DOCUMENT RESUME

ED 057 990 24 RE 003 897

AUTHOR Cronnell, Bruce
TITLE Designing a Reading Program Based on Research Findings in Orthography.
INSTITUTION Southwest Regional Educational Lab., Inglewood, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO PP-14
BUREAU NO BR-6-2865
PUB DATE 26 Jul 71
NOTE 13p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Beginning Reading; *Computers; Grade 1; Grade 2; Grade 3; Graphemes; Kindergarten; Phonics; *Reading Programs; *Sequential Learning; Spelling; *Word Lists

ABSTRACT

A lexicon of approximately 9,000 words classified into levels from kindergarten to grade 3 was compiled as a base for studying the spelling-to-sound correspondences needed in beginning reading. The spelling-to-sound correspondences were established for 69 grapheme units, resulting in a total of 166 correspondences for the one- and Two-syllable words in the lexicon. These words were coded for the correspondence number of each grapheme unit, stress, syllable division, form-class, grade level, and special word conditions. Then 6,000 coded words were computer-processed and arranged by correspondences in a sequence for use in an instructional program. This systematic arrangement of words provided for the gradual increasing of word length and for organized introduction of vowels in unstressed syllables. The sequencing also allowed each one- or two-syllable words to appear as an exemplar of only one correspondence, after all other correspondences in the word have been introduced. The uncoded words of three or more syllables were arranged in the sequence by hand, bringing the total number of sequenced correspondences to 186. Finally, all lexicon words, including exceptions to the correspondences and proper names, were arranged according to the sequence. References are included. (AW)

RE

BR6-2865

PA24

SWRL

ST REGIONAL LABORATORY FOR EDUCATIONAL RESEARCH & DEVELOPMENT

ng: A Reading Program Based On ch Findings In Orthography

DESIGNING A READING PROGRAM BASED ON RESEARCH FINDINGS IN ORTHOGRAPHY

Bruce Cronnell

ABSTRACT

Procedures are described leading to the selection and organization of a lexicon appropriate for reading instruction at the kindergarten through third grade level based on an analysis of spelling-to-sound correspondences.

DESIGNING A READING PROGRAM BASED ON RESEARCH FINDINGS IN ORTHOGRAPHY

Bruce Cronnell

The Southwest Regional Laboratory for Educational Research and Development (SWRL) is developing a Communication Skills Program for kindergarten through grade three. A major component of this total language program is a phonics-based reading program. A basic premise of this program is that in order to read words not previously introduced in reading instruction, children must have an understanding of the spelling-to-sound correspondences of English orthography and must know how to apply these correspondences appropriately. In the design of a reading program which would provide for this knowledge, two objectives were set:

1. to select a lexicon appropriate for children at the kindergarten through third grade (K-3) level, and to establish spelling-to-sound correspondences for this lexicon;

2. to organize these correspondences and the lexicon for use in beginning reading.

This paper describes the work undertaken to fulfill these objectives.[1]

The Lexicon

The first task was to select a lexicon appropriate for children at the K-3 level. The major source of lexicon words was the Rinsland (1945) list of words used by elementary school children; this was supplemented by additional published word lists and by words from children's books (including reading series), songs, and TV programs. (A complete list of sources is found in Berdiansky, Cronnell, and Koehler, 1969, pp.7-9). Obsolete or old-fashioned words were deleted (e.g., marcel), as were objectionable terms (e.g., nigger) and words unknown to the staff (e.g., pung). Contractions and abbreviations were removed and listed separately because of their unique orthographic forms which are generally not amenable to direct spelling-scund analysis. In general, only base forms of words were used (e.g., boy, but not boys, boy's, or boys'); inflected forms were included when irregular (e.g., men) or semantically unusual (e.g., glasses). Proper nouns, other than a few pertaining to geography (e.g., Pacific Ocean), were excluded. However, a parallel study was made later of nearly 500 proper names, using the same procedures as outlined below.

---

[1] A complete description of objective (1) is found in Berdiansky, Cronnell, and Koehler (1969); a preliminary description of objective (2) is found in Desberg and Cronnell (1969).

The resulting lexicon consisted of approximately 9000 words likely to be in the recognition, if not active, vocabulary of K-3 children. This provided an adequate base for studying the spelling-to-sound correspondences needed in beginning reading, as well as an ample supply of words for actual use in reading instruction. Based on source information and staff judgments, these words were classed as being known at the kindergarten and first-grade level, or only at the second- and third-grade level (approximately two-thirds and one-third, respectively).[2]

Spelling-To-Sound Correspondences

Having selected a lexicon appropriate for K-3 children, spelling-to-sound correspondences were established for the words in this lexicon. These correspondences were based on Venezky's (1967, 1970) work, which is the most complete and theoretically valid study of English spelling-to-sound correspondences. Using a computer, Venezky thoroughly analyzed the relationships between the spelling and pronunciation of 20,000 words. However, some of his conclusions appear not to be applicable to beginning reading and were not used in the study reported here. Moreover, because of interest in the early stages of reading, the study of spelling-to-sound correspondences was limited to the 6000 one- and two-syllable words in the SWRL lexicon, which were believed to be easier and more useful for beginning readers.

Correspondences were established for a variety of grapheme units:

1. primary vowels, i.e., the six single vowel letters, a, e, i, o, u, and y;
2. secondary vowels, i.e., two or more vowel letters with a single corresponding pronunciation, e.g., au, ea;
3. single consonants, e.g., m, t;
4. double consonants, e.g., mm, tt;
5. consonant digraphs, i.e., two consonant letters, the pronunciation of which cannot be determined from the correspondences for the individual letters, e.g., ch, th;
6. other strings of letters which commonly function together as units, e.g., ck, le.

---

[2]In actual practice, words were classed by age-level rather than grade-level, using 6-7 years for kindergarten and first grade and 8-9 for second and third grades. For convenience, in this paper only grade-level will be referred to, since this is the practical significance of the classification.

A total of 69 grapheme units of the types described above were identified for the lexicon (see Table 1 for a list of these grapheme units). There are actually a few other grapheme units in English, both within the one- and two-syllable words studied and within the language as a whole (e.g., rh, eau). However, correspondences were not established for these grapheme units because of their rarity.

A spelling-to-sound correspondence is defined as the relationship between a particular grapheme unit and a particular pronunciation, with the surrounding word environment indicated where necessary. A correspondence was generally established when it has at least ten word exemplars. Correspondences which did not meet this criterion were included when they were part of general correspondences concerning all primary vowels (for example, while relatively infrequent for individual letters, there is a common generalization that vowels have their "long" [letter name] pronunciation when followed by a consonant and le, e.g., able, idle). In addition, low-frequency correspondences were included when the words to which they applied occur very frequently in language use (for example, the correspondence for final e as in he, she, me, we, etc.).

A total of 166 correspondences were established for the one- and two-syllable words in the lexicon. (See Berdiansky et al., 1969, Section III, and Cronnell, 1971, for a complete description of the correspondences used in this project). For nearly two-thirds of the 69 grapheme units one correspondence was sufficient; for the remainder, two to five were needed, except for the primary vowels, where up to 16 correspondences were used for one letter. This latter result is not surprising considering the well-known variability of English vowel spellings. (Table 1 lists the number of correspondences per grapheme unit.) A few comments are in order concerning the number of correspondences used in this project.

1. Many correspondences were generalizable across several grapheme units; this was particularly true of primary vowels where there were 50 individual letter correspondences which could be reduced to ten general vowel correspondences (e.g., correspondences for final VCe were established separately for each vowel although a single general correspondence could be used for all vowels).

2. More general, but more complex, correspondences were broken down into several simpler ones. For example, the correspondence for "long o" could be stated as o → [o] /_C({r/l})V; that is, the letter o is pronounced [o] when followed by a consonant, an optional r or l, and a vowel. While this was the most general statement of the "long o" correspondence, it was believed to be too complex for beginning reading instruction. Thus the correspondence was broken down into four parts:

   a. o → [o] /_Ce#, e.g., home, smoke; this is common "VCe" rule;

   b. o → [o]/_C{r/l}e#, e.g., ogre, noble;

TABLE 1

Grapheme Units and Number of Correspondences*

Grapheme unit	Number of correspondences	Grapheme unit	Number of correspondences
Primary vowels		Consonants	
A	18	B	2
E	15	BB	1
I	18	C	2
O	17	CC	3
U	12	CH	2
Y	11	CK	1
		D	1
Secondary vowels		DD	1
		F	1
AI	2	FF	1
AU	1	G	4
AW	1	GG	1
AY	1	GH	1
EA	3	H	2
EE	1	J	1
EI	2	K	2
EW	1	L	2
EY	2	LE	1
IE	3	LL	1
OA	1	M	1
OE	1	MM	1
OI	1	N	2
OO	2	NG	1
OU	5	NN	1
OW	2	P	1
OY	1	PP	1
UE	1	PH	1
UI	2	QU	1
		R	1
		RR	1
		S	4
		SH	1
		SS	1
		T	2
		TCH	1
		TH	3
		TT	1
		V	1
		W	2
		WH	1
		X	1
		Y	1
		Z	1
		ZZ	1

*Total number of correspondences: 186 (Only 166 correspondences were identified in the original analysis [Berdiansky et al., 1969]; when they were sequenced, 20 more correspondences were added).

c. o → [o] /_CV, e.g., notice, odor;

d. o → [o] /_C{f}V, e.g., okra, only.

This kind of simplification or correspondences, while reducing complexity, also increased the number of correspondences.

3. Correspondences were established for each double consonant, although in most cases the pronunciation is the same as for the single consonant (e.g., correspondences were separately established for f and ff, although the pronunciation is the same for both).

4. While environmental conditions differ for each correspondence for a particular grapheme unit, the number of different pronunciations per grapheme unit is relatively small; for example, while the letter i has fourteen correspondences, only five different pronunciations are involved.

Thus, the large number of correspondences does not necessarily indicate a proportionately great amount of learning difficulty.

In spite of the number of different correspondences, they did not account for all individual grapheme units in words. Over 10% of the 6000 one- and two-syllable words studied contained at least one grapheme unit not accounted for by the correspondences. This was, however, only 3% of the 27,000 individual grapheme units in all the words. These exceptions, in the opinion of SWRL staff, are insufficient to disprove the overall regularity of English orthography. A few points should be noted about exceptions:

1. Some of these exceptions are exemplars of additional correspondences which can be used with a larger lexicon including multisyllable words.

2. Further study of spelling-sound relations may reveal additional correspondences which account for some exceptions.

3. Even when there are exceptions to established correspondences (primarily due to environmental constraints), new or unusual pronunciations for a grapheme unit are rarely involved (for example, while the pronunciation of i in pint is an exception to the regular correspondence [cf, hint, lint, mint], it is still a common pronunciation of i [cf, pine, line, mine]).

4. Exceptions include a number of simple base words (e.g., to) plus their derived forms, which are otherwise regular (e.g., into, onto, unto). A relatively small number of such base words plus their derived forms can account for a considerable number of exceptions.

If beginning readers can learn a group of basic sight words and can have a set for diversity, they should be able to decode most words containing exceptions.

Coding and Processing

Having established the set of correspondences appropriate to the lexicon, the one- and two-syllable words were coded as follows (see example):

1. Each grapheme unit in a word was listed with the number of its correspondence (the number indicates pronunciation and, where necessary, environment); exceptions to established correspondences were numbered 40.

2. Stress was indicated (e.g., 0 = one-syllable word; 1 = two-syllable word with first syllable stressed; 2 = two-syllable word with second syllable stressed).

3. Position of syllable division was indicated, but only when such division affected the correspondences within a word (e.g., in create 34 indicates syllable division between the third and fourth letters since e and a are separate grapheme units and not one unit ea as in cream).

4. Up to two form-class (part of speech) designations were listed for each word (e.g., V = verb, N = noun, C = conjunction, PN = pronoun).

5. A few other symbols were used to indicate special word conditions (e.g., P = a possible alternate pronunciation).

6. Grade-level was indicated.

Example of Coding

Word	Correspondences	Stress	Syl	FC	Other	Grade
ACT	A15,C12,T10	0		V,N		K-1
CREATE	C12,R10,E26,A11, T10,E18	2	34	V		2-3
EITHER	EI10,TH12,E21,R10	1		C,PN	P	K-1

After all the 6000 one- and two-syllable words were coded in this fashion, they were processed by computer and arranged by correspondences. This initial processing permitted correction of errors and establishment of a few additional correspondences where needed. The words were again processed by computer, which provided printouts of words listed by grade, by individual correspondences, by exceptions, by form class, and by stress. Using the computer-generated data, counts were made of the frequency of individual correspondences. These frequency counts provided information on the usefulness of correspondences and were employed in the sequencing of correspondences.

Sequencing Spelling-To-Sound Correspondences

In order to use spelling-to-sound correspondences in an instructional program, they must first be sequenced. This section of the paper describes the guidelines and procedures used in the sequencing process.

The major criterion for sequencing correspondences was frequency of occurrence in the SWRL lexicon. More frequent correspondences are more useful since they provide more word exemplars and greater opportunity for transfer. To complement frequency considerations, several other general sequencing guidelines were established:

1. Correspondences for single letters precede correspondences for grapheme units of two or more letters (e.g., w before wh).

2. Double consonants are sequenced with the corresponding single consonants (e.g., s and ss together).

3. Grapheme units with only one correspondence precede those with two or more (e.g., ee before ea).

4. Correspondences with no environmental constraints precede those with environmental constraints (e.g., d before c).

5. Correspondences in one-syllable words precede those in two-syllable words (e.g, ll before le).

6. Correspondences with words occurring commonly in language production may be sequenced earlier because of their usefulness (e.g., the correspondence for final y as in by, fly, and my).

The guidelines outlined above provided the basis for the overall sequencing of correspondences. To enhance the use of this sequence for developing an instructional program, groups of correspondences were combined to form blocks; the blocks were themselves ordered (i.e., sequenced), as well as the correspondences within them. Correspondences were grouped into blocks to promote generalizations through similarity. Some of the criteria for grouping correspondences into blocks were as follows:

1. type of grapheme unit (e.g., primary vowels vs. secondary vowels);
2. generalizability of correspondences across grapheme units (e.g., VCe correspondences);
3. phonological similarity or identity (e.g., correspondences for oi and oy);
4. environmental similarity (e.g., correspondences for a and o before l).

Blocks were further subdivided into units, smaller groups for instructional purposes, based on the criteria outlined above.

In addition, provisions were made in the sequence for introducing words of two, three, or four syllables at successive points. The general steps involved in sequencing words of more than one syllable were as follows (using two-syllable words as examples):

1. compounds (e.g., bedtime),
2. affixed forms (e.g, bigger),
3. words arranged by stress pattern (e.g., first syllable stressed: rabbit; second syllable stressed: until).

This systematic arrangement of words of two or more syllables provided for the gradual increasing of word length and for organized introduction of vowels in unstressed syllables.

When the sequence of correspondences was established, the coded one- and two-syllable words were processed by computer and arranged in the sequence, such that each word appeared as an exemplar of only one correspondence, after all other correspondences in the word had been previously introduced. In this way, there could be cumulative learning of correspondences in words with controlled introduction of new material. In addition, the uncoded words of three or more syllables were arranged in the sequence by hand. When these words were included, several more correspondences were established to account for additional spelling-to-sound relations found in the larger lexicon; this brought the total number of sequenced correspondences to 186 (see Table 1).

The number of sequenced word exemplars per correspondence varied greatly. When the number was small, the exemplars were simply listed alphabetically. However, whenever possible, exemplars for each correspondence were organized for presumed ease of learning. This within-correspondence sequencing was based primarily on word length, morphological complexity, and position of the correspondence within the word. Organization of words according to these criteria should make easier the introduction and learning of new correspondences.

Words containing exceptions to the established spelling-to-sound correspondences were organized separately, but in relation to the sequence described above. These words were not introduced until all regular correspondences within them had appeared in the sequence. Exception words were sequenced to coincide with correspondences to which they were similar or to which they were exceptions. It was recognized, however, that many exceptions were high frequency function words, useful and necessary for reading-text syntax, and must be used earlier in the sequence for naturalness in written materials.

As noted previously, a parallel study was made of spelling-to-sound correspondences in nearly 500 common proper names (primarily first names, but with some high frequency surnames). These names (both regular and exceptions) were arranged according to the sequence and principles described above.

Conclusions and Implications

The procedures described in this paper have resulted in the following:

1. a lexicon of 9000 words suitable for children in the K-3 range;

2. a set of spelling-to-sound correspondences appropriate for beginning reading;

3. one- and two-syllable words from the lexicon coded for and arranged by spelling-to-sound correspondences;

4. a sequence of spelling-to-sound correspondences, ordered for beginning reading instruction;

5. all lexicon words, including exceptions to the correspondences, plus proper names, arranged according to the sequence.

In addition, the activity has generated a considerable amount of data on spelling-to-sound correspondences.

However, except in sequencing the lexicon, the primary concern was with one- and two-syllable words. There is a need to expand the lexicon and to study polysyllabic words, in order to provide additional and more comprehensive data on spelling-to-sound correspondences, and to provide the basis for extending instruction beyond the beginning stages. Nonetheless, the work described here has established a comprehensive design for a phonics-based beginning reading program based on a thorough analysis of English spelling-to-sound correspondences.

References

Berdiansky, B., Cronnell, B., & Koehler, J., Jr. Spelling-sound relations and primary form-class descriptions for speech-comprehension vocabularies of 6-9 year-olds. Technical Report No. 15, 1969, Southwest Regional Laboratory, Inglewood, California.

Cronnell, B. Annotated spelling-to-sound correspondence rules. Technical Report No. 32, June 7, 1971, Southwest Regional Laboratory, Inglewood, California.

Desberg, P., & Cronnell, B. An instructional sequence for spelling-to-sound correspondences for the one- and two-syllable words in vocabularies of 6-9 year-olds. Technical Report No. 16, 1969, Southwest Regional Laboratory, Inglewood, California.

Rinsland, H. D. A basic vocabulary of elementary school children. New York: Macmillan Company, 1945.

Venezky, R. L. English orthography: Its graphical structure and its relation to sound. Reading Research Quarterly, 1967, 2, 75-106.

Venezky, R. L. The structure of English orthography. The Hague: Mouton, 1970.