DOCUMENT RESUME

ED 073 443	24	CS 000 424

AUTHOR	Venezky, Richard L.; And Others
TITLE	The Development of Letter-Sound Generalizations from Second through Sixth Grade. Technical Report.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Research and Development Centers Branch.
REPORT NO	WRDCCL-TR-231
BUREAU NO	BR-5-0216
PUB DATE	Sep 72
CONTRACT	OEC-5-10-154
NOTE	26p.; Report from the Project on Reading and Related Language Arts Basic Prereading Skills: Identification and Improvement
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Elementary School Students; *Graphemes; *Phonemes; Reading; Reading Ability; Reading Instruction; *Reading Research; Word Recognition
IDENTIFIERS	*Letter Sound Patterns

ABSTRACT

The purpose of this study was to determine the development of four specific letter-sound patterns from second through sixth grade: invariant consonants, long and short vowels, "c," and "g." A 69-item list was presented to second-, fourth-, and sixth-grade subjects in one of two random orders. Oral responses were tape recorded, transcribed by graduate students trained in phonics, and coded as correct, plausible, or incorrect. Separate analyses of variance were run on each of the four pattern categories. Some of the results indicated that although ability to generalize each of the four patterns increased from second through sixth grade, there were striking differences in the acquisition of the patterns. At the second- and, to a lesser degree, the fourth-grade levels, there was little difference in ability between the best and worst readers to generalize invariant consonant correspondences when the letters occurred in initial position, but large differences were evident for medial and final positions, with the poorer readers showing the least ability. For the "c" patterns, learning of the correct pronunciation for "c" before "e," "i," or "y" occurred slowly and failed to reach 60 percent by sixth grade. For the "c" pronounced as /s/ pattern the correct responses were high at all grades. (Author/WR)

Technical Report No. 231

THE DEVELOPMENT OF LETTER-SOUND GENERALIZATIONS FROM SECOND THROUGH SIXTH GRADE

by

Richard L. Venezky, Robin S. Chapman, and Robert C. Calfee

Report from the
Project on Reading and Related Language Arts
Basic Prereading Skills: Identification and Improvement

Richard L. Venezky, Principal Investigator

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

September 1972

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

Statement of Focus

Individually Guided Education (IGE) is a new comprehensive system of elementary education. The following components of the IGE system are in varying stages of development and implementation: a new organization for instruction and related administrative arrangements; a model of instructional programing for the individual student; and curriculum components in prereading, reading, mathematics, motivation, and environmental education. The development of other curriculum components, of a system for managing instruction by computer, and of instructional strategies is needed to complete the system. Continuing programmatic research is required to provide a sound knowledge base for the components under development and for improved second generation components. Finally, systematic implementation is essential so that the products will function properly in the IGE schools.

The Center plans and carries out the research, development, and implementation components of its IGE program in this sequence: (1) identify the needs and delimit the component problem area; (2) assess the possible constraints—financial resources and availability of staff; (3) formulate general plans and specific procedures for solving the problems; (4) secure and allocate human and material resources to carry out the plans; (5) provide for effective communication among personnel and efficient management of activities and resources; and (6) evaluate the effectiveness of each activity and its contribution to the total program and correct any difficulties through feedback mechanisms and appropriate management techniques.

A self-renewing system of elementary education is projected in each participating elementary school, i.e., one which is less dependent on external sources for direction and is more responsive to the needs of the children attending each particular school. In the IGE schools, Center-developed and other curriculum products compatible with the Center's instructional programing model will lead to higher student achievement and self-direction in learning and in conduct and also to higher morale and job satisfaction among educational personnel. Each developmental product makes its unique contribution to IGE as it is implemented in the schools. The various research components add to the knowledge of Center practitioners, developers, and theorists.

Contents

List of Tables and Figures

Abstract

Letter-sound generalizations play an indispensable role in the acquisition of reading ability, regardless of the pedagogical methods employed. Their main functions are (a) to provide a self-generated check for the reader on words he thinks he has identified correctly from context or visual cues, but is not altogether sure of; and (b) to give a phonological approximation to a word which may be in a reader's listening vocabulary, but has not yet been encountered in reading. For both of these procedures, perfectly predictable letter-sound correspondences are not a necessity, since the reader has other cues to use for word identification.

In earlier studies (Johnson, 1970; Calfee, Venezky, & Chapman, 1969) it was shown that children begin to acquire letter-sound generalizations in the first year of reading instruction and continue to acquire them far beyond the time that basic reading skills are emphasized in reading programs. The concern of the study reported here was to determine the development of four specific letter-sound patterns from second through sixth grade: invariant consonants, long and short vowels, *c*, and *g*.

A 69-item list of synthetic words was prepared and presented to second-, fourth-, and sixth-grade *S*s in one of two random orders. Oral responses to each test item were tape recorded and transcribed by graduate students trained in phonetics. Responses were then coded as *correct*, *plausible*, or *incorrect*, and separate analyses of variance were run on each of the four pattern categories.

The results showed that although ability to generalize each of the four patterns increased from second through sixth grade, there were striking differences in the acquisitions of the patterns. At the second- and, to a lesser degree, the fourth-grade levels, there was little difference in ability between the best and worst readers to generalize invariant consonant correspondences when the letters (e.g., *b*, *m*, *n*) occurred in initial position, but large differences were evident for medial and final position, with the poorer readers showing the least ability. This indicates a qualitative difference in word attack strategies, with the poor readers using only the first few letters plus perhaps length for making identifications. Long- and short-vowel correspondences were acquired at nearly equal rates, with the differences between quartiles (for reading ability) being considerably less marked than in the responses to invariant consonants.

For the *c* patterns, which are theoretically more predictable than the vowel patterns, learning of the correct pronunciation for *c* before *e*, *i*, or *y* occurred extremely slowly and failed to reach 60% correct by sixth grade. For all other environments, however, generalization of the correct response started at approximately 73% correct in second grade and reached 88% correct by sixth grade. For the *c* → /s/ pattern, the percentage of correct plus plausible responses was extremely high at all grade levels, indicating a strong response bias for /k/ which we hypothesize is due to the overwhelming preponderance of *c* → /k/ exemplars in primary-level reading materials.

For the g patterns, the percentage correct for g before e, i, or y reached only about 35% by sixth grade. This, we hypothesize, is due to two effects: first, that the majority of the g words which occur in primary-level reading texts have the /g/ pronunciation for g, and second, that the first and most common initial g/e i y words which the beginning reader encounters are exceptions to the general pattern (e.g., give, get, girl).

x

I
Introduction

Background

For a skilled reader, letter-sound generalizations are of limited practical use, being invoked only for that small number of visually unfamiliar words which are met in everyday reading. But for the child learning to read English—or any other language with an alphabetic writing system—such generalizations play an indispensable role in the acquisition of reading ability. Regardless of the teaching procedures employed, all children who learn to read successfully acquire letter-sound generalizations and continue to strengthen them well past the time of emphasis on letters, sounds, and the other paraphernalia of beginning reading instruction.

Venezky (in press) speculates that letter-sound generalizations play two related roles in learning to read: first, in providing a self-generated check for the reader on words he thinks he has identified correctly from context or visual cues, but is not altogether sure of; and second, in [illegible]ing a phonological approximation to a word which may be in the reader's listening vocabulary, but has not been encountered before in reading. For both of these procedures, perfectly predictable letter-sound correspondences are not a necessity, since the reader has other cues to use for word identification. What is important is that he produce something which is close enough to the target so that by alteration of one or more sounds a correct match is obtained.

Support for the assertion that letter-sound generalizations are acquired in learning to read is found in studies by Calfee, Venezky, and Chapman (1969) and Johnson (1970), both of which involved pronunciations of synthetic words by readers of different scholastic levels. Johnson showed, inter alia, that by the spring of the first year of reading instruction, children could give plausible responses to the vowel digraphs in forms like cleese, thead, and bouse from letter-sound correspondences alone; Calfee, Venezky, and Chapman found that for certain variant but predictable letter-sound patterns, such as the /s/ and /k/ pronunciations of initial c (e.g., city and carpet), response accuracy continued to improve from third grade through the junior year of high school. In these studies, good readers tended, when giving incorrect responses, to choose from the set of responses which are plausible for the stimulus patterns, while poor readers gave a significantly higher number of "wild" responses. It is evident, therefore, that letter-sound generalizations are formed by readers, even for the variant patterns, since in both studies children were reading words they had almost certainly never seen before.

The development of some of the more important letter-sound generalizations, however, cannot be determined from these studies because of limitations on subjects (Calfee, Venezky, and Chapman tested students no younger than third grade) or stimulus materials (Johnson tested first, second and third graders, but only on digraph vowel spellings). Therefore, several additional studies were planned, the first of which is reported here.

Purpose

The present study was designed to explore further the development of four sets of letter-sound patterns:

1. Invariant consonants
2. c pronounced /k/ and /s/
3. g pronounced /g/ and /ǰ/
4. the long and short pronunciations of the simple vowel spellings (a, e, i, o, u, y)

The invariant consonants are those consonant letters or letter strings which typically have a single, invariant pronunciation. Actually, almost all consonants admit exceptions; for example, n is silent in hymn, f is pronounced /v/ in of, and l combines with o to give /ər/ in colonel. But for practical purposes, the subset of consonants selected for this study (b, d, l, m, n, and p) admit so few exceptions, almost none of which are introduced in elementary-school readers, that the label of invariant consonant can be applied. Because of this invariance we hypothesized that the pronunciations for these consonants would be learned early in reading instruction—and learned to a high degree of accuracy, regardless of their position in a word.[1]

The c patterns, in environments in which only /s/ or /k/ could occur, should also be learned to a high degree of accuracy since the pronunciations are highly predictable, but should be learned more slowly than the invariant consonants, due to the need to learn the proper environments for each pronunciation.[2]

Long and short vowel pronunciations for the simple vowel spellings (a, e, i, o, u, and y) are predictable on the basis of succeeding graphemic environment,[3] but admit numerous exceptions. For the final e-nonfinal e environments used in this study (e.g., dape, dap), there are exceptions for a after w (e.g., wad), o (e.g., ton, won), and u (e.g., put). We therefore hypothesized that these patterns would be learned more slowly than the c patterns and to a lower degree of accuracy.

And finally the g patterns, which parallel the c patterns in environmental conditioning,[4] should be learned most slowly and to the lowest level of accuracy, because the g → /ǰ/ pattern has more high-frequency exceptions (e.g., gear, geese, get, gift, giggle, girl) than either the c or the vowel patterns.

II
Method

Stimulus Materials

A 69-item list of synthetic words (words following English spelling patterns) was constructed to test a variety of letter-sound correspondences. The words used to test the four major patterns described in Section I are listed in Table 1. Synthetic words constructed to test the c pattern included instances of initial c before the single vowel letters a, e, i, o, u, y, which in turn were followed by a consonant cluster (CC) or consonant plus final e (CE); medial c followed by the six single vowel letters plus another consonant; and final c preceded by the six single vowel letters.

Synthetic words were constructed to test the g pattern in a similar fashion, except that y was omitted from the single vowel letter contexts. The vowel letters a, e, i, o, u were each tested twice in CVC (consonant + vowel + consonant [e.g., ran]) and CVCe (consonant + vowel + consonant + e [e.g., sane]) contexts. (However, the letter e was inadvertently tested in only one C_C context.) The six consonant letters with a single common pronunciation which occurred two or more times in each word position (initial, medial, final) were assigned to the invariant category.

Procedures

Two random orders of the 69-item list, preceded by five practice items (chung, foll, goot, lech, veeg), were prepared. The words were presented individually to Ss in one of three forms: slides made of the single items lettered in sans-serif capitals; 5" x 8" cards on which one item per card was typed, centered, in sans-serif 14-point capitals; or 5" x 8" cards typed in 14-point lower case.[5]

A Kodak Carousel slide projector was used to present the slides at low intensity on a wall or screen in the test room; minimum inter-item interval was 5 seconds. Words on cards were bound in notebooks; E turned the page as S pronounced each item.

Ss were told that the words had been made up and were asked to pronounce them as if they were real English words. Responses were recorded on a Uher 5000 tape recorder at 3 3/4 i.p.s. with a Shure lavaliere microphone. Assignment to list order and upper or lower case (for cards) was random.

Subjects

Participating in the study were 493 Ss attending second, fourth, or sixth grade in three elementary schools in Madison and Marshall, Wisconsin. The two elementary schools in Madison, Wisconsin, draw students from upper-middle class and middle class socioeconomic strata, respectively. The school located in Marshall, Wisconsin, draws from a heterogeneous rural and semirural population. All children in a grade were tested; protocols were dropped for those who did not have achievement test data. All subjects were tested in November and December of 1968.[6]

Intelligence and achievement test scores from school-administered tests were recorded for each S; since these tests differed according to school and grade, quartile splits on a summary achievement measure were computed separately for each school-grade group. Ss were assigned, post-hoc, to one of the four quartile groups defined for their school and grade.

The distribution of Ss by school, grade, sex, and achievement quartile is presented in Table 2. Of the Ss, 248 were male and 245 female.

TABLE 1
SYNTHETIC WORDS, GROUPED BY MAJOR PATTERNS

	C Pattern		G Pattern		Invariant Consonants					
	c→/k/	c→/s/	g→/g/	g→/j/	d	b	l	m	p	n
Initial	cade	cefe	gade	geme	dac	bim	lebe	mec	ponge	nade
	carg	cemp	gand	geft	dag	bime	legute	mecal	pog	nuc
	cobe	cipe	gon	gite	dap		licul			nacom
	corb	cilf	gope	gink	dape		lig			noge
	cuse	cyfe	golb							nad
	cung	cylm	gube							necy
			gurk							nigome
Medial	mecal	roce	hugan	noge	cade	cobe	cilf	agime	cipe	vune
	nacom	hacen	nigom	porge	gade	fube	cylm	bime	gope	gand
	licul	recilt	legute	agime	jode	gube	golb	cemp	dape	
		necy			nade	lebe	recilt	geme	kipe	
Final	dac		dag		gand	corb	licul	bim	cemp	vun
	mec		feg		jod	fub	mecal	cylm	kip	fon
	jic		lig		nad	golb		nacom	dap	hugan
	woc		pog			wab		nigom		gon
	ruc		sug							hacen
	zyc		carg							

Vowel in Final -e Pattern

Long Vowel	Short Vowel
dape	dap
nade	nad
kete	ket
lebe	---
bime	bim
kipe	kip
jode	jod
wome	wom
fube	fub
vune	vun

TABLE 2
DISTRIBUTION OF Ss BY SCHOOL, GRADE, SEX, AND ACHIEVEMENT QUARTILE

Grade	Quartile	Sherman (Middle) M	Sherman (Middle) F	Nakoma (Upper-Middle) M	Nakoma (Upper-Middle) F	Marshall (Rural) M	Marshall (Rural) F	Total
2	1	8	6	2	6	3	3	28
	2	9	5	3	5	10	5	37
	3	9	8	2	8	5	7	39
	4	9	10	4	6	5	9	43
	Total	35	29	11	25	23	24	
4	1	9	5	9	4	6	3	36
	2	18	4	9	6	7	6	50
	3	10	7	7	5	6	5	40
	4	10	11	7	12	4	11	55
	Total	47	27	32	27	23	25	
6	1	10	5	5	7	4	3	34
	2	10	9	4	8	5	3	39
	3	5	10	8	8	5	7	43
	4	8	13	4	11	9	4	49
	Total	33	37	21	34	23	17	

Transcription Procedures

Responses were transcribed phonemically in IPA (International Phonetic Alphabet) by one of five students, all of whom were trained in phonetic transcription. Tapes were played over Ampex 140 stereophonic headphones on an Ampex 1100 recorder. Every tenth S was independently transcribed by a second listener and disagreements noted; a third checker resolved disagreements.[7]

Coding Procedures

Each response was mated with its stimulus word and the phoneme corresponding to each letter was coded. For instance, if the stimulus word was cobe and the response /korb/, /k/ was coded for c, /o/ for o, /r/ as an insertion, /b/ for b, and silence for e. Consonant phonemes were matched only with consonant letters, and vowel phonemes with vowel letters. In the few cases where two consonant phonemes corresponded to a single consonant letter, the phoneme most similar in articulation to the expected pronunciation was matched and the other coded as an insertion. Deletion was indicated by coding silence for the letter. For order reversals (e.g., /klaim/ for cylm), phonemes were matched with the appropriate letter (e.g., /ai/ with y) and a reversal code attached. Special codes were used for no response and for aberrant responses (e.g., "peanut butter").

Coding was checked in two ways. For every 25 card images coded, two successive lines were chosen at random and independently coded by a second coder; the lines were then checked for disagreement. (Disagreements were noted about once for every 25 lines checked.) Once the data were punched, a clean-up program checked for inappropriate format; too few, too many or repeated stimulus words, incorrect number of letter matches for an item; and consonant-vowel mismatches with letters.

Categorization of Pronunciations

Correct responses for the letters c, g, and the vowels were defined as the pronunciations appropriate to the vowel letter context; the alternate common pronunciation of the letter was defined as plausible (see Table 3). Either /u/ (as in lute) or /ju/ (as in cube) was treated as a long pronunciation of u. No plausible responses were defined for the invariant category.

TABLE 3
CORRECT AND PLAUSIBLE RESPONSES

	Correct	Plausible
c Before e, i, or y	/s/	/k/
Elsewhere	/k/	/s/
g Before e or i	/ǰ/	/g/
Elsewhere	/g/	/ǰ/
Long Vowels		
a	/e/	/æ/
e	/i/	/ɛ/
i	/aɪ/	/ɪ/
o	/o/	/a/
u	/ju/ or /u/	/ə/
Short Vowels		
a	/æ/	/e/
e	/ɛ/	/i/
i	/ɪ/	/aɪ/
o	/a/	/o/
u	/ə/	/ju/ or /u/
Invariant Consonants		
b	/b/	
d	/d/	
l	/l/	none defined
m	/m/	
n	/n/	
p	/p/	

III
Results and Discussion

Invariant Consonants

For the six consonant letters which occurred two or more times in each word position (initial, medial, final), mean percentage correct scores by grade and quartile are shown in Table 4. A 3 x 3 x 2 x 4 x 2 unequal-n analysis of variance with repeated measures on the last factor, school by grade by sex by reading achievement quartile by position of the invariant letter (initial vs. medial or final), was run on the percentage scores for these 67 invariant consonants.

The results, shown in Table 5, show that the largest between-subject effects are associated with grade and quartile. Letter position was a major within-subject effect, as was the grade by letter position interaction. Smaller but significant differences were found for the grade by quartile interaction and for the three-way interaction of grade by quartile by letter position.

TABLE 4
MEAN PERCENTAGE CORRECT FOR SELECTED INVARIANT CONSONANTS

Position	Quartile	Grade 2	Grade 4	Grade 6	Average
Initial	Q1	91	95	98	95
	Q2	91	98	98	96
	Q3	97	99	100	98
	Q4	98	99	100	99
	Average	95	98	99	97
Medial and Final	Q1	66	84	93	82
	Q2	72	87	92	84
	Q3	84	95	95	92
	Q4	93	96	98	96
	Average	80	91	95	89
Averaged Across Position	Q1	78	90	96	89
	Q2	82	92	95	90
	Q3	90	97	98	95
	Q4	96	98	99	98
	Average	88	94	97	93

TABLE 5
ANALYSIS OF VARIANCE SUMMARY FOR INVARIANT CONSONANTS
(significant sources only, $p < .01$)

Source	df	MS	F
Grade (G)	2	0.72	60.36
Quartile (Q)	3	0.44	37.05
School x Q	6	0.04	3.11
G x Q	6	0.07	5.87
School x G x Sex	4	0.04	3.80
School x G x Q	12	0.03	2.39
(Between Ss	414	0.01)	
Letter Position (P)	1	1.59	344.94
P x School	2	0.02	5.26
P x G	2	0.22	48.15
P x Q	3	0.13	28.16
P x School x Q	6	0.03	6.53
P x G x Q	6	0.28	5.96
P x School x G x Q	12	0.01	2.57
(Within Ss	414	0.005)	

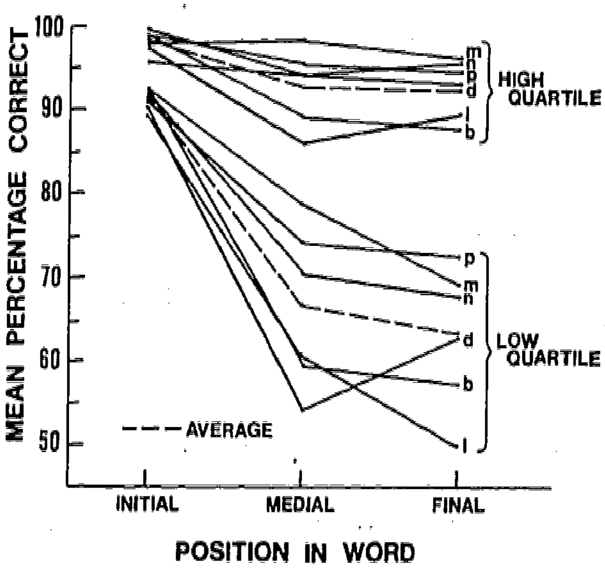

Fig. 1. Correct responses to each invariant consonant by position for the lowest and highest quartiles of second grade.

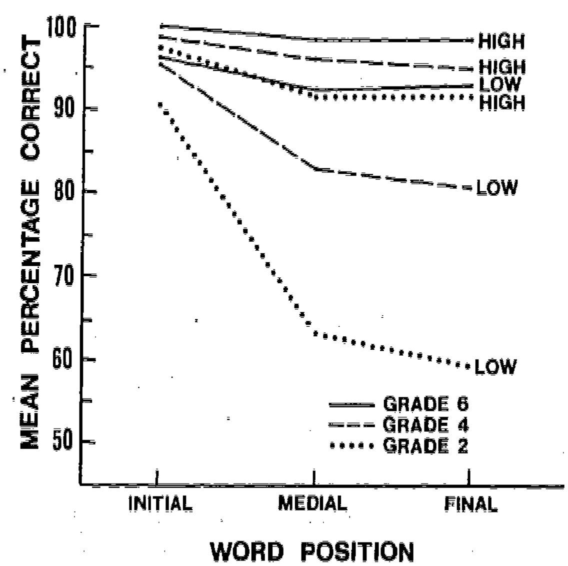

Fig. 2. Correct responses to invariant consonants by grade and position for the highest and lowest achievement quartiles.

Mean correct increased moderately from second through sixth grade, from the lowest to the highest quartiles, and from medial or final position to initial position. The interactions arise because these main effects were attenuated in the higher grades, the higher quartiles, and in initial position. Thus, more striking differences were found for quartile and position in second grade, as shown in Figure 1. Invariant consonants in initial portion are misread slightly more often by children in the lower quartiles, but the differences are modest. The same consonants in medial and final positions are misread substantially more often by the lower quartiles in the second (and fourth) grades. The fact that the same child who produces the correct letter-sound correspondence for a letter in initial position mispronounces the letter in medial position suggests that only a portion of the information is processed. For whatever reason, the children in the lower quartiles are not consistently decoding the letters to the right of the initial position. The data for the highest and lowest quartile of each grade are summarized in Figure 2.

This difference between the highest and lowest quartiles may also reflect a difference in word attack strategies. The best readers tend to observe the entire visual pattern of a word before responding to it and thus observe letter-sound correspondences equally well in all word positions. But the poorer readers, as shown by studies of oral reading errors (Swanson, 1937; Weber, 1970), tend to rely heavily upon the initial letters of a word plus perhaps word length for identification cues. Thus, correspondences for letters in medial and final position would be observed less often than those for letters in initial position. What is surprising is that due to the restricted word lists which occur in primary-school reading materials, an initial-letters plus length-identification strategy would be highly effective for words in context. (In a study by Durr [n.d.] of 80 books frequently read by primary-grade children, it was found that only 188 different words accounted for almost 70% of all running words. Within this list, 64.4% of the words could be uniquely identified [assuming a closed set] on the basis of the first two letters plus length. If context were considered, most of the remaining words could also be identified.)

TABLE 6
MEAN PERCENTAGE CORRECT FOR VOWEL PRONUNCIATIONS

Vowel Context	Quartile	Grade 2	Grade 4	Grade 6	Average
Long	Q1	36	49	68	52
	Q2	44	63	69	59
	Q3	45	77	82	69
	Q4	70	83	85	80
	Average	50	69	77	65
Short	Q1	38	56	65	54
	Q2	45	67	70	61
	Q3	63	78	74	72
	Q4	79	86	86	84
	Average	59	73	75	68
Averaged Over Context	Q1	37	52	66	53
	Q2	44	65	70	60
	Q3	54	78	78	70
	Q4	74	84	86	82
	Average	54	66	76	66

TABLE 7
ANALYSIS OF VARIANCE SUMMARY FOR LONG AND SHORT VOWELS
(significant sources only, $p < .01$)

Source	df	MS	F
School (S)	2	2.26	47.26
Grade (G)	2	3.47	72.69
Quartile (Q)	3	3.89	81.40
S x Q	6	0.28	5.80
G x Q	6	0.18	3.73
S x G x Sex	4	0.16	3.44
(Between Ss	414	0.05)	
Vowel Pattern (V)	1	0.22	9.04
V x G	2	0.22	9.35
V x G x Q	6	0.08	3.44
(Within Ss	414	0.02)	

Long and Short Vowels

A summary of the percentages correct on the vowel patterns by grade and reading achievement quartile is shown in Table 6. A 3 x 3 x 2 x 4 x 2 unequal-n repeated measures analysis of variance was performed on percentage correct scores: school by grade by sex by reading quartile by vowel pattern (long vs. short). The results (Table 7) show that quartile, grade, and school were, in that order, substantial sources of between-subject variance. The grade by quartile interaction was statistically significant, but accounted for a miniscule proportion of the overall variance.

The school effect arises from an increase in mean percentage correct from the rural school (62%), to the middle-class school (65%), to the upper-middle class school (78%). (This trend appeared in all four major patterns, but was not significant for the invariant consonants.) Grade and quartile effects result from the expected increases in mean proportion correct with increasing grade and increasing reading ability.

The overall difference in correct pronunciations between long- and short-vowel spellings was small though significant: 65% correct for long-vowel spellings, as opposed to 68% correct for short-vowel spellings.

The superiority of short vowels over long vowels in the lower grades was probably due to two factors: (a) the tendency of most reading programs to introduce the short-vowel correspondences before the long-vowel correspondences, and (b) the prepunderance of

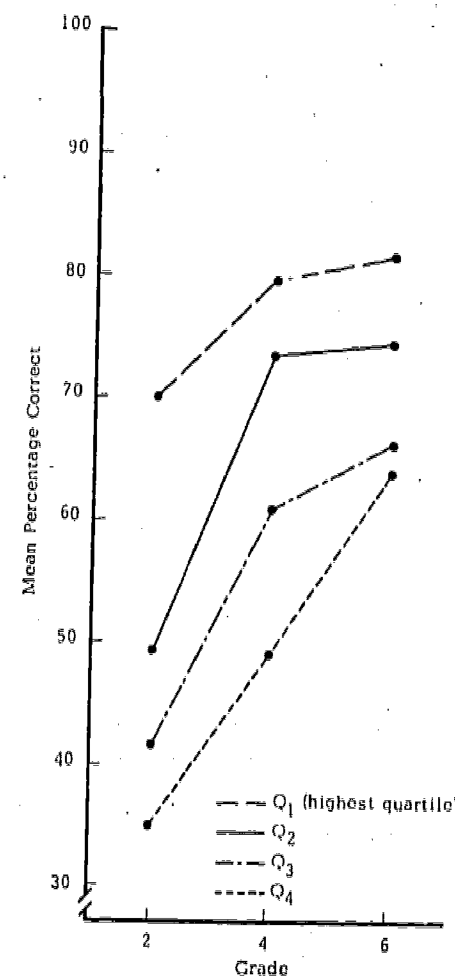

Fig. 3. Correct responses by grade and quartile to the long- and short-vowel patterns.

short-vowel words over long-vowel words in the vocabulary of introductory reading materials.

Differences Between Good and Poor Readers

Development by quartiles of correct responses for the long and short patterns combined is shown in Figure 3. The difference in mean correct between the highest (Q1) and lowest (Q4) quartile, which was large at all three grade levels, was only half as great at sixth grade as it was at second grade. Thus, rather than revealing a cumulative deficit, these data show just the opposite; i.e., that the poorer readers tended to improve relative to the better readers in letter-sound learning. In fact, from fourth to sixth grade almost all of the improvement was accounted for by the poorer readers.

Differences Among the Vowels

Judging from the correspondences for the vowel letters a, e, i, o, and u which occur

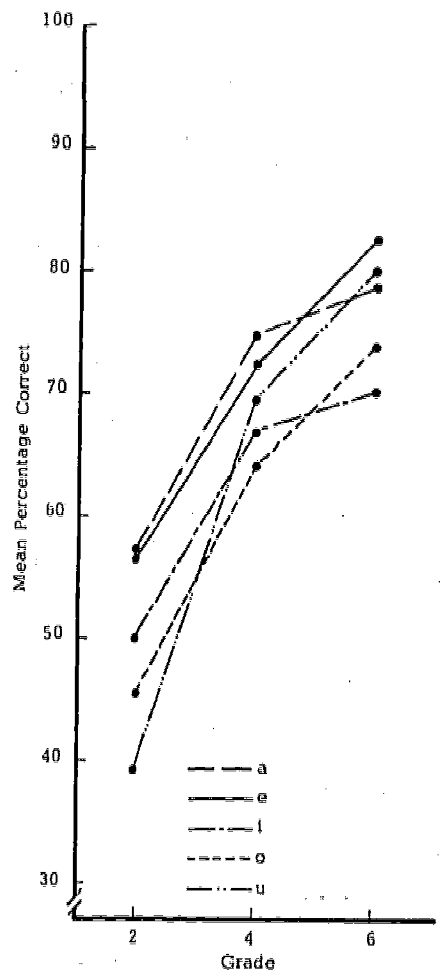

Fig. 4. Correct responses by grade to each of the long-vowel patterns.

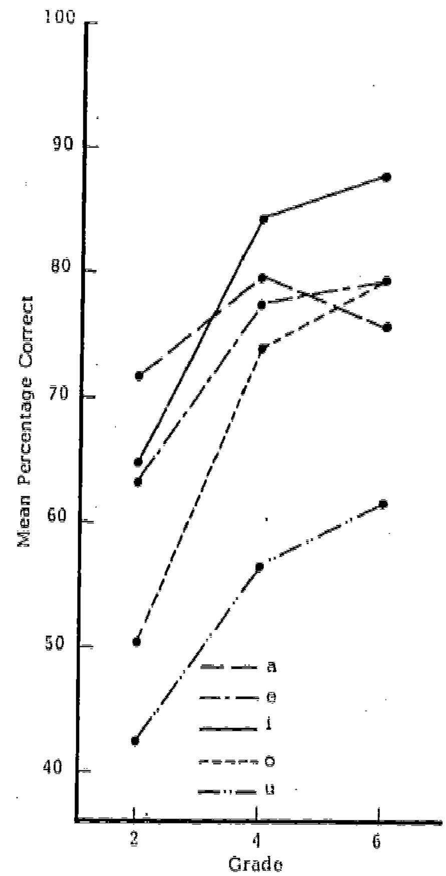

Fig. 5. Correct responses by grade to each of the short vowels.

in English words, we would anticipate a wide variation in the consistency of responses given to these letters in synthetic words. The letter o, for example, has several frequent correspondences in common English words besides the long and short ones, as shown in the example below.[8]

[u]: approve, improve, lose, move, prove, whose, do, to, two, who

[ə]: above, come, cover, lover, money

[ɔ]: off, offer, office, across, cloth, cost

The letter u, on the other hand, is almost invariantly /(j)u/ in the long vowel environment.[9] And e has few exemplars of the pattern -eCe in the vocabularies sanctified for introductory readers.

Yet the development patterns for correct responses to each of the long and short vowels, as shown in Figures 4 and 5, bore no consistent relationship to these counts. For example, correct responses to long u were statistically no different from those for long o at sixth grade, although by the examples just given we would anticipate considerably more correct responses to u. There were, nevertheless, tendencies which might be related to type counts for real words, particularly among the short-vowel pronunciations (Figure 5). The letter o, which admits the greatest number of exceptions in its short-vowel environments, consistently received a significantly lower number of mean correct responses than the other short vowels, and short i, which is nearly invariant for stressed occurrences in real words, had a substantially higher mean correct than the other vowels in both fourth and sixth grades.

Error Analyses

In the highest quartile the responses which lie outside of the correct plus plausible range were generally distributed over a variety of other vowels. The only exception was for short o at the second-grade level, where /ə/ occurred more than twice as frequently as the plausible response, /a/ (15.3% /ɔ/, 7.4% /a/). In the lowest quartile, there were relatively high percentages of deviant responses for both long and short e (both pronounced as /ɪ/) and for long and short o (both pronounced as /ɔ/).

The persistence of the /ɪ/ pronunciation to both long and short e is not traceable to correspondences which occur in English words. Except for some regional pronunciations of standard /ɛ/ as /ɪ/ before /n/ (e.g., pen, when, men), e→/ɪ/ does not occur in English. But none of the stimulus items for e in this experiment contained n following e. /ɔ/ pronunciations, on the other hand, are expected on the basis of words like boss, cost, off, and long.

The c Patterns

Scores for the c patterns by position (initial, medial), graphemic context (e, i, y; a, o, u), and grade are shown in Table 8. A repeated measures analysis of variance, school by grade by sex by reading achievement quartile by letter position (initial or medial) by graphemic context, was run on mean proportion correct scores. The results (Table 9) show that grade and quartile sources accounted for the largest proportion of the between-subject variance; smaller though statistically significant contributions were made by sex and school factors. Graphemic context was the major within-subject source, with lesser contributions from the interaction of context with grade, position, and grade by position.

The school, grade, and quartile effects showed the same tendencies as in the other major patterns, with increases in mean correct corresponding to increases in grade, reading ability, and SES of schools (from rural to middle-class to upper middle-class). In addition, girls performed somewhat better than boys (.68 vs. .62).

There was a strong response bias in all positions for /k/, giving high correct scores for c when not followed by e, i, or y and low correct scores otherwise. In initial position (see Figure 6) the percentage of correct re-

TABLE 8
MEAN PERCENTAGE CORRECT FOR c PRONUNCIATIONS

Position of c	Graphemic Context	Grade 2	Grade 4	Grade 6
Initial	a o u	82	88	92
	e i y	22	41	59
Medial	a o u	64	81	85
	e i y	39	53	64
Averaged Over Position	a o u	73	84	88
	e i y	30	47	62
	Average	52	66	75

TABLE 9
ANALYSIS OF VARIANCE SUMMARY FOR c PATTERNS
(significant sources only, $p < .01$)

Source	df	MS	F
School (S)	2	5.63	16.73
Grade (G)	2	30.99	92.10
Sex (E)	1	5.79	17.21
Quartile (Q)	3	11.94	35.48
S x G x E	4	1.16	3.45
(Between Ss	414	0.34)	
Position x G x Q	6	.16	4.21
(Within Ss	414	0.04)	
Context (C)	1	61.11	558.80
C x G	2	0.99	9.05
C x Position	1	1.39	112.05
(Within Ss	414	0.11)	
C x G x Position	2	0.13	10.86
(Within Ss	414	0.12)	

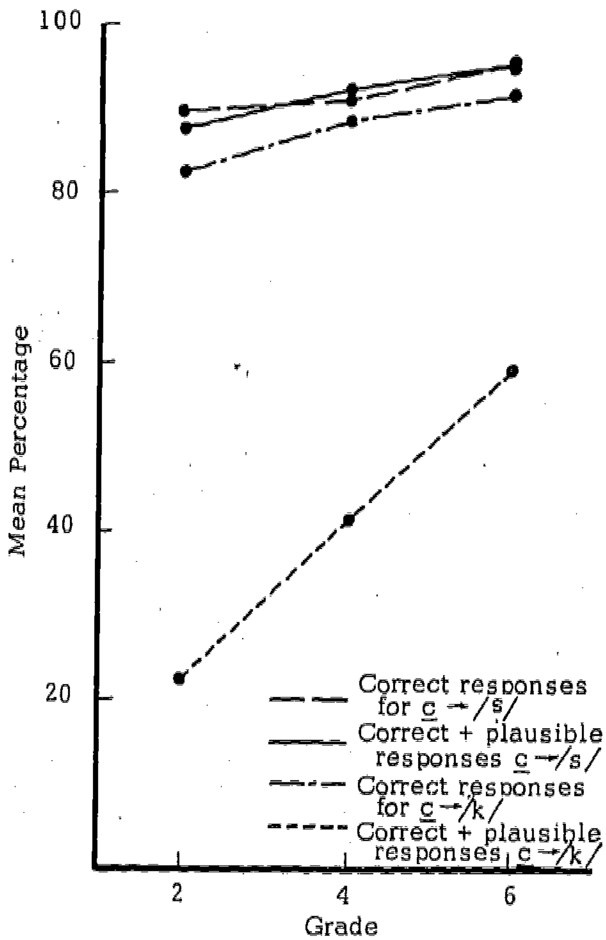

Fig. 6. Correct and plausible responses by grade to initial c.

sponses for c in /k/ environments began at approximately 82% at second grade and rose to 91.6% correct by sixth grade. (If plausible responses [/s/] are summed with the correct responses, the corresponding figures are 89.9% and 95.3%.) This development mirrors, at a slightly lower level of accuracy, the development of correct responses for invariant consonants.

But c before e, i, or y showed a radically different development; the correct responses began at 22.4% in second grade and reached only 58.9% in sixth grade. The extraordinary response bias for /k/ is shown by summing correct (/s/) and plausible (/k/) responses, the resulting curve being nearly indistinguishable in shape and position from the correct plus plausible curve for c in /k/ environments (see Figure 6). Furthermore, the increase in correct responses for c before e, i, or y from fourth to sixth grade was nearly as large as the increase from second to fourth grade.

The c→/k/ correspondence in initial position began at a relatively high level (82%) in second grade and changed relatively little, mirroring the progression for invariant consonants. The same correspondence in medial position was misread more often, largely because of errors in the lower quartiles. The c→/s/ correspondence, interestingly, showed a contrasting pattern. In initial position, particularly in second grade, it was pronounced as

/k/, while in medial position it was pronounced /s/ twice as often in second grade.

A response bias for c→/k/, especially in initial position, is not unexpected, in that almost all of the initial c words introduced in primary-grade reading programs have the /k/ pronunciation for c. Thus, while the variant pronunciations of c are highly predictable, the pattern is not acquired in the lower grades at all and only to a limited degree by sixth grade. Apparently, attempts to teach this pattern deductively fail, due to the lack of sufficient exemplars. The superiority of medial c→/s/ learning over initial c→/s/ learning is also attributable to the primary-level vocabularies, which show a significantly higher number of medial c→/s/ words, especially in c before a final vowel, than initial c→/s/ words (e.g., ice, icy, dance, face, race). For the two test items in which c occurred before a final vowel (roce, necy), the mean percentage correct at sixth grade was 75.5%, as compared to a mean percentage correct of 51.9% for the two test items with embedded c (hacen, recit). This difference, however, may be due to both a position effect, as was observed for invariant consonants, and to an exemplar effect.

It should also be noted that although a single rule based on following letter could be acquired to handle both initial and medial c, separate rules based on position appear to have been acquired by the children tested. This implies that not only is the teaching of reading ineffective for acquiring certain patterns, but it also causes extraneous learning of position where it is not required.

The g Patterns

Scores for the g patterns by position, graphemic context, and grade level are shown in Table 10. An analysis of variance identical to the one run for c patterns (Table 11) showed that the major between-subject sources of variance were grade and quartile, and to a lesser degree, school. Graphemic context and letter position were the major contributors to within-subject variance. Grade, quartile, and school trends were similar to those reported for the other patterns.

With respect to graphemic context, g patterns are identical to c: a sibilant (actually, affricate) pronunciation occurs before a spelling for a mid- or high-front vowel and a stop pronunciation in all other environments. However, g differs from c in one important feature: g admits a large number of exceptions to this affricate-stop variation and many of the exceptions are common words, e.g., get, girl, tiger.

The development of correct responses for the stop pronunciation of g was nearly identical to the corresponding development for c for all three word positions, as shown in Figure 7. But the development of the /ǰ/ pronunciation differed radically from that of c→/s/. For initial g before e, i, or y, the percentage of correct responses rose from 12.6% at second grade to 21.8% at fourth grade, and to 23.6% at the sixth grade. Clearly there was no generalization of this pattern, even though the number of words which obey the pattern in initial position outnumbers the exceptions.[10] In medial position, performances on g→/ǰ/ showed some improvement, but mainly in the second- to fourth-grade interval, where they rose from 26.1% correct to 42.2% correct. Thereafter only a marginal gain was made.

The failure to acquire the g→/ǰ/ pattern is, like the c→/s/ pattern, attributable to the words commonly introduced in primary-level readers, only here we are dealing with exceptions more than with imbalance of exemplars. For g→/ǰ/, the words most commonly introduced first in reading programs are the exceptions to the rule: get, girl, give. Only in the higher elementary grades are exemplars of the correct correspondence encountered. How-

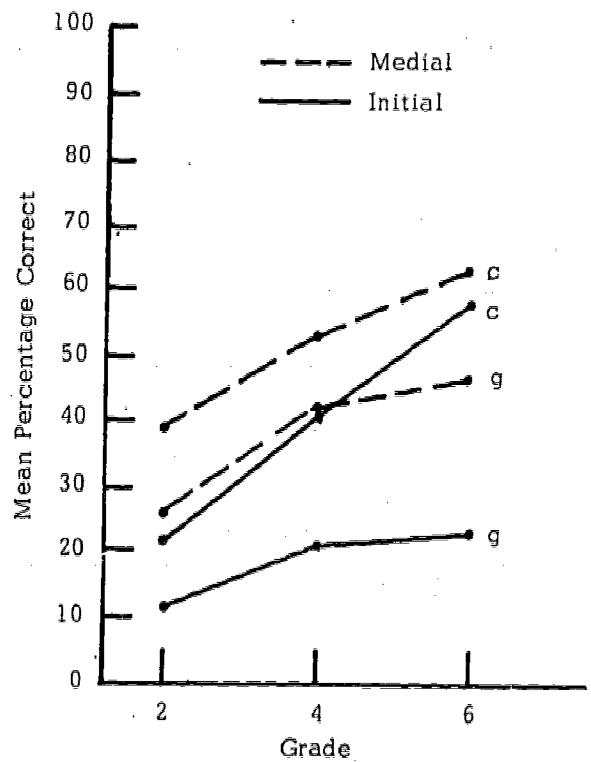

Fig. 7. Correct responses to initial and medial c and g before e, i, or y.

TABLE 10
MEAN PERCENTAGE CORRECT FOR g PRONUNCIATIONS

Position of g	Graphemic Context	Grade 2	Grade 4	Grade 6
Initial	a o u	86	91	95
	e i y	13	22	24
Medial	a o u	63	79	86
	e i y	26	42	47
Averaged Over Position	a o u	74	85	91
	e i y	20	32	35
Average		47	58	68

TABLE 11
ANALYSIS OF VARIANCE SUMMARY FOR g PATTERNS
(significant sources only, $p < .01$)

Source	df	MS	F
School (S)	2	2.39	10.37
Grade (G)	2	15.79	68.56
Quartile (Q)	3	7.71	33.47
G x Q	6	1.74	7.54
G x Sex x Q	6	0.82	3.55
(Between Ss	414	0.23)	
Letter Position (P)	1	0.34	8.65
P x G	2	0.55	13.89
P x Sex	1	0.32	8.08
P x Q	3	0.65	16.48
(Within Ss	414	0.04)	
Context (C)	1	143.64	1823.78
C x S x G	4	0.27	3.41
(Within Ss	414	0.08)	
P x C	1	3.39	288.33
(Within Ss	414	0.01)	

ever, the tokens for the exceptions continue to outnumber the tokens for exemplars, since the exceptions are among the very high frequency words for English. This raises a serious question about the g→/ǰ/ pattern: Should it be taught at all? Although by type counts, the exemplars outnumber the exceptions, neither the total numbers nor the difference is large. Thus, the utility of a g→/ǰ/ rule is quite low. Learning the initial g→/ǰ/ words as exceptions to a general g→/g/ rule may be as efficient as learning both a g→/ǰ/ and a g→/g/ pattern plus the exceptions to g→/ǰ/.

IV
Conclusions

Analyses of letter-sound correspondences which occur in common English words led us originally to the following assumptions: (a) that the invariant consonants would be learned early and to a high degree of accuracy in all word positions, (b) that the c pattern, especially in initial position before e, i, or y, would be learned more slowly, but equally well by the end of the primary grades, (c) that the long and short vowels would be learned more slowly and less accurately than the c pattern, and (d) that the g pattern for initial position before e, i, or y would develop slowest of all patterns tested, but would eventually reach a moderate level of accuracy.

The experimental results were hardly consonant with these assumptions. Position of the letter within the word was an important factor in the responses to almost all patterns. For the invariant consonants, the lowest quartiles in the second and fourth grades nearly equalled the performances of the highest quartiles for initial position, but fell drastically below for medial and final position. For the c→/s/ and g→/ǰ/ patterns, the positional influence was reversed, in that the best performances were found for medial, not initial, position. In addition, the long and short vowels, which admit numerous exceptions, were acquired faster and to a higher degree of accuracy than the c→/s/ pattern which admits almost no exceptions, and the initial g→/ǰ/ pattern which was never acquired as a pattern.

Behind these discrepancies between hypotheses and results we see one consistent factor, and that is the number of real-word exemplars which are introduced in reading programs for each pattern. For the long and short vowels, most of the popular reading programs offer a hefty stock of examples, and hence sufficient materials are available for generalization. But for c→/s/ and g→/ǰ/ almost no examples are included in the reading programs for the first two primary grades and only a few for third grade. In one popular series, for example, all initial c words in first grade are pronounced /k/; at second grade, 91% are pronounced this way; and in third grade, 95% are. (The total number of examples of initial c→/s/ words for second and third grades is five each.) What is remarkable in light of these data is that sixth-grade students pronounced any of the initial c's before e, i, or y correctly. Whatever process is established from formal teaching for generalizing letter-sound correspondences continues to operate long after formal instruction in reading ends. However, there also appears to be a critical period for generalization of letter-sound correspondences—probably during the first two to three years of instruction—after which generalization is slowed. The various pronunciations for initial c are almost perfectly predictable, yet the bias toward /k/ which characterized the responses of the early grades is far from eliminated by sixth grade. Initial g→/ǰ/ exemplars apparently appear too late to even dent the early-acquired bias for /g/.

For the quartile by position interaction for invariant consonants, we hypothesize that the poorer readers often learn to differentiate words on the basis of initial letters plus length, thus learning initial position correspondences quite well, but not medial and final ones.[11]

We do not find a general set for invariance in letter-sound learning—the competence of even the poorer-reading second graders in handling two different pronunciations for the simple vowel spellings showed that variant pronunciations were not ipso facto difficult. What appears to be important is the time lag between the introduction of examples of each pronunciation; the greater the time between the two, the less chance there is to overcome a response bias toward the first. On the question of successive versus concurrent presentation of variant, predictable pronunciations, this experiment offers no support for either alternative, nor does it give any indication of the relative influence of types versus tokens. These questions, along with the question of how to minimize the influence of exceptions to patterns, remain to be investigated.

Notes

1. We distinguish three word positions for spelling units: initial (first in the word), final (last in the word), and medial (everything that is neither initial nor final).

2. There are few exceptions to the rule that c before e, i, or y is /s/, and before any other letter, /k/. Cello is the only word in the 20,000 word Thorndike-Lorge (Thorndike, 1941) list which varies from this rule in initial position. In medial position, facade (without the cedilla) is an exception. In addition, medial c can be /š/, but this is also predictable (see Venezky, 1970, pp. 92ff.).

3. For the pattern VsCsV—where Vs is any simple vowel spelling (a, e, i, o, u, y), Cs is a simple consonant spelling (any single letter or digraph consonant spelling except x, wh, dg, or ck), and V is any vowel spelling, including final e—Vs has a checked (short) pronunciation (see Venezky, 1970, 101ff.).

4. That is, a sibilant (actually, an affricate) pronunciation (/ǰ/) before i, e, or y, and a stop pronunciation (/g/) otherwise.

5. Slide presentation was used for one elementary school and the sixth grade of another. Because of the high background noise, this procedure was not used again.

6. Although this study was concerned with the development of certain letter-sound patterns, this development was ascertained by cross-sectional rather than longitudinal testing, and hence variations in teaching methods and in materials across the schools and classes represent possible sources of distortion in the results. Due to the large sample size, however, and to the diversity of schools involved, such distortion is probably small.

7. One transcriber showed a tendency to shift low vowels upward; tapes done originally by this person were completely recycled through the transcription process.

8. In the first 5,000 words in the Thorndike list (Thorndike, 1941), in which o occurs in a stressed position, 79% of these words have either the long (/o/) or short pronunciation (/a/). The remaining 21% are divided among four other pronunciations.

9. An exception is /u/ as in bull, bullet, bushel, and pull. These represent approximately 7% by types of the stressed short-vowel correspondences in the corpus described in note 2.

10. These counts are based upon the 20,000 most common words in English, as designated by the Thorndike-Century Senior Dictionary (Thorndike, 1941).

11. This explanation is not in conflict with the superior scores of medial over initial position for c→/s/ and g→/ǰ/, in that the percentage correct scores for medial c→/s/ and g→/ǰ/ are still lower than the medial position scores for invariant consonants. This does imply, however, that the medial c and g scores would be even higher if the medial deficit did not occur.

References

Calfee, R., Venezky, R., & Chapman, R. Pronunciation of synthetic words with predictable and unpredictable letter-sound correspondences. Wisconsin Research and Development Center for Cognitive Learning, Technical Report No. 71, 1969.

Durr, W. K. A computer study of high-frequency words in popular trade juveniles. Mimeograph, Michigan State University [n.d.].

Johnson, D. Factors related to the pronunciation of vowel clusters. Wisconsin Research and Development Center for Cognitive Learning, Technical Report No. 149, 1970.

Swanson, D. E. Common elements in silent and oral reading. Psychological Monographs, 1937, 48, 36-60.

Thorndike, E. (Ed.) Thorndike-Century Senior Dictionary. New York: Appleton-Century-Crofts, 1941.

Venezky, R. Language and cognition in reading. In B. Spolsky (Ed.), Current Trends in Educational Linguistics. The Hague: Mouton, in press.

Venezky, R. The Structure of English Orthography. The Hague: Mouton, 1970.

Weber, R. M. First graders' use of grammatical context in reading. In H. Levin & J. Williams (Eds.), Basic Studies in Reading. New York: Harper and Row, 1970.