+			
Bormuth et al.	4						+			+ +	
Fagan	4,5,6	+		+	+	+	+	+		+ +	
Slobin	K,2,4,6 Adult						+		+		
Takahashi	9		+	+				+			+
Gough	Univ.						+		+		
Savin and Perchonock	Univ.								+		
Sentence-craft Program		+	+	+	+	+		+	+	+ +	+

difficult to comprehend. At the bottom of the table is a list showing those structures which were practiced in the treatment. It was hypothesized that the treatment should result in an improvement in reading comprehension, since it provided students with practice in ten of the structures identified as difficult by researchers.

One purpose of the pilot study was to determine the appropriateness of the instructional materials for grade nine students. The two teachers involved reported that the workbook exercises were neither too easy nor too difficult. They indicated, as did O'Hare (1973), Fisher (1973), and Combs (1975), that students enjoyed doing the exercises. Some student comments were:

> The books were a lot of fun and I wouldn't mind doing some more like them.
>
> Some of the sentences were long and made me think about them. They were hard to understand at times but I managed to make it.
>
> I think that the program helped people to learn how to join sentences and not to repeat themselves.
>
> The sentences, in most cases, were very enjoyable and had great variety. I felt the program could have been extended in the length of time we had to complete the book.

Writing assignments within the English classes during the seventeen weeks of the experiment were limited to approximately one loose-leaf page per week. This limit included the writing on weeks in which writing tests were given. It was not possible to control the amount of writing practice done outside of English classes; however, it was not a major confounding variable since the classes were not intact across other subject areas and thus random effect acted as a control.

The control group undertook the same amount of writing as the

experimental group. During the time devoted to the experimental treatment, the control groups studied other writing-related areas such as spelling, punctuation, usage, and parsing. They did not undertake any kind of sentence-combining activity.

The in-school reading activities for the students were also limited in that neither the experimental nor the control groups undertook any kind of formal reading instruction during the time of the experiment.

Instrumentation

Writing. The instrument for measuring syntactic ability consisted of a single theme written on each testing occasion. The validity of this method of obtaining measures of group achievement has been supported by Diederich (1946) and Kincaid (1953). Kincaid (1953) studied factors which caused variation in the quality of student writing and concluded:

> An evaluation of the overall, or average group improvement resulting from a writing course may be obtained from a single pre-test theme and a single post-test theme. (p. 95)

Braddock (1963) in his review of research in composition also suggested methods for researching writing. He pointed out the necessity for controlling the variables of topic, mode, and time. (p. 17) He indicated that this control was especially imperative if one was conducting frequency counts.

As early as 1946, Diederich indicated that the topic must be within the student's comprehension because

> even the better students write badly when the topic is beyond them. Their struggles are apparent not only in lack of organization and in vagueness of statement, but in the very structure of their sentences. (p. 585)

Hunt (1977, p. 95) demonstrated that topic did in fact influence the kind of syntactic structures used by the subjects. Thus, the topics used in the present study were taken from a list suggested by the students and approved by all the teachers involved. These topics were used in the pilot study and students reacted positively to all topics, according to teacher report. Pilot results also showed no significant difference between topics, as measured by total number of words and words per T-unit. However, to further control for the possible influence of topic on the type of syntactic structures used, one-sixth of each group wrote on each topic on each test occasion. This was achieved by random assignment to the six different topic sequences as indicated in Appendix C. The topics were:

1. Education should be compulsory.
2. I wouldn't want to live in any other country.
3. The world won't survive past the year 2000.
4. Marijuana should be legalized.
5. Advertising is bad for people.
6. Everybody needs friends.

Topics were chosen to provide for a wide range in content, personal involvement, and thought-provoking capability.

All the topics on which the students wrote were in the argumentative mode. Diederich (1946) commented: "Test assignments must represent the kind of writing that students may be expected to use later in life . . . .A persuasive argument probably represents one

of the main types of writing which most students will have to do." (p. 585) However, this was not the primary reason for choosing argumentative topics for this experiment. As was indicated earlier, Perron (1976) and Crowhurst (1977) amply demonstrated that even by grade five and six the argumentative mode provided students with the best opportunity to display their syntactic abilities. In both studies the argumentative mode resulted in significantly more complex syntax than did the other modes as measured by words per T-unit, words per clause, and clauses per T-unit. This finding was borne out by the pilot study of the present experiment. The two grade-nine classes, on the pre-test, averaged 13.32 words per T-unit, 7.93 words per clause, and 1.72 clauses per T-unit. Table 1, (p.18) shows that these frequencies are close to grade 12 achievement as indicated by Loban's (1976) and Hunt's (1965) results. These results confirmed the argumentative mode as the mode for providing students with the best opportunity to display their syntactic skills. Therefore, this mode was used in this experiment.

The third variable which had to be controlled was the time allotted for completing the writing assignments. Subjects in this experiment were given fifty-five minutes in class to write their compositions. Pilot study results indicated that this was sufficient. Braddock (1963) in his "Suggested Methods for Research" indicated that junior high school students should be afforded at least fifty minutes for writing a composition. (p. 8)

Reading. The main variables which had to be controlled to measure gains in reading comprehension level and speed were semantics

and syntax. These could not be controlled if one simply used different forms on the different testing occasions. Farr (1969) pointed out the difficulties involved in measuring reading improvement.

> Even if statistical equivalency of test forms could be established, there would still be unanswered questions about the content equivalency of any two forms. It would be impossible for a test developer to control all the variables on a reading test from one form to another. The difficulty of the vocabulary, the content of the material, and the sentence length and complexity are all variables which most test authors attempt to control, but for each factor that is controlled, there are several others which are uncontrolled. (p. 141)

Therefore, it appeared that the only way to control these variables was to have one-third of the subjects write each form on each testing occasion. To do this, it was necessary to find a test with three roughly equivalent forms. The test also had to have reliable comprehension level and speed measures built in. A standardized test was used because any meaningful improvement in reading comprehension level or speed should have been evident in standardized test results. "Reading comprehension speed" rather than "reading speed" was measured because as Farr (1969, p. 66) noted, the speed with which one reads words is not as important as the speed with which one can comprehend the material that he is reading.

The Davis Reading Test, Series 2 was selected for use in the present study based on these criteria:

1. Three roughly equivalent forms were needed.
2. The test had to be suitable for grade nine.
3. The test had to contain reliable measures of reading comprehension power and reading comprehension speed.

4. The test had to receive good reviews.

The Davis Reading Test has four forms for Series 2 of which the three most statistically equivalent are forms A, B, and C. The means and standard deviations in Raw Scores for the three forms for grade nine (N=288) are:

Form		Level	Speed
A	Mean	22.7	36.5
	S.D.	9.7	18.6
B	Mean	20.1	33.6
	S.D.	9.9	18.7
C	Mean	19.4	34.4
	S.D	9.8	19.5

The average standard error of measurement in Scaled Scores across all forms is 3.3 for Level of Comprehension and 2.3 for Speed. These findings were based on a sample of 1152 grade nine students. The mean expressed in scaled scores is 67.7 for both Level and Speed. Secondly, Coffman (1968) in reviewing this test in Reading Tests and Reviews stated: "It is doubtful that one can find a better reading test for use in grades 8-13." (pp. 291-2) Thirdly, the average reliability coefficients for 1152 grade nine students in fifty-two schools and twenty-eight communities were 0.84 for reading comprehension level and 0.91 for reading comprehension speed for Series 2 forms. (Davis, 1962, p. 15) Fourthly, Coffman (1968) in reviewing the test wrote:

> There is a freshness about the stimulus passages and a challenge to all questions which makes the task of marking answers an encounter with an interesting adversary. On a subjective basis alone, one could conclude that these are unusually effective collections of passages and questions for assessing the ability of adolescents to garner meaning from the printed page.
>
> Each form of the test consists of two parallel halves of forty questions each. Almost all students complete the first half and almost nobody completes the second half. Therefore, the score on the first half is taken as a measure of accuracy or depth of comprehension while the score on the total test measures both speed and accuracy. (pp. 291-2).

Based on the fact that the test met the criteria set, the Davis Reading Test, Forms 2A, 2B, and 2C, were used to measure reading improvement.

Intelligence Quotient. The Canadian Lorge-Thorndike Intelligence Test Level F, Verbal Battery was administered to all students involved in the study in early February, 1978. This multilevel test is recommended for students through grade nine. A major portion of the Canadian norming was carried out in the Edmonton Separate School System. Using a group of 278 grade eight students in 1968-69, a reliability of 0.85 was obtained. The correlation between this Verbal Battery and the Otis Mental Ability Test was 0.82; with the Stanford-Binet, it was 0.78. (Technical Supplement, 1972, pp. 9-22) The subtests measure word knowledge, sentence completion, verbal classification, verbal analogies, and arithmetic reasoning. Because the main interest of the study was language-based, the Non-Verbal Battery was omitted. The Verbal Battery was used to measure possible interaction of treatment and IQ.

Data Collection Procedures

Writing. Students were given some introductory notes on argumentation along with a list of writing sequences and the selected topics as shown in Appendix C. One-sixth of the students in each class were randomly assigned to each topic sequence. On the first administration, students were given twenty-five minutes to read over the instructions with the classroom teacher and raise any questions. The six writing topics were introduced as a unit on argumentative writing and were not linked in any way to the sentence-combining treatment in order that the results would be unbiased. Teachers agreed to mark these compositions as usual, that is, for such things as content, organization, spelling, punctuation, and grammar. All essays were photocopied by the researcher and returned to the teachers within one day. The teachers agreed to be punctual in their marking in order that momentum could be maintained.

The compositions were written at two-week intervals except for occasion two which was three weeks after the pre-test. (See Appendix D) They were written in the first two periods on Monday or Tuesday mornings. Students had fifty-five minutes to complete each composition. In order to diminish the problem of "what to say", students were given five minutes at the beginning of the period to discuss their topic with the other students who were writing on that topic. They were then given fifty minutes to write. They could rewrite if they so desired, and they were encouraged to reread their compositions to eliminate as many weaknesses as possible. To further alleviate

the problem of "what to say" students were also permitted to argue for both sides of the topic as directed in the introductory notes to the unit. The compositions were collected at the end of fifty minutes. They were marked by the teacher and returned to the students for their perusal and comment. They were then re-collected and held until after the delayed post-test in order to eliminate opportunities for students to copy the work of others.

Reading. As indicated in the "Overview" on page 75, the Davis Reading Test was administered concurrently with writing occasions one, five, and six. Students wrote these tests in one of three form sequences: ABC, BCA, or CAB. One-third of the students in each class were randomly assigned to each form on each occasion. The "Directions for Administration" were strictly adhered to as outlined in the test manual. The test takes forty minutes of actual writing time and another five for administrative detail. The post-test and delayed post-test administrations were presented to students as an attempt to get a better overall picture of their reading ability as well as an attempt to determine how much their reading had improved.

Scoring Procedures

Writing. The compositions written by the experimental and control groups were "scored" by counting these frequencies: total number of words, words per T-unit, words per clause, and clauses per T-unit. In addition, the remaining structures listed in Tables 1 and 2 (pp.18-19) were counted for the pre-test, post-test, and delayed post-test

compositions of a number of randomly selected subjects in the experimental and control groups. Loban's (1967) method which involved using a random group (N=30) in lieu of the total group was followed here. The reason was the same as that given by Loban, "the time-consuming nature of the analysis". (Loban, 1967, p. 140)

The T-unit segmentation rules followed were those used by O'Hare (1973, p. 48) and Combs (1975, pp. 54-6). The T-unit as explained in the "Definition of Terms" was defined as each independent clause with its modifiers. The count for words per T-unit was determined by dividing the total number of words in the writing sample by the total number of T-units. Words per clause was determined by dividing the total number of words in the writing sample by the total number of clauses. Dependent clauses per T-unit was determined by dividing the total number of dependent clauses by the total number of T-units. Other counts followed traditional parsing procedures. However, there were a few additional decisions to be made and again the modus operandi of Combs and O'Hare was followed.

Words -- dates were counted as one word as were phrasal proper nouns.

-- compound words were considered as one word, unless hyphenated, in which case they were counted as two; contractions were also counted as two words.

-- omitted subjects were counted in word totals

-- interjections such as "Hi!" were not included in the total word counts.

Discourse -- indirect discourse was analyzed according to traditional sentence parsing rules.

-- direct discourse, because there was not a significant amount, was omitted from the totals as suggested by Hunt (1965).

Unintentional Omissions -- obviously unintentional omissions were supplied by the researcher and counted in the word total since some students did not have time to proofread their papers. (See O'Hare, 1973, pp. 48-9)

Garbles -- Any group of words that could not be understood by the investigator was omitted from all structure counts.

Conjunctions -- "Except" was treated as a coordinating conjunction when it occurred between two clauses.

-- "So" and "then" were treated as coordinating or subordinating conjunctions depending on context.

A sample of analyzed composition is given in Appendix E.

Reading. The reading test results were optically scored, and both reading comprehension level scores and reading comprehension speed scores were calculated. These raw scores were converted to scaled scores to standardize them and alleviate the problem of slight differences in the mean scores among forms further refining the control achieved by random assignment to test form.

Conclusion

The design for the present study controlled as many of the variables as possible. Previous studies carried out in the area have played an important part in eliminating design problems in this experiment.

In short, the design of the study allowed for maximum possible control of the variables affecting both the reading and writing scores chosen for analysis.

CHAPTER IV

ANALYSIS OF DATA

The present experiment was designed to determine whether the Sentencecraft program of Frank O'Hare would induce significant changes on various indices used to measure writing complexity in the argumentative compositions of grade nine students. The effect of the program on level and speed of comprehension as measured by the Davis Reading Test was also studied. Major writing indices were calculated for both experimental (n=75) and control group (n=68) compositions on pre-test, three intervening occasions, post-test and delayed post-test occasions. These indices were: the number of words per essay; the number of T-units; the number of clauses; the number of words per T-unit; the number of words per clause; and the number of clauses per T-unit. Secondary counts were made on a number of specific sentence structures appearing in the writing of random subsamples of thirty experimental and thirty control group students on the three main testing occasions. Reading level and speed of comprehension measures were taken on the three main testing occasions for the complete experimental and control groups as well.

The analysis of results is therefore divided into two major areas: writing and reading. The first part of the writing analysis was carried out to determine: 1) whether the experimental and control groups were significantly different on the pre-test occasion according to the instruments used in the data-gathering; 2) whether there

were any significant differences between essay topics according to the major indices. The former analysis was especially important because intact classes were used. The latter analysis was carried out as a check on the effectiveness of the topic-randomization procedures. The six major writing indices were used in both these procedures. IQ scores were only used in comparing the two groups on the pre-test. One-way analysis of variance was the statistical program to which the data were submitted. This analysis was followed by a two-way analysis of variance with repeated measures using the six indices across the six writing occasions. The procedure was used to determine whether any significant differences between the experimental and control groups were detectable as a result of the Sentencecraft treatment when these six indices were used as the dependent variables.

The writing analysis next consisted of a comparison of the experimental and control subsamples. The first step was an attempt to determine whether the subsamples were significantly different from the remainder of their samples on the pre-test occasion. The Chi-square test was used. The six major writing indices were again used in this procedure. The same analysis was done to determine whether the experimental and control group subsamples were significantly different from each other on the pre-test occasion when the same six indices were tested. Following this preliminary subsample analysis, the fifteen new indices were used in determining whether the two subsamples were significantly different over the course of the treatment.

The reading results were then analyzed. The experimental and control group scores were subjected to a one-way analysis of variance

to determine whether there were any significant differences between the scores of the experimental and control groups on the pre-test occasion. The two reading indices used were Level of Comprehension and Speed of Comprehension as measured by the Davis Reading Test. These reading indices were used to determine whether any significant differences resulted from the treatment when the scores on the pre-test, post-test, and delayed post-test were compared.

IQ: Experimental vs Control

The Canadian Lorge-Thorndike Intelligence Test, Level F, (Verbal Subtest) was administered one week before the experiment began. Descriptive statistics for group comparisons are given in Table 5. The results of the one-way analysis of variance are reported in Table 6.

Table 5

Experimental and Control Group Comparison: IQ

Group	Cases,	Mean	Variance	Standard Deviation
Experimental	75	100.88	137.06	11.70
Control	68	101.07	167.86	12.95

Table 6

One-way Analysis of Variance: IQ

Source	SS	MS	df	F	p
Groups	1.00	1.00	1	0.01	0.935
Error	21389.00	151.70	141		

A study of these two tables reveals no significant differences between the two groups on the basis of these IQ scores.

Writing Analysis: Experimental versus Control Samples

Hypothesis: There will be no significant difference between the mean scores of classes of ninth grade students who have and who have not received sentence-combining treatment as measured by argumentative compositions written across the six occasions using the following indices: Number of words, number of T-units, number of dependent clauses, words per T-unit, words per clause, and dependent clauses per T-unit.

Topic comparison. Prior to the experimental-control group comparison, a check was made to determine whether the topics were roughly equivalent. Despite the random assignment of students to topic on pre-test, this additional analysis was carried out to determine whether the teachers' intuitions about their equivalency were correct. The one-way analysis of variance results for the six indices are shown in Table 7. A significant difference among topic means was found only on the words per clause index. A Scheffé-Multiple Comparison of means shows that the difference between topics was only

significant between topics three and six. Topic three was "The World Won't Survive Past the Year 2000"; topic six was "Everybody Needs Friends". It is difficult to pinpoint the most probable cause for this difference. In view of the random assignment to topic sequence which was followed in the study, this finding is more of interest than consequence. Nevertheless, the null hypothesis is rejected, for that index.

Table 7

Comparison of Topics: Pre-Test

Topics	1 (n=23)	2 (n=22)	3 (n=28)	4 (n=24)	5 (n=23)	6 (n=23)	p
Words	221.87	228.63	217.04	203.50	208.83	198.48	0.466
T-units	16.22	18.27	15.43	15.33	15.22	15.65	0.405
Clauses	27.83	27.82	25.86	27.00	26.65	27.87	0.928
Dep.Cl.	11.61	9.55	10.43	11.67	11.44	12.22	0.296
W/TU	14.11	13.25	14.49	13.92	13.95	13.09	0.456
W/Cl	8.11	8.30	8.50	7.79	8.01	7.23	0.020 *
Dep.Cl/TU	0.75	0.60	0.71	0.82	0.77	0.82	0.200

*Significant difference beyond the .05 level

Words per essay. A two-way analysis of variance with repeated measures was carried out to determine whether there were any significant differences between the experimental and control groups on the

pre-test and across the six writing occasions as measured by the number of words written on the pre-test. The cell means are listed in Table 8 and the results of the analysis in Table 9.

Table 8

Mean Number of Words, Repeated Measures

Occasions	1	2	3	4	5	6
Experimental	221.7	234.5	232.6	227.3	227.5	242.0
Control	203.4	192.0	185.5	193.9	211.2	189.1

Table 9

Treatment Effect on Number of Words Written

Source	SS	df	MS	F	p
Between subjects		142			
"A" main (between treatments)	263894.313	1	263894.313	13.155	.0004*
Subjects within groups		141	20059.914		
Within subjects		715			
"B" main (across occasions)	9607.098	5	1921.419	0.778	.568
"AxB" interaction	41170.098	5	8234.020	3.333	.006 *
'B' x subject within groups		705	2470.718		

*Significant beyond the .05 level

The interaction between treatment and occasion is significant beyond the ,05 level. Group effect was also significant but not consistently so, across the six occasions. When interaction is significant, Winer (1971, p. 529) states that tests on simple main effects is the preferred procedure. Tests on simple group effects showed no significant differences between pre-test and post-test means. There were significant differences on occasions 2, 3, 4 and on the delayed post-test. However, a glance at Table 8 shows that the decline in control group scores when compared to their pre-test scores is responsible for the significant differences in group means on these occasions. A test on the simple main effects of occasion bears this out. The increase in the number of words written by the experimental group in the course of the treatment is not significant, but the decline in the control group means for this index is significant beyond the 0.01 level.

In summary, although the experimental group does increase the number of words per essay that they write, the increase is not significant at the ,05 level. A further test using the analysis of covariance procedure produced almost identical results (See Appendix F). Therefore, the null hypothesis is not rejected.

Number of T-units. The mean number of T-units written by each group on the six occasions is displayed in Table 10. Table 11 displays the results of the two-way analysis of variance which indicated that group effect was significant. A Newman-Keuls comparison of means showed a significant difference between groups on the pre-test which carried across all occasions. The group/treatment main effect for both groups was not significant. In fact, all experimental means are significantly

higher than the control group means. In summary, the experimental treatment did not result in a significant increase in the mean number of T-units written by the experimental group. Therefore, the null hypothesis was not rejected.

Table 10

Mean Number of T-units Across Occasions

Occasion	1	2	3	4	5	6
Experimental	16.97	16.80	17.01	16.49	16.79	17.75
Control	14.88	13.74	14.03	14.03	15.01	14.16

Table 11

Treatment Effect on Number of T-units

Source	SS	df	MS	F	p
Between subjects	17722.125	142			
'A' main	1514.100	1	1514.100	13.172	.0004*
Subjects within groups	16208.063	141	114.951		
Within subjects	12612.688	715			
'B' main	77.598	5	15.520	0.878	.496
'AxB' interaction	81.394	5	16.279	0.921	.468
'B' x subject within groups	12455.000	705	17.667		

*Significant beyond the .05 level

Number of dependent clauses. The mean number of dependent clauses written by each group on each occasion is listed in Table 12.

Table 12

Mean Number of Dependent Clauses

Occasion	1	2	3	4	5	6
Experimental	11.16	11.44	11.75	11.88	11.51	12.25
Control	11.12	9.68	9.44	9.87	10.00	9.10

The analysis of variance results shown in Table 13 reveal a treatment effect and an interaction effect.

Table 13

Treatment Effect on Number of Dependent Clauses

Source	SS	df	MS	F	p
Between subjects	8477.563	142			
'A' main	690.823	1	690.823	12.509	.0006*
Subjects within groups	7786.750	141	55.225		
Within subjects	11514.500	715			
'B' main	32.965	5	6.593	0.411	.842
'AxB' interaction	189.189	5	37.838	2.361	.038 *
'B" x subject within groups	11297.813	705	16.025		

*Significant beyond the .05 level

Because of the significant interaction, tests were done on simple main effects. Group/treatment means were significantly different on occasions 2, 3, 4, and 6. Again, there is a noticeable decline in the mean number of dependent clauses written by the control group, whereas the experimental group means increase slightly for all occasions except occasion 5.

Despite this increase, the mean number of dependent clauses written by each group did not change significantly across the six occasions. The null hypothesis was not rejected.

Words per T-unit. Tables 14 and 15 display the mean number of words per T-unit written by the groups across the six occasions and the analysis of variance results.

Table 14

Mean Number of Words Per T-unit

Occasion	1	2	3	4	5	6
Experimental	13.64	14.38	13.96	14.38	13.89	14.03
Control	14.03	14.47	13.76	14.43	14.54	14.13

The analysis of variance results reveal no significant differences between treatment or occasion means.

Table 15

Treatment Effect on Words per T-unit

Source	SS	df	MS	F	p
Between subjects	2777.500	142			
'A' main	6.983	1	6.983	.355	.552
Subjects within groups	2770.563	141	19.649		
Within subjects	4825.313	715			
'B' main	47.349	5	9.470	1.402	.220
'AxB' interaction	16.178	5	3.236	.479	.794
'B' x subject within groups	4761.875	705	6.754		

In summary, no significant mean differences were found between groups or across occasions. The null hypothesis was not rejected.

Words per clause. The mean number of words per clause is displayed in Table 16.

Table 16

Mean Number of Words Per Clause

Occasion	1	2	3	4	5	6
Experimental	7.968	8.462	8.186	8.122	8.177	8.124
Control	8.038	8.252	8.010	8.172	8.520	8.248

The analysis of variance results revealed no significant differences as a result of the treatments administered to the two groups. Table 17 outlines the results of this analysis. As these tables show, there were no significant differences found in the mean number of words per clause written by the experimental and control groups across the six writing occasions of the experiment. The null hypothesis was not rejected.

Table 17

Treatment Effect on Words Per Clause

Source	SS	df	MS	F	p
Between subjects	587.177	142			
'A' main	.261	1	0.261	.063	.803
Subjects within groups	586.844	141	4.162		
Within subjects	1068.363	715			
'B' main	14.036	5	2.807	1.890	.093
'AxB' interaction	7.453	5	1.491	1.004	.415
'B' x subject within groups	1046.953	705	1.485		

Dependent clauses per T-unit. The mean number of dependent clauses written by the experimental and control groups across the six occasions is given in Table 18.

Table 18

Mean Number of Dependent Clauses Per T-unit

Occasion	1	2	3	4	5	6
Experimental	.724	.717	.720	.784	.722	.748
Control	.773	.794	.741	.781	.715	.731

The mean number of clauses per T-unit did not increase significantly in the writing of either group, nor were any differences significant on any of the six occasions as is shown by Table 19.

Table 19

Treatment Effect on Dependent Clauses Per T-unit

Source	SS	df	MS	F	p
Between subjects	35.919	142			
'A' main effects	.084	1	.084	.330	.566
Subjects within groups	35.834	141	.254		
Within subjects	66.239	715			
'B' main effects	.354	5	.071	.760	.581
'AxB' interaction	.243	5	.049	.523	.762

In summary, the treatment did not produce any significant differences between the mean numbers of dependent clauses written by the experimental and control groups or across the six writing occasions. The

null hypothesis was not rejected.

Writing Analysis: Experimental vs Control Subsamples

Hypothesis: There will be no significant difference between the mean scores of samples of ninth grade students who have and who have not received sentence-combining treatment as measured by argumentative compositions, written on pre-test, post-test, and delayed post-test occasions, using the following indices:

a) the mean number of noun, adjective, and adverb clauses per T-unit;
b) the percentage of short (8 words or less), medium (9-20 words) and long (21 or more words) T-units;
c) the mean number of nominalized verbals (gerunds and factive infinitives) per clause;
d) the mean number of modal auxiliaries, perfect forms, and passives per clause;
e) the mean number of prepositional phrases as noun modifiers per clause;
f) the mean number of verb forms (infinitives, present participles, and past participles) as noun modifiers per clause;
g) the mean number of genitives (inflected and phrasal) per clause;
h) the mean number of adjectives per clause;

The six major indices, as shown in the previous discussion, revealed no significant differences between groups as a result of the experimental treatment. Therefore, an examination of the structures within the clauses was undertaken. Twenty-one indices were analyzed, some of which are combinations of the variables that were counted.

Prior to the main analysis a check was made to ensure that the subsamples (n=30) randomly chosen to represent each group were not significantly different from the samples from which they were chosen. Tests for goodness-of-fit were used to compare each subsample's count with those of the remaining members of the samples for each group. The probability that the two distributions are of a similar shape are given in Table 20. The six major writing indices were used in these tests of Chi-square.

Table 20

Chi-Square: Subsamples versus Remainder of Samples

Index	Group		n	$\bar{X}$	SD	χ^2	df	p
Words	Experimental:	Subsample	30	244.33	62.25	1.046	3	.790
		Remainder	45	220.02	54.72			
	Control:	Subsample	30	210.03	66.31	1.478	3	.687
		Remainder	38	198.16	44.73			
T-units	Experimental:	Subsample	30	17.87	6.86	3.955	3	.266
		Remainder	45	16.38	5.18			
	Control:	Subsample	30	15.10	5.63	2.280	3	.516
		Remainder	38	14.71	3.78			
Dependent clauses	Experimental:	Subsample	30	10.73	3.90	1.621	3	.655
		Remainder	45	11.44	4.22			
	Control:	Subsample	30	11.10	5.54	10.941	3	.012*
		Remainder	38	11.13	3.44			

Table 20(continued)

Index	Group		n	$\bar{X}$	SD	χ^2	df	p
Words per T-unit	Experimental:	Subsample	30	13.21	3.06	4.699	3	.195
		Remainder	45	13.93	2.69			
	Control:	Subsample	30	14.40	2.86	2.214	3	.529
		Remainder	38	13.75	2.27			
Words per clause	Experimental:	Subsample	30	7.93	1.13	1.396	3	.707
		Remainder	45	8.00	1.23			
	Control:	Subsample	30	8.41	1.79	8.197	3	.042*
		Remainder	38	7.74	1.19			
Dependent clauses per T-unit	Experimental:	Subsample	30	0.67	0.34	2.091	3	.550
		Remainder	45	0.76	0.33			
	Control:	Subsample	30	0.75	0.35	6.950	3	.070
		Remainder	38	0.79	0.28			

* Significant beyond the .05 level

The Chi-square tests showed the number of words per clause on the pre-test and the number of dependent clauses per essay to be distributed significantly differently for the control group subsample and the remainder of the control group. Because the majority of the syntactic indices analyzed in this section are tested as "per clause" counts, the two indices which are significant should not cause much concern. However, the subsamples' comparisons on these two indices will be omitted from the following discussion.

In the next step of the analysis, an analysis of variance test was performed comparing the two subsamples using the other four major indices. In short, no significant differences were found between the means of the experimental subsample and the control subsample across the three writing occasions. The closest any index came to being significant was $p = 0.19$. The results of these analyses are given in Appendix F. (p.177) These results confirm the subsamples as being relatively representative of the population samples from which they were drawn when counts on the important major indices are used as measures. The results of the two-way analysis of variance procedure with repeated measures for the twenty-two additional counts are outlined in the following section. It is useful to note that these figures represent proportions (e.g. x noun clauses per T-unit or x present participles per clause). They can also be seen as percentages, that is, they indicate the mean percentage of occasions on which a student will write a present participle in his T-units or clauses depending on which index is being used. The analysis of variance results for these indices are also provided in Appendix F.

Noun clauses per T-unit. The two-way analysis of variance with repeated measures was used to test differences between the experimental and control subsamples with noun clauses per T-unit as the dependent variable. The means are provided in Table 21. The complete results are provided in Appendix F.

Table 21.

Mean Number of Noun Clauses per T-unit

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.190	0.202	0.203
Control subsample	30	0.247	0.158	0.215

No significant differences were found between groups or across the three testing occasions. All differences were attributable to chance. The null hypothesis was not rejected.

Adjective clauses per T-unit. The analysis of variance used to test for significant changes in the mean number of adjective clauses used by the two groups showed no significant differences. Table 22 displays the mean number of these constructions used by the two treatment groups across the three occasions.

Table 22

Mean Number of Adjective Clauses per T-unit

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.175	0.219	0.170
Control subsample	30	0.165	0.186	0.149

None of the differences were significant, and the null hypothesis was not rejected.

Adverb clauses per T-unit. Again the analysis of variance testing the mean number of adverb clauses per T-unit used by the two subsamples revealed no significant differences between the two groups across the three occasions. The means are shown in Table 23.

Table 23

Mean Number of Adverb Clauses per T-unit

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.280	0.313	0.378
Control subsample	30	0.339	0.387	0.317

The null hypothesis was not rejected.

Percentage of short T-units (8 words or less). The analysis of variance results for the test on the mean percentage of short t-units written by the two groups revealed no significant differences. The means are given in Table 24.

Table 24

Mean Percentage of Short T-units

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	29.1	21.1	20.8
Control subsample	30	21.3	20.7	22.6

None of the differences were significant and the null hypothesis was not rejected.

Percentage of medium T-units (9-20 words). The analysis of variance results for the test on the mean percentage of medium T-units written by the two groups revealed no significant differences. The means are given in Table 25.

Table 25

Mean Percentage of Medium T-units

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	56.5	64.2	64.1
Control subsample	30	60.5	60.8	61.4

None of the differences were significant, and the null hypothesis was not rejected.

Percentage of long T-units (21 words or more). The analysis of variance results for the test on the mean percentage of long T-units written by the two groups revealed no significant differences. The mean percentages are given in Table 26.

Table 26

Mean Percentage of Long T-units

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	14.3	14.7	15.1
Control subsample	30	18.5	19.0	13.6

None of the differences were significant and the null hypothesis was not rejected.

Gerunds per clause. The analysis of variance on the mean number of gerunds per clause also revealed no significant differences between groups across the three occasions. Table 27 provides the means.

Table 27

Mean Number of Gerunds per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.106	0.102	0.126
Control subsample	30	0.130	0.062	0.078

None of the differences were significant and the null hypothesis was not rejected.

Factive infinitives per clause. The analysis of variance detected no significant differences in the mean number of factive infinitives used by the two groups across the three occasions. The means are provided in Table 28.

Table 28

Mean Number of Factive Infinitives Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.096	0.143	0.146
Control subsample	30	0.138	0.119	0.120

None of the differences were significant and the null hypothesis was not rejected.

Nominalized verbals per clause (gerunds + factive infinitives). The analysis showed no significant increase in the mean number of nominalized verbals used by the experimental and control subsamples. This result was not unexpected in view of the fact that neither single index showed significant change. The mean number of nominalized verbals per clause is given in Table 29.

Table 29

Mean Number of Nominalized Verbals Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.208	0.245	0.272
Control subsample	30	0.268	0.181	0.200

None of the differences were significant and the null hypothesis was not rejected.

Modal auxiliaries per clause. The analysis of variance detected no significant differences between the mean number of modal auxiliaries per clause as they were used by the two subsamples. The means are given in Table 30.

Table 30

Mean Number of Modal Auxiliaries Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.359	0.347	0.365
Control subsample	30	0.320	0.343	0.338

None of the differences were signficant and the null hypothesis was not rejected.

Number of perfect forms per clause. The analysis of variance to test mean differences between experimental and control subsamples detected no significant differences. The means are given in Table 31.

Table 31

Mean Number of Perfect Forms Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.028	0.039	0.023
Control subsample	30	0.029	0.019	0.030

None of the differences were significant, and the null hypothesis was not rejected.

Number of passives per clause. The analysis of variance to test mean differences between the experimental and control subsamples detected no significant differences. The means are given in Table 32.

Table 32

Mean Number of Passives Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.044	0.062	0.067
Control subsample	30	0.063	0.073	0.075

None of the differences were significant and the null hypothesis was not rejected.

Verb forms (perfect forms, passives, and modals) per clause. The analysis of variance testing for mean differences between the three combined verb forms (perfect forms, passives, and modals) detected no significant differences. The means are given in Table 33.

Table 33

Mean Number of Verb Forms Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.430	0.448	0.455
Control subsample	30	0.412	0.435	0.443

None of the differences were significant, and the null hypothesis was not rejected.

Prepositional phrases as noun modifiers per clause. The analysis of variance test for mean differences in the number of prepositional phrases used as noun modifiers revealed no significant differences between groups across the three occasions. Table 34 displays the means for the two subsamples.

Table 34

Mean Number of Prepositional Phrases as Noun Modifiers Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.118	0.123	0.105
Control subsample	30	0.156	0.118	0.144

All differences were non-significant and the null hypothesis was not rejected.

Infinitives as noun modifiers per clause. The analysis of variance test for mean differences on the number of infinitives as noun modifiers per clause revealed no significant differences. The means are given in Table 35.

Table 35

Mean Number of Infinitives as Noun Modifiers per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.043	0.047	0.039
Control subsample	30	0.036	0.062	0.039

All mean differences were non-significant and thus the null hypothesis was not rejected.

Present participles as noun modifiers per clause. The analysis of variance testing for mean differences on the number of present participles as noun modifiers per clause revealed no significant differences. The means are displayed in Table 36.

Table 36

Mean Number of Present Participles as Noun Modifiers per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.034	0.062	0.029
Control subsample	30	0.040	0.022	0.034

Because all differences were non-significant, the null hypothesis was not rejected.

Past participles as noun modifiers per clause. The analysis of variance testing the mean differences between treatment subsamples across occasions, using the mean number of past participles as noun modifiers per clause, revealed no significant differences. The means are given in Table 37.

Table 37

Mean Number of Past Participles as Noun Modifiers per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.042	0.039	0.053
Control subsample	30	0.022	0.034	0.035

All differences were non-significant and the null hypothesis was not rejected.

Verb forms (infinitives, present participles, and past participles) as noun modifiers per clause. The combined verb forms used as noun modifiers were tested using analysis of variance to check for significant differences between means. The means are given in Table 38.

Table 38

Mean Number of Verb Forms as Noun Modifiers per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.114	0.145	0.114
Control subsample	30	0.094	0.110	0.101

All differences were non-significant and the null hypothesis was not rejected.

Inflected genitives per clause. The analysis of variance test for significant differences between subsamples across the three testing occasions using inflected genitives as the dependent variable revealed no significant differences. The means are given in Table 39.

Table 39

Mean Number of Inflected Genitives Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.071	0.088	0.072
Control subsample	30	0.093	0.109	0.081

All differences were non-significant and the null hypothesis was not rejected.

Phrasal genitives per clause. The analysis of variance test for significant differences between subsamples across the three testing occasions for phrasal genitives revealed no significant differences. The means are given in Table 40.

Table 40

Mean Number of Phrasal Genitives Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.045	0.025	0.039
Control subsample	30	0.060	0.068	0.044

None of the differences were significant and the null hypothesis was not rejected.

Total genitives per clause. The analysis of variance test for significant differences between groups across the three testing occasions for total genitives revealed no significant differences. The means are given in Table 41.

Table 41

Mean Number of Genitives Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.116	0.113	0.112
Control subsample	30	0.152	0.178	0.124

None of the increases or declines in means were significant. Therefore, the null hypothesis was not rejected.

Adjectives per clause. The two-way analysis of variance with repeated measures for testing for significant differences in group performance revealed a significant treatment/group effect on the adjectives per clause index. The means are given in Table 42.

Table 42

Mean Number of Adjectives Per Clause

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental subsample	30	0.442	0.459	0.415
Control subsample	30	0.491	0.604	0.558

A Newman-Keuls comparison of means showed that the two group means on pre-test occasion were not significantly different. None of the experimental subsample changes were significant. However, by the

post-test occasion the control group had improved significantly ($p < .05$). They maintained this significant increase on the delayed post-test despite a small decline on that occasion. The mean scores for the control subsample were significantly higher than the experimental subsample mean scores on both the post-test and the delayed post-test. The null hypothesis was therefore rejected.

Summary. In short, the treatment did not result in significant changes in the grade nine experimental subsample's use of the syntactic structures which Hunt (1965) indicated were most highly correlated with mature writing despite the fact that the experimental group practiced using these structures in the course of the treatment.

Reading Analysis

The main analyses of reading scores was carried out using the same two-way analysis of variance with repeated measures procedure that was used in analyzing student writing. Previous to this, however, the three forms of the Davis Reading Test (2A, 2B, 2C) were subjected to a one-way analysis of variance to test their comparability. This was done as a secondary check on the reading test; the random assignment of students to test form sequence was the primary method used to control the effect of reading form on group differences.

Form comparability: Level. The one-way analysis of variance to test mean differences between reading level test scores revealed no significant differences. The means are listed in Table 43.

Table 43

Davis Reading Test Level Scores Compared By Form

Form	n	$\bar{X}$	SD
2A	49	66.57	35.04
2B	45	68.93	44.93
2C	49	66.57	32.67

A Scheffé test of multiple means indicated no significant differences between any set. Tests comparing 2A to 2B and 2B to 2C showed a p-value of 0.177. The 2A to 2C comparison showed a p-value of 1.000. This check on form comparability along with the randomization procedure provided assurance that the influence of test form on group Level means was not significant.

Form comparability: Speed. The one-way analysis of variance to test mean differences between reading Speed test scores revealed no significant differences. The means are listed in Table 44.

Table 44

Davis Reading Test Speed Scores Compared By Form

Form	n	$\bar{X}$	SD
2A	49	67.02	41.15
2B	45	70.13	47.89
2C	49	67.47	41.68

A Scheffé test of multiple means indicated no significant differences between any set. Tests comparing 2A to 2B showed a p-value of 0.077. The 2B to 2C comparison had a p-value of 0.151. The 2A to 2C comparison p-value was 0.945. Overall probability of finding such mean differences was 0.052. These findings were not disturbing as the random assignment alone should have reduced form effect to insignificance.

Reading level: group comparisons. Hypothesis: There will be no significant differences between the mean scores of classes of ninth grade students who have and who have not received the sentence-combining treatment as measured by the "Level" subtest of the Davis Reading Test across three different occasions: pre-test, post-test, and delayed post-test.

The analysis of variance procedure used to test mean differences between reading level means across the three test occasions showed occasions to be significant. The means are displayed in Table 45, and the analysis of variance results are given in Table 46.

Table 45

Mean Reading Level Across Occasions

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental	75	67.33	69.29	68.49
Control	68	67.29	68.35	69.12

Table 46

Reading Level: Analysis of Variance Results

Source	SS	df	MS	F	p
Between subjects	16961.000	142			
'A' Main	2.229	1	2.229	0.019	0.892
Subjects within groups	16960.000	142	120.284		
Within subjects	3712.000	286			
'B' main effects	215.658	2	107.829	8.801	0.0002*
'AxB' interaction	42.630	2	21.315	1.740	0.1774
'B' x subject within groups	3455.000	282	12.252		

*Significant beyond the .05 level

Neither interaction nor treatment effect was significant. An analysis of differences between multiple means using the Newman-Keuls comparison revealed significant differences between pre-test and post-test, and pre-test and delayed post-test means of the experimental group. In other words, the decline that took place over the delay period did not decrease gains so drastically that the initial significant increase became non-significant. However, the control group also showed significant mean increases between pre-test and delayed post-test. The increase from pre-test to post-test was not significant. As indicated by Table 46, no significant differences between occasional means were detected. The experimental and control groups taken individually responded in predictable fashion on the three testing occasions. Because the treatment means were not significantly different, the null hypothesis was not rejected.

> Reading speed: group comparisons. Hypothesis: There will be no significant differences between the mean scores of classes of ninth grade students who have and who have not received the sentence-combining treatment as measured by the "Speed" sub-test of the Davis Reading Test across three different occasions: pre-test, post-test, and delayed post-test.

The analysis of variance results comparing treatment means across the three testing occasions showed significant differences for occasions and for interaction effect. The means are given in Table 47.

Table 47

Mean Reading Speed Across Occasions

Group	n	Pre-test	Post-test	Delayed Post-test
Experimental	75	67.61	69.76	69.11
Control	68	68.75	69.22	69.04

It is evident from a study of Table 47 that the experimental group showed the larger increase in mean scores. The analysis of variance results are given in Table 48.

Table 48

Reading Speed: Analysis of Variance Results

Source	SS	df	MS	F	p
Between subjects	16543.000	142			
'A' Main	4.458	1	4.458	0.038	0.8457
Subjects within	16541.000	141	117.312		
Within subjects	2150.000	286			
'B' Main	128.447	2	64.224	9.240	0.0001 *
'AxB' interaction	51.546	2	25.773	3.708	0.0257 *
'B' x subjects within groups	1960.000	282	6.950		

*Significant beyond the .05 level

Because the interaction was significant a test of simple occasion main effects ('B') was made. This test showed a significant difference between the experimental group means which was significant beyond the .01 level. The test of simple main effects for the control group across the three occasions revealed no significant differences. (Table 49) A one-way analysis of variance was carried out to compare the mean occasional scores for the experimental group. The Newman-Keuls comparison which was performed following the analysis showed the mean differences between occasions one and two, and one and three for this group were significant. (Table 50)

Table 49

Reading Speed: Test of Simple Treatment Effects

Treatment	Observed Value	Critical Value	df	Significance
Experimental	13.06	3.00	2,296	.01
Control	1.30	3.00	2,296	NS

Table 50

Reading Speed: Newman-Keuls Comparison -- Experimental Group

Occasion 1		3	2	df	Critical Value
1	—	1.49*	2.15*	3,148	1.02
3	—		.653	2,148	1.27

*Significant beyond the .05 level Multiplier = .3087

The experimental group thus showed significant growth between pre-test and post-test and maintained this significance over the course of the experiment despite a decline during the delay period. The null hypothesis was rejected.

Summary. The treatment appeared to elicit significant growth in the speed of comprehension of an experimental group of grade nine students when the Davis Reading Test was used as the measuring instrument and a control group was used for comparison. However, the treatment did not appear to elicit significant growth in the level of comprehension of the same students when the Davis Reading Test was used as the measuring instrument and the control group was used for comparison.

CHAPTER V

DISCUSSION OF RESULTS

Writing: Experimental vs Control

No significant mean increases were found on the six major writing indices across the six writing occasions as a result of the Sentencecraft program. Profiles of the group means on the six occasions are given in Figures 1 to 6 and provide a visual comparison to complement the tables of the previous chapter. It should be noted that, despite the lack of significant mean changes, the experimental group's overall performance improves for each of these six indices over the course of the study. The same is not true for the control group. The Hawthorne effect may be at work in the experimental-group performance, and in complementary fashion, the control groups may have lacked the motivation to continue writing on the same topics although they were not repeated for any individual. This lack of motivation may account in large part for their pre-test to delayed post-test decline on four of the six indices.

Writing: Pre-test Means and Comparisons with Previous Studies--Major Indices

The first three indices were seen as representing a measure of fluency in writing in this experiment. The mean counts for the total group (n=143) on the pre-test occasion are provided in Table 51. These means represent a fairly reliable indicator of what grade nine students are syntactically capable in an argumentative composition

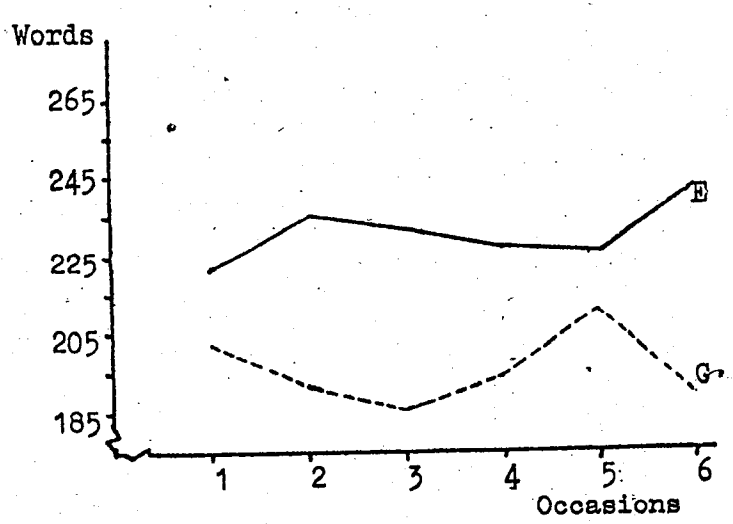

FIGURE 1: Words

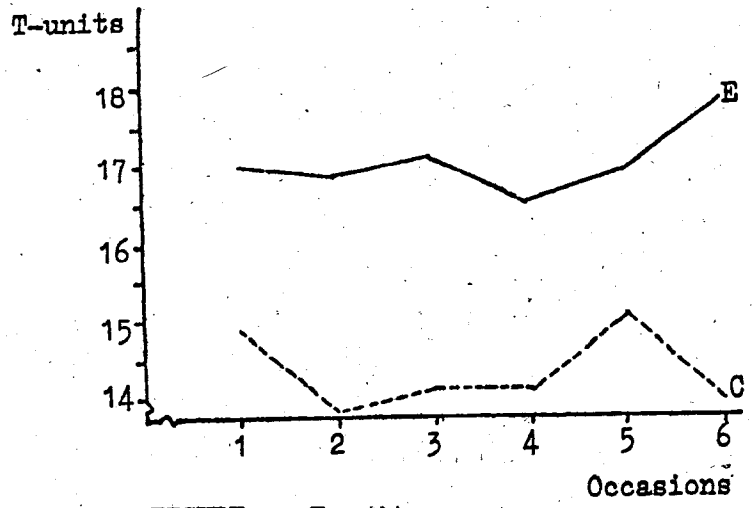

FIGURE 2: T-units

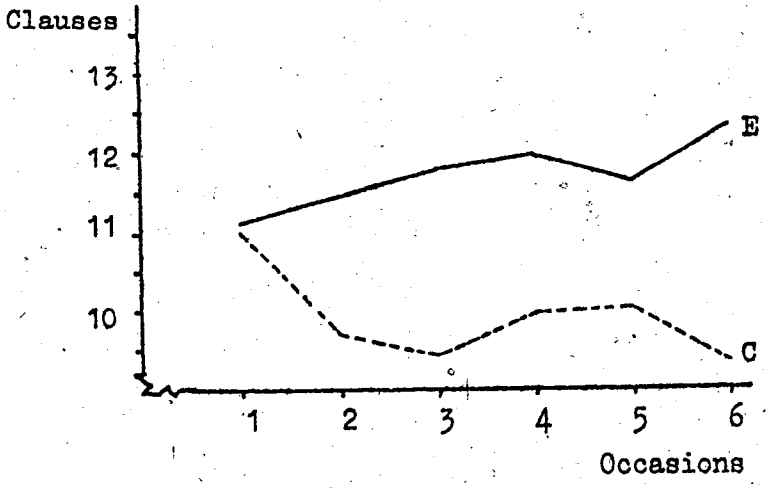

FIGURE 3: Clauses

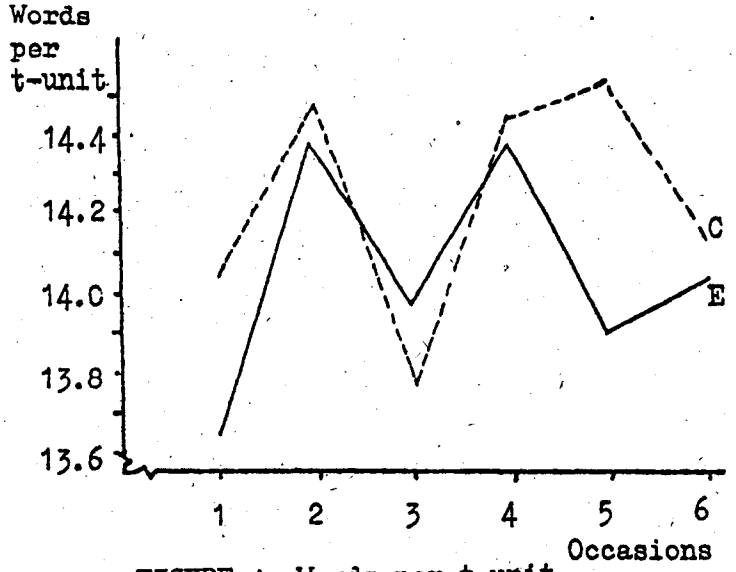

FIGURE 4: Words per t-unit

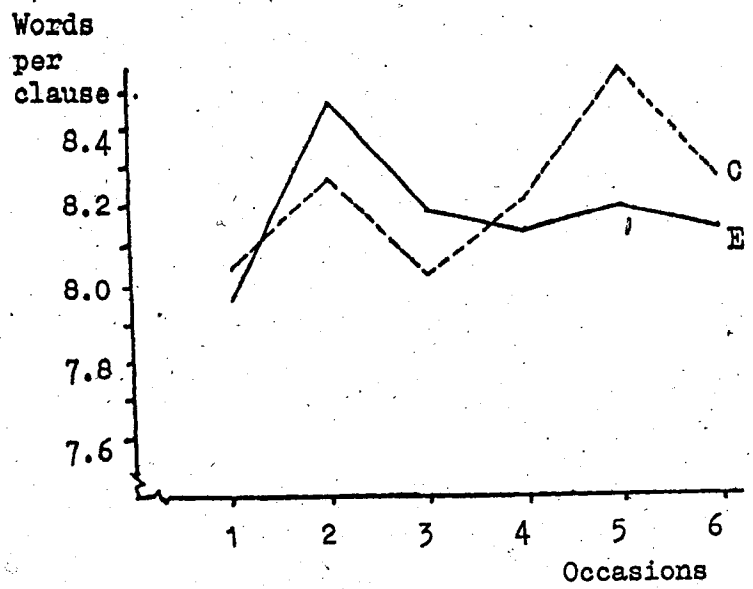

FIGURE 5 : Words per clause

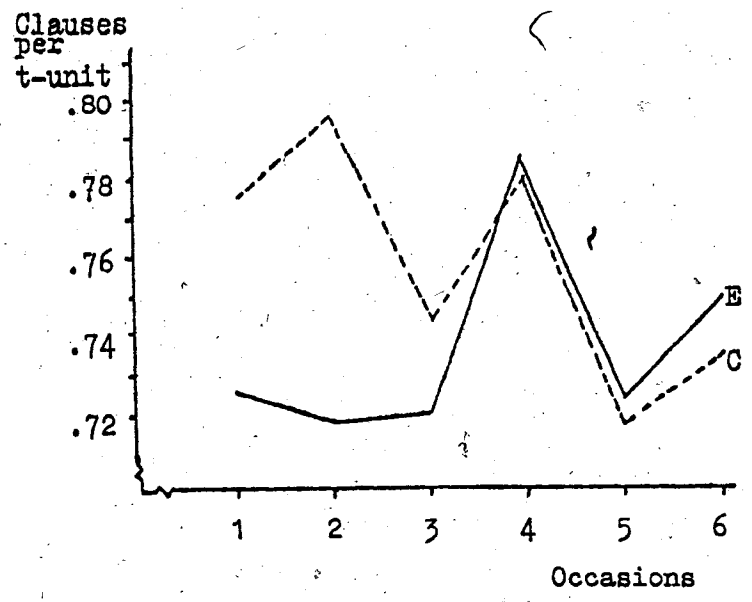

FIGURE 6: Clauses per t-unit

written in forty-five minutes. The pre-test means are provided as benchmarks for future researchers. These means were chosen because they are the one occasion least affected by treatment motivation, or other confounding variables.

Table 51

Mean Number of Words, T-units, and Clauses Written: Pre-test

Index	n	$\bar{X}$	SD
Words	143	213.00	56.38
T-units	143	15.98	4.88
Dependent Clauses	143	11.14	4.26

The next three major indices, words per T-unit , words per clause, and clauses per T-unit, are the indices for which counts have been made most often in previous studies. Comparisons of the mean scores attained by students in this study with those in previous studies are informative and should be helpful to future researchers in that they represent a benchmark indicator for argumentative writing by grade nine students. (Table 52)

Table 52

Comparison of Results of Various Studies: W/TU, W/Cl, CL/TU

Index	Author	Grade 5	7	8	9	10	11	12	Superior Adults
W/TU	Hunt			11.5				14.4	20.3
	Loban		8.94	10.37	10.05	11.79	10.69	13.27	
	Perron	13.06							
	O'Hare*		15.75						
	Crowhurst					14.26			
	MacNeill				13.83				
W/CL	Hunt			8.1				8.6	11.5
	Loban		6.98	6.91	6.84	7.76	7.37	8.29	
	Perron	7.72							
	O'Hare*		8.55						
	Crowhurst					8.88			
	MacNeill				8.00				

*Post-test scores

Table 52(continued)

Index	Author	Grades 5	7	8	9	10	11	12	Superior Adults
C1/TU	Hunt			.42				.68	.74
	Loban		.28	.50	.47	.52	.45	.60	
	Perron	.72							
	O'Hare*		.84						
	Crowhurst					.73			
	MacNeill				.75				

*Post-test scores

As can be seen from Table 52 the effect of mode on student writing is very significant. The results found by Crowhurst (1977) and the present study confirm Perron's (1976) findings with grade five students. Argumentative writing does seem to provide students with the opportunity to write more syntactically mature sentences than other descriptive studies have shown. For example, the students in the present study surpass the achievement levels found by Loban (1976) on words per T-unit and clauses per T-unit. Their mean on words per clause is close to Loban's grade twelve mean. Crowhurst's group surpassed Loban's on all three indices. The fact that Perron's grade five sample means are close to Loban's grade twelve means on two of the indices and surpass his grade twelve mean on the third suggests that there may in fact be a topping out at this level. Further research using argumentative writing with higher grades should clarify the situation. Nevertheless, results of studies which have ignored the question of mode of discourse in establishing norms are very suspect.

O'Hare's findings of a great increase in words per T-unit is also somewhat misleading. Hunt's mature adults, Loban's grade twelves, Crowhurst's grade tens, and the present study's grade nines never use subordination in more than seventy-five percent of their sentences. O'Hare's results showed a mean use of eighty-four percent. This appears to represent an overuse of a relatively unsophisticated syntactic option.

The results on these indices in the present study also seem to indicate that the Sentencecraft program is perhaps too unsophisticated syntactically to produce significant increases in the relatively high

mean scores of grade nine students on these indices. Another possible explanation is that grade nine students have reached a plateau in these three areas as Loban (1976, p. 60) suggested, and the significant increases which O'Hare and Combs found with grade seven students are no longer so easily attained. More light can be shed on this question by a comparison of the syntactic structure counts of the subsamples with the results of the Loban and Hunt studies. It should be kept in mind that neither the Hunt nor the Loban descriptive studies insisted on the argumentative mode when taking writing samples.

Writing: Pre-test Subsample Means Compared with Previous Studies--Minor Indices

Due to the lack of consistent significant increases on the minor indices by either subsample across the six occasions, this discussion will concentrate on comparing the pre-test means of the combined subsamples with those of Hunt (1965).

The frequency counts carried out to determine what was happening within the clauses of the experimental and control subsamples' argumentative compositions reached no significant, consistent increases by either group. The conclusion can be drawn that the Sentencecraft program did not induce consistently significant increases in the writing of the experimental group exposed to the treatment. However, some caution is necessary in any further interpretation of the results. Comparisons of the mean frequencies for the minor indices with those of Hunt (1965) show that these grade nine students in the present study are already quite adept in the use of these syntactic structures.

The number of counts that were made in the present study, should provide future researchers with a new set of benchmarks when studying the syntactic capabilities of grade nine students. Again pre-test means are reported because they should be least affected by treatment, motivation or other confounding variables. The control and experimental subsamples' scores are combined in this analysis (n=60).

The first three counts were concerned with the percentage of short, medium, and long T-units that these sixty students used in their argumentative essays. The results are reported and compared in Table 53. The results clearly show that these grade nine students writing in the argumentative mode are closer to grade twelve percentages than to grade-eight. In fact, they use a slightly higher percentage of long T-units than did the grade twelves in Hunt's study. However, they use a slightly higher percentage of short T-units as well. In total, they are very close to Hunt's twelves. It is entirely possible that they are forced into using longer T-units by the very nature of argumentation. This mode seems to provide them with the best opportunity for displaying their written syntactic ability.

Table 53

Comparison of Hunt (1970) and the Present Study on Percentage of Short, Medium, and Long T-units

Index	Author	Grade 8	9	12	Adult
%Short	Hunt	39		23	19
	MacNeill		25.0		
%Medium	Hunt	54		61	43
	MacNeill		58.6		
%Long	Hunt	7		16	38
	MacNeill		16.4		

The second set of indices that will be studied here are the three kinds of subordinate clauses. The means are listed in Table 54. Again, it should be noted that the mean percentage for using noun, adjective, or adverb clauses in their writing is shown as being above the Loban percentages for grade nine students.

Table 54

Comparison of Results of Hunt (1970), Loban (1976) and the Present Study: Noun, Adjective, and Adverb Clauses

Index	Author	Grade 7	8	9	10	11	12	Superior Adults
Noun clauses/ T-unit	Hunt	.16					.29	.23
	Loban	.08	.19	.15	.16	.15	.18	
	MacNeill			.22				
Adjective clauses/T-unit	Hunt		.09				.17	.25
	Loban	.05	.11	.13	.16	.18	.20	
	MacNeill			.17				
Adverb clauses/ T-unit	Hunt							
	Loban	.15	.21	.19	.18	.12	.22	
	MacNeill			.31				

Adverbial clause use (31%) within T-units by the grade nines in this study is much higher than the use of these structures by grade twelves in Loban's study. This increased use, however, does not suggest maturity, as adverbial clauses tend to be the most easily adopted clause form. It is possible and reasonable to suggest that such clauses are induced by the mode of writing. Place, time, manner, and reason are needed in forming arguments. It is noteworthy here that the largest clausal increase by the experimental group in this study was adverb clauses. The frequency increased from twenty-eight percent of the

time to thirty-eight percent between pre-test and delayed post-test. It is possible that Sentencecraft promotes the use of this relatively early-maturing structure. In a similar way, noun clauses are used more frequently than in the writing of Loban's grade twelves, although less frequently than Hunt's grade twelves. It should again be pointed out that this is not a particularly sophisticated sentence structure. Loban's grade eights already use them in nineteen percent of their T-units. Use declines slightly from that point and remains below twenty percent. On the other hand, adjective clauses, which are late-blooming structures, do not reach in the present study the percentage figure which Loban's grade elevens set. The mean for the grade nine students in the present study rests between Loban's grade ten and eleven norms. Again this could be due to the mode of writing. Descriptive writing should probably be expected to elicit more adjective clauses than the argumentative mode used in the present study.

The remaining sixteen counts can only be compared to Hunt's norms which were based on compositions for which the mode of writing was not controlled. However, he is the only researcher who made these complex counts. For some of the indices, Hunt's norms could not be calculated from the data he supplied. Norms from the present study are included anyway because the contingency coefficients which Hunt calculated suggest that all these structures are significant indicators of syntactic maturity. The contingency coefficients are supplied in Table 2, page 19.

The first set of these counts were on nominalized verbals per clause. These included gerunds and factive infinitives. It is noteworthy that the grade nines in the present study used over one and a

half times as many of these structures as Hunt's grade twelves did. The means are provided in Table 55. This is a rather interesting finding because gerunds and factive infinitives are both better indicators of syntactic maturity than adjective clauses or the use of verb forms as noun modifiers according to Hunt (1965, p. 149). They represent significant development in the use of syntactic options. It is also noteworthy that every time a clause is reduced to a noun or adjective form, the writer may, in fact, be cutting the mean number of words per T-unit which he would have attained if he merely subordinated. It appears that the argumentative mode elicits the use of these relatively mature structures.

Use of the six modals, perfect forms of the verb, and passives also suggest a high degree of writing sophistication. Yet, these students use these forms in forty-two percent of their clauses. This suggests that they already have, at this stage, a good control of various verb forms. It also suggests that the argumentative mode may well demand finer distinctions in meaning than other modes and thus elicit more adept manipulation of these verb forms.

Prepositional phrases as noun modifiers, in contrast, would appear to be more frequently used in descriptive or narrative writing in which short descriptive phrases are much in demand. Although the grade nine percentages fall between Hunt's grade eights and twelves, they are nearest the higher grade. In short, students in this study and writing in this mode are relatively adept in their use of this structure.

Another set of noun modifiers, namely verb forms, also show the

Table 55

Mean Number of Syntactic Structures Per Clause

Index	Author	Grade 8	9	12
Total nominalized verbals/clause	Hunt	.07		.14
	MacNeill		.24	
-Gerunds	Hunt	.03		.[illegible]
	MacNeill		.12	
-Factive Infinitives	Hunt	.04		.07
	MacNeill		.12	
Total selected verb forms/clause	MacNeill		.42	
-Modal auxiliaries	MacNeill		.34	
-Perfect forms	MacNeill		.03	
-Passives	MacNeill		.05	
Prepositional phrases as noun modifiers/ clause	Hunt	.10		.15
	MacNeill		.14	
Verb forms as noun modifiers/clause	Hunt	.07		.09
	MacNeill		.10	
-Past participles	MacNeill		.03	
-Present participles	MacNeill		.04	
-Infinitives	MacNeill		.04	
Total genitives/clause	Hunt	.29		.38
	MacNeill		.13	
-Inflected genitives	MacNeill		.08	
-Phrasal genitives	MacNeill		.05	
Adjectives/clause	Hunt	.39		.43
	MacNeill		.47	

present study's grade nines as using present participles, past participles and infinitives more than Hunt's grade twelves. Again, these are clause reductions and so are a fairly sophisticated syntactic form.

The group's performance on genitives is difficult to explain. The percentage occurrence of these structures (13.4%) is less than one-half that of Hunt's grade eights and one-third that of his grade twelve sample. It is easy to see that descriptive writing would again be more conducive to use of these structures. The differences found seem rather large to attribute to mode alone but the number of genetives does not increase over the course of the experiment. This would seem to indicate that, in fact, mode may well be the limiting factor.

The number of adjectives per clause, like most of the other secondary counts, is higher than Hunt's grade twelve mean. What appears to be happening here is that these grade nines are using more simple adjectives and using fewer prepositional adjective phrases. If one totals these two sets of modifiers, the grade nines in the present study still outperform Hunt's grade twelves by two and a half percentage points.

Reading: Experimental vs Control

The hypothesis that reading "Level of Comprehension" scores as defined by Davis (1962), would not differ significantly for treatment groups was not rejected. The fact that the differences in mean scores of both groups were significant between pre-test and delayed post-test with no significant differences on occasions suggests that the

experimental treatment was no more effective than the control activities in improving students' level of comprehension. This is, in fact, what was predicted would happen. The mean Level scores, it was thought, would not increase significantly because there were no structures being practiced that grade nine students could not interpret, given an unlimited amount of time. Thus, it was predicted, the mean differences between groups would not attain significance. Perhaps a more difficult or complex set of sentence-combining problems, using more difficult structures, would produce significant Level differences. It is hard to account for the control groups' significant increases over the course of the experiment considering it only lasted seventeen weeks. Practice effect could be cited as a possible uncontrolled factor.

The hypothesis that reading "Speed of Comprehension" scores, as defined by Davis (1962) would not differ significantly for treatment groups was rejected. The experimental group does in fact show a significant mean increase between pre-test and post-test, and between pre-test and delayed post-test. None of the control group changes were significant. This result was predicted because the practice of sentence combining, it was thought, would facilitate speed in comprehending sentences using the syntactic structures practised in the Sentencecraft program despite the fact that students could already understand them, given a non-speeded situation. It appears that students in grade nine can significantly improve their speed of comprehension by undergoing the Sentencecraft program or in all probability any similar program in which the same structures are practiced.

CHAPTER VI

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Summary

The Sentencecraft program was used as the treatment in the present experiment which was designed to measure the effect of a sentence-combining program on selected writing and reading skills of a heterogeneous group of ninth grade students. Six argumentative essays were used as the writing measure and the Davis Reading Test (Form 2) was used to measure reading comprehension. The treatment lasted for nine weeks as outlined by O'Hare (1973) in the "Teacher's Guide" to Sentencecraft (See Appendix A). Reading and writing measurements were taken on pre-test, post-test, and delayed post-test occasions; writing measurements were also taken on three other occasions between the pre-test and the post-test.

The experimental group did not achieve significantly higher mean increases than the control group on any of the six major writing indices which were used: the number of words written, the number of T-units written; the number of dependent clauses used; the number of words per T-unit; the number of words per clause; or the number of clauses per T-unit. Nor did the experimental group subsamples achieve significantly higher mean increases than the control group subsample on any of the sixteen additional indices which consisted of syntactic structure counts carried out on the compositions written on the pre-test, post-test, and delayed post-test occasions. However, comparison with earlier studies (Hunt, 1965; Loban, 1976) confirmed that the

students in the present study already used many of these structures more often than grade eleven and twelve students in the earlier studies.

The experimental group did not achieve significantly higher mean increases than the control group on the reading comprehension "Level" index as measured by the Davis Reading Test. The experimental group did, however, achieve significantly higher mean increases than the control group on the reading comprehension "Speed" index as measured by the Davis Reading Test. This increase was significant beyond the .01 level.

Conclusions

The results of this study showed that the Sentencecraft program did not result in significant increases in the mean number of selected words, phrases, or clauses that grade nine students wrote in argumentative compositions. A number of possibilities could explain these results: 1) the argumentative mode itself may actually elicit the greatest complexity or maturity of which grade nine students are syntactically capable; 2) the Sentencecraft program may not provide grade nine students with practice in structures that are sophisticated enough to significantly increase the occurrence of such structures in their writing; or 3) grade nine students may well have reached a plateau in their writing development as Loban (1976) has suggested. Whatever the case, the significant increases achieved by the O'Hare (1973) and Combs (1975) grade seven groups were not found in the argumentative writing of the grade nine students in the present study.

However, the argumentative mode did appear to provide student writers with the opportunity to display this syntactic tendency . They scored higher than Loban's grade nines on the majority of writing indices and close to his grade eleven and twelve students on many of these indices. The results supported the findings of researchers such as Perron and Crowhurst who pointed out the importance of controlling mode in writing research. The norms for grade nine argumentative writing which the present study established should be viewed as representing optimum syntactic ability for average grade nine students replacing those of Loban (1976).

Further, the results of this study showed that the Sentencecraft program was not effective in eliciting significant growth in "Level" of comprehension of grade nine students as measured by the Davis Reading Test. This may, in fact, have been due to the possible pretest ability of students to decode any of the structures practiced in the program if given sufficient time.

However, from the results of this study it is reasonable to conclude that the "Speed" of comprehension of grade nine students can be significantly improved by having them complete the Sentencecraft program.

Implications

The present study has several implications for future research on the effects of sentence combining on writing, for research which attempts to set norms for syntactic ability, as well as for research using sentence combining to improve reading comprehension.

1. Researchers can no longer ignore the influence of mode on the syntactic elements of student writing. This study supports the findings of Perron (1976) and Crowhurst (1977) which indicated that argumentative writing influenced the three major writing indices of words per T-unit, words per clause, and clauses per T-unit. The present study supports their findings on these indices but goes one step further in pointing out the probable influence of mode on the use of mature within-clause structures. Together, these pieces of research confirm that the argumentative mode must be used if researchers are attempting to provide students with the opportunity to display their full range of syntactic skills.

2. A new set of norms based on argumentative writing will have to be constructed to take the place of the Hunt and Loban norms if the concern is with measuring highest capability rather than performance in the "usual writing" of students.

3. It is imperative that more researchers look within the clause in order to determine whether the three indices of words per T-unit, words per clause, and clauses per T-unit are by themselves accurate indicators of syntactic maturity.

4. More research is needed to determine whether more sophisticated syntactic structures than are practiced in Sentencecraft will induce continued growth in the argumentative writing of grade nine students.

5. Teacher effect and ability level as influences on the effectiveness of the Sentencecraft program were not controllable in the present study. Studies are needed in which teacher effect is controlled so that any differential effect of this program as it

relates to ability level can be determined.

6. Sentence-combining problems such as are used in Sentencecraft should be tested at other levels to determine whether such practice can consistently induce growth in speed of comprehension.

Selected Bibliography

ALBANESE, DOROTHY A study of the complexity of sentence structure as a factor in reading comprehension. Doctoral dissertation, Fordham University, 1972.

BATEMAN, D.R. & ZIDONIS, F.J. The Effect of a Study of Transformational Grammar on the Writing of Ninth and Tenth Graders. Champaign, Illinois: National Council of Teachers of English, 1966.

BORMUTH, J.R. Readability: a new approach. Reading Research Quarterly, 1: 79-132, 1966.

BORMUTH, J.R., CARR, J., MANNING, J., & PEARSON, D. Children's comprehension of between- and within- sentence syntactic structures. Journal of Educational Psychology 61(5): 349-357, 1970.

BRADDOCK, R., LLOYD-JONES, R. & SCHOER, L. Research in Written Composition. Champaign, Illinois: National Council of Teachers of English, 1963.

BRENEMAN, ELIZABETH The relationship between linguistic awareness and reading comprehension among fourth, sixth, and eighth graders. Doctoral dissertation, University of Colorado, 1975.

BRINTON, J.E., & DANIELSON, W.A. A factor analysis of language elements affecting readability. Journalism Quarterly 35: 420-426, 1958.

BROOKS, G.M. An investigation of the relationship of syntactic complexity to reading comprehension. Doctoral dissertation, Purdue University, 1974.

CAMPBELL, D.T. & STANLEY, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1966.

CANADIAN LORGE-THORNDIKE INTELLIGENCE TEST (Level F), Manual for Administration by Irving Lorge, R.L. Thorndike, Elizabeth Hagen, and E.N. Wright. Toronto: Thomas Nelson and Sons, 1967.

CARROLL, J.B. Comprehension by third, sixth, and ninth graders of words having multiple grammatical functions. Washington, D.C.: U.S. Office of Education, 1970.

CHOMSKY, N. Aspects of the Theory of Syntax. Cambridge, Massachusetts: M.I.T. Press; 1965.

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

CHRISTENSEN, FRANCIS Notes Toward a New Rhetoric. New York: Harper and Row, 1967.

CHRISTENSEN, FRANCIS The problem of defining a mature style. English Journal 57(4): 572-576, 1968.

COFFMAN, WILLIAM E. Review of The Davis Reading Test (Series 1 and 2) In Oscar K. Buros (Ed), Reading Tests and Reviews. Highland Park, N.J.: The Gryphon Press, 1968.

COMBS, WARREN E. Explaining the Fit of Practice in Syntactic Manipulation and Scores of Reading Comprehension. In Donald A. Daiker et al. (Eds.), Sentence-Combining and the Teaching of Writing. Conway, Arkansas: L and S Books, 1979.

COOMBS, W.E. Some further effects of sentence-combining exercises for the secondary Language Arts curriculum. Doctoral dissertation, University of Minnesota, 1975.

CROWHURST, MARION AND PICHE, GENE L. Audience and Mode of Discourse Effects on Syntactic Complexity in Writing at Two Grade Levels. Research in the Teaching of English, 13: 101-110, 1979.

CROWHURST, MARION The Effect of Audience and Mode of Discourse on the Syntactic Complexity of the Writing of Sixth and Tenth Graders. Unpublished doctoral dissertation, University of Minnesota, 1977.

CROWHURST, MARION Syntactic Complexity and Teachers' Quality Ratings of Narrations and Arguments. Research in the Teaching of English 14: 223-231, 1980.

DAVIS READING TEST (Series 2) Manual by Frederick B. Davis and Charlotte Croon Davis. New York: The Psychological Corporation, 1962.

DAVIS F.B. Research in comprehension in reading. Reading Research Quarterly 3: 499-545, 1968.

DIEDERICH, P.B. The measurement of skill in writing. School Review 54: 584-592, 1946.

DIEDERICH, P.B. Measuring Growth in English. Champaign, Illinois: National Council of Teachers of English, 1974.

EVANECHKO, P., OLLILA, L. & ARMSTRONG, R. An investigation of the relationships between children's performance in written language and their reading ability. Research in the Teaching of English. 8(3): 315-326, 1974.

EVANS, RONALD V. The Relationship between the reading and writing of syntactic structures. Research in the Teaching of English 13: 129-135, 1979.

FAGAN, WILLIAN T. Transformations and comprehension. The Reading Teacher 25: 169-172, 1971.

FARR, ROGER Reading: What Can Be Measured. Newark, Delaware: International Reading Association, 1969.

FISHER, K.D. An investigation to determine if selected exercises in sentence-combining can improve reading and writing. Doctoral dissertation, Indiana University, 1973.

FREEDLE, ROY & CARROLL, J.B. Language comprehension and the acquisition of knowledge. In Roy O. Freedle and John B. Carroll (Eds.), Language Comprehension and the Acquisition of Knowledge. Washington, D.C.: V.H. Winston and Sons, 1972.

GAJADHARSINGH, J.L. A study of the effects of instruction in the rhetoric of the sentence on the written composition of junior high school students. Doctoral dissertation, University of Alberta, 1970.

GAMLIN, PETER The interaction of meaning and syntax in sentence comprehension. Child Study 30(4): 3-13, 1968-69.

GILLILAND, J. Readability. London: University of London Press, 1972.

GOLUB, LESTER S. AND KIDDER, CAROLE Syntactic density and the computer. Elementary English 51(8): 1128-1131, 1974.

GOUGH, P.B. Grammatical transformations and speed of understanding. Journal of Verbal Learning and Verbal Behavior 4: 107-111,

HAKE, ROSEMARY Sentence expanding: Not can, or how, but when. In Donald A. Daiker et al. (Eds.), Sentence-Combining and the Teaching of Writing. Conway, Arkansas: L and S Books, 1979.

HUGHES, THEONE Sentence-combining: a means of increasing reading comprehension. Doctoral dissertation, Western Michigan University, 1975.

HUNT, K.W. Grammatical Structures Written at Three Grade Levels. Champaign, Illinois: National Council of Teachers of English, 1965.

HUNT, K.W. Sentence structures used by students in grades four, eight, and twelve, and by superior adults. Tallahasee, Florida: Florida State University, 1966.

HUNT, K.W. Syntactic development of schoolchildren and adults. Monographs of the Society for Research in Child Development 35: 1, 1970.

HUNT, K.W. Early blooming and late blooming structures. In Charles R. Cooper & Lee Odell (Eds), Evaluating Writing. Champaign, Illinois: National Council of Teachers of English, 1977, pp. 90-132.

JENKINSON, M.D. Selected processes and difficulties of reading comprehension. Doctoral dissertation, University of Chicago, 1957.

KEREK, ANDREW et al. The effects of intensive sentence-combining on the writing ability of college freshmen. In Donald McQuade (Eds.), Linguistics, Stylistics and the Teaching of Composition. New York: City University Press, 1979.

KINCAID, GERALD. Some factors affecting variation in the quality of student writing. Doctoral dissertation, Michigan State University, 1953. As reported in Richard L. Braddock et al., Research in Written Composition. Champaign, Illinois: National Council of Teachers of English, 1963, pp. 83-95.

KINNEAVY, JAMES L. Sentence-combining in a comprehensive language framework. In Donald A. Daiker et al. (Eds.) Sentence-Combining and the Teaching of Writing. Conway, Arkansas: L and S Books, 1979.

KLARE, GEORGE R. The Measurement of Readability. Ances, Iowa: Iowa State University Press, 1963.

KUNTZ, M.H. The relationship between written syntactic attainment and reading ability in seventh grade. Doctoral dissertation, University of Pittsburgh, 1975.

LABRANT, LOU A study of certain language developments of children in grades four to twelve, inclusive. Genetic Psychology Monographs 14(5): 387-489, 1933.

LATHAM, ROSS O. Cognitive synthesis and the comprehension of written language. Doctoral dissertation, University of Alberta, 1973.

LEFEVRE, CARL A. Linguistics and the Teaching of Reading. Toronto: McGraw-Hill Book Company, 1964.

LEFEVRE, CARL A. Linguistics, English, and the Language Arts. Boston: Allyn and Bacon, 1970.

LENNON, ROBERT What can be measured? Reading Teacher 15(5): 326-337, 1962.

LESGOLD, A.M. Variability in children's comprehension of syntactic structures. Journal of Educational Psychology 66(3): 333-338, 1974.

LETSON, C.T. Speed and comprehension in reading. Journal of Educational Research 52(2): 49-53, 1958.

LITTLE, P.S. An investigation into the relationship between structural ambiguity and reading comprehension. Master's thesis, University of Alberta, 1972.

LOBAN, WALTER Language ability: grades seven, eight, and nine. Cooperative Research Monograph, Number 18. Washington: U.S. Office of Education, 1966.

LOBAN, WALTER Language Ability: grades ten, eleven, and twelve. Berkeley, California: University of California, 1967.

LOBAN, WALTER Language Development: Kindergarten through Grade Twelve. Urbana, Illinois: National Council of Teachers of English, 1976.

MANNING, ANDREW R. Sentence-combining and writing: Research designs. Canadian Society for Studies in Education Conference Paper, 1979.

MELLON, J.C. Transformational Sentence-Combining: A Method for Enhancing the Development of Syntactic Fluency in English Composition. Champaign, Illinois: National Council of Teachers of English, 1969.

MCNEILL, DAVID The Acquisition of Language. New York: Harper and Row, 1970.

MILLMAN, JASON A review of the Gates-McGinitie Reading Test (Survey E). In Oscar K. Buros (Ed.), Reading Tests and Reviews. Highland Park, New Jersey: The Gryphon Press, 1975.

MOFFETT, JAMES Teaching the Universe of Discourse. Boston: Houghton Mifflin, 1968.

NEY, JAMES W. Notes toward a psycholinguistic model of the writing process. Research in the Teaching of English 8(2): 158-168, 1974.

O'DONNELL, R.C. The relationship between awareness of structural relationships in English and ability in reading comprehension. Doctoral dissertation, Auburn University, 1961.

O'DONNELL, R.C. A critique of some indices of syntactic maturity. Research in the Teaching of English 10: 31-38, 1976(a).

O'DONNELL, R.C. On linguistic structure and reading comprehension. Paper presented at the Annual Meeting of the National Conference on Language Arts in the Elementary School, 1976 (b). EDRS 123592.

O'DONNELL, R.C. Language development: kindergarten through grade twelve: a review. Research in the Teaching of English 11(1): 49-53, 1977.

O'DONNELL, R.C., GRIFFIN, W.J. & NORRIS, R.C. A transformational analysis of oral and written grammatical structures in grades three, five, and seven. The Journal of Educational Research 6(1): 35-39, 1967.

O'DONNELL, R.C. AND KING, F.J. An exploration of deep structure recovery and reading comprehension skills. Research in the Teaching of English 8(3): 327-338, 1974.

O'DONNELL, R.C. & SMITH, W.L. Increasing ninth grade students' awareness of syntactic structure through direct instruction. Research in the Teaching of English 9(3): 257-262, 1975.

O'HARE, FRANK The effect of sentence-combining practice not dependent on formal knowledge of a grammar on the writing of seventh graders. Doctoral dissertation, Florida State University, 1971.

O'HARE, FRANK Sentence-combining. Urbana, Illinois: National Council of Teachers of English, 1973.

O'HARE, FRANK Sentencecraft. New York: Ginn and Company, 1975.

OLSON, G.M. Developmental changes in memory and the acquisition of language. In Timothy E. Moore, Cognitive Development and the Acquisition of Language. New York: Academic Press, 1973.

PAVLAK, STEPHEN A. Reading comprehension -- a critical analysis of selected factors affecting comprehension. Doctoral dissertation, University of Pittsburgh, 1973.

PEDERSEN, ELRAY Improving syntactic and semantic fluency in the writing of language arts students through extended practice in sentence-combining. Doctoral dissertation, University of Minnesota, 1977. University Microfilms, 7802703.

PEDERSEN, ELRAY Prospects for sentence-combining. In Donald A. Daiker et al. (Eds.) Sentence-Combining and the Teaching of Writing. Conway, Arkansas: L and S Books, 1979.

PERRON, J.D. The impact of mode on written syntactic complexity (Part 3, Grade 5). In Studies in Language Education. Report Number 27. Athens, Georgia: The University of Georgia, 1976.

ROBINSON, SAM Sentence-combining: An inward look. Canadian Society for Studies in Education Conference Paper, 1979.

RUDDELL, ROBERT The effect of the similarity of oral and written language structure on reading comprehension. Elementary English 42: 403-410, 1965.

RUTLAND, J.W. An investigation into relative syntactic difficulty and memory strategies in prose comprehension. Master's thesis, University of Alberta, 1975.

SAVIN, H.B. & PERCHONOCK, ELLEN Grammatical structure and the immediate recall of English sentences. Journal of Verbal Learning and Verbal Behavior 4: 348-353, 1965.

SHOCKLEY, SANDRA J. An investigation into the effects of training in syntax on reading comprehension. Doctoral dissertation, University of Georgia, 1974.

SIEDOW, MARY D. Relationships between syntactic maturity in oral and written language and reading comprehension materials of varying syntactic complexity. Doctoral dissertation, University of Indiana, 1973.

SIMISON, DIANE L. The Relation Between the Grammatical and Psychological Complexity of Some English Sentences. Doctoral dissertation, Indiana University, 1970.

SLOBIN, D.I. Grammatical transformations and sentence comprehension in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior 5: 219-227, 1966.

SMART, W. DOUGLAS & OLLILA, LLOYD O. The effect of sentence-combining practice on written compositions and reading comprehension. The Alberta Journal of Educational Research 24: 113-120, 1978.

SMITH, W.L. The effects of syntax on reading. Doctoral dissertation, university Microfilms, 1969.

SMITH, W.L. & MASON, G.E. Syntactic control involving better comprehension. Journal of Reading 15(5): 355-358, 1972.

SPEARRIT, DONALD Identification of subskills of reading comprehension by maximum likelihood factor analysis. Reading Research Quarterly 8(1): 92-111, 1972.

STEWART, MURRAY F. Syntactic maturity from high school to university: a first look. Research in the Teaching of English 12(1): 37-46, 1978.

STOLUROW, L.M. & NEWMAN, J.R. A factorial analysis of objective features of printed language presumably related to reading difficulty. Journal of Educational Research 52(7): 243-251, 1959.

STOTSKY, SANDRA L. Sentence-combining as a curricular activity: its effect on written language development and reading comprehension. Research in the Teaching of English 9(1): 30-71, 1975.

STRAW, STANLEY B. Measuring the effect of sentence-combining instruction on reading comprehension. In Donald A. Daiker et al. (Eds.), Sentence-Combining and the Teaching of Writing. Conway, Arkansas: L and S Books, 1979.

SULLIVAN, MAUREEN A. Parallel sentence-combining studies in grades nine and eleven. In Donald A. Daiker et al. Sentence-Combining and the Teaching of Writing. Conway, Arkansas: L and S Books, 1979.

SWAN, MARY BEVERLY The effects of instruction in transformational sentence-combining on the syntactic complexity and quality of college level writing. Doctoral dissertation, Boston University, 1977. Dissertation Abstracts International, 34: 4556A, 1974.

TAKAHASHI, BARBARA Comprehension of written syntactic structures by good readers and slow readers. Master's thesis, Rutgers University, 1975.

WATTS, A.F. The Language and Mental Development of Children. London: George C. Harrap, 1944.

PHONES: 963-2231

Appendix A

COUNTY OF PARKLAND NO. 31
OFFICE OF THE SUPERINTENDENT

STONY PLAIN, ALBERTA
T0E 2G0

October 21, 1977

Mr. Thomas Mac Neill
511 Michener Park
Edmonton, Alberta

Dear Mr. Mac Neill:

Last Friday you visited my office to discuss your proposal for the Cooperative Activities Program. I indicated at that time that I would contact the school to determine the teachers' interest in cooperating with this project. I have spoken to both the Principal and the Assistant Principal who will be in charge of the project. Mr. Tkachuk, the Principal, has agreed to proceed with the project as has Mrs. Pat Penner, the Assistant Principal. I might add that Mrs. Pat Penner was very enthusiastic about the project.

You have our permission to proceed with the project and all future contacts may be made directly with the school.

Because of my own personal interest in the project I would appreciate your keeping me in touch. I will make an attempt to visit the classroom in which this project is being conducted.

We are pleased that you will be conducting this project in the County of Parkland.

Yours truly,

Robert J. Toews
Assistant Superintendent
of Schools (Personnel)

RJT:dl

c.c. Mr. M. Tkachuk
Mrs. Pat Penner

Revised Design

Cooperative Activities Program

1. Nature of Activity (Check One)

Student Teaching Internship _____ Demonstration/Experimentation__

Special Practicum_____ Research Experimental

2. Organization to be Involved

Edmonton Public School System_____ County of Strathcona_____

Edmonton Separate School System_____ St.Albert Protestant/Separate school system _____

N.A.I.T._____

U. of A. Faculty of _____ Other County of Parkland, #31

3. Requestor(staff member)

Name Dr. J. Oster Position Professor Date Oct.14,1977

Request made on behalf of Thomas B. MacNeill

4. Description of Activity - Include title, objectives, procedure, evaluation, techniques, etc.

Title: A study to determine the effect of sentence-combining practice on the development of comprehension and writing abilities in grade nine students.

Objectives: To determine if sentence-combining(SC) practice results in an increase in syntactic maturity in student writing. To determine if SC practice results in an increase in students' ability to understand complex syntactic structures.

To achieve some insight into the acquisition of complex syntactic structures as indicated by student writing samples.

To determine if there is any significant change in writing and comprehension abilities as indicated by a delayed post-test.

Procedure: The study is a pre-test, experiment, post-test, delayed post-test design. Each testing period will involve the writing of a composition and the taking of the Davis Reading Test.

Each teacher participating should be teaching two classes; one will be used for the experiment and the other will act as a control group.

The classes involved should be "non-streamed", should not be studying a second language, and should not be taking a reading instruction program as these situations would complicate the results.

No syntax instruction of any other type can take place during the 17 weeks of the experiment other than the sentence-combining practice. This is a case for the experimental and the control groups. The experimental classes must forego syntax instruction for six to eight additional weeks until the delayed post-test. If necessary, the control groups may resume such instruction or begin the sentence-combining program immediately following the post-tests. But, if necessary, the control groups could undertake the sentence-combining program the following year.

Writing exercises for every third week are included in the workbook. Both the experimental and control groups will undertake these as well as any additional writing which the teachers deem appropriate.

The experimental groups will practice sentence-combining using Frank O'Hare's workbook, Sentencecraft, which will be provided by the researcher. The control groups during these periods can be doing any other work required in the course of studies except direct instruction in syntax or reading. All students will study other writing-related material such as usage, punctuation, spelling, etc.

Evaluation: Testing should not require more than two periods at each of the testing dates. The writing tests can be treated as regular writing assignments. The Davis Reading Test will require one period to administer.

The writing done in the tri-weekly sessions will be used to study the students' acquisition of these complex syntactic structures.

Time: It is proposed that the pre-test be done at the beginning of January; the post-test at the end of March, just before mid-term break; and the delayed post-test about May 15.

5. Anticipated value to requestor The results of the experiment are to be used for my Doctoral Dissertation. I am also very interested in the curricular application of this new technique for improving student writing.

6. Anticipated value to cooperating organization Because the proposed Junior High Language Arts Curriculum does not prescribe any grammar, it is felt that if the results are positive, this program would warrant serious consideration as a replacement.

7. Estimate of cost (see remuneration guidelines)

8. Suggested personnel, schools and times

School: Spruce Grove Composite High School

Personnel: Miss Adell M. Nyberg

Mrs. Pat Penner

Miss Pam Smith

Ms. Shirley Chmilar

(The teachers listed above have indicated some interest in participating)

For office use only

Approved by_____ Division of Field Experiences

Approved by_____ Date_____

Subject to the following conditions:

(a) A report of the results of findings of this project is required by the cooperating school system(Check one)
yes / no

(b) Other

TO: JR. HIGH ENGLISH TEACHERS TO ENCOURAGE YOU TO PARTICIPATE IN THE STUDY

FROM: Thom MacNeill, Doctoral Student, English Education, U.of A.

RE: PROPOSED EXPERIMENT IN YOUR SCHOOL TO STUDY THE EFFECT OF SENTENCE-COMBINING PRACTICE ON THE DEVELOPMENT OF WRITING AND READING COMPREHENSION ABILITIES

What is sentence-combining?

It is a simple exercise in which the student learns to transform short, choppy sentences, which are often inefficient, into a variety of more complex and efficient structures by manipulation of syntactical elements such as words, phrases, and clauses.

This sentence manipulation is supplementaed by a series of writing exercises(composition writing) in which both experimental and control classes will participate.

Why sentence-combining?

From a practical standpoint, it has at least two pluses:

1. "It is tremendously simple for students to learn, partly because it does not involve the learning of grammatical terminology and partly because it is fun to do. A number of researchers have commented on the fact that "students are impressed by the maturity of the sentences they produce and often claim credit for them. . . (Working on these sentence-combining problems gives students confidence with sentence manipulation) . . . . Seeing sentences "click together", as one student put it, is a positive reinforcement of the sentence-

writing process."(Sentencecraft Teachers' Guide, p.6)

It is also relatively easy for teachers to learn. As O'Hare notes: "Sentence-combining exercises and the signal system that is an integral part of them are difficult to describe but very easy to understand after seeing some examples [cf. last pages of this brochure]. The system is so simple and practical that teachers are comfortable with it after about 30 minutes. . . . Teachers in sentence-combining workshops have estimated that it took them between two and three hours to become expert with the system." (Sentencecraft Teachers' Guide, p.5)

The second plus and the most important is that such programs have consistently resulted in "better" student writing. Students write more mature sentences(see Teachers' Guide p.6 for other research studies) and the quality of the writing(overall impression) as judged by experienced English teachers with no knowledge of the nature of the experiments has been consistently rated superior to that of equivalent non-treatment groups. Such findings led O'Hare to suggest that perhaps form can, in some sense, generate content. "The students used much more detail and seemed to 'see' more clearly. Perhaps the manipulative skill they had developed, because it entailed a wider practical set of syntactic alternatives, invited or attracted detail. Perhaps knowing how does help to create what." (Teacher's Guide, p.7)

The answer to the "Why" question can be summed up in Baretta's comment "Cause it works".

Sentence-Combining and the Curriculum

The new Junior High Language Arts Curriculum Guide does not impose a program of grammar study. If Sentence-Combining works well

in the three Junior High grades, if it significantly increases writing maturity as well as reading comprehension, perhaps school systems could take a close look at adopting some form of sentence-combining program.

But couldn't the time be better spent on traditional grammar?

The answer from research is a resounding no. Traditional grammar has been shown to have little effect on student writing. The main reason seems to be that analysis of sentences does not help students to produce these sentences in their writing. The author of the workbook Sentencecraft did one of the first experiments and had his grade sevens writing at a grade twelve level. Whether it does as much for grade eights and nines has not yet been explored. Also, the effect on comprehension remains unexplored.

If the method works as well as research seems to indicate, it may provide us beleaguered teachers of writing with one practical tool for helping students to improve their writing. For once, research seems to suggest a concrete "how-to" rather than simply pointing out the lack in student writing achievement.

For further information call me at 436-1484, evenings. Leave a message if I'm not in and I'll get back to you as soon as possible.

Appendix B

Pilot Study

A pilot study was carried out to obtain information for the proposed experiment. No control group was used, and the delayed post-test was omitted as was the test for reading comprehension. This pilot involved forty grade nine students in the school. According to the administration, these two classes were average ability grade nines. Two regular classroom teachers carried out the experimental program which lasted for five weeks.

The pilot was designed to provide feedback on the suitability of composition topics, and sentence-combining materials as well as to provide some preliminary information about the effect of the program on such major indices as words per T-unit, words per clause, and clauses per T-unit.

Appendix C

Argumentative Writing Handout

Argumentation: An Evaluation of Grade Nine Writing Ability

Between now and _____* we will practice writing essays in a mode which should be fairly interesting for you. We want to see how good you are as writers, and research shows that Grade 9's write their best when they're arguing. So here's a brief introduction to argumentative writing. (The other "modes", by the way, are description, narration, and exposition).

---

* You will write on your last topic sometime in _____, to give you a final chance to show your ability as Grade Nine "arguers". PLEASE BRING THESE NOTES TO EACH WRITING PERIOD SO THAT YOU CAN REFRESH YOUR MEMORY BY REFERRING TO THEM.

Argumentative Writing: is the kind of writing that attempts to convince a person to do something or to change his attitude. Political speeches, editorials, ads, sermons, debates, and sales talks are some of the many forms of argumentative writing. Arguments are usually a blend of emotional and logical appeals. (The Writer's Handbook, 1975) p.69) All argumentative writing has two elements — a "what" and a "why". The what is called the "thesis"; e.g. "Life is more fun when you're physically fit." The why is the reasons an arguer gives to support his position: e.g. (1) "You have more energy to spend in doing the things you like to do". THIS IS A LOGICAL ARGUMENT, AND EASY

TO PROVE. (2) "Energetic people are more popul[illegible]." THIS IS AN EMOTIONAL APPEAL — Hard to prove.

You, as the arguer, pretend that your audience is hostile to, or undecided about your thesis(position) and aggressively put forward all your arguments in an attempt to convince them that your position is the right or best one. You are also trying to disprove or refute the opposing position by showing the harmful effects of adopting that position or by showing in any way you can that your position is the right or best one.

You can, for example — cite statistics; cite authorities; make comparisons; show the good or bad consequences of adopting a position; or just plain argue for your thesis. For these essays, you can fabricate or make up statistics, authorities, consequences, etc.

REMEMBER: Your essay must be mainly a defence of your position or an attack on the opposite one. You are not writing description, telling a complete story, or explaining how something is done or works.

DIRECTIONS: In this set of argumentative essays you will be [illegible] to write on six topics. Before each essay [illegible] have five minutes to discuss the topic with the four or fi[illegible] students writing on your topic. Please share your ideas free[illegible] We want you to have enough id[illegible] to write a little more than a page on each topic. It is difficult to make a proper evaluation without about 250 words(one page of normal-sized writing). We are more interested in how well you express yourself in writing than we are in seeing ear[illegible] making arguments. However, interesting arguments would be much appreciated.

You will all get to write on each topic before we conclude this argumentative writing section.

ASSUME THAT YOU ARE ARGUING WITH A GROUP OF HIGH SCHOOL STUDENTS. TRY TO CONVINCE THEM OF YOUR POSITION.

In other words, your "audience" is a group of high school students (Grade 9-12).

WRITING TOPICS

These topics were suggested by some of your fellow grade 9 students and by teachers. Two classes have tried them and wrote very well. ***IT IS IMPORTANT THAT YOU WRITE ON THE TOPIC ASSIGNED TO YOU FOR THAT WEEK. Remember that you will get a chance to write on all the topics before the end of this section of the writing program.

You will notice in the chart below that there are six topic sequences (A,B,C, etc). It is very boring to read 20 or 30 essays on the same topic, so to make things a little more interesting for you and us, we have some of you writing on each topic at each evaluation period.

As you can see from the previous page, ARGUMENTATION involves taking a position on a topic and defending it.
YOU CAN TAKE EITHER SIDE OF THE ARGUMENTS AS STATED, OR BOTH SIDES. BUT COMPLETE YOUR ARGUMENTS FOR ONE SIDE BEFORE BEGINNING THE ARGUMENTS FOR THE OTHER.
e.g. PRO: "Education should be compulsory"(i.e. as it is).
or CON: "Education should NOT be compulsory."

****

For your information, the topics are: (TOPIC NUMBERS ARE TO BE USED WITH THE CHART BELOW).

1. Education should be compulsory.
2. I wouldn't want to live in any other country.
3. The world won't survive past the year 2000.
4. Marijuana should be legalized.
5. Advertising is bad for people.
6. Everybody needs friends.

Occasion		1	2	3	4	5	6
Sequence:	A	1	2	3	4	5	6
	B	2	3	4	5	6	1
	C	3	4	5	6	1	2
	D	4	5	6	1	2	3
	E	5	6	1	2	3	4
		6	1	2	3	4	5

(e.g. Students given sequence "B", write on topic 2 on the first writing occasion).

Appendix D

Sentence-Combining Study Schedule + Writing

*Writing Test	Week	Complete to End of Page	Number of Pages
FEB. 13	1	15	12
	2	30	14
	3	44	13
MAR. 6	4	61	15
	5	76	13
MAR. 20	6	87	11
	7	103	15
APR. 10	8	117	14
	9	136	17

**********END OF TREATMENT**********

APR. 24 10 WRITING ASSIGNMENT #5

****

JUNE 14, 15, or 16 LAST WRITING ASSIGNMENT # 6

* GIVEN ON MONDAYS AS LISTED OR NEXT POSSIBLE DAY

Appendix E

Composition Sample

Structures	Counts
/ T-unit marker	22
* noun clause	5
+ adjective clause	3
# adverb clause	7
(S)(in margin) short T-unit(8 words or less)	4
(M)(in margin) medium T-unit(9-20 words)	17
(L)(in margin) long T-unit(21 words or more)	1
..... gerund	1
::::: factive infinitive	2
.•.-.. modal auxiliaries	10
__•__ perfect form	2
.-.-.-. passive	4
,,,,, prepositional phrase as noun modifier	4
= = = = infinitive as noun modifier	1
=.=.=.= present participle as noun modifier	2
.==.==. past participle as noun modifier	1
.,.,., inflected genitive	1
;;;;;;; phrasal genitive	3
_____ adjective	25

Marijuana Should Be Legalized

The legalization of marijuana is a highly controversial
;;;;;;;;;;;;;

(M) topic/. There are many reasons as to# why marijuana should
..-...

(M) remain illegal/ but even more-so there are good reasons as to#

(M) why it should be legalized./
..-.. .-.-.-.-.-.-.-

Firstly,# when you compare marijuana's being illegal to#
.,.,.,.,.,.,.

(M) when liquor was illegal, you come out with the same problems/.

Secondly, the government could make a fortune# if they capita-
.. ..

(M) lized on this drug/. Lastly, the people would have the
..-..

government guarantee+ that the marijuana would be pure and
..-..

(L) not mixed with other dangerous drugs/.
.=.=.

The marijuana problem is similar in many ways to the

(M) illegality of liquor in the 1920's/. Then, the government
,,,,,,,,

(M) found out* that it was impossible to stop all liquor dri[redacted]
::::::

Now, the government knows* that it is too late to suppress all
::::::::::

(M) activities involving marijuana/.
=.=.=.=.=

It is obvious* that the people find very little wrong

(M) with marijuana/. # If they did, the law regarding it would

(M) have been changed years ago/. Today the courts are becoming
crowded with more and more people+ who have been caught with

(M) marijuana on them/. The courts were made for more useful

(M) purposes than small fines/. Why doesn't the government just

(M) give in and make marijuana legal/.

(M) There is a fortune to be made in the legalization of drugs/.

(S) The government should take advantage of it/.

If the government legalized marijuana, the people smoking it would know* that it contained no other harmful

(M) drugs/. This is part of the reason+ that people are so

(M) (S) afraid of marijuana/. No-one knows for sure* what's in it/.

Both the government and the people would benefit from

(M) (S) legalizing this drug/. Let's face it/; the suppression of

(S) this can last no longer/.

Appendix F

Analysis of Covariance: Adjusted Mean Number of Words, Repeated Measures

Occasion	1	2	3	4	5	6
Experimental	215.6	228.4	226.5	221.2	221.4	235.9
Control	210.2	198.7	192.3	200.6	217.9	195.8

Treatment Effect on Number of Words Written, Analysis of Covariance

Source	SS	df	MS	F	p
Mean	229549.456	1	229549.456	21.37	
'A' Main	103089.240	1	103089.240	9.60	0.002
Covariate	1324665.889	1	1324665.889	123.32	
1--Error	1503856.036	140	10741.829		
'B' Main	9599.451	5	1919.89	0.78	0.566
'AxB' interaction	41151.237	5	8230.247	3.33	0.006
2--Error	1741749.563	705	2470.567		

Beta Estimate = 0.702

Subsample Comparison: Treatment Effect on Number of Words Written, Three Repeated Measures

Source	Error Term	Sum of Squares	D.F.	Mean Square	F	p
Mean	S(T)	7999016.8	1	7999016.8	620.10	
T	S(T)	22378.1	1	22378.1	1.73	0.19
O	SO(T)	3955.4	2	1977.7	0.85	0.43
S(T)		748176.1	58	12899.6		
TO	SO(T)	2074.8	2	1037.4	0.45	0.64
SO(T)		269157.8	116	2320.3		

Subsample Comparison: Treatment Effect on Number of T-units Written, Three Repeated Measures

Source	Error Term	Sum of Squares	D.F.	Mean Square	F	p
T	S(T)	182.0	1	182.0	2.64	0.11
O	SO(T)	83.1	2	41.5	1.90	0.15
S(T)		3992.3	58	68.8		
TO	SO(T)	13.2	2	6.6	0.30	0.74

Subsample Comparison: Treatment Effect on Number of Words per T-unit, Three Repeated Measures

Source	Error Term	Sum of Squares	D.F.	Mean Square	F	p
Mean	S(T)	35651.847	1	35651.847	2724.25	
T	S(T)	9.068	1	9.068	0.69	0.41
O	SO(T)	16.095	2	8.048	1.24	0.29
S(T)		759.036	58	13.087		
TO	SO(T)	14.201	2	7.100	1.09	0.34
SO(T)		753.956	116	6.500		

Subsample Comparison: Treatment Effect on Clauses per T-unit, Three Repeated Measures

Source	Error Term	Sum of Squares	D.F.	Mean Square	F	p
			1			
Mean	S(T)	96.629	1	92.629	623.63	
T	S(T)	0.000	1	0.000	0.00	0.97
O	SO(T)	0.015	2	0.008	0.08	0.93
S(T)		8.615	58	0.149		
TO	SO(T)	0.164	2	0.082	0.84	0.44
SO(T)		11.352	116	0.098		

Analysis of Variance for Noun Clauses per T-unit

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	7.381	1	7.381		
2	T	S(T)	0.003	1	0.003	0.10	0.75
3	O	SO(T)	0.050	2	0.025	1.18	0.31
4	S(T)		1.842	58	0.032		
5	TO	SO(T)	0.077	2	0.038	1.81	0.17
6	SO(T)		2.459	116	0.021		

Analysis of Variance for Adjective Clauses per T-unit

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	5.664	1	5.664	228.82	
2	T	S(T)	0.021	1	0.021	0.84	0.36
3	O	SO(T)	0.061	2	0.031	2.29	0.11
4	S(T)		1.436	58	0.025		
5	TO	SO(T)	0.004	2	0.001	0.15	0.86
6	SO(T)		1.543	116	0.013		

Analysis of Variance for Adverb Clauses per T-unit

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	20.268	1	20.268	292.90	
2	T	S(T)	0.026	1	0.026	0.38	0.54
3	O	SO(T)	0.061	2	0.031	0.77	0.47
4	S(T)		4.013	58	0.069		
5	TO	SO(T)	0.166	2	0.083	2.08	0.13
6	SO(T)		4.631	116	0.040		

Analysis of Variance for Percentage of Short T-units

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	9.191	1	9.191		
2	T	S(T)	0.020	1	0.020	0.692	0.40
3	O	SO(T)	0.064	2	0.032	1.297	0.28
4	S(T)		1.679	58	0.029		
5	TO	SO(T)	0.076	2	0.038	1.552	0.22
6	SO(T)		2.841	116	0.024		

Analysis of Variance for Percentage of Medium T-units

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	67.574	1	67.574		
2	T	S(T)	0.002	1	0.002	0.077	0.78
3	O	SO(T)	0.068	2	0.034	1.227	0.30
4	S(T)		1.771	58	0.031		
5	TO	SO(T)	0.050	2	0.025	0.905	0.41
6	SO(T)		3.201	116	0.028		

Analysis of Variance for Percentage of Long T-units

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	4.540	1	4.540		
2	T	S(T)	0.024	1	0.024	0.830	0.37
3	O	SO(T)	0.021	2	0.011	0.827	0.44
4	S(T)		1.707	58	0.029		
5	TO	SO(T)	0.034	2	0.017	1.328	0.27
6	SO(T)		1.480	116	0.013		

Analysis of Variance for Gerunds per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	1.824	1	1.824	115.35	0.00
2	T	S(T)	0.020	1	0.020	1.27	0.26
3	O	SO(T)	0.039	2	0.020	2.10	0.13
4	S(T)		0.917	58	0.016		
5	TO	SO(T)	0.046	2	0.023	2.48	0.09
6	SO(T)		1.079	116	0.009		

Analysis of Variance for Factive Infinitives per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	2.907	1	2.907	173.99	0.00
2	T	S(T)	0.001	1	0.001	0.02	0.90
3	O	SO(T)	0.009	2	0.005	0.31	0.73
4	S(T)		0.969	58	0.017		
5	TO	SO(T)	0.046	2	0.023	1.54	0.22
6	SO(T)		1.740	116	0.015		

Analysis of Variance for Nominalized Verbals per Clause(Gerunds and Factive Infinitives)

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	9.441	1	9.441	311.05	0.00
2	T	S(T)	0.028	1	0.029	0.94	0.34
3	O	SO(T)	0.023	2	0.012	0.49	0.61
4	S(T)		1.760	58	0.030		
5	TO	SO(T)	0.166	2	0.083	3.55	0.03
6	SO(T)		2.718	116	0.023		

Analysis of Variance for Modal Auxiliaries per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Squares	F	Prob.
1	Mean	S(T)	21.441	1	21.441	644.79	0.00
2	T	S(T)	0.024	1	0.024	0.73	0.40
3	O	SO(T)	0.004	2	0.002	0.08	0.92
4	S(T)		1.927	58	0.033		
5	TO	SO(T)	0.009	2	0.005	0.18	0.84
6	SO(T)		3.075	116	0.027		

Analysis of Variance for Perfect Forms per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	0.140	1	0.140	74.91	0.00
2	T	S(T)	0.001	1	0.001	0.34	0.56
3	O	SO(T)	0.001	2	0.001	0.05	0.95
4	S(T)		0.108	58	0.002		
5	TO	SO(T)	0.006	2	0.003	1.85	0.16
6	SO(T)		0.190	116	0.002		

Analysis of Variance for Passives per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	0.741	1	0.741	111.03	0.00
2	T	S(T)	0.007	1	0.007	1.09	0.30
3	O	SO(T)	0.011	2	0.005	0.73	0.49
4	S(T)		0.387	58	0.007		
5	TO	SO(T)	0.001	2	0.000	0.06	0.94
6	SO(T)		0.851	116	0.007		

Analysis of Variance for Perfect Forms, Passives, and Modals Per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	34.396	1	34.396	945.05	0.00
2	T	S(T)	0.009	1	0.009	0.25	0.62
3	O	SO(T)	0.026	2	0.013	0.30	0.74
4	S(T)		2.111	58	0.036		
5	TO	SO(T)	0.001	2	0.001	0.00	0.99
6	SO(T)		4.990	116	0.043		

Analysis of Variance for Prepositional Phrases as Noun Modifiers per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	2.922	1	2.922	201.61	0.00
2	T	S(T)	0.026	1	0.026	1.82	0.18
3	O	SO(T)	0.009	2	0.004	0.60	0.55
4	S(T)		0.841	58	0.015		
5	TO	SO(T)	0.019	2	0.009	1.26	0.29
6	SO(T)		0.853	116	0.007		

Analysis of Variance for Infinitives as Noun Modifiers per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	0.353	1	0.353	111.53	0.00
2	T	S(T)	0.000	1	0.000	0.10	0.76
3	O	SO(T)	0.009	2	0.005	1.30	0.28
4	S(T)		0.184	58	0.003		
5	TO	SO(T)	0.004	2	0.002	0.50	0.61
6	SO(T)		0.411	116	0.004		

Analysis of Variance for Present Participles as Noun Modifiers per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	0.245	1	0.245	43.26	0.00
2	T	S(T)	0.004	1	0.004	0.71	0.40
3	O	SO(T)	0.003	2	0.002	0.54	0.59
4	S(T)		0.329	58	0.006		
5	TO	SO(T)	0.021	2	0.011	3.35	0.04
6	SO(T)		0.368	116	0.003		

Analysis of Variance for Past Participles as Noun Modifiers per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	0.252	1	0.252	68.04	0.00
2	T	S(T)	0.009	1	0.009	2.45	0.12
3	O	SO(T)	0.005	2	0.002	0.67	0.51
4	S(T)		0.215	58	0.004		
5	TO	SO(T)	0.002	2	0.001	0.32	0.73
6	SO(T)		0.394	116	0.003		

Analysis of Variance for Total Verb Forms as Noun Modifiers per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	2.299	1	2.299	156.77	0.00
2	T	S(T)	0.023	1	0.023	1.57	0.22
3	O	SO(T)	0.020	2	0.010	0.95	0.39
4	S(T)		0.850	58	0.015		
5	TO	SO(T)	0.004	2	0.002	0.18	0.84
6	SO(T)		1.189	116	0.010		

Analysis of Variance for Inflected Genitives per Clauses

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	1.321	1	1.321	155.06	0.00
2	T	S(T)	0.013	1	0.013	1.54	0.22
3	O	SO(T)	0.016	2	0.008	1.66	0.20
4	S(T)		0.494	58	0.009		
5	TO	SO(T)	0.002	2	0.001	0.18	0.83
6	SO(T)		0.551	116	0.005		

Analysis of Variance for Phrasal Genitives per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	0.395	1	0.395	58.95	0.00
2	T	S(T)	0.019	1	0.019	2.91	0.09
3	O	SO(T)	0.004	2	0.002	0.41	0.66
4	S(T)		0.388	58	0.007		
5	TO	SO(T)	0.012	2	0.006	1.45	0.24
6	SO(T)		0.491	116	0.004		

Analysis of Variance for Total Genitives per Clause(Inflected and Phrasal)

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	3.160	1	3.160	148.23	0.00
2	T	S(T)	0.065	1	0.065	3.03	0.09
3	O	SO(T)	0.023	2	0.011	1.15	0.32
4	S(T)		1.236	58	0.021		
5	TO	SO(T)	0.020	2	0.019	1.02	0.36
6	SO(T)		1.137	116	0.010		

Analysis of Variance for Adjectives per Clause

Source		Error Term	Sum of Squares	D.F.	Mean Square	F	Prob.
1	Mean	S(T)	44.077	1	44.077	631.07	0.00
2	T	S(T)	0.568	1	0.568	8.13	0.00
3	O	SO(T)	0.134	2	0.067	1.42	0.25
4	S(T)		4.051	58	0.070		
5	TO	SO(T)	0.091	2	0.045	0.96	0.38
6	SO(T)		5.459	116	0.047		

Appendix G

Additional Correspondence

511 Michener Park
Edmonton, Alberta
T6H 4M5
June 23, 1977

Ginn and Company
Educational Publishers
3771 Victoria Park Avenue
Scarborough, Ontario
M1W 2P9

Dear Sir:

Enclosed herein is a proposal for my Doctoral study to be carried out in the Edmonton Separate School System. I have recently received permission from the Board and have met with the Principal and English teachers at St. Cecilia Junior High School. They are willing to undertake the study and have assured me of their full cooperation. The study is now in the final planning stages. It is slated to run from October to December, 1977.

The request which I make of you is this. I require permission to duplicate 100 copies of the sentence-combining exercises from Sentencecraft: An Elective Course in Writing by Frank O'Hare. The eventual benefit to your company through increased sales, is fairly

certain. There have been few experiments in such activities which have not resulted in a two to four year growth in syntactic maturity. The newly-revised Alberta Junior High Language Arts Curriculum does not impose any grammar study. It seems to me that if this program works in a middle-class Alberta school, the number of teachers and Boards who will replace traditional grammar study with this sentence-combining program would more than offset the loss in revenue from the 100 workbooks which would be required for the experiment.

The main reason for this request is that I do not want the "Writing Workshops" to be integrated with the sentence-combining activities. Both control and experimental classes deliberately practicing their sentence-combining skills in the workshops BECAUSE they can see that it is an integral part of the workbook program. By keeping both the controls and the experimentals ignorant of the direct tie-in with sentence-combining practice, it is expected that they will write more naturally and that this writing will make it possible to study and compare the rate and sequence of student acquisition of various syntactical forms.

In summary then, I believe that the study outlined in the attached document could result in increased Sentencecraft sales and also enable me to make a contribution to the scientific study of this promising writing activity. I therefore request permission to duplicate the exercises as mentioned above.

Please advise me as soon as possible.

Yours truly,

Thomas B. MacNeill

Ginn and Company

191 Spring Street
Lexington, Massachusetts 02173
Telephone: (617) 861-1670
Cable: EGINN

XOX

July 12, 1977

Mr. Thomas B. MacNeill
511 Michener Park
Edmonton, Alberta T6H 4M5
Canada

Dear Mr. MacNeill:

Your letter of June 23 to our Canadian office has been forwarded to us here in the United States, as publisher of O'Hare: SENTENCECRAFT.

For permission to use the sentences on pages 4-9 of our book, which are from "The Eighty-Yard Run" by Irwin Shaw from his book SELECTED SHORT STORIES, copyright 1955, 1961 by Irwin Shaw, it will be necessary for you to write to the publisher Random House, Inc. (201 East 50th Street, New York, New York 10022).

You may have permission to duplicate the remaining sentence-combining exercises in our book for use as described in your letter, with the following provisions:

(1) that you print the following acknowledgment on each copy:

From SENTENCECRAFT (An Elective Course in Writing), by Frank O'Hare, © Copyright, 1975, by Ginn and Company (Xerox Corporation). Used with permission.

(2) that you make only 100 copies and that after the experiment is concluded they be destroyed

(3) that you send us a copy of the results of your study.

If you ever intend to publish any of our material, it will be necessary to reapply for permission to use it.

Sincerely yours,

Darwin M. Newton

Darwin M. Newton
Vice President and
Executive Managing Editor

DMN:JM

511 Michener Park
Edmonton, Alberta
T6H 4M5
October 2, 1977

Random House Inc.
201 East 50th Street
New York, N.Y.
10022

Dear Sir:

Enclosed is a copy of a letter I sent to Ginn and Company requesting permission to duplicate their workbook Sentencecraft by Frank O'Hare.(Copyright, 1975) The request was made to enable me to separate the sentence-combining exercises from the writing exercises of the study.

For permission to duplicate pp. 4-9 of Sentencecraft(Based on "The Eighty-Yard Run" by Irwin Shaw; Selected Short Stories c 1955) I was advised to contact your company. This request is the purpose for this letter. I need to duplicate one hundred copies and these will be destroyed upon conclusion of the study. The study is to begin January 1, 1978 not as indicated in the letter to Ginn.

I thank you for your attention to this matter and await your reply.

Yours truly,

Thomas B. MacNeill

RANDOM HOUSE, INC. ALFRED A. KNOPF, INC.

201 EAST 50TH STREET, NEW YORK, N.Y. 10022 · (212) 751-2600

October 7, 1977

Mr. Thomas B. MacNeill
511 Michener Park
Edmonton, Alberta
T6H 4M5
Canada

Dear Mr. MacNeill:

Thank you for your letter of October 2nd.

We are pleased to grant you permission for use of the specified material from "The Eighty-Yard Run," from SELECTED SHORT STORIES OF IRWIN SHAW, as it appeared in SENTENCECRAFT: An Elective Course in Writing, by Frank O'Hare (published by Ginn and Company), in 100 copies for your Doctoral study, provided that the copies be destroyed upon conclusion of the study and the following acknowledgment appears on each copy:

Copyright 1940 and renewed 1968 by Irwin Shaw.
Reprinted from SELECTED SHORT STORIES OF IRWIN SHAW,
by permission of Random House, Inc.

Sincerely yours,

Debra Maltzman

Debra Maltzman
Permissions Department

VITA

NAME: Thomas Bernard MacNeill

PLACE OF BIRTH: Simpson, Saskatchewan

YEAR OF BIRTH: 1942

POST-SECONDARY EDUCATION AND DEGREES:

University of Windsor
Windsor, Ontario
1961-2, 1963-5 B.A.

University of Saskatchewan
Saskatoon, Saskatchewan
1966-1968 B.Ed.

University of Alberta
Edmonton, Alberta
1975-1977 M.Ed.

RELATED WORK EXPERIENCE

Teaching Assistant
University of Alberta
1976-78

Assistant Professor
University of Saskatchewan
1978-1980

Assistant Professor
Brandon University
1980-Present

PUBLICATIONS:

(1) "Poetry and Fresh Grass for Adolescents", Skylark, 15: 3: 1-4

(2) "The Blight of Term Appointments", CSSE News, 8: 5: 3-4