DOCUMENT RESUME

ED 038 387   TE 001 191

AUTHOR   Hodges, Richard E.; Rudorf, E. Hugh
TITLE   Searching Linguistics for Cues for the Teaching of
              Spelling.
INSTITUTION   National Council of Teachers of English, Champaign,
              Ill.
PUB DATE   May 65
NOTE   7p.
JOURNAL CIT   Elementary English; v42 n5 p527-33 May 1965

EDRS PRICE   EDRS Price MF-$0.25 HC-$0.45
DESCRIPTORS   *American English, Comparative Analysis,
              Componential Analysis, Educational Research,
              Elementary Education, Graphemes, *Linguistic
              Patterns, Linguistics, Morphology (Languages),
              *Orthographic Symbols, Phonemes, Phonology,
              *Spelling, *Spelling Instruction, Syntax

ABSTRACT

              Although spelling instruction has traditionally
treated almost every spelling word as requiring a separate learning
act, recent research indicates that over 90% of American-English
orthography actually is determined by a set of rules for unit
phoneme-grapheme relationships, based with decreasing productivity
upon three levels of analysis--phonological, morphological, and
syntactical. The phonological level of consistent sound-to-letter
relationships can be further examined in terms of the components of
position, stress, and environmental factors; and the morphological
level can be subdivided into components of compounding, affixation,
and word families. These new insights into the nature of
American-English orthography, provided by statistical analysis,
modern computer technology, and field work in linguistics, should
make a profound difference in a linguistically oriented spelling
program, both in what is learned and in how that learning is
accomplished. (See ED 026 361 for related papers.) (JB)

ED038387

RICHARD E. HODGES and
E. HUGH RUDORF

Searching Linguistics for Cues for the Teaching of Spelling*

(The following article states the overall design for the research approach on spelling improvement reported in this series of five articles. The first article in this series by Hodges and Rudorf reports the research phases on phoneme-grapheme correspondences completed in December, 1964. Other phases of the overall research design herein discussed are underway or are being planned for extensive field testing.)

The relationship of linguistics to spelling instruction.

Linguistic approaches to spelling instruction can be traced back well over a quarter of a century (1); however, the general introduction of linguistic principles into the school spelling curriculum has not been widespread in the English-speaking world. Typically, the teaching of spelling has been predicated on the assumption that there is little relationship between the way words are said and how they are spelled so that each spelling word requires a separate act of learning. Consequently, lists of spelling words for class study have been selected largely on the basis of the utility of these words in children's and adults' writings (3, 11).

Statistical analysis of phoneme-grapheme correspondences, on the other hand, suggests a considerably different rationale for spelling instruction. An early research into the consistency with which the 3,000 most frequently used words in children's writing are spelled was initiated by Paul h. Hanna of Stanford University in 1950. This research revealed that the phonemes (sounds) of the 3,000 words are regularly represented by certain graphemes (letters) approximately 80 percent of the time (12). More recently, with the advent of computer technology, other investigators have attempted to analyze the orthography by linguistic techniques for their own particular purposes; and these studies, too, have indicated that large numbers of words have relatively consistent phoneme-grapheme (sound-to-letter) relationships (4, 8).

What are some of the linguistic assumptions which underlie these kinds of investigations, and what do these investigations imply for the teaching and learning of American-English spelling?

First of all, the American-English orthography is an *alphabetically* constructed sys-

Dr. Hodges is an Assistant Professor of Education in the Graduate School of Education at the University of Chicago. Dr. Rudorf is an Assistant Professor of Education at the University of Delaware, Newark.

tem for the writing of spoken words. Many languages use this type of orthography in which each of the phonemes (sounds) of the spoken code has from one to several graphemes (letter symbols) which represent it when spoken words are encoded (translated) into written forms. Ideally, an alphabetic orthography would have one, and only one, grapheme to represent each phoneme. Thus, if a spoken language used forty different phonemes, the written code would also have exactly forty different graphemes. Some languages (e.g., Hawaiian, Finnish, and Italian) come close to achieving this ideal. The American-English language, however, does not attain this criterion (10).* Through the processes of borrowing words (including their spellings) from other languages, through changes in the way sounds are pronounced without changing the way they are spelled, and through historical accidents of printers' preferences or dictionary-makers' errors, the orthography has acquired many more letter representations for phonemes than are necessary.

The problem of learning to spell in most spelling classes centers on the assumption that there are very few useful rules to determine which graphemes do in fact represent the sounds of spoken words. Thus, a child learning to spell cannot with certainty predict how a particular sound will be spelled when it occurs in a specific word; hence, he needs to be helped to learn the spellings of words largely by principles other than the basic principles of sound-to-letter correspondences (6).**

These assumptions have been widely held, largely because there was no massive evidence to support the contention that most American-English phonemes are spelled with reasonable consistency. The Hanna-Moore study of 3,000 words seemed too narrow a sample of the American-English lexicon (the total stock of words existing in the language) (7). Examining more closely additional thousands of words, it was suggested, would verify that the orthography was inconsistent to the point that the Hanna-Moore findings would be deemed unreliable. Other investigators offered findings which were disparate with the conclusions of Hanna and Moore. Bost, for example, applied Moore's phoneme data to 1,148 representative words from Books 3 and 6 of the Horn-Ashbaugh series and found lower percentages of consistency, e.g., 45.7 percent for vowels and 35.7 percent for all phonemes (2).

A study of phonological relationships between sound and letter.

To clarify this and related issues, an intensive study was launched in 1962 at Stanford University of the relationships between phonemes and graphemes in over 17,000 different words (13). Under the direction of Hanna, and with the collaboration of the authors, this research sought not only to examine the degree of consistency of phoneme-grapheme relationships in these 17,000 words, but to analyze the structure of the American-English orthography in general. Using modern computer technology, it became possible to examine the structure of the orthography to a degree never before attempted nor possible

---

*For a further discussion of past and current efforts to revise the orthography so that there is a more consistent "fit" between the phonemes of speech and the graphemes of writing, see: Richard E. Hodges, "A Short History of Spelling Reform in the United States," Phi Delta Kappan, 7 (April, 1964), 330-332.

**See Jean S. and Paul R. Hanna, "Spelling as a School Subject: A Brief History," The National Elementary Principal, 38 (May, 1959), 8-23, for an elaboration of various ways in which weekly spelling lists have been developed in order to stress similarities among words other than phoneme-grapheme correspondences.

by the hand-analysis methods of previous studies of the orthography.

What kinds of insights into the American-English orthography were found? Most important, perhaps, it was demonstrated in Phase I (9)* that the orthography is actually a far more consistent reflection of spoken language than had been assumed, particularly when the several components of the phonology (sound system) underlying the orthography are examined. It is true that most phonemes have more than one way of being represented in writing. And it is equally true that, taking into account the way phonemes are spelled in large numbers of *different* words, it is difficult to sort out measures of consistency. But phonemes occupy *positions* in syllables and in monosyllabic words, and when phoneme-grapheme correspondences are tabulated in terms of their occurrences in these positions, a remarkable amount of consistency is found. Furthermore, when the amount of stress given to syllables in these 17,000 words is considered, even more consistency between phonemes and their graphemic representations is evident.

This statistical examination of the orthography, Phase I, does not necessarily presume that the results obtained are in themselves adequate to justify a firm claim for a linguistic approach to spelling instruction. In the first place, the fact that a phoneme is represented by a given grapheme over 80 percent of the time in some position in stressed and unstressed syllables does not tell how useful this information may be in the spelling of *words*. Secondly, increasingly restricting the tabulations of phoneme-grapheme correspondences to particular positions in stressed and unstressed

syllables means that the obtained results are generalizable to fewer numbers of words.

Beyond these restrictions, the statistical examination made in the course of the study ascertained that the great majority of phonemes in spoken American-English are indeed consistently represented in writing when the main phonological factors underlying the orthography are taken into consideration: 1) position in syllables, 2) syllabic stress, and 3) internal constraints. In addition, this thorough analysis of the relationship between phonemes and graphemes indicates that other kinds of linguistic factors are determinants of the ways in which some words are spelled. And further, the evidence obtained from the Phase I investigation made it possible to design a second computer program which takes the findings of this first study and uses them to *predict* the spellings of some 17,000 different words.

Predicting the spelling of American-English words.

This second computer program, Phase II (14),** it should be emphasized, relies upon phonological factors alone for its spelling "rules." Three factors which determine the choice of a graphemic option are: 1) the simple phoneme-grapheme relationships, 2) the effect of position of a phoneme in a syllable, and 3) the effect of syllabic stress upon choice of graphemic option. A fourth phonological factor is utilized, a factor designated "internal constraints" or "environmental factors." For example, while the spelling of the phoneme /f/*** can be

---

*This dissertation, referred to in this article as Phase I, is the first of a series of studies to be completed as part of a continuing research project in spelling initiated at Stanford University; it will be available from USOE as part of the Project 1991 report.

**This dissertation, referred to in this article as Phase II, is the second of a series of studies to be completed as part of a continuing research project in spelling initiated at Stanford University; it will be available from USOE as part of the Project 1991 report.

*** // indicates a phoneme (sound); < > indicates a grapheme (letter).

predicted only some 74 percent of the time on the basis of the first three factors, it is seen from the data in the Hodges study that when this phoneme follows the phoneme /s/, it is *always* spelled <ph> rather than <f> (e.g., *sphere, sphinx*). Thus, the immediate environment of the phoneme limits the choice of graphemic options which may represent it.

An algorithm (a set of rules or symbols defining a process) was therefore developed which utilizes the data from the Hodges study and adds observable factors of internal constraints. For each phoneme a set of rules was constructed which indicated which spelling of that phoneme should be used under various conditions of position, stress, and environment.

The algorithm was then utilized to process the 17,000 words from their phonemicization to their graphemic representation. This processing was expected to show: 1) how many and which words in the corpus could be spelled accurately by the use of oral-aural cues alone; and 2) how many of the words could not be so spelled. Further, the program was constructed to list these words according to the number of spelling errors made and to identify the particular phonemes producing the misspellings.

What are some of the results obtained from this computer run? Of the total number of words, 8,346 (49 percent) were spelled correctly. An additional 6,332 (37.2 percent) of the words were spelled with only one error, 1,941 (11.4 percent) with two errors, and 390 (2.3 percent) with three or more errors.

Morphological and syntactical elements of spelling.

The power of the algorithm, and the phonological approach to spelling, is strengthened when the error list is examined. A glance at these words and types of errors

involved indicates that many of these errors may not constitute a serious spelling problem. Many of them could be obviated with the mastery of simple morphological rules (morphology is the study of word formation—the combination of phonemes into meaningful units: roots, affixes, and inflection). For example, the factor of compounding in the formation of words obscures certain rules with regard to position. One rule which this study confirms states that when the long /a/ sound occurs in final position in a word, it is in almost all cases spelled <ay>. But in spelling the word *playground* on phonological cues alone, we obtained the spelling *plaground*. *Play*, however, was spelled correctly, as was *ground*. Because it can be assumed that a child who can spell both of these words can also spell the compound word *playground*, this type of error in the phonological spelling may be discounted. However, field tests of such assumptions which involve children have not yet been reported. Other morphological factors such as affixation and assimilation can also be taught as additional spelling cues which, when combined with a sense of the phonological base of the orthography, should help the child to spell correctly many hundreds, if not thousands, of the words contained in the printout of error lists.

One further morphological factor which may be utilized in producing correct spelling can also be identified from preliminary scanning of the error lists. Misspellings of certain phonemes can be seen which form a pattern, and these patterns can often be related to the origin of the root word. Families of words from French, Spanish, Italian, or Greek and Latin can be identified.

The teaching of etymology has been generally omitted in the elementary school spelling program. The research here re-

ported lends weight to the suggestion that it might well be a fruitful area of investigation. The evidence indicates that the bulk of the words in a typical elementary school program can be spelled on a phonological basis and a smaller, but still significant, number of words can be spelled correctly by combining phonological and morphological factors such as compounding and affixation. It seems a reasonable hypothesis that an analysis of the relatively few words remaining to be learned by reliance upon other cues might indicate that knowledge of a few important roots from various foreign languages could be a significant factor to enable the child to spell additional numbers of words. For example, a child who learns the spellings and meanings of *phono, photo,* and *graph* can spell additional numbers of words in which these root forms are included.

Finally, of course, as was expected, there does remain a residue of words that must simply be mastered by eye and hand learning methods. These words fall into two broad categories: 1) certain words, a limited number, whose graphemic correspondence to the phonemes is so irregular that they cannot be attacked by phonological or morphological means—words such as *one, acre, iron,* and some of the nautical terms like *forecastle;* and 2) the homonyms or homophones such as *bear* and *bare.* Quite obviously, there is nothing in either phonology or morphology which can help one to distinguish between the spellings of two different words with the same pronunciation. Here we must proceed to a third primary source of information, the syntactic or semantic level of language.

A model of American-English orthography.

Thus, out of this Stanford research project there begins to appear a basis for analyzing the structure of the orthography of the American-English language. We see how such a structure emerges on empirical grounds; it is also quite defensible upon a logical basis. Linguists have long emphasized the fact that what we refer to as a *language* is a system of *oral* symbols. Writing, the orthography, is a surrogate for the oral language; it is, in effect, a symbol for a symbol. Therefore, the structure of the oral language should be reflected in the orthography.

Linguists typically analyze the structure of language on three levels: phonology, morphology, and syntax. Thus, an orthography will reflect phonological, morphological, and syntactical components. The influence of each of these components will depend upon the nature of the written form of the language; that is, whether it is logographic (word) writing, syllabic (syllable) writing, or alphabetic (sound-to-letter) writing. A word-writing system (such as the Chinese) would depend primarily upon morphological and syntactical factors, while an alphabetic writing system would, by definition, be determined primarily at the phonological level. Thus, we can give a definitional model for the spelling of American-English: *The orthography of American-English is determined by a set of rules for unit phoneme-grapheme relationships based, with decreasing productivity, upon three levels of analysis—phonological, morphological, and syntactical.* The phonological level can be further divided into the components of position, stress, and environmental factors; the morphological level can be subdivided into components of compounding, affixation, and word families. This model may be summarized in tabular form as follows:

Phoneme-grapheme relationships determined by:

1. Phonological factors
   1.1 Position

1.2 Stress

1.3 Environmental factors

2. Morphological factors
   2.1 Compounding
   2.2 Affixation
   2.3 Word families

3. Syntax

That the assumptions upon which this model is based are sound has been demonstrated by the Stanford spelling research project (13). Individual phoneme-grapheme relationships, though not in terms of whole *words*, can be predicted with an accuracy of 89.6 percent by use of the phonological cues contained in the algorithm. Equally interesting is the statistical evidence that eight phonemes (/â/ as in *care*, /ē/ as in *here*, /o͞o/ as in *food*, /o͝o/ as in *foot*, /û/ as in *urn*, /ə/ as in *circus*, syllabic /'n/ as in *b· 'ton* and /z/ as in *zebra* can be identified .ch cause a large majority of the spelling problems. When these are considered separately, the percentage of predictability of the remainder rises to over 91 percent. The implication of this for development of a spelling curriculum is obvious.

It must be emphasized that neither the definitional model nor the algorithm is intended to be solely prescriptive of a spelling curriculum. What has been demonstrated at this stage of the research is that the orthography reflects the structure of the oral language upon which it is based. It suggests that regularities exist in the relationship between phonological elements in the oral language and their graphemic representations in the orthography, and that a pedagogical method based upon oral-aural cues to spelling may well prove to be more efficient and powerful than present methods which rely primarily upon visual and hand learning approaches. The next stage of research is to compare the effect of a linguistic approach on learning to spell with other methods.

Summary and implications.

We have seen that by relying upon phonological cues alone we can spell over 8,300 words correctly from the research list of 17,000 words. Consider this in relation to the typical spelling program for the elementary school which contains in a series of textbooks from grade two through grade eight some 3,000 words which are in the main taught as separate learning acts.

Greene and Petty in the 1963 edition of their *Developing Language Skills in the Elementary School* state that "... the ability to spell one word is distinct from the ability to spell other words ..." (5). From these Stanford research studies, one evidently can hypothesize that even a limited knowledge of the phonological relationships between the sounds and the letters of the orthography can provide the power to spell literally thousands of words and that other abilities relating to morphology and syntax may give pupils the ability to spell the vast majority of the words in their oral vocabularies.

Much work yet needs to be done. The algorithm must be examined to determine how words should be selected to help the pupil to arrive inductively at the generalizations that would help him to translate oral cues into writing.

The error lists need to be examined to determine what morphological and morphophonemic factors can be utilized in a spelling curriculum to add to the pupil's ability to combine meaningful units into words for his writing needs.

Finally, the words which the pupil needs that depart markedly from the basic alphabetic nature of the orthography need to be identified and introduced into the curriculum at appropriate points with a heavy reliance upon visual and haptical learning techniques.

These new insights into the nature of the

American-English orthography are currently being matched by increasing insights into the nature of the language learning process. Developers of spelling programs will need to take into account the best available generalizations regarding both the content of the curriculum and appropriate instructional processes; that is, the selection of words which best exemplify the alphabetic principles underlying the orthography and methods of teaching-learning which most effectively help children to apply their learnings to their writings.

In addition, material changes in the conventional means of evaluating children's spelling abilities will undoubtedly need to be made, because both what is learned and how this learning is accomplished may be quite different in a linguistically-oriented spelling program.

Bibliography

1. Bloomfield, Leonard, *Language*. New York: Henry Holt and Co., 1933. 564 pp.
2. Bost, Theda, "An Application of Moore's Phoneme Data to the Spelling of 1,148 Words," unpublished study. Austin: The University of Texas, 1952 (unpaged).
3. Fitzgerald, James A., *A Basic Life Spelling Vocabulary*. Milwaukee: The Bruce Publishing Company, 1951. 161 pp.
4. Garvin, Paul L. and Edith C. Trager, "The Conversion of Phonetic into Orthographic English: A Machine Translation Approach to the Problem," *Phonetica* (to be published).
5. Greene, Harry A. and Walter T. Petty, *Developing Language Skills in the Elementary School*. Second edition. Boston: Allyn and Bacon, Inc., 1963. 572 pp.
6. Hanna, Jean S. and Paul R. Hanna, "Spelling as a School Subject: A Brief History," *The National Elementary Principal*, 38 (May, 1959), 8-23.
7. Hanna, Paul R. and James T. Moore, Jr., "Spelling — From Spoken Word to Written Symbol," *Elementary School Journal*, 53 (February, 1953), 329-37.
8. Higginbottom, Eleanor, "A Study of the Representations of English Vowel Phonemes in the Orthography," *Language and Speech*, 5 (1962), 67-117.
9. Hodges, Richard E., "An Analysis of the Phonological Structure of American-English Orthography," unpublished doctoral dissertation, Stanford, California: Stanford University, 1964. 625 pp.
10. ―――――, "A Short History of Spelling Reform in the United States," *Phi Delta Kappan*, 7 (April, 1964), 330-32.
11. Horn, Ernest, *A Basic Writing Vocabulary*. Iowa City: University of Iowa Monographs in Education, First Series, No. 4, 1926.
12. Moore, James T., Jr., "Phonetic Elements Appearing in a Three Thousand Word Spelling Vocabulary," unpublished doctoral dissertation. Stanford, California: Stanford University, 1951. 69 pp., plus appendices.
13. *Phoneme-Grapheme Relationships Basic to Cues for Improvement of Spelling*. USOE Cooperative Research Project No. 1991. Stanford, California: Stanford University (to be available from the United States Office of Education).
14. Rudorf, E. Hugh, "The Development of an Algorithm for American-English Spelling," unpublished doctoral dissertation. Stanford, California: Stanford University, 1964. 974 pp.