DOCUMENT RESUME

ED 046 626 24 RE 003 194

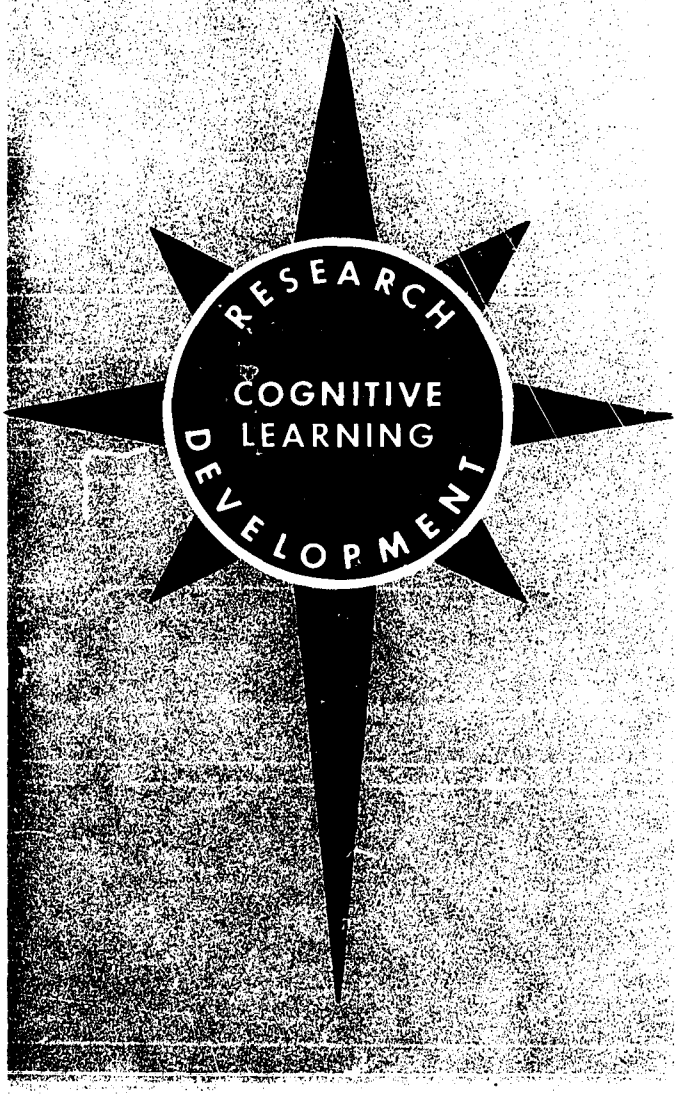
AUTHOR Johnson, Dale; Venezky, Richard
TITLE An Instrument for Testing Pronunciations of Vowel Clusters. Part I: Linguistic Background.
INSTITUTION Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO TR-137
PUB DATE Jul 70
CONTRACT OEC-5-10-154
NOTE 34p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Child Language, Elementary School Students, *Language Research, Language Tests, Linguistics, Phonetics, *Pronunciation, Pronunciation Instruction, *Reading Research, Spelling, Test Construction, *Vowels

ABSTRACT

Since pronunciations of vowel clusters are among the most unpredictable letter-sound correspondences in English and therefore children learning to read must often rely on something other than spelling as a clue to pronunciation of vowel cluster words, data relating to pronunciation frequencies of certain vowel clusters were gathered for this study, and a rationale for selection of words containing these clusters was developed. An instrument devised for use in assessing children's initial pronunciations of vowel clusters in unfamiliar words is described, and a study using it is proposed. Results of the study will be presented in a second report. (MS)

AN INSTRUMENT FOR TESTING PRONUNCIATIONS OF VOWEL CLUSTERS



WISCONSIN RESEARCH AND DEVELOPMENT

CENTER FOR
COGNITIVE LEARNING

Technical Report No. 137

AN INSTRUMENT FOR TESTING PRONUNCIATIONS OF VOWEL CLUSTERS

PART I: LINGUISTIC BACKGROUND

By Dale Johnson and Richard Venezky

Report from the Project on Basic
Pre-Reading Skills: Identification and Improvement

Robert C. Calfee and Richard L. Venezky, Principal Investigators

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

July 1970

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Launor F. Carter
Senior Vice President on Technology and Development
System Development Corporation

Francis S. Chase
Professor
Department of Education
University of Chicago

Henry Chauncey
President
Educational Testing Service

Martin Deutsch
Director, Institute for Developmental Studies
New York Medical College

Jack Edling
Director, Teaching Research Division
Oregon State System of Higher Education

Elizabeth Koontz
Wage and Labor Standards Administration, U.S. Department of Labor, Washington

Roderick McPhee
President
Punahou School, Honolulu

G. Wesley Sowards
Director, Elementary Education
Florida State University

Patrick Suppes
Professor
Department of Mathematics
Stanford University

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

RESEARCH AND DEVELOPMENT CENTER POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

Robert E. Grinder
Chairman
Department of Educational Psychology

Russell J. Hosler
Professor, Curriculum and Instruction

Clauston Jenkins
Assistant Director
Coordinating Committee for Higher Education

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Stephen C. Kleene
Dean, College of Letters and Science

Donald J. McCarty
Dean
School of Education

Ira Sharkansky
Associate Professor of Political Science

B. Robert Tabachnick
Chairman, Department of Curriculum and Instruction

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

M. Crawford Young
Associate Dean
The Graduate School

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of Sociology

Anne E. Buchanan
Project Specialist
R & D Center

Robin S. Chapman
Research Associate
R & D Center

Robert E. Davidson
Assistant Professor, Educational Psychology

Frank H. Farley
Associate Professor, Educational Psychology

Russell J. Hosler
Professor of Curriculum and Instruction and of Business

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Robert G. Petzold
Associate Dean of the School of Education
Professor of Curriculum and Instruction and of Music

FACULTY OF PRINCIPAL INVESTIGATORS

Vernon L. Allen
Professor of Psychology

Ted Czajkowski
Assistant Professor of Curriculum and Instruction

Robert E. Davidson
Assistant Professor of Educational Psychology

Gary A. Davis
Associate Professor of Educational Psychology

M. Vere DeVault
Professor of Curriculum and Instruction (Mathematics)

Frank H. Farley
Associate Professor of Educational Psychology

Lester S. Golub
Lecturer in Curriculum and Instruction and in English

John G. Harvey
Associate Professor of Mathematics and of Curriculum and Instruction

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Donald Lange
Assistant Professor of Curriculum and Instruction

James Moser
Assistant Professor of Mathematics Education; Visiting Scholar

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Milton O. Pella
Professor of Curriculum and Instruction (Science)

Thomas A. Romberg
Associate Director, R & D Center
Professor of Mathematics and of Curriculum and Instruction

B. Robert Tabachnick
Chairman, Department of Curriculum and Instruction

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

Alan Voelker
Assistant Professor of Curriculum and Instruction

Larry Wilder
Assistant Professor of Curriculum and Instruction

Peter Wolff
Assistant Professor of Educational Psychology

MANAGEMENT COUNCIL

Herbert J. Klausmeier
Director, R & D Center
V.A.C. Henmon Professor of Educational Psychology

Mary R. Quilling
Director
Technical Development Program

Thomas A. Romberg
Associate Director

James Walter
Director
Dissemination Program

Dan G. Woolpert
Director
Operations and Business

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from the Project on Basic Pre-Reading Skills: Identification and Improvement in Program 1. General objectives of the Program are to generate new knowledge about concept learning and cognitive skills, to synthesize existing knowledge, and to develop educational materials suggested by the prior activities. Contributing to these Program objectives, this project's basic goal is to determine the processes by which children aged four to seven learn to read and to identify the specific reasons why many children fail to acquire this ability. Later studies will be conducted to find experimental techniques and tests for optimizing the acquisition of skills needed for learning to read.

00005

CONTENTS

LIST OF TABLES

ABSTRACT

The pronunciations associated with vowel cluster spellings are among the most unpredictable letter-sound correspondences in English. If learning to read includes learning to translate from spelling to sound, then vowel clusters should pose a particularly difficult problem for children. The manner in which children solve this problem, i.e., the factors influencing children's pronunciation of vowel clusters in unfamiliar words, may shed more general light on information attended to in the selection of pronunciations for spellings.

The present report is the first of two describing a study of factors influencing children's pronunciations of vowel clusters at different ages. In this report are presented data on vowel cluster pronunciations in English, the linguistic rationale for the selection of vowel digraphs used in the study, and a description of the testing instrument devised. Results of the study will be presented in a second report.

00009

I
INTRODUCTION

It has been said that the act of learning to read is perhaps the greatest intellectual feat of anyone's lifetime; yet there is no generally accepted definition of what reading is.

Of all the skills of language and thought, perhaps the only one unique to beginning reading is the ability to translate what is written to oral language already possessed. Symbols represent sounds; unfortunately, some sounds are represented by many symbols in English. Recent studies, aided by computer technology (Venezky, 1963; Hanna, Hanna, Hodges, & Rudorf, 1966), have tabulated the relationships between spelling and sound and sound and spelling in common English words. However, little research has been done to determine whether or not these relationships are actually used by competent readers—and if they are, how children acquire them.

Since reading includes the translation from spelling to sound, it is important to know how children acquire this behavior. When a child encounters an unfamiliar word, what factors influence his choice of pronunciation?

00010

II

THE PRONUNCIATION OF VOWEL CLUSTERS IN ENGLISH

Vowel clusters are perhaps the most complex and unpredictable components of the letter-sound correspondence code. Vowel cluster spellings differ from single vowel spellings in several ways. They rarely appear before geminate consonant clusters; some, such as ai and au, occur infrequently in word final position, while others, such as oa and ie, rarely begin a word in English.

Some vowel clusters have a major phonemic correspondent and several minor correspondents. For example, the major correspondent of ai is /e/ as in bait; and it represents this sound 85% of the time that it occurs. It represents /ə/, villain; /aɪ/, aisle; /ɛ/, again; /æ/, plaid, and others much less frequently. Other vowel clusters have two or more major correspondents, as well as minor correspondents. The vowel cluster ow is /o/ as in own 51% of the time and /au/ as in owl 48%. Its only minor correspondent is /a/ as in knowledge. By contrast, all single vowel spellings have two major correspondences (e.g., a is /e/ or /æ/ as in rate and rat) plus several minor correspondences.

Given that this variation in letter-sound correspondences exists in English (that is, in common English words many vowel clusters have six or more pronunciations), what factors influence children's pronunciation of unfamiliar words containing vowel clusters? A study was designed to investigate the relationships between the pronunciations of vowel clusters by a representative sample of elementary school children and the letter-sound correspondences of vowel clusters in a large corpus of common English words.

The present report will discuss the development of a test to be used in this study. A later report will present a summary and analysis of the research findings.

The first requirement for the development of a vowel cluster pronunciation test was an analysis of the letter-sound correspondences of vowel clusters in common English words. As part of an interdisciplinary study of the reading process begun at Cornell University in 1961, Venezky developed a computer program to derive and tabulate letter-sound correspondences in a corpus of 20,000 common English words (1963). The 20,000-word corpus was a modification of the most common 20,000 words according to the Thorndike Frequency Count (1941). Venezky omitted many archaic and low-frequency words, particularly proper nouns, and added a number of words in their place. Along with other information, the computer analysis provided an inclusive tabulation of letter-sound correspondences found in the corpus as well as totals and percentages for each pronunciation in each word position, and a complete word list for each correspondence. A Pronouncing Dictionary of American English (Kenyon & Knott, 1953) was used to determine the pronunciation of most words in the corpus.

The principal purpose of this analysis and later research by Weir (1964), Venezky and Weir (1966), and Venezky (1967a, 1967b) was, ". . . to construct a theoretical framework for deriving sound from spelling and to search for the most plausible linkages for fitting these relationships into the total language structure" (Venezky, 1967, p. 80). Later work (Calfee, Venezky, & Chapman, 1969) was concerned with whether or not readers use these theoretical patterns of symbol-sound relationships when reading.

Venezky's unpublished computer print-out (1965) of spelling-to-sound correspondences in 20,000 English words was analyzed to determine letter-sound correspondences for digraph vowels. This analysis disclosed the following:

1. There are 61 vowels clusters (including those containing the semi-vowels w and y) in the corpus.

2. These 61 vowel clusters represent 92 different single vowel phonemes and phoneme strings producing more than 300 symbol-sound correspondences. For example, oa represents /o/, /ɔ/, /oa/ and other phonemes and phoneme strings. Yet each of these and others, are represented by a variety of spellings. Consequently, there are over 300 symbol-sound correspondences.

3. These 61 vowel clusters appear more than 6,000 times in the 20,000-word corpus.

4. There is great variance in the frequency of the 61 vowel clusters. As shown in Table 1, one occurs in more than 1,000 words while 26 occur in three words or less.

5. Vowel clusters vary greatly in the number of individual phonemes or phoneme strings they represent. Table 2 indicates that some represent only one sound while one represents 17 sounds.

6. Most vowel cluster pronunciations are unpredictable from their spellings.*

7. Of the 61 vowel clusters, 30 occur in 10 or more words in the corpus. Of these 30, 23 occur in words in which the vowel cluster is sometimes disyllabic. Only six of these vowel clusters are disyllabic more often than monosyllabic. Thus, these 30 vowel clusters, occurring in more than 6,000 words, represent single vowel phonemes about 80% of the time and two or more phonemes about 20%. This is shown in Table 3.

Perhaps the best way to exemplify the variety of possible pronunciations of the vowel clusters is to list the most common clusters and their most common pronunciations. Tables 4 through 20 list the 17 vowel clusters which occur in more than 100 words each in the corpus. For each cluster the four most common pronunciations are included.

Table 1

Frequency of Vowel Clusters in 20,000 Word Corpus

Number of Clusters	Number of Words
1	over 1000
2	500 - 999
14	100 - 499
9	50 - 99
4	10 - 49
6	4 - 9
26	1 - 3

Table 2

Frequency of Occurrences of Vowel Clusters and the Number of Sounds they Represent

Vowel Cluster	Number of Words in Which it Occurs	Number of Sounds it Represents
io	1293	10
ea	599	17
ia	581	15
ou	475	11
ee	319	6
oo	312	7
ai	303	9
ie	274	15
ow	256	3
au	191	6
ay	159	8
iou	139	5
oi	130	7
oa	125	7
ue	108	16
ua	104	13

*Some letter-sound correspondences are invariant or nearly invariant; therefore the sound can be derived from the symbol regardless of contextual restraints. Other sound correspondences are variant but are considered predictable because the correspondence can be determined by some feature within the word, such as a consonant environment. For example, c is usually /k/ before a, c and u, as in cat, cot, and cup. On the other hand, ea may be either /i/, /ɛ/, or /e/ before /t/ as in heat, threat, and great, and both /i/ and /ɛ/ after /h/ as in heat and head. Therefore, since features within a word do not signal the pronunciation of ea, it is considered unpredictable.

Table 2 (cont.)

Vowel Cluster	Number of Words in Which it Occurs	Number of Sounds it Represents
ui	102	8
ei	94	8
ey	92	5
aw	88	3
ew	82	3
eo	75	13
iu	56	4
oy	56	2
oe	52	10
eu	51	8
eou	33	2
uou	27	3
ae	21	7
eau	14	3
ao	6	3
ieu	5	2
iew	5	1
oui	5	3
aeo	4	4
uo	4	3
uy	3	1
uoy	3	1
aa	2	1
oia	2	1
uay	2	1
eea	1	1
aea	1	1
eia	1	1
iaow	1	1
ii	1	1
oau	1	1
eow	1	1
ioa	1	1
uia	1	1
eoi	1	1
eei	1	1
oeu	1	1

Table 2 (cont.)

Vowel Cluster	Number of Words in Which it Occurs	Number of Sounds it Represents
oie	1	1
oua	1	1
eue	1	1
aiia	1	1
aii	1	1
aie	1	1
oue	1	1
uu	1	1

Table 3

Monosyllabic and Disyllabic Status of the 30 Most Common Vowel Clusters

Vowel Cluster	One Syllable		Disyllabic	
	Number of Words	Per Cent	Number of Words	Per Cent
ae	18	85.7	3	14.3
ai	298	98.3	5	1.7
au	191	100.0	0	0.0
aw	86	97.7	2	2.3
ay	158	99.4	1	0.6
ea	486	81.1	113	18.9
eau	14	100.0	0	0.0
ee	310	97.2	9	2.8
ei	68	72.3	26	27.7
eo	19	25.3	56	74.6
eou	8	24.2	25	75.8
eu	37	72.5	14	27.5
ew	82	100.0	0	0.0
ey	92	100.0	0	0.0
ia	150	25.8	431	74.2
ie	184	67.1	90	32.9
io	1141	88.2	152	11.8
iou	79	56.8	60	43.2
iu	4	7.1	52	92.9

Table 3 (cont.)

Vowel Cluster	One Syllable		Disyllabic	
	Number of Words	Per Cent	Number of Words	Per Cent
oa	104	83.2	21	16.8
oe	30	57.7	22	42.3
oi	108	83.1	22	16.9
oo	305	97.8	7	2.2
ou	475	100.0	0	0.0
ow	256	100.0	0	0.0
oy	56	100.0	0	0.0
ua	1	1.0	103	99.0
ue	76	70.3	32	29.7
ui	68	62.7	34	37.3
uou	0	0.0	27	100.0
TOTAL	4904		1307	

Table 4

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ai

Phoneme	Number of Words	Percentage	Example
/e/	260	85.8%	bait
/ə/	20	6.6%	villain
/ɪ/	6	2.0%	captain
/ai/	5	1.7%	aisle
5 others	12	3.9%	plaid
Total Occurrences - 303 words			

Table 5

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster au

Phoneme	Number of Words	Percentage	Example
/ɔ/	175	91.6%	cause
/o/	6	3.1%	chauffeur
/æ/	5	2.6%	laugh
/au/	3	1.6%	sauerkraut
2 others	2	1.1%	gauge
Total Occurrences - 191 words			

Table 6

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ay

Phoneme	Number of Words	Percentage	Example
/e/	142	89.3%	day
/ɪ/	10	6.3%	always
/ai/	2	1.3%	aye
/ɛ/	1	0.7%	says
4 Others	4	2.5%	picayune
Total Occurrence - 159 words			

Table 7

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ea

Phoneme	Number of Words	Percentage	Example
/i/	318	53.1%	each
/ɛ/	135	22.6%	bread
/iə/	45	7.5%	cereal
/ɪə/	24	4.0%	area
13 Others	77	12.8%	ocean, great
Total Occurrence - 599 words			

Table 8

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ee

Phoneme	Number of Words	Percentage	Example
/i/	293	91.8%	bleed
/ɪ/	12	3.8%	creek
/iɛ/	8	2.5%	preempt
/e/	3	1.0%	matinee
2 Others	3	0.9%	reelection
Total Occurrence - 319 words			

Table 9

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ia

Phoneme	Number of Words	Percentage	Example
/ɪə/	238	41.0%	alias
/ə/	124	21.3%	special
/ɪe/	77	13.2%	humiliate
/aiə/	56	9.7%	giant
11 Others	86	14.8%	piano

Total Occurrence - 581 words

Table 10

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ie

Phoneme	Number of Words	Percentage	Example
/i/	73	26.7%	movie
/ɪ/	42	15.3%	sieve
/ɪə/	33	12.0%	audience
/aiə/	27	9.9%	diet
11 Others	99	35.1%	friend, lie

Total Occurrence - 274 words

Table 11

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster io

Phoneme	Number of Words	Percentage	Example
/ə/	1138	88.0%	action
/ɪə/	33	2.5%	idiot
/jə/	30	2.3%	onion
/aiə/	29	2.2%	lion
8 Others	63	5.5%	trio

Total Occurrence - 1,293 words

Table 12

Frequency of Occurrence of the 5 Most Common Pronunciations of the Vowel Cluster iou

Phoneme	Number of Words	Percentage	Example
/ə/	75	54.0%	delicious
/ɪə/	59	42.4%	furious
/jə/	3	2.2%	rebellious
/u/	1	0.7%	Sioux
/aijə/	1	0.7%	pious

Total Occurrence - 139 words

Table 13

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster oa

Phoneme	Number of Words	Percentage	Example
/o/	94	75.2%	oat
/oə/	13	10.4%	coalition
/ɔ/	9	7.2%	broad
/oæ/	6	4.8%	coagulate
3 Others	3	2.4%	oasis

Total Occurrence - 125 words

Table 14

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster oi

Phoneme	Number of Words	Percentage	Example
/ɔi/	104	80.0%	coin
/oɪ/	18	13.8%	coincide
/ə/	3	2.3%	porpoise
/ui/	2	1.5%	doing
3 Others	3	2.4%	chamois

Total Occurrence - 130 words

Table 15

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster oo

Phoneme	Number of Words	Percentage	Example
/u/	194	62.2%	boot
/ʊ/	84	26.9%	foot
/ə/	23	7.4%	flood
/oa/	6	1.9%	zoology
3 Others	5	1.6%	brooch

Total Occurrence - 312 words

Table 16

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ou

Phoneme	Number of Words	Percentage	Example
/au/	238	50.1%	ounce
/ə/	181	38.2%	touch
/u/	30	6.3%	soup
/o/	13	2.7%	soul
6 Others	13	2.7%	should

Total Occurrence - 475 words

Table 17

Frequency of Occurrence of the 3 Most Common Pronunciations of the Vowel Cluster ow

Phoneme	Number of Words	Percentage	Example
/o/	131	51.2%	own
/au/	122	47.7%	cow
/a/	3	1.1%	knowledge

Total Occurrence - 256 words

Table 18

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ua

Phoneme	Number of Words	Percentage	Example
/uə/	44	42.3%	actual
/ue/	14	13.5%	fluctuate
/jue/	13	12.5%	evacuate
/juə/	11	10.6%	annual
9 Others	22	21.1%	language

Total Occurrence - 104 words

Table 19

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ue

Phoneme	Number of Words	Percentage	Example
/u/	25	23.1%	blue
/ju/	24	22.2%	value
/#/	23	21.3%	tongue
/uə/	14	13.0%	cruel
12 Others	22	20.4%	guess

Total Occurrence - 108 words

Table 20

Frequency of Occurrence of the 4 Most Common Pronunciations of the Vowel Cluster ui

Phoneme	Number of Words	Percentage	Example
/juɪ/	21	20.6%	ambiguity
/u/	19	18.6%	fruit
/ɪ/	18	17.6%	build
/wɪ/	18	17.6%	penguin
4 Others	26	25.6%	ruin

Total Occurrence - 102 words

III

SELECTION OF VOWEL CLUSTERS FOR STUDY

It was decided that rather than testing all vowel clusters, a representative subset of the total array of vowel clusters would permit sufficient analysis of children's vowel cluster pronunciation behavior. The two principal criteria used for selection of vowel clusters were frequency of occurrence and phonemic representation.

All vowel clusters occurring in fewer than 100 words were eliminated; these totaled 44. The remaining 17 were analyzed to determine the range of their sound correspondences. To test children's pronunciations of the spectrum of vowel clusters it was deemed necessary to include: (1) some clusters which have one principal pronunciation, such as ai→/e/ (gain), oa→/o/ (boat), and au→/ɔ/ (pause); (2) clusters which have two principal pronunciations such as oo→/u/ (food) or /ʊ/ (good), and ow /o/ (grow) and /au/ (plow); and (3) clusters with more than two main pronunciations such as ie→/i/ (movie), /ɪ/ (sieve) and /ai/ (die); and ou→/au/ (proud), /u/ (group) and /ə/ (famous). The cluster ay was included because of its alternation with ai in word position, and ea was included because of its frequency. Though io is the most frequent vowel cluster, it was omitted because nearly 90% of the time it occurs in /ʃən/ syllables as in nation and passion.

Based upon the preceding criteria, these nine vowel clusters, ai, au, ay, ea, ie, oa, oo, ou, and ow, appeared to comprise a representative cross-section of all vowel clusters. Further, they account for nearly half of all the occurrences of the original 61 vowel clusters in the 20,000-word corpus. By testing each of them in a variety of environments, a manageable instrument could be constructed.

IV

LETTER-SOUND CORRESPONDENCES OF THE NINE SELECTED VOWEL CLUSTERS IN THE 1000 MOST FREQUENT ENGLISH WORDS

The pronunciation frequencies in the 20,000-word corpus discussed and tabled previously are based on word types. That is, each word was counted only once regardless of its frequency. Common vowel cluster words such as would, could, and should affected the pronunciation proportions no more than such rarely used words as brooch and ooze.

To provide another basis for the analysis of children's pronunciations of vowel clusters in relation to actual pronunciation frequencies, an analysis of word frequencies was required. One possible influence on children's pronunciations could be the employment of either matching or maximizing strategies related to pronunciations of common English words. For example ou is /au/ (ounce) in 50% of the words in which it occurs but is /u/ (could) in only 1%. Yet the /u/ pronunciation occurs in three highly frequent words. Would children's pronunciations of vowel clusters in unfamiliar words be more likely to reflect the type count or token count?

In 1967 Kučera and Francis published an exhaustive computational analysis of English words. The corpus consists of 1,014, 232 words of natural-language text in 15 different genre, and includes 50,406 distinct words (types). Their analysis ranks these 50,000 words on the basis of their frequencies in the more than a million running words. For example, the is the most frequent word, occurring 69,971 times, while accordion is one of the most infrequent, occurring only once.

The present investigators analyzed the 1000 most frequent words in the Kučera and Francis corpus to determine the frequency of pronunciation of the nine vowel clusters based on tokens. That is, each word was multiplied by its number of occurrences. For example, in the 1000-word corpus, there are five words which contain the au spelling. Of these five, four have the /ɔ/ pronunciation (because, etc.), while one has the /æ/ pronunciation (laugh). If the pronunciation frequencies were based on types, /ɔ/ would equal 80% and /æ/ 20%. With a token description, based on number of occurrences of each word in a million running words, /ɔ/ is equal to 91.43% and /æ/ to 8.57%. Pronunciations were derived from Kenyon and Knott's A Pronouncing Dictionary of American English (1953). Words in this subset of 1000 words occurred from 106 to 69,971 times per million running words. It was found that approximately 20% of the words in this corpus contain vowel clusters, compared with a third of the words in the 20,000-word (type) corpus.

Table 21 compares the pronunciation proportions of the type and token corpora for the vowel clusters selected for this study.

The following table shows that for some vowel clusters (e.g., ow, au) there is little difference between type and token pronunciation frequencies, while for others (e.g., ou, ie) the differences are considerable.

Table 21

A Comparison of Vowel Cluster Pronunciation Proportions
in the 20,000-Word and 1,000-Word Corpora

Vowel Cluster	Phoneme	Type Corpus (20,000)		Token Corpus (1,000)		Example
		Words	Per Cent	Words	Per Cent	
ai	/e/	260	85.8%	10	27.4%	bait
	/ə/	20	6.6%	2	9.1%	villain
	/ɪ/	6	2.0%	2	26.1%	captain
	/ɛ/	3	1.0%	1	38.3%	said
	others	10	4.6%	0	0.0%	
ay	/e/	142	89.3%	14	89.5%	days
	/ɪ/	10	6.3%	1	7.3%	always
	/ai/	2	1.3%	0	0.0%	aye
	/ɛ/	1	0.7%	1	3.2%	says
	others	4	2.5%	0	0.0%	
au	/ɔ/	175	91.6%	4	91.4%	cause
	/o/	6	3.1%	0	0.0%	chauffeur
	/æ/	5	2.6%	1	8.6%	laugh
	/au/	3	1.6%	0	0.0%	sauerkraut
	others	2	1.1%	0	0.0%	
ea	/i/	318	53.1%	22	57.4%	each
	/ɛ/	135	22.6%	9	22.8%	bread
	/iə/	45	7.5%	3	8.1%	cereal
	/e/	12	2.0%	2	10.2%	great
	others	89	14.8%	1	1.5%	
ie	/i/	73	26.7%	8	46.6%	movie
	/ai/	25	9.1%	2	10.7%	lie
	/ɛ/	5	1.8%	2	11.4%	friend
	/aiə/	4	1.5%	2	14.3%	diet
	others	166	60.9%	3	16.7%	
oa	/o/	94	75.2%	2	100.0%	boat
	/oə/	13	10.4%	0	0.0%	coalition
	/ɔ/	9	7.2%	0	0.0%	broad
	/oæ/	6	4.8%	0	0.0%	coagulate
	others	3	2.4%	0	0.0%	
oo	/u/	194	62.2%	8	47.8%	boot
	/ʊ/	84	26.9%	7	50.0%	foot
	/ə/	23	7.4%	1	2.2%	flood
	/oa/	6	1.9%	0	0.0%	zoology
	others	5	1.6%	0	0.0%	
ou	/au/	238	50.1%	15	36.4%	ounce
	/ə/	181	38.2%	7	7.9%	touch
	/u/	30	6.3%	4	22.5%	soup
	/ʊ/	6	1.3%	5	25.9%	should
	others	20	4.1%	4	7.2%	
ow	/o/	131	51.2%	15	46.7%	own
	/au/	122	47.7%	6	51.4%	cow
	/a/	3	1.1%	1	1.9%	knowledge
	others	0	0.0%	1	0.0%	

V
DESCRIPTION OF THE INSTRUMENT

To measure children's pronunciations of vowel clusters in unfamiliar words, it was essential that real words not be used. If real words had been used it is likely that most subjects would have been familiar with some of them and consequently the test results would have revealed little more than a child's pronunciation of words with which he was familiar. The dependent variable, pronunciation of familiar vowel clusters in unfamiliar contexts, could only be assessed by constructing synthetic words containing the nine vowel clusters. The principal guideline followed in the construction of these words was linguistic plausibility. It was essential that the synthetic words resemble real words in both appearance and potential pronunciation sound. For example, many consonant clusters appear only in initial word position in modern English spelling; dr, fl, fr, gl, gr, sm, etc., while others occur only in final position; ck, nt, ll, etc. To be plausible, synthetic words had to be consonant with the vagaries of English spelling. However, it was important not to choose forms with possible pronunciations so close to real English words that they might bias the responses.

The word positions of the vowel clusters included were controlled to reflect their position frequencies in the 20,000-word corpus. These positions were ascertained from the analysis presented in Table 22. In the construction of the synthetic words, the choice of preceding and following consonants was based on further examination of the 20,000-word corpus. Since ee is never followed by g nor is ie preceded by c in English, such sequences were avoided.

The first draft of the synthetic word list, containing 10 synthetic words for each of the nine vowel clusters was submitted to a linguist, a psycholinguist, a reading specialist,

Table 22

Word Positions of the 9 Selected Vowel Clusters in the 20,000 Word Corpus

Vowel Cluster	Number of Words	Initial Position		Medial Position		Final Position	
		Words	Per Cent	Words	Per Cent	Words	Per Cent
ai	303	7	2.3%	294	97.0%	2	0.7%
ay	159	1	0.6%	48	30.2%	110	69.2%
au	191	53	27.7%	136	71.2%	2	1.0%
ea	599	16	2.7%	559	93.3%	24	4.0%
ie	274	0	0.0%	247	88.0%	33	12.0%
oa	125	7	6.0%	112	90.0%	6	4.0%
oo	312	2	0.6%	294	94.2%	16	5.1%
ou	475	37	7.8%	433	91.2%	5	1.0%
ow	256	4	1.6%	161	62.9%	91	35.6%

and a psychologist as a further check on content validity. As a result of their evaluation, several items were deleted because of their similarity to real words in either appearance or sound, and additional synthetic words were added.

In addition to the 90 synthetic words containing vowel clusters (ten each of the nine vowel clusters: ai, ay, au, ea, ie, oa, oo, ou, and ow) 10 check items were included to determine children's attentiveness to the task. Five of these were real words and five were synthetic words with predictable letter-sound correspondences (e.g., pid, p→/p/).

The 100 items were divided into two halves (A and B), each composed of five synthetic words for each vowel cluster and five check items. Using a table of random numbers, each 50-item subtest was arranged in two orderings. The four orderings were designated A1, A2, B1, and B2.

Two methods of testing were used in the study (the results will be reported in a later monograph): oral pronunciation of the synthetic words and a multiple-choice paper-and-pencil test containing synthetic words and real words similar in sound to the pronunciations of the vowel cluster being tested (see Appendix). Both forms of the test, oral and multiple-choice, contained the same words in the same sequences. On the four oral tests, each synthetic word was printed on a flash card; on the multiple-choice tests, the synthetic words and response choices were duplicated on two pages.

Three real words were offered as multiple-choice response items for each synthetic word. These response words contained at least two of the most frequent pronunciations of the vowel cluster in the 20,000-word corpus. The response words were selected from Clarence R. Stone's Revision of the Dale List, 769 Easy Words (Spache, 1960), words which, purportedly, most children can read by the end of the First Grade. In no case were the vowel sounds in the real words spelled the same way as the vowel cluster in the synthetic word being tested. To control for order effects, the response items for each vowel cluster were randomly assigned to each subtest ordering. As an example, Table 23 presents two synthetic words used to test the vowel cluster ea, and shows their test form, item number, and response sequence.

In summary, there were 100 test items of which 90 were synthetic words (ten for each of the nine vowel clusters), five were real words, and five were synthetic words with predictable letter-sound correspondences. The five real words and the five predictable synthetic words were included as control items.

Table 23

An Example of Test Form, Item Number, and Response Sequence of Two Synthetic Words

Test Form	Item Number	Synthetic Word	Response Sequence		
A1	3	polead	be	bed	baby
A2	50	polead	be	baby	bed
B1	28	deach	bed	be	baby
B2	15	deach	baby	bed	be

The research study (which will be reported later) was conducted in three stages: a pilot study, Investigation One, and Investigation Two. During the pilot study and Investigation One, both oral and multiple-choice forms of the test were used. On the multiple-choice test, each synthetic word was followed by three real word response items. All three response items for each synthetic word contained phonemes represented by that vowel cluster in the 20,000-word corpus.

Each test item was typed on a 5" x 7" flash card in primary type lower case letters. The flash cards were arranged in sequences identical to tests A1, A2, B1, and B2.

The multiple-choice test was designed to be administered either individually or to groups. The pupil's task was to circle a real word from a choice of three whose underlined letters were, he felt, closest in sound to the underlined letters in the synthetic word. The oral pronunciation test was administered individually. Each subject viewed each synthetic word on a flash card and pronounced it into a tape recorder. Later, phonemic transcriptions of the tape recording were made.

Prior to Investigation Two the response words for each item were modified to account for oral pronunciations given during Investigation One. To accomplish this, four response words were offered for each synthetic word, instead of three (see Appendix). The test was not designed to see whether children pronounced vowel clusters in synthetic words correctly or incorrectly, but rather, which of the correct pronunciations they preferred. This would permit an analysis of factors which possibly influence children's pronunciation of vowel clusters.

A discussion of the findings of Investigations One and Two will be reported after completion of the analysis.

00021

REFERENCES

Calfee, R. C., Venezky, R. L., & Chapman, R. S. Pronunciation of Synthetic Words with Predictable and Unpredictable Letter-Sound Correspondences. Technical Report No. 71, Madison: The University of Wisconsin, 1969.

Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudorf, E. H., Jr. Phoneme-Grapheme Correspondences as Cues to Spelling Improvement. USOE Research Program Monograph 16, Washington, D. C.: Government Printing Office, 1966.

Kenyon, J. S., & Knott, T. A. A Pronouncing Dictionary of American English. Massachusetts: G. & C. Merriam Co., 1953.

Kučera, H., & Francis, W. Computational Analysis of Present-Day American English. Providence, R. I.: Brown University Press, 1967.

Spache, G. Good Reading for Poor Readers. Champaign, Illinois: Gavard Press, 1960, pp. 120-122.

Thorndike, E. Thorndike-Century Senior Dictionary. New York: D. Appleton-Century, 1941.

Venezky, R. L. "A Computer Program for Deriving Spelling-to-Sound Correlations." In H. Levin (Ed.), A Basic Research Program on Reading. Ithaca: Cornell University, 1963.

Venezky, R. Unpublished Computer Print-Out. Stanford University, 1965.

Venezky, R. L. & Weir, R. H. A Study of Selected Spelling-to-Sound Correspondence Patterns. (Final Report, Cooperative Research Project No. 3090). Stanford: Stanford University, 1966.

Venezky, R. L. "The Basis of English Orthography." Acta Linguistica Hafniensia, Vol. X, No. 2, Copenhagen, 1967. (a)

Venezky, R. L. "English Orthography: Its Graphical Structure and Its Relation to Sound." Reading Research Quarterly, II/3, (Spring, 1967), pp. 75-105. (b)

Weir, R. H. Formulation of Grapheme-Phoneme Correspondence Rules to Aid in the Teaching of Reading. (Final Report, Cooperative Research Project, No. S-039) Stanford: Stanford University, 1964.

Weir, R. H. & Venezky, R. L. Rules to Aid in the Teaching of Reading. (Final Report, Cooperative Research Project No. 2584). Stanford: Stanford University, 1965.

00022

APPENDIX

GPO 818-670-4

00023

COMPOUND VOWELS - MULTIPLE-CHOICE TEST - LIST A-1

This is a test of how you pronounce unfamiliar English words. At the far left of each page, in every row of the test, there is a short English-like word that you probably have never seen before. On the right, in each row, there are four words that you already know how to say. First decide how you would say the new word on the left, and notice what sound you make for the underlined letter. Then circle the word on the right that has that same sound for its underlined letters.

Here are three examples:

A.	mip	him	bed	go	say
B.	gan	let	call	dad	so
C.	pode	baby	no	cow	me

As you work through the test be careful to check which letters are underlined in each word. It could be any one of them or two together.

These are the words you know that will be on the test. Listen to the sound of the underlined letters as you read these words with me.

put	gave	no	but	be
ran	cow	my	ball	top
out	bed	to	him	

00024

1.	thaim	my	bed	him	gave
2.	chause	ran	ball	no	cow
3.	polead	but	be	bed	gave
4.	tay	my	him	gave	bed
5.	coad	ball	cow	top	no
6.	pid	him	my	put	but
7.	abiek	my	bed	be	him
8.	monood	no	to	put	but
9.	coudry	cow	but	to	put
10.	frowl	ball	top	out	no
11.	man	him	ball	gave	ran
12.	pokay	my	bed	gave	him
13.	kaidon	gave	my	him	bed
14.	saut	cow	ran	ball	no
15.	dease	be	but	bed	gave
16.	mullow	ball	out	top	no
17.	manous	to	but	cow	put
18.	yook	to	no	put	but
19.	yiet	my	we	him	bed
20.	smoal	top	cow	no	ball
21.	droon	no	but	to	put
22.	slaum	no	cow	ball	ran
23.	oan	ball	cow	top	no
24.	dat	put	bed	no	ran
25.	bease	be	but	gave	bed
26.	ploub	put	to	cow	but
27.	comiel	my	be	him	bed
28.	gowl	no	top	out	ball
29.	chaig	bed	my	gave	him
30.	betay	gave	my	him	bed
31.	feel	him	be	bed	my
32.	baish	him	gave	my	bed
33.	trould	to	but	put	cow
34.	vayt	bed	gave	my	him
35.	wies	my	him	bed	be
36.	theat	gave	be	but	bed
37.	aclow	top	out	ball	no
38.	broam	cow	ball	no	top
39.	naugh	ran	no	ball	cow
40.	frool	no	to	but	put

41.	goag	top	cow	ball	no
42.	cote	no	top	to	out
43.	porie	my	be	bed	him
44.	dusaig	him	bed	my	gave
45.	bloose	to	put	no	but
46.	ays	bed	my	gave	him
47.	ealod	bed	be	gave	but
48.	fough	to	but	cow	put
49.	haup	no	ball	ran	cow
50.	zown	out	no	ball	top

00026

COMPOUND VOWELS - MULTIPLE-CHOICE TEST - LIST A-2

This is a test of how you pronounce unfamiliar English words. At the far left of each page, in every row of the test, there is a short English-like word that you probably have never seen before. On the right, in each row, there are four words that you already know how to say. First decide how you would say the new word on the left, and notice what sound you make for the underlined letter. Then circle the word on the right that has that same sound for its underlined letters.

Here are three examples:

A.	mip	him	bed	go	say
B.	gan	let	call	dad	so
C.	pode	baby	no	cow	me

As you work through the test be careful to check which letters are underlined in each word. It could be any one of them or two together.

These are the words you know that will be on the test. Listen to the sound of the underlined letters as you read these words with me.

put	gave	no	but	be
ran	cow	my	ball	top
out	bed	to	him	

00027

1.	saut	no	cow	ball	ran
2.	porie	be	bed	my	him
3.	ays	my	gave	him	bed
4.	coad	cow	ball	no	top
5.	frowl	top	out	ball	no
6.	oan	so	ball	top	cow
7.	dease	bed	be	but	gave
8.	smoal	cow	ball	top	no
9.	pid	but	him	put	my
10.	ealod	but	gave	bed	be
11.	manous	put	to	cow	but
12.	yook	but	to	no	put
13.	thaim	him	bed	gave	my
14.	fough	put	cow	but	to
15.	bloose	put	but	to	no
16.	vayt	him	gave	bed	my
17.	zown	top	ball	out	no
18.	dat	bed	no	put	cow
19.	baish	gave	my	bed	him
20.	frool	no	put	but	to
21.	theat	but	bed	be	gave
22.	aclow	no	top	out	ball
23.	haup	ball	cow	ran	no
24.	wies	him	be	my	bed
25.	gowl	top	out	ball	no
26.	kaidon	bed	my	gave	him
27.	ploub	cow	put	but	to
28.	droon	to	but	put	no
29.	abiek	him	bed	my	be
30.	cote	cow	ran	no	to
31.	coudry	put	cow	but	to
32.	naugh	cow	ball	ran	no
33.	pokay	my	gave	him	bed
34.	broam	no	top	ball	cow
35.	slaum	ball	no	cow	ran
36.	mullow	out	top	no	ball
37.	feel	to	be	bed	him
38.	betay	him	bed	my	gave
39.	bease	gave	we	but	bed
40.	goag	cow	ball	no	top

00028

41.	chaig	my	him	bed	gave
42.	trould	put	but	cow	to
43.	dusaig	bed	him	gave	my
44.	chause	cow	ball	ran	no
45.	comiel	be	my	bed	him
46.	monood	put	but	no	to
47.	man	gave	ran	ball	him
48.	yiet	him	bed	my	be
49.	tay	him	my	bed	gave
50.	polead	be	gave	but	bed

00029

COMPOUND VOWELS - MULTIPLE-CHOICE TEST - LIST B-1

This is a test of how you pronounce unfamiliar English words. At the far left of each page, in every row of the test, there is a short English-like word that you probably have never seen before. On the right, in each row, there are four words that you already know how to say. First decide how you would say the new word on the left, and notice what sound you make for the underlined letter. Then circle the word on the right that has that same sound for its underlined letters.

Here are three examples:

A.	mip	him	bed	go	say
B.	gan	let	call	dad	so
C.	pode	baby	no	cow	me

As you work through the test be careful to check which letters are underlined in each word. It could be any one of them or two together.

These are the words you know that will be on the test. Listen to the sound of the underlined letters as you read these words with me.

put	gave	no	but	be
ran	cow	my	ball	top
out	bed	to	him	

00030

1.	boys	gave	bed	him	to
2.	loat	cow	ball	no	top
3.	chays	him	my	gave	bed
4.	sprow	out	top	ball	no
5.	houn	put	cow	to	but
6.	areak	bed	be	gave	but
7.	mauf	no	cow	ball	ran
8.	thood	but	no	put	to
9.	mief	be	my	bed	him
10.	ogaim	my	bed	gave	him
11.	stappow	top	ball	out	no
12.	trak	put	cow	gave	no
13.	gies	him	bed	be	my
14.	poad	no	ball	top	cow
15.	poom	put	but	to	no
16.	konay	my	him	gave	bed
17.	koump	cow	put	but	to
18.	laip	my	bed	gave	him
19.	fead	but	bed	be	gave
20.	paud	ball	ran	cow	no
21.	blay	him	bed	my	gave
22.	yoap	cow	ball	no	top
23.	blue	put	to	no	ball
24.	poup	but	to	cow	put
25.	mook	no	to	but	put
26.	prient	bed	my	be	him
27.	trown	top	out	ball	no
28.	deach	bed	but	be	gave
29.	aucol	cow	ball	no	ran
30.	taise	gave	him	my	bed
31.	frean	but	bed	be	gave
32.	wouth	cow	but	put	to
33.	sunt	ran	be	him	put
34.	spows	no	ball	top	out
35.	hauge	ball	no	ran	cow
36.	sloot	no	to	but	put
37.	blaing	bed	my	gave	him
38.	toang	no	ball	top	cow
39.	ziegle	bed	my	be	him
40.	onchay	him	bed	my	gave

GPO 818-670-3

41.	fowt	no	ball	out	top
42.	biesh	him	bed	my	be
43.	toul	put	but	cow	to
44.	saw	ball	no	to	ran
45.	boase	no	cow	ball	top
46.	ducay	my	him	bed	gave
47.	raitel	gave	bed	him	my
48.	voop	put	but	no	to
49.	aup	no	ball	cow	ran
50.	cheam	but	be	gave	bed

00032

COMPOUND VOWELS - MULTIPLE-CHOICE TEST - LIST B-2

This is a test of how you pronounce unfamiliar English words. At the far left of each page, in every row of the test, there is a short English-like word that you probably have never seen before. On the right, in each row, there are four words that you already know how to say. First decide how you would say the new word on the left, and notice what sound you make for the underlined letter. Then circle the word on the right that has that same sound for its underlined letters.

Here are three examples:

A.	mip	him	bed	go	say
B.	gan	let	call	dad	so
C.	pode	baby	no	cow	me

As you work through the test be careful to check which letters are underlined in each word. It could be any one of them or two together.

These are the words you know that will be on the test. Listen to the sound of the underlined letters as you read these words with me.

put	gave	no	but	be
ran	cow	my	ball	bed
to	him	out	bed	

00033

1.	onchay	gave	him	my	bed
2.	aucol	no	ran	cow	ball
3.	spows	out	ball	top	no
4.	ogaim	him	my	bed	gave
5.	prient	my	bed	him	be
6.	frean	be	but	gave	bed
7.	biesh	my	be	bed	him
8.	trak	gave	no	put	cow
9.	stappow	no	top	ball	out
10.	paud	cow	no	ball	ran
11.	raitel	him	my	gave	bed
12.	mauf	ball	ran	cow	no
13.	trown	ball	no	out	top
14.	thood	put	to	no	but
15.	deach	gave	but	bed	be
16.	ducay	gave	my	him	bed
17.	blue	no	ball	to	put
18.	voop	but	no	to	put
19.	aup	ran	ball	no	cow
20.	toul	to	put	cow	but
21.	blaing	my	him	bed	gave
22.	fead	be	gave	bed	but
23.	chars	bed	gave	him	my
24.	loat	ball	top	no	cow
25.	sloot	put	to	no	but
26.	konay	gave	my	bed	him
27.	sunt	him	put	be	ran
28.	taise	my	gave	bed	him
29.	poad	ball	cow	top	no
30.	wouth	put	to	cow	but
31.	cheam	bed	gave	be	but
32.	mook	to	put	no	but
33.	hauge	cow	no	ran	ball
34.	toang	ball	no	cow	top
35.	mief	him	be	bed	my
36.	boys	to	gave	him	bed
37.	fowt	ball	top	no	out
38.	areak	be	gave	bed	but
39.	saw	ran	no	ball	to
40.	poup	cow	put	but	to

00034

41.	boase	top	no	cow	ball
42.	blay	my	him	bed	gave
43.	poom	no	but	to	put
44.	koump	but	to	put	cow
45.	gies	my	him	be	bed
46.	houn	cow	to	put	but
47.	yoap	ball	top	no	cow
48.	laip	bed	gave	my	him
49.	ziegle	him	my	be	bed
50.	sprow	no	ball	top	out

GPO 816-670-2