VOCABULARY LOAD OF BEGINNING READERS AUTHORIZED FOR
BRITISH COLUMBIA SCHOOLS, 1872 TO 1977

by
Joyce Isobel Patterson

A THESIS SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
in the Faculty
of
GRADUATE STUDIES
Department of Reading Education

We accept this thesis as conforming
to the required standard

© JOYCE ISOBEL PATTERTON, 1977
UNIVERSITY OF BRITISH COLUMBIA
August 1977
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Reading Education

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date 27 AUGUST 1977
ABSTRACT

Identifies and compares three dimensions of the vocabulary load of the nine readers and reader series prescribed for beginning reading in the Province of British Columbia since 1872. The dimensions explored were: number of running words; number of new words introduced; and average number and range of repetitions. Periods of prescription were designated by the first year of prescription, namely: 1872, 1884, 1901, 1915, 1923, 1935, 1948, 1964, and 1968. (The 1964 and 1968 materials are currently both authorized by the Ministry of Education.) Tabulations were made for each dimension in each time period. The number of running words in each time period was 3,495, 1,468, 2,540, 2,021, 7,252, 12,766, 20,577, 23,177, and 7,531 respectively. The number of new words introduced in each time period was 392, 367, 405, 447, 633, 601, 341, 500, and 860 in each case. The average number of repetitions found in a sample of the first 500 words in each time period was, respectively: 6.1, 5.3, 3.3, 3.0, 5.8, 6.3, 22.7, 14.2, and 4.6. The range of repetitions in the same 500-word sample was as follows in each case: 1-43, 1-75, 1-34, 1-31, 1-55, 1-38, 5-60, 1-54, and 1-38. The method did not account for repetitions of words having similar patterns, an aspect of the vocabulary system observed in the 1872, 1923, and 1968 materials. Concludes that the differences between the early time periods and the recent time periods of 1935, 1948, and 1964 are related to the use of vocabulary control which provides for the introduction of a limited number of new words and their systematic repetition. Recommends (1) investigation of methods of determining vocabulary load of beginning readers that gives weight to word patterns; (2) content analysis of related teachers manuals and other documents to ascertain teaching methods; and (3) content analysis of the readers
to determine attitudes expressed. Appendix describes procedures in identifying and verifying the periods of prescription and the prescribed readers.
ACKNOWLEDGEMENTS

The writer wishes to acknowledge the very generous assistance given by Dr. H. M. Covell, chairman of the supervisory committee, whose guidance was most helpful at all stages in the development of this thesis. Appreciation is also expressed to Dr. Jane Catterson for her constructive criticism at various times during the preparation of the text.

Also, the writer is indebted to Dr. E. Summers, whose interest and help were a significant contribution to this study.

Grateful acknowledgement is made to those who made the readers available: to the Special Collections Division, University of British Columbia, for the texts on the authorized lists of 1872, 1884, 1901, 1915, and 1935; to Dr. H. M. Covell for the text of 1923 (now in the Reading Resources Centre, Faculty of Education, University of British Columbia); and to Miss Dorothy Sharrock of the Reading Resources Centre for the materials prescribed in 1948, 1964, and 1968.
# TABLE OF CONTENTS

## CHAPTER  

<table>
<thead>
<tr>
<th>I.  THE PROBLEM .................................................</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction ...................................................................</td>
<td>1</td>
</tr>
<tr>
<td>Purpose of the Study ..............................................</td>
<td>2</td>
</tr>
<tr>
<td>Source Material ....................................................</td>
<td>2</td>
</tr>
<tr>
<td>Definition of Terms ................................................</td>
<td>2</td>
</tr>
<tr>
<td>The Prescribed Beginning Readers and Reader Series ........</td>
<td>3</td>
</tr>
<tr>
<td>Limitations of the Study .........................................</td>
<td>4</td>
</tr>
<tr>
<td>SUMMARY OF CHAPTER ................................................</td>
<td>4</td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE ................................</td>
<td>6</td>
</tr>
<tr>
<td>CONTENT ANALYSIS AS A RESEARCH TECHNIQUE .....................</td>
<td>6</td>
</tr>
<tr>
<td>VOCABULARY STUDIES OF BEGINNING READERS .......................</td>
<td>7</td>
</tr>
<tr>
<td>Word Lists ....................................................................</td>
<td>7</td>
</tr>
<tr>
<td>Vocabulary Control ..................................................</td>
<td>7</td>
</tr>
<tr>
<td>Definition ....................................................................</td>
<td>7</td>
</tr>
<tr>
<td>Reasons for vocabulary control ..................................</td>
<td>8</td>
</tr>
<tr>
<td>Historical survey of vocabulary control .......................</td>
<td>9</td>
</tr>
<tr>
<td>Pros and cons of vocabulary control ............................</td>
<td>14</td>
</tr>
<tr>
<td>SUMMARY OF CHAPTER ................................................</td>
<td>15</td>
</tr>
<tr>
<td>III. METHODOLOGY .......................................................</td>
<td>16</td>
</tr>
<tr>
<td>IV. RESULTS ............................................................</td>
<td>18</td>
</tr>
<tr>
<td>THE TOTAL SAMPLES .....................................................</td>
<td>18</td>
</tr>
<tr>
<td>Number of Running Words ..........................................</td>
<td>18</td>
</tr>
<tr>
<td>Number of New Words ...............................................</td>
<td>18</td>
</tr>
<tr>
<td>Comparison of Running Words and New Words ..................</td>
<td>21</td>
</tr>
<tr>
<td>CONTENT</td>
<td>PAGE</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>THE 500-WORD SAMPLES</td>
<td>21</td>
</tr>
<tr>
<td>Average Number and Range of Repetitions</td>
<td>21</td>
</tr>
<tr>
<td>Distribution of Word Repetitions</td>
<td>26</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>26</td>
</tr>
<tr>
<td>Number of Running Words</td>
<td>26</td>
</tr>
<tr>
<td>Number of New Words</td>
<td>27</td>
</tr>
<tr>
<td>Repetitions</td>
<td>29</td>
</tr>
<tr>
<td>SUMMARY OF CHAPTER</td>
<td>30</td>
</tr>
<tr>
<td>V. SUMMARY, CONCLUSIONS, RECOMMENDATIONS</td>
<td>32</td>
</tr>
<tr>
<td>SUMMARY</td>
<td>32</td>
</tr>
<tr>
<td>Related Literature</td>
<td>32</td>
</tr>
<tr>
<td>Methodology</td>
<td>34</td>
</tr>
<tr>
<td>Findings</td>
<td>35</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>36</td>
</tr>
<tr>
<td>FUTURE RESEARCH</td>
<td>37</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>39</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>43</td>
</tr>
</tbody>
</table>
I. THE PROBLEM

Introduction

Nine different readers and reader series have been prescribed for beginning reading in British Columbia since the inception of public schooling in the Province in 1872. Since new texts have replaced the old at regular intervals, this body of material lends itself to a study of changes in beginning readers over the past one hundred years. For this reason, and because the texts had not previously been brought together and described, a search was made for a technique which would serve both to describe the readers and to identify changes in the materials over time.

The search revealed many applications of systematic content analysis procedures in textbook analysis, readability studies, and historical descriptions of readers. This process, which usually results in quantitative information, permits valid inferences about relationships between the data.

It was also found that there has been a marked interest in this century in the vocabulary load of children's readers, specifically, in the number of new words introduced and the number of repetitions of words. This has led to many vocabulary studies of beginning reading texts for evaluative purposes. Moreover, because words are a constant factor across texts of different formats, they have also provided a useful unit for the description and comparative analysis of readers in different time periods.

Since content analysis of all possible components of the prescribed beginning reading materials was too broad a task for an exploratory study, it was decided to limit the preliminary description to an analysis of the vocabulary load of the readers in each period of prescription, and to identify changes in the readers by comparing the findings for the different time periods.
Purpose of the Study

The purpose of this study was to describe and compare the vocabulary load of the nine beginning readers and reader series prescribed by the Province of British Columbia from 1872 to the present. Three dimensions of vocabulary load were explored:

1. Number of running words.
2. Number of new words introduced.
3. Average number and range of word repetitions among the first 500 running words.

Source Material

Primary sources for the study consisted of: the reading materials prescribed for Grade I (or equivalent prior to 1923) in the Province of British Columbia; and publications of the Province of British Columbia which were used to verify dates of prescription and ascertain titles of the reading materials, i.e., Annual Reports of the Public Schools, Programmes of Studies for the Elementary Schools, and Division of Curriculum materials. Reference was also made to the teacher's manuals which accompanied the prescribed readers in recent periods of prescription.

Secondary materials comprised books, articles, and unpublished papers related to: the history of school readers and reading instruction; vocabulary studies of readers; and content analysis theory, procedures, and studies of children's reading materials.

Definition of Terms

Reader - A book for learning and practising reading.

Reader series - In this study, a collective term to describe beginning reading materials in periods of prescription which included more than one beginning reader: for example, the Primer and First Reader authorized in 1935.

Beginning reader - In this study, a general term for any reader intended for
Prescribed or authorized texts - Textbooks stipulated for use in the public schools of the Province and provided to students under various giving and lending systems over the years. As is the case of the readers of this study, the texts are not generally produced by the Department (now Ministry) of Education, but are chosen from materials written by authorities in the field and available from textbook publishing firms.

Time period - In this study, a term referring to the period of time a reader or reader series is prescribed; for convenience, this period is identified by the first year of prescription.

Vocabulary, controlled - A system of introducing and repeating a restricted number of words in a reader series, and providing for the gradual increase in the number and difficulty of words.

New word - A word not previously introduced in a given reader or reader series. In this study, and following the usual practice, variant forms are not new words when s or 's is added to or dropped from known nouns and when s, ed, d and ing are added to known verbs; nor is a compound a new word when formed from two known words.

Repetitions - In this study, the number of times a word is repeated in the first 500 running words of a reader, or reader series where these occur. More generally, this term refers to the number of times a given word is repeated in a reader.

The Prescribed Beginning Readers and Reader Series

The first reader on the 1872 list of authorized texts for British Columbia schools and the first reader on the new list of 1884 were considered to be the beginning readers prior to 1901.

From 1901 to 1923 the organization of the elementary schools consisted of
three grades, Junior, Intermediate, and Senior. The Junior grade represented the first three years of school. Therefore the First Primer in the three readers prescribed for the Junior grade in 1901, and the Beginner's Reader in the new series prescribed in 1915 were considered to be the beginning or first-year readers for this study.

The present elementary school grading system began in 1923 with a division into Grades I to VIII. Beginning in 1923, therefore, the readers prescribed for Grade I were the materials of study.

Procedures and sources used to identify and verify the periods of prescription and the prescribed readers are described in Appendix A. The nine readers and reader series of the study are listed chronologically by first year of prescription in the Bibliography.

Limitations of the Study

The study will provide a numerical description of the vocabulary load of the beginning reading texts in each period of prescription. The analysis will not take into account factors such as size of type, sentence length, illustrations, word difficulty, or subject matter of the materials. Nor will the study attempt to analyze the larger social, economic, and technological changes which have affected education, and therefore children's readers, in British Columbia during the period of study.

SUMMARY OF CHAPTER

The texts authorized for beginning reading in British Columbia schools since 1872 provide a body of material suited to an examination of changes which have occurred in first-year readers over the past one hundred years. As these materials have not hitherto been described, a search was made for a research technique which would provide basic descriptive data and probe aspects of change in the readers.
This investigation revealed that concern for the vocabulary load of children's readers in the present century has produced many word counts of beginning reading materials. It was also observed that this form of analysis provides numerical data that can be used to describe and compare readers of different historical periods. Consequently, it was decided to tabulate the vocabulary load of the readers in each period of prescription and compare the resulting data in order to identify changes which may have taken place in this area of the texts over the years.
II. REVIEW OF RELATED LITERATURE

Two areas of the literature related to the present study are: content analysis as a research technique; and vocabulary studies of beginning readers.

CONTENT ANALYSIS AS A RESEARCH TECHNIQUE

Content analysis is a research technique for the objective, systematic, numerical description of printed or other human communication. The method is based on the premise that what is said, written, pictured, or otherwise communicated can be analyzed into units that can be quantified or counted. Thus, within a sample of material, when all the occurrences of a unit of one kind have been recorded and counted, it is possible to make a numerical comparison with another unit of a different kind which has been similarly tabulated in the material. Comparisons can also be made with other material that has been analyzed in the same way, or with other data expressed in numerical terms.

A standard reference on the theory and use of content analysis is by Berelson (1952). Holsti (1969) describes more recent theoretical concepts and applications in the field. Requirements of the procedure are: precise specification of the content universe or sample of material to be analyzed; explanation of the unit of analysis (e.g., word, title, page, story); operational definition of the categories for analysis; and systematic coding of all the relevant occurrences of the categories in the sample. Validity is dependent upon the relevance of the categories to the intent of the analysis; in modern content analysis, when inferences are made about the sources or receivers of the communication it is usually assumed that the content analysis data will be compared with an independent or noncontent index of the attributes inferred. Reliability requires consistency in coding occurrences of the categories; when judgements are involved, inter-coder reliability should be measured and reported.
VOCABULARY STUDIES OF BEGINNING READERS

Word Lists

In preparing basal readers, authors have attempted to make use of words of high functional value, as revealed by word counts of reading content and children's language. To establish graded lists of the most frequently used words in reading vocabularies, word counts were made by Thorndike and many others. Thorndike (1921) identified the 20,000 words occurring most frequently in the reading vocabularies of children's literature, the Bible and English classics, elementary-school textbooks, tradebooks, daily newspapers, and correspondence. Forty-one different sources were used and the sample comprised 4,565,000 words, of which 3,000,000 were from the Bible and English classics.

A Reading Vocabulary for the Primary Grades was constructed by Gates (1926) according to standards of "utility for childhood", interest, and difficulty. The selection criteria was applied to the 2500 words of highest frequency in the Thorndike list and to an additional 281 words from a selection of literature for young children, a series of readers for the primary grades, and the Horn list of 1,003 words used most frequently by kindergarten children. In 1936 Dolch prepared a basic sight vocabulary consisting of 220 words, excepting nouns, common to three word lists (Dolch, 1936).

Vocabulary Control

Definition. As defined by Gray (1960a), vocabulary control in basic reading programmes is "a planned, sequential pattern of introducing and maintaining vocabulary" which helps to ensure mastery of a stock of sight words (words instantaneously recognized) and, later, of new words that pupils analyze independently or look up in the dictionary (p. 18). The procedure is to introduce one or a very small number of new words on a page; once a word has been introduced it is repeated
at spaced intervals throughout the book and the word is also repeated and
maintained in succeeding books in the series.

Schonell (1959) discusses the vocabulary of the Happy Venture Reading Scheme
under the following headings: everyday words used; meaningful words of different
visual patterns; controlled vocabulary; confusable words kept apart; adequate
repetition of new words; and sentence form helps word recognition. In the section
headed "Controlled Vocabulary", the following explanation is given:

In order to provide success and so sustain interest, the grading of words
used in each book of the series is controlled. It is now a well-established
teaching principle that the amount and the difficulty of material in any
school subject should be so graded as to suit the child's age and learning
capacity. Now this is just what the Happy Venture Readers do by controlling
the vocabulary, that is, by restricting the number of words used in each
book. (pp. 16-17)

Under "Adequate Repetition of New Words", Schonell again refers to controlled
vocabulary:

Now, intimately related to grading of words is their adequate repetition.
It is not sufficient merely to control the number of words per page, they
must also be repeated on many pages, embodied in familiar phrases, and in a
variety of contexts. It is word control plus repetition that makes for
effective word recognition. Thus in Introductory Book there are only 44
new words and these are used in a total of 553 running words covering 27
pages, so that on an average a word is repeated about 12 times throughout
the book. (p. 21)

Reasons for vocabulary control. Controlling the vocabulary burden of early
reading texts was a response to the "heavy" and unregulated word loads occurring
in readers, and avoids the disabling effects of discouragement and word-by-word
reading (Schonell, 1951; Spache and Spache, 1977). As pointed out below, its
development is related to vocabulary studies of reading materials in the 1920s
(Gray, 1925); this was also a time when seriously high failure rates in the
primary grades were reported (McIntosh and others, 1960).

The advantages of the principle of vocabulary control according to Schonell
(1951) are: gradation ensures a good initial attitude to reading by giving every
child a chance to make progress, and allows brighter children to get on more
quickly with new books; repetition permits maximum use of learning through discrimination of visual patterns of words; lists of the new words that appear on each page permit teachers to make preparatory and revision work, while the gradation of vocabulary in a series allows the use of books which suit particular groups; carrying on the vocabulary from book to book maintains the child's acquaintance with the early reading material.

**Historical survey of vocabulary control.** The vocabulary of American school readers from colonial times was surveyed by Smith (1965). Before 1776 the rate of introduction of new words in the first five pages of primers ranged from 20 to 100 new words per page, proceeding from one-syllable words to two- and so on up to words of five and six syllables.

In the succeeding period, 1776-1840, spelling books as well as readers were used for reading instruction, and Noah Webster's *Blue-back Speller*, first published in 1790, had a range of 86 to 197 new words and syllables per page on the first ten pages. Smith describes the organization of the reader as follows:

The first twenty-five pages of the book were given over to rules and instructions. Page 26, the first which the child was supposed to read, contained the alphabet, syllables, and consonant combinations. The second page for the child to read contained 197 syllables. The succeeding several pages were devoted to lists of words arranged in order by their number of syllables, and further organized into lists according to the similarity of phonetic elements. (p. 46)

Provision for repetition and a striking decrease in the number of new words per page occurred in primers of 1840-1890, *McGuffey's First Reader* being notable for introducing only 10 to 12 new words per page with no new words in review lessons. The first page of this reader was given over to the alphabet in large and small letters; then came the picture alphabet, which occupied three pages; and then Lesson I based on two-letter words which were presented in isolation for spelling and in sentences for reading. By Lesson LXV, words of three and four letters were being used. Referring to the vocabulary of the stories, Smith found
that the sentences were usually subservient to the phonetic elements which McGuffey selected for drill purposes.

A reader published at the end of the period 1890-1910, which, according to Smith, was a time of emphasis on reading as a cultural asset, was found to have a range of 4 to 6 new words per page in the first 10 pages, with abundant repetition provided in the cumulative features of the folktale content. However, the initial ten pages in the first book of a silent reader series published at the end of the period 1910-1925 contained a range of 2 to 30 new words per page.

Smith reported that during the period 1925-1935 vocabulary load in reading materials was a topic of interest in research, standard word lists were used as a basis for selecting vocabulary, and the total vocabulary of primers averaged 269 new words, with a trend toward higher repetitions of words.

Beginning readers continued to show vocabulary reduction in the years 1935-1950: in six basal series published between 1940 and 1950 examined by Smith, the average combined vocabulary of primers and pre-primers was 174 words. Authors continued to check word selection for early readers against "scientifically determined vocabulary lists", all the words in the pre-primers and primers were repeated in the first reader, and there was a higher number and careful account of repetitions.

In the last period examined by Smith, 1950-1965, there was evidence of a further decrease in vocabulary: the pre-primer programme had an average of 59 new words and the primer 89, or a total of 148. Smith reported that authors continued to check vocabulary against word lists in these years, except for one group who said such a practice has many fallacies (Smith, 1965, p. 328).

Smith reviewed the sequential development of applying linguistic theory to the teaching of reading which occurred during the last period of her study, and presented the content from pages of two sets of reading materials for children based on linguistic principles. The first publication for children's use
appeared in 1961 and was titled *Let's Read*. Approximately 5000 words were introduced in the book as a whole. The learning sequence proceeded from the alphabet to words having similar patterns, followed in each case with sentences containing these words. The emphasis was on symbol-sound correspondences, as in *a bag, a rag, a rag bag*. The concern of these authors, said Smith, "is not phonics, but rather word patterns" (p. 388). *A Basic Reading Series Developed Upon Linguistic Principles* published in 1965 was "the first graded series of basal readers prepared for the express purpose of applying linguistic theory" (p. 389); a beginning reading programme similar to the earlier *Let's Read* was followed.

Edmund B. Huey's study of reading, *The Psychology and Pedagogy of Reading* (Huey, 1908), gives an indication of approaches to the vocabulary of children's readers early in the present century. Huey quoted a Professor Burke's preface to a reader written by children in California. This preface makes reference to vocabulary load in the concluding sentences:

"... Much care has been used to keep the stories within a limited vocabulary. Less than 750 different words are used in the entire series (two volumes), and these, excepting the necessary geographical names are all of the commonest use among children." (Burke quoted by Huey, 1908, p. 540)

Otherwise, Huey's references to vocabulary were concerned with the use that should be made of context in recognizing unknown words. For example, he quoted from a description of beginning reading lessons at the Horace Mann School, Teachers College, Columbia University, where "the vocabulary is 'not limited to a very few words,' the pupil gaining many words from the context" (p. 291). Describing his own method, Huey said that the child "enlarges his vocabulary for himself by the use of the context" (p. 333), and that the "most natural and real meanings (of words) dawn upon the reader as he feels the part that is left for them to take in the various contexts in which they occur" (p. 348).
Studies made in the early 1920s by United States investigators of the extent of the vocabulary in different sets of readers revealed such a surprisingly large vocabulary in the readers for each grade that a demand developed for a so-called "basic" vocabulary for mastery at each level (Gray, 1925). In 1922 Selke and Selke tabulated the frequencies of the words used in 12 beginning readers and found a total of 1,636 different words, with a range of from 157 to 630 words in individual books; they also found a very limited frequency of a large number of words within books and between books (Gray, 1925). A 680-word mastery vocabulary which had been developed for the three primary grades was criticized as being too small, a point justified by a study of textbooks for the first three grades prescribed by Oregon State where a total of 5,190 different words was found to occur in the primers, readers, spellers, and arithmetic, with 30 per cent of the words appearing only once (Gray, 1925). Commenting on these and similar findings, Gray said that problems related to "the control of vocabulary difficulties in the organization of reading material" and the development of "methods of vocabulary mastery" were in urgent need of careful study (p. 190).

In 1925 Gates and others counted the number of different words in pre-primer, primer, and other first-grade materials of 21 systems of beginning reading and found a range of from 90 or less to 900 words in the various systems; for these investigators, the findings indicated a need for standards of selection in beginning reading vocabularies (Gray, 1926).

E. W. Dolch compared modern readers with the historical McGuffey Readers of the same grade level (Dolch, 1945). He found that the modern primers and first readers had about 2750 more running words than the McGuffey materials and that modern second and third readers contained about 10,500 and 13,000 more in each case. Both had the same percentage of difficult words, but the McGuffey Readers introduced them a year earlier. The principles underlying the contemporary organization of vocabulary in readers were expressed by Gray in a comment on Dolch's
study: "Obviously modern readers provide an easier introduction to reading, much more reading material, and more frequent contact with specific words" (Gray, 1960b, p. 1120).

The Happy Venture Readers were described by their author, Fred J. Schonell, as the "first vocabulary-controlled reading books specifically prepared for English-speaking children and made available in British Commonwealth countries" (Schonell, 1959, p. v). An account of the theoretical and research background of the scheme is to be found in The Psychology and Teaching of Reading (Schonell, 1951). The first-year vocabulary of the series was 209 words, but the author stated that the reading vocabulary of "the majority of pupils at the end of first year will vary from 200 words up to 380 or even 500 or more words, according to the amount of phonic family work and supplementary reading they have done" (Schonell, 1959, p. 66). As evidence of other current vocabulary content, Schonell reported word counts made by the Scottish Council for Research in Education of four different Infant Series: these showed a range of 466 to 1022 words in the first two (or three) books (Schonell, 1951).

In 1962 the authors of a Canadian primary reading programme (McIntosh (Ed.) and others, 1969) compared the conservatism of the 250-vocabulary of the most widely used United States (and Canadian) first-grade readers with findings of a 1950 study of seven series commonly used in Scotland. The average vocabulary of the Scottish readers was 520, the easiest series introducing 370 new words, the most difficult 728. These authors also referred to a second 1950 study which showed that the average Scottish child, who begins school at age five, one year earlier than North American children, made more than a year's progress in the first grade in terms of the Metropolitan Reading Test. The evidence of these Scottish studies was a factor in the rationale for the "more demanding programme" based upon 503 words provided by McIntosh and his associates. (This series is one of the prescribed reader series in the present study, i.e., the 1964 series.)
A survey of vocabulary in American basal readers by Spache and Spache (1977) showed a trend toward greater repetition and smaller vocabularies at every level from 1930 to 1965. In findings similar to those reported by Smith (1965), they found that the average vocabulary in primers in 1930-1931 was 304, but by 1965 it was an estimated 113 to 173 words. These authors point out that a reversal in the trend toward smaller vocabularies in American first-grade readers is suggested by a 1974 study by Leo V. Rodenborn and Earlene Washburn: based on four leading American series published between 1960 and 1970, first-grade vocabulary was found to range from 305 to 675 words.

Pros and cons of vocabulary control. In a recent discussion of arguments both for and against vocabulary control, Spache and Spache (1977) gave the following reasons why basal vocabulary control is not justified: standards for number of repetitions and rate of introducing new words are not precisely known; children spontaneously learn many words other than the basal vocabulary; except for a few hundred service words which occur frequently in practically all reading materials, there is hardly any such entity as a core vocabulary which overlaps various basal series; studies by Gates in 1961 and 1962 showed that even the poorest readers in Grades two and three recognize a large proportion of the words in the next grade-level of their basal series; the need for repetition varies from child to child; repetition does not ensure learning, for errors in words introduced in primary grades persist into college level.

In support of less rigorous vocabulary control these authors also point out that today's children are more verbal, more widely travelled, and have larger vocabularies as a result of televiewing experiences (Spache and Spache, 1977). Similar factors resulting from changes in children's environment were cited by McIntosh and associates (1962) to justify a reduced readiness or pre-reading period in their first-year programme; however, an additional influence in children's lives
mentioned by these authorities was the "smooth upward gradient" provided by beginning reading programmes.

SUMMARY OF CHAPTER

The procedure used in vocabulary studies is content analysis, a research technique for quantifying and comparing the material of human communication. Analysis of reading vocabularies in the form of word counts began in the United States in the 1920s for the purpose of creating functional word lists and evaluating the difficulty of reading materials. Word counts were also made of children's speaking vocabularies.

Findings from these and similar studies underlie the method of controlled vocabulary that has been generally used in North American basal readers since that time. A similar system of controlling the vocabulary burden of early readers was developed in the United Kingdom.

Vocabulary studies of beginning readers indicate that different systems of presenting words have existed in different historical periods. For many years word introduction was related to the graded presentation of words according to number of syllables; then this method was combined with attention to phonic similarities. A prototype of modern vocabulary control may be seen in the McGuffey Reader of the mid-nineteenth century where a limited number of new words graded in difficulty was introduced and provision was made for the repetition of words. There is evidence that early in the present century consideration was given to the use of context in deciding on the meaning of words and to the repetitive nature of folktales in the organization of beginning reading vocabularies.

In recent years, the linguistic method of introducing words based on word patterns has been used in some basal series. A principle common to all these systems is the graded presentation of words or word parts, proceeding from some standard of least to more difficult.
III. METHODOLOGY

The primary sources for analysis were the nine readers and reader series prescribed for Grade I (or equivalent prior to 1923) in British Columbia schools since 1872. The time periods of prescription or authorization were identified by the first year of authorization by the British Columbia Department of Education, viz., 1872, 1884, 1901, 1915, 1923, 1935, 1948, 1964, and 1968. The 1964 and 1968 series are currently both authorized by what is now called the Ministry of Education.

Tabulations were made to determine:

1. Number of running words in the reader(s) in each time period.
2. Number of new words introduced in the reader(s) in each time period.
3. Average number and range of word repetitions among the first 500 words of the reader or reader series in each time period.

The sample for the first and second tabulations comprised all the material in the texts with the exception of: words presented in isolation; exercises, e.g., fill in the blanks; and poems provided for enrichment. This sample was called the Total Sample in each case; thus there were nine Total Samples, one for each time period.

The sample for the third tabulation was limited to the first 500 words of each Total Sample. This sample was called the 500-Word Sample in each case; thus there were nine 500-Word Samples, one for each time period.

The texts were analyzed three times; in cases of series in the Total Sample, separate tabulations were made for each text in the series. In the first analysis, the running words in each Total Sample were counted and totalled. In the second analysis, the new words introduced in each Total Sample were listed and their numbers totalled; in cases of series, the count for each book excluded words.
previously used in the same series. Finally, for each 500-Word Sample, the frequency of occurrence of each word introduced was recorded; using these data, the average number and range of word repetitions were computed.
IV. RESULTS

THE TOTAL SAMPLES

Number of Running Words

The number of running words in each reader and the total for each time period are shown in Table 1. The beginning readers for 1872, 1884, 1901, and 1915, which have 3,495, 1,468, 2,540, and 2,021 running words respectively, provide relatively small amounts of reading material in comparison with totals for subsequent time periods. Moreover, among these early texts, that for 1872 has more than twice as many words as the 1884 reader and approximately one-third as many as the 1901 and 1915 readers.

With 7,252 running words, the 1923 period shows a marked increase in amount of reading material over the early time periods. A further increase to a total of 12,766 running words occurs in the 1935 material, and the trend towards larger amounts continues in the 1948 and 1964 series, which have totals of 20,577 and 23,177 running words respectively. However, the total running words in the 1968 series is 7,531, a figure comparable to the total for 1923.

Number of New Words

Table 2 shows the number of new words introduced in each reader and the total number of new words introduced in each time period. The 1872, 1884, 1901, and 1915 materials have similar word burdens, with totals of 392, 367, 405, and 447 new words respectively. Increased vocabularies of 633 and 601 words occur in the 1923 and 1935 materials. After a sharp drop to 341 new words in the 1948 series, the word burden rises to 500 in 1964. The heaviest word load occurs in the 1968 series, which introduces 860 new words.
Table 1

Number of Running Words in Each Total Sample

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Pre-Primer</th>
<th>Pre-Primer</th>
<th>Pre-Primer</th>
<th>Primer</th>
<th>First Reader</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872</td>
<td></td>
<td></td>
<td></td>
<td>3,495</td>
<td>3,495^a</td>
<td></td>
</tr>
<tr>
<td>1884</td>
<td></td>
<td></td>
<td>1,468</td>
<td></td>
<td>1,468^a</td>
<td></td>
</tr>
<tr>
<td>1901</td>
<td></td>
<td>2,540</td>
<td></td>
<td></td>
<td>2,540^a</td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td></td>
<td>2,021</td>
<td></td>
<td></td>
<td>2,021^a</td>
<td></td>
</tr>
<tr>
<td>1923</td>
<td>2,327</td>
<td></td>
<td>4,925</td>
<td></td>
<td>7,252</td>
<td></td>
</tr>
<tr>
<td>1935</td>
<td>4,461</td>
<td>8,305</td>
<td></td>
<td></td>
<td>12,766</td>
<td></td>
</tr>
<tr>
<td>1948</td>
<td>401</td>
<td>975</td>
<td>1,591</td>
<td>6,061</td>
<td>11,549</td>
<td>20,577</td>
</tr>
<tr>
<td>1964</td>
<td>2,169</td>
<td>7,385</td>
<td></td>
<td>13,623</td>
<td>23,177</td>
<td></td>
</tr>
<tr>
<td>1968^b</td>
<td></td>
<td>537^c</td>
<td>2,644</td>
<td>4,350</td>
<td>7,531</td>
<td></td>
</tr>
</tbody>
</table>

^aThis figure represents the word count for the first reading text on the authorized list for this time period; the total first-year programme is not precisely known.

^bResults are shown for three readers in a series of three readers and one novelette.

^cFirst reader.
Table 2

Number of New Words in Each Total Sample

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Pre-Primer</th>
<th>Pre-Primer</th>
<th>Pre-Primer</th>
<th>Primer</th>
<th>First Reader</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1872</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>392</td>
<td>392^a</td>
</tr>
<tr>
<td>1884</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>367</td>
<td>367^a</td>
</tr>
<tr>
<td>1901</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>405</td>
<td>405^a</td>
</tr>
<tr>
<td>1915</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>447</td>
<td>447^a</td>
</tr>
<tr>
<td>1923</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>273</td>
<td>360</td>
</tr>
<tr>
<td>1935</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>284</td>
<td>317</td>
</tr>
<tr>
<td>1948</td>
<td>17</td>
<td>22</td>
<td>21</td>
<td>99</td>
<td>182</td>
<td>341</td>
</tr>
<tr>
<td>1964</td>
<td>83</td>
<td>183</td>
<td>234</td>
<td></td>
<td>500</td>
<td></td>
</tr>
<tr>
<td>1968^b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>111^c</td>
<td>312</td>
</tr>
</tbody>
</table>

^aThis figure represents the word count for the first reading text on the authorized list for this time period; the total first-year programme is not precisely known.

^bResults are shown for three readers in a series of three readers and one novelette.

^cFirst reader.
Comparison of Running Words and New Words

One way of comparing the number of running words and number of new words in the periods of study is to delineate graphically the two sets of data over the nine time periods (Figure 1). Except for 1968, this shows the general trend toward greater amounts of reading material since 1923 and the dramatic increases over the periods 1935 to 1964. At the same time, and again with the exception of 1968, the number of new words or word burden in the materials is smallest in the 1948 period and highest in the periods 1923 and 1935.

Another way of comparing running words and new words in the different time periods is to compare the ratio of new words to 100 running words in each time period (Figure 2). For 1872 the ratio is 11.2:100 or 11.2 new words for every 100 running words. The results for the remaining time periods similarly are (1884) 25.0:100, (1901) 15.9:100, (1915) 22.1:100, (1923) 8.7:100, (1935) 4.7:100, (1948) 1.7:100, (1964) 2.2:100, (1968) 11.4:100. These data show that among the early texts the ratios of new words to running words in the 1884 and 1915 materials were notably high. A general trend towards increasingly smaller ratios of new words to 100 running words occurs beginning in 1923 with the 1948 series having the smallest ratio. Again an exception exists in the 1968 data.

THE 500-WORD SAMPLES

Average Number and Range of Repetitions

The average number and range of repetitions in the first 500 running words in each time period are shown in Table 3. These data show that the average number of repetitions in the samples ranges from 3.0 to 6.3 in the first six time periods, then rises sharply in 1948 to 22.7. The subsequent time period, 1964, has an average of 14.2 repetitions in the sample, while the 1968 material has an average of 4.6, a number similar to the earlier periods. On the other hand, the range of
Figure 1. Number of running words and number of new words in each Total Sample.
Figure 2. Ratio of new words to 100 running words in each Total Sample.
Table 3

Average Number and Range of Repetitions in each 500-Word Sample

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Repetitions</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average Number</td>
<td>Range</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>1872</td>
<td>6.1</td>
<td>1 - 43</td>
<td></td>
</tr>
<tr>
<td>1884</td>
<td>5.3</td>
<td>1 - 75</td>
<td></td>
</tr>
<tr>
<td>1901</td>
<td>3.3</td>
<td>1 - 34</td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td>3.0</td>
<td>1 - 31</td>
<td></td>
</tr>
<tr>
<td>1923</td>
<td>5.8</td>
<td>1 - 55</td>
<td></td>
</tr>
<tr>
<td>1935</td>
<td>6.3</td>
<td>1 - 38</td>
<td></td>
</tr>
<tr>
<td>1948</td>
<td>22.7</td>
<td>5 - 60</td>
<td></td>
</tr>
<tr>
<td>1964</td>
<td>14.2</td>
<td>1 - 54</td>
<td></td>
</tr>
<tr>
<td>1968</td>
<td>4.6</td>
<td>1 - 38</td>
<td></td>
</tr>
</tbody>
</table>
Table 4

Distribution of Repetitions in Each 500-Word Sample

<table>
<thead>
<tr>
<th>Number of Occurrences</th>
<th>1872</th>
<th>1884</th>
<th>1901</th>
<th>1915</th>
<th>1923</th>
<th>1935</th>
<th>1948</th>
<th>1964</th>
<th>1968</th>
</tr>
</thead>
<tbody>
<tr>
<td>71 - 80</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61 - 70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51 - 60</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41 - 50</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31 - 40</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 - 30</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>7</td>
<td>9</td>
<td>10</td>
<td></td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>16</td>
<td>9</td>
<td>16</td>
<td>9</td>
<td>6</td>
<td></td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>13</td>
<td>21</td>
<td>30</td>
<td>7</td>
<td>10</td>
<td></td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>33</td>
<td>85</td>
<td>86</td>
<td>32</td>
<td>15</td>
<td></td>
<td>3</td>
<td>38</td>
</tr>
<tr>
<td>Totals</td>
<td>81</td>
<td>95</td>
<td>152</td>
<td>167</td>
<td>86</td>
<td>79</td>
<td>22</td>
<td>35</td>
<td>108</td>
</tr>
</tbody>
</table>
repetitions is greatest in the 1884 time period, which had a range of 1-75, and least in 1915 and 1901, where the range of repetitions was 1-31 and 1-34 respectively.

Distribution of Word Repetitions

The distribution of word repetitions in each 500-Word Sample is shown in Table 4. The mode or most typical number of occurrences of words is 1 in all time periods except 1948 and 1964, where the modes fall in the interval 21 to 30.

DISCUSSION

Number of Running Words

The increasing amounts of actual reading material found in the readers of recent periods compared with those of earlier periods generally conforms with the results of Dolch's comparison of McGuffey Readers and modern readers (Dolch, 1945). Two reasons suggested by Dolch for the quantitative differences between the late nineteenth-century readers of his study and those of more recent times may be equally applicable to the materials of the present study:

First, books were more expensive in previous years than they are now in proportion to other necessities of life. There was also a conviction that "thorough reading" practically meant memorizing. A child read and re-read a story until he could repeat it almost verbatim, and he might even read the whole book again and again until he knew it by heart. Nowadays we believe in repetition in different contexts, and the words are repeated over and over in different ways in the book. Thus we hope to get word recognition without memorization of stories. (p. 98)

The earlier approach to repetition is illustrated by a note in the 1872 reader of the present study recommending that the "back lessons may with profit be frequently reviewed". The subsequent provision for word repetition in the texts themselves is shown by the increasingly smaller proportions of new words to running words in the readers from 1935 to 1964.

The 1948 material appears to be directly related to the concern for vocabulary
control which began in the late 1920s in the United States. The 1948 texts of this study are the readers of the First-Grade Program of the *Curriculum Foundation Series*. The major author of this series was William S. Gray, who for many years was editor of the annual *Summaries of Reading Investigations*, and was author of many reading studies, including the methods text *On Their Own in Reading*. What is evidently the first edition of the *Curriculum Foundation Series* is cited in Nila Banton Smith's history of American reading instruction under the year 1927 (Smith, 1965). Many revisions followed over the years, and the popular use of the series is attested to by the fact that the names of the two main child characters, Dick and Jane, have virtually entered the language. As pointed out by Topping (1968), this material became the first American-authored reading series prescribed for use in British Columbia schools.

The relatively small number of running words in the 1968 series represents a break in the trend towards larger amounts of reading material. One explanation for this change may be the distinctive nature of the programme (linguistic method). It will be recalled that the first children's reader based on linguistic principles appeared in the United States as recently as 1961 (Smith, 1965).

**Number of New Words**

The 1884 *First Primer*, which is assumed to represent the first-year text in that period, has a high proportion of new words to running words (as well as a small amount of total reading material) when compared with the 1872 reader. The 1884 text is composed of numbered sentences, many of which are not sequential, rhymes, and a small amount of narrative material. The lessons were designed to teach the single letters "in one and only one of their powers... then the double letters - double vowels and double consonants, initial and final" (Preface to the 1884 *First Primer*). The 1872 text, on the other hand, contains lengthy sequences (story-like passages) in which words having similar patterns repeatedly occur.
In the *Annual Report of the Public Schools* for the year 1883-84, the following reasons were given for the changeover to the series to which the 1884 *First Primer* belongs:

"Among the many advantages gained by the introduction of the new series the following features are noticeable: - the excellency of the typographical execution; the judicious selection and gradation of the elementary combinations; the association of the written word-sign with the pictorial; ... the careful gradation of the matter from lesson to lesson and from book to book ... ." (Excerpt from the *Annual Report of the Public Schools of the Province of British Columbia, 1883-84*, quoted by Green, 1938)

Apparently what would today be considered a relatively heavy word burden was not a factor in the choice of materials at that time.

The 1915 *Beginner's Reader* also has a proportionately high number of new words to running words. In a manner somewhat similar to the 1884 text, short sentences which act as a vehicle to introduce phonic elements in words comprise the body of the material. A statement which occurs in a phonic manual of the period may provide an explanation for this device. According to this manual, which was prepared as a companion to a phonic primer, "... the short sentence aids in gaining quick word recognition, inasmuch as it admits of more practice on the new letter than does the long sentence or connected story" (Macmillan's, 1917).

Like the 1884 and 1915 materials, the 1901 *First Primer* contains a number of non-sequential sentences which serve to introduce phonic elements in words. However in this case the sentences are confined to the first few pages of the text, and the body of the primer (about three-quarters) consists of short narrative passages, some of which are associated with nursery rhymes. Beginning with the 1923 materials, isolated sentences do not occur and the narrative mode is dominant.

The very high number of new words in the 1968 series is related to a system of word-introduction based on word patterns and a synthetic phonics system. As indicated by the claims of the editors, the word burden of this material is not intended for comparison with earlier basal readers:
... children no longer need to be limited to a 40–60 word reading vocabulary in their initial 4 months of formal reading instruction, but in fact can quite easily synthesize (blend) up to 1,000 words in that time period. The vast increase in reading and writing vocabulary through this [multi-sensory and sound structure] approach is a distinctive break with the traditional belief that the vocabulary provided by basic readers is the limit of the reading vocabulary to be acquired by children in this period. (Linn (Ed.), Language Patterns Teacher's Guide, Part I, 1968, p.7)

The vocabulary system used in these linguistic materials indicates that in order to determine the word burden it would be more appropriate to count the number of different word patterns rather than the number of different words.

**Repetitions**

Marked increases occur in the average number of repetitions of words in the 500-Word Samples for 1948 and 1964 in comparison with the samples from other periods. As these series occurred in a time of increasing use of vocabulary control in readers in the United States, and as the 1948 series are in fact American materials, the treatment of repetitions in these readers is explained most simply as a reflection of the then current practice.

The relatively low average number of repetitions and the small differences between the averages for the samples in the periods other than 1948 and 1964 raise the question as to whether systematic repetition of words was used at all in these materials; in other words, it is possible that equivalent levels of repetition might be found in any sample of reading material written for young children during the same time periods. However, it became apparent during the tabulation of the data that methods of providing repetition existed in some of the materials prior to the intensive use of vocabulary control.

A prevalent device for repetition in the early materials, and one similar to that used in the recent 1968 series, is repetition of word patterns. In the 1872 material, phonograms with a consistent sound–symbol relationship are repeated in different words in the early lessons, e.g., *at, fat, Pat;* at the
same time, later lessons present orthographic variants of the same sound, e.g., be, bee, pea, or oh, mow, doe, oar, which indicates a rapidly increasing level of difficulty in present-day terms. In the 1923 material (and occasionally in the 1901 text), repetition of phonograms is provided through the rhyming words of nursery rhymes, and these word patterns are repeated in prose retellings of the rhymes; repetition also occurs in the traditional nursery tales which are used in this text. These observations suggest that a simple word count does not provide a complete picture of methods of repetition in some of the early readers.

The system of repetition in the 1935 readers appears to be based on the use of the same words in different contexts, the choice of words being determined by reference to the Gates Primary Word List. The method of repeating words in different contexts is the one used in the 1948 and 1964 materials; here, however, and particularly in the 1948 series, there is also reliance on what might be called serial repetition, e.g., Oh, oh, oh or Look! Look!; this device does not occur in the 1935 materials.

**SUMMARY OF CHAPTER**

While the total number of running words in the readers in the four earliest periods of the study, 1872, 1884, 1901, and 1915, varied somewhat in quantity, the numbers were relatively small. In the next period, 1923, approximately three times as many running words occurred as in the 1915 text, and a trend towards increasingly larger amounts of material in first-grade readers continued in each of the following three periods, 1935, 1948, and 1964. Over the same time periods, the number of new words introduced in the readers increased only moderately, though comparatively large numbers were found in the 1923 and 1935 periods and an unusually small number in the 1948 time period. An exception to these general tendencies occurred in the findings for the 1968 material where the number of running words was comparable to the 1923 text, and the number of new words introduced was higher.
than in any of the other time periods. Except for the 1968 series, the relatively extreme change in quantities of reading material and the smaller change in number of new words introduced in the readers over the periods of study resulted in a reduced ratio of new words to running words in the later periods; this characteristic was particularly marked in the 1948 and 1964 materials.

Similar average numbers of repetitions were found in the 500-Word Samples of the first six time periods and the 1968 period, but averages for the 1948 and 1964 samples were considerably higher. The range of repetitions followed a somewhat different pattern, the greatest range occurring in the 1884 sample and the least in the samples for 1901 and 1915.

Two reasons which have been advanced for the quantitative changes in modern readers in the United States in comparison with the McGuffey Readers of the nineteenth century, namely, the smaller relative cost of readers and the incorporation of word repetitions in the texts themselves, seem equally applicable to the changes observed over the periods of the present study. The growing abundance of word repetitions in the 1935, 1948, and 1964 series indicates increasing use of then current practices in modern vocabulary control. However, it was noted that the vocabulary burden of the 1968 linguistic series seemed less adequately described by a word count than it might be by a count of word patterns; systems of repetition based on word patterns were also observed in the readers for 1872 and 1923.
V. SUMMARY, CONCLUSIONS, RECOMMENDATIONS

SUMMARY

The problem of this study was to describe the vocabulary load of the texts authorized for beginning reading in British Columbia from 1872 to the present, and to identify changes which have occurred in this area of the readers over this time span. In order to determine the historical texts and ascertain periods of authorization, the lists of authorized texts and other pertinent sources in the history of British Columbia education were examined. Because the assignment of first-year textbooks prior to the establishment in 1923 of the present system of grade classification is not generally clear, the first reading text on the list of authorized texts in each period of prescription prior to 1923 was arbitrarily designated as the beginning reader in each case. The study was a descriptive word count, and word difficulty, and production and thematic differences in the texts were not taken into account; in addition, historical investigation was limited to systems of vocabulary introduction in readers, and no attempt was made to treat other factors which have affected education in the Province over the period of study.

Related Literature

Research relative to this project was found in the areas of content analysis and vocabulary studies of beginning readers. The technique of content analysis, which requires the precise definition of the universe of content or "population" to be sampled, and the tabulation of all relevant instances of the categories of analysis, served as a procedural guide to the quantitative description of the reading vocabularies, and provided the rationale underlying the comparison of the resulting data.

Vocabulary studies in the research literature offered historical and current
information about systems of vocabulary introduction in readers, and served to identify the dimensions of vocabulary load that were analyzed in the present study. Examinations of vocabulary systems in historical readers reported by Nila Banton Smith in her history of American reading instruction show that different systems of presenting words in readers have existed in different time periods (Smith, 1965); however, a factor common to all the examples described was the presentation of words or word parts graded according to some perceived standard of difficulty. E. W. Dolch (1945) compared the vocabularies of nineteenth-century McGuffey Readers and modern American readers and found that the latter provided considerably more reading material, fewer new words, and less difficult words than the former, but that there was considerable agreement between steps of difficulty in the two sets of readers. References to reading vocabularies in Edmund B. Huey's study of reading published early in the present century (Huey, 1908) indicate a concern for how the context of words can enlarge the vocabulary of a child rather than an interest in limited numbers of new words.

Early in this century, frequency counts of reading vocabularies in the United States led to the establishment of functional word lists graded in difficulty, and these came to be used as reference criteria for authors in the preparation of children's readers. Also in this period, researchers observed many inconsistencies in beginning reading vocabularies and stressed the need for standards of selection and control.

The system of controlled vocabulary which has since come into general use in North American basal readers is the planned introduction of a limited number of words, graded as to difficulty and interest, the repetition of these words in different contexts at spaced intervals in the text, and the carry-over of these words to succeeding books in the series. A similar method of vocabulary control was developed in the United Kingdom.
A trend towards increasingly smaller vocabularies in American primers from the 1930s to the 1960s has been observed by reading authorities in the United States. On the other hand, Scottish studies reported in the early 1950s show widely ranging vocabulary burdens in beginning readers. Data from Scottish vocabulary studies of this period were also used to justify a larger word burden in the primary programme of a Canadian reading series in contrast to the then small vocabularies of American beginning reading texts (McIntosh (Ed.) and others, 1960).

The importance of controlled vocabularies in reading texts as a solution to learning problems engendered by unregulated and heavy word loads, and the advantages of the system to both the child and the teacher have been cogently demonstrated in the writings of such leading reading authorities as Gray and Schonell. Recently, however, and in the light of the wider experiences of today's children, some researchers have questioned the too rigorous application of the principle. There is also evidence that the vocabulary burden of United States primers is increasing in the 1970s.

Methodology

The universe of content for this analysis was the nine readers and reader series authorized for beginning reading in the Province of British Columbia since 1872. The first year of prescription was used to designate the time period in each case, namely: 1872, 1884, 1901, 1915, 1923, 1935, 1948, 1964, and 1968. The 1964 and 1968 reader series are both currently prescribed.

Criteria were specified for a Total Sample (with minor exceptions, this comprised all the reading content in each text) and a 500-Word Sample (the first 500 words in the Total Sample). Thus there were nine Total Samples and nine 500-Word Samples. For each Total Sample, the analysis consisted of tabulations of the total number of running words and the total number of new words introduced. For each 500-Word Sample, the average number and range of word repetitions were computed.
Findings

The total numbers of running words in the 1872, 1884, 1901, and 1915 Total Samples were 3,495, 1,468, 2,540, and 2,021 respectively. In the 1923 Total Sample, however, 7,252 running words were tabulated, or more than three times as many as in 1915, the immediately preceding period. The trend towards increasingly larger amounts of reading material continued in 1935, 1948, and 1964, which had 12,766, 20,577, and 23,177 running words in the Total Samples respectively. Over the same time periods, the numbers of new words introduced in the readers varied only moderately; among these findings, the highest numbers of new words were 633 and 601 in the 1923 and 1935 Total Samples, and the lowest number of new words was 341 in the 1948 Total Sample. Exceptions to the general trends in numbers of running words and numbers of new words were found in the 1968 Total Sample: specifically, the number of running words in the 1968 material was 7,531, an amount comparable to that for 1923, while the number of new words in the 1968 material was 860, or higher than in any of the other time periods. Compared with the four early periods, the large increases in amounts of reading material and the lesser change in numbers of new words introduced in the readers in the later periods (except for 1968) resulted in a reduced ratio of new words to running words. Small ratios of new words to running words were particularly evident in the 1948 and 1964 materials.

The average number of repetitions found in the 500-Word Samples of the first six time periods and the 1968 time period ranged from 3.0 to 6.3. Average repetitions for the 1948 and 1964 samples were considerably higher, namely 22.7 and 14.2 in each case. The greatest range of repetitions occurred in the 1884 sample and the least in the samples for 1901 and 1915.
CONCLUSIONS

Within the scope of the dimensions analyzed, the results of this study provide a numerical description of the vocabulary load of each of the nine beginning readers prescribed for use in British Columbia schools since 1872. This description shows the relationship between number of running words and number of new words in each period of study, and indicates the extent of word repetition in each case. However, during the tabulation of the data it was observed that the dimensions of word introduction and word repetition tended to disguise the sequential introduction and repetition of words having similar patterns which occurred in the 1872 and 1923 readers and in the 1968 linguistic series.

It seems reasonable to conclude that the differences observed in the results for the early and recent time periods illustrate the increasing application of principles of modern vocabulary control in the 1935 and 1948 time periods, and its continued intensive use in the 1964 period. There is evidence that a movement towards controlled vocabularies in beginning readers began in the United States in the 1920s; moreover, vocabulary studies of American beginning reading texts in the thirty-year period from 1930 show a continuing decrease in numbers of new words. The striking changes beginning in the 1935 time period of the present study in numbers of new words and the proportion of new words to running words suggest by inference the relative immediacy and pervasiveness of the influence of American vocabulary practices in these years of reading education in British Columbia.
FUTURE RESEARCH

Future study of the readers could usefully proceed in two main directions. First, more information is needed about methods of determining the vocabulary load of first-year readers. The present study has been concerned with relationships between aggregate numbers of words. However, the introduction and repetition of words having similar patterns observed in some of the readers of the sample point to a method of analysis which gives weight to this dimension.

Secondly, the patterns of change observed in the readers over the nine periods of prescription indicate that content analysis techniques could provide data for other meaningful comparisons. One kind of comparison that could be made is a comparison of instructional practices over the different time periods. Classifiable data about methods of instruction occur in the notes to teachers in the early texts, and in the manuals that are associated with the readers from 1923 onwards. Data from the present vocabulary study could be a contributing variable in such an analysis. An additional variable, and one which could be documented from historical sources, is how reading was evaluated in each time period.

Another kind of comparison based on content analysis procedures is a comparison of attitudes expressed in the reading content. An attitude scale developed by Gaston E. Blom, Lawrence Wiberg, and others has been used to identify measurable differences in the attitudes expressed in a cross-cultural sample of primers from thirteen countries (Blom & Wiberg, 1973). This instrument might be used for analyzing the attitude content of the beginning readers in the different time periods and thus provide data for inferences about cultural change over the one-hundred year span of the materials. A purely descriptive study which explores whether or not measurable differences in attitudes
expressed occur in the readers of different time periods could be a first step in such an investigation. Some of the problems of content analysis of children's reading materials have been examined (Patterson, 1976), and the reader is referred to detailed references by Berelson (1952) and Holsti (1969).
BIBLIOGRAPHY

The Readers of the Study by Period of Prescription


- Pre-reader 1: We look and see
- Pre-reader 2: We work and play
- Pre-reader 3: We come and go
- Primer: Fun with Dick and Jane
- First reader: Our new friends


- Pre-primer: Off to school
- Primer: Come along with me
- First reader: It's story time


- Listening letters (First reader)
- Laughing letters
- Magic letters

Teacher's Manuals


39.


**Government Documents**


Province of British Columbia Department of Education. *Programme of studies for the elementary schools of British Columbia, 1925-1926*. Victoria: King's Printer, 1925.


Articles, Books, and Unpublished Papers


Covell, H. M. The development of reading in Canada. Unpublished manuscript, The University of British Columbia. (No date)


Dolch, Edward W. How hard were the McGuffey Readers? Elementary School Journal, 1945, 46, 97-100.


Mathews, Mitford M. Teaching to read: historically considered. Chicago:


APPENDIX A: PERIODS OF PRESCRIPTION AND THE PRESCRIBED READERS

Secondary Sources

An unpublished survey of the eight different reading series prescribed for use in British Columbia schools from 1900 to 1967 (Topping, 1968) provided dates of prescription and titles of the series used over those years. Reference was also made to unpublished findings by Covell (undated) and Boyce (1949). Using these three sources, it was possible to tentatively identify some of the early readers and the intervals of authorization.

Since it was necessary to verify the dates of prescription and ascertain the beginning reader or readers in each case, a search was made to find the original lists of authorized textbooks. A study by Green (1938) on the development of the curriculum in the elementary schools of British Columbia prior to 1936 showed that the lists of authorized texts were published in different sources at different times, namely:

1872-1893 and 1911: Annual Reports of the Public Schools of the Province of British Columbia

1893-1916: Manuals of School Law

After 1916: separate booklets under various titles

Green's study also provided useful information about: text authorization in the periods 1884, 1900-1901, and 1923; the classification system used in the schools prior to 1923; and the changeover to the present grade system in 1923.

Primary Sources

1872. The first reading text on the list of the authorized textbooks published in the First Annual Report of the Public Schools of the Province of British Columbia, 1871-72 (1872) is: Canadian First Reader - Part I. Examination of bibliographic and other data in the readers in the Special Collections Division of the University
of British Columbia, and references in a history of Ontario prescribed texts (Parvin, 1965) indicated that this text was:


For purposes of the study, this book was arbitrarily designated the first-year reader in the 1872 period of authorization.

1884. References to the change in texts which occurred in 1884 were found in Green's study (1938), but the *Annual Report* for 1883-84 was not available for examination. However, the first reading text on the list of authorized textbooks in the *Fourteenth Annual Report of the Public Schools of the Province of British Columbia, 1884-85* (1885) is: *Gage's First Primer, Pt. I.* There seems little question but that this text was:


For purposes of the study, this book was arbitrarily designated the first-year reader in the 1884 period of authorization. (This text is based on a series prepared by J. M. D. Maiklejohn; for a reference to this author's work, see Mathews, 1966, p. 154.)

1901. Since the *Manuals of School Law* were not available, the description by Green (1938) of the changeover to the *Twentieth Century Edition* of the *New Canadian Readers* in 1901 was used to confirm the date and ascertain the appropriate text for this period of prescription, namely:


Following the rule for selection used in the two previous time periods, and on the basis of information about assignment of the readers in the Junior Division (Green, 1938), this text was designated the first-year reader in the 1901 period of prescription.

1915. Although *The British Columbia Readers: First Reader* of 1915 was
available to the researcher, and although other sources pointed to a change in readers in this year, no reference to changes in readers in 1915 was found in Green's study (1938); moreover, the Manuals of School Law were not available. However, extracts from Circulars of Instructions in the Annual Reports of 1914-15 and 1915-16 (Forty-Fourth Annual, 1916; Forty-Fifth Annual, 1917) confirmed the gradual displacement beginning in 1915 of the First and Second Primers and First and Second Readers of the previous series. The first two texts listed in the extract for 1915 appeared in reverse order the following year, 1916 (the change is stipulated, but not explained); thus the 1915 materials, using the 1916 order for the first two texts, were as follows: B. C. Beginner's Reader (replacing the former First Primer); B. C. Phonic Primer (replacing the former Second Primer); B. C. First Reader (replacing the former First Reader).

On the basis of this listing, and as a result of a further search of the Special Collections Division sources, the following text was identified as the first reader in this period of authorization:


1923. In addition to information provided by Green (1938) regarding the change-over in reading texts which occurred in 1923, further bibliographic verification was found in the List of Authorized Text-Books for Public Schools published in the Programme of Studies for the Elementary Schools of British Columbia, 1925-1926 (Province of B. C., 1925). The authorized text for Grade I in this period was:


1935. The next change in reading texts, which occurred in 1935, was verified by comparing the lists of free textbooks issued in the Annual Reports for 1933-34 and 1934-35 (Sixty-Third Annual, 1934; Sixty-Fourth Annual, 1935) respectively. The texts designated for Grade I were also confirmed by reference to the Programme.
of Studies for the Elementary Schools of British Columbia, Grades I to VI.

Bulletin I (Province of B. C., 1936). Specifically, the Grade I texts were:

Highroads to Reading: Jerry and Jane (Primer); Book One. Toronto: Ryerson Press, 1934.

1948. Since original sources documenting the change in reading texts in 1948 were not available, secondary sources cited above (Topping, 1968; Covell, undated) were used to identify the year of authorization and the authorized texts.

The newly authorized series was:

Curriculum Foundation Series First-Grade Program: Pre-Readers 1, 2, and 3; Primer; First Reader. Chicago: Scott, Foresman and Company, 1946.

1964. The 1964 revision in the authorized list of reading texts for Grade I was verified by comparing the 1963-64 and 1964-65 lists of Prescribed Textbooks (Province of B. C., 1963, 1964). The new texts on the 1964-65 list were:


1968. The addition of authorized alternate reading material for Grade I in 1968 was confirmed by reference to the British Columbia Language Arts Guide - Primary Levels, 1968 (Province of B. C., 1968). Specifically, the texts were:


As the first three texts in the foregoing series are readers and as the fourth text is a novelette, it was decided to limit the sample for this period of authorization to the three readers.

Prescribed readers 1976-77. The two series currently prescribed for Grade I remain the same as those authorized in 1964 and 1968, and are listed in Prescribed Textbooks 1976-77 Grades K-XII (Province of B. C., 1976).