This paper analyzes evidence related to the effectiveness of vowel generalization from two viewpoints: the nature of letter-sound and sound-letter relationships as derived from Dewey, Hanna, and Hillerich, and the nature of these relationships as derived from results of teaching vowel generalizations to children. From the evidence and the research studies reviewed the author concludes that generalizations about vowels can be grouped into two categories: generalizations which hold true most of the time but which include too few words to be worth teaching, and those which apply to many words but which are so unreliable that they are not worth teaching. Finally, the author suggests an alternative to the direct teaching of generalizations about vowels, which emphasizes the exploration of language and learning to read through reading.

(Author/WR)
"The Truth About Vowels"

Phonics and Reading

Thursday, May 2, 1974
10:45-11:45 a.m.
Young children learn to read through the development and application of certain skills as well as through practice in reading. Pertinent to this paper is the implication that, within the limited time available for reading instruction, that time must not be squandered with drill on non-essential "skills" at the expense of experience in reading.

Activities dealing with vowel generalizations clearly fall into the category of non-essentials in the view of this writer. Many children do not perform well on drill sheets relating to vowel generalizations; hence they are given additional work on these generalizations instead of moving on to more important activities. Evidence follows to suggest that the teaching of vowel generalizations is ineffective, both from analyses of the nature of English orthography and from studies of the effectiveness of emphasizing vowel generalizations to primary children.

The Nature of English Orthography

In the past ten years, considerable research has been published relating to the "utility of phonic generalizations." Undoubtedly the best known is that of Clymer (6), who examined four basal reading programs at the primary level. He drew out 45 generalizations that were to be taught in the four programs and then checked the "utility" of these generalizations against the vocabularies of the same four programs. Of 24 vowel generalizations, only six reached Clymer's criterion of 75% utility. Of the six, two are so vague as to be of doubtful help in reading.

Bailey (1), Emans (8), and Burroughs (4) applied the Clymer generalizations to vocabularies beyond the primary level and found essentially the same lack of utility. Rather than settle the issue, such findings have led educators to two quite different conclusions: some, including
this writer, recognize that vowel generalizations are faulty and
therefore to suggest that time can be better spent on other, more
effective reading activities; others also recognize that the vowel
generalizations are faulty, but seek to find more accurate generali-
izations.

Taking the latter direction, Fry (11), Emans (9), and Burmeister (3)
have suggested a proliferation of rules to replace the general digraph
rule ("When two vowels go walking. . ."), which Clymer found to be true
45% of the time. Dealing with each digraph separately does insure
greater "truth value" for some, but it also reduces the applicability
in terms of numbers of words. For example, one might teach children
rules for ee-e, ei-e, and eo, all three of which would be true 100%
of the time. Unfortunately, each rule would apply to one word in
Fitzgerald's list (10) of 2,650 words.

Children will eventually leave basal reader vocabularies. Are
there elements in the real world of English that lend credibility to
some vowel generalizations? In an effort to answer this question,
Hillerich (17) charted the vowel letter-sound (reading) and sound-
letter relationships (spelling) as found in three studies: Hillerich's
analysis (14) of Fitzgerald's 2,650 words (10), Hanna's Stanford study
of 17,000 words (13), and Dewey's study of over 10,000 words (7).

Both the Dewey and Hanna analyses were from a spelling viewpoint, and
are, therefore, not directly pertinent to reading. However, by developing
two-way charts from these studies, this author has made them applicable
both to spelling and to reading.

The Hanna study has been interpreted in various ways. For example,
one finding was that the computer -- programmed for 203 rules -- was
able to spell the 17,000 words with 83.82% accuracy. This finding has led some to the conclusion that English sound-letter relationships are about 84% "regular." Often overlooked, however, is the fact that this finding was based on feeding the 17,000 words into the computer phoneme by phoneme. When whole words were examined, the computer was only 49% accurate, i.e., in terms of whole words rather than phonemes, English was found to be 49% "regular" in its sound-letter relationships.

This great difference between the "consistency" of sound representations in phonemes as opposed to whole words was also borne out by Groff (12) when he analyzed 1,101 words from the New Iowa Spelling Scale. Despite the fact that he considered both c's (cut, cent), both g's (give, gem), and both oo's (foot, too) as "phonetic," Groff found 75% of the words "non-phonetic;" on the other hand, he reported only 18.8% of the letters "non-phonetic."

Such discrepancies between the "regularity" of phonemes or graphemes and whole words should not be too surprising. Words in English have about two consonant phonemes for every vowel phoneme. The consonants tend to more of a one-to-one correspondence, especially from the reading viewpoint where geminate consonants are not a problem. In practice, one must consider the whole word rather than individual phonemes in speaking of "regular" correspondences: no one credits the reader or speller as being "75% correct" in reading or spelling a word.

It is not our purpose here to criticize the Hanna study, since that has been done elsewhere (5, 20). To the point, however, any attempt to analyze the Hanna findings must first reorganize his classifications. Hanna used a twenty-two-vowel system as opposed to the usual dictionary classification of about fifteen vowel sounds. As a
result, there are eight different classes of graphemes (really pre-organized for spelling) for the schwa. Vowel sounds in off, all, arm, and pot are in four different classes, whereas they normally would be in two (/o/ and /a/). Secondly, Hanna used the very formal "spelling pronunciation" of words, a pronunciation style used only in that "other world" of English.

Space and copyright do not permit reproduction of charts of the three studies here, but some highlights follow. In general the studies are in agreement, once the Hanna data are reorganized into a standard fifteen-vowel classification. Most remaining variations among the three studies can be accounted for in terms of vocabulary size (2, 10, and 17 thousand words) and dialect variation.

Table 1 summarizes, from a reading viewpoint, some of the data out of the three studies.

Table 1. Summary of Three Studies Based on a Fifteen-Vowel-Sound Classification

<table>
<thead>
<tr>
<th></th>
<th>Dewey</th>
<th>Hanna</th>
<th>Hillerich</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Words</td>
<td>10,119</td>
<td>17,000</td>
<td>2,650</td>
</tr>
<tr>
<td>Number of Different Graphemes</td>
<td>87</td>
<td>69</td>
<td>57</td>
</tr>
<tr>
<td>Average Number of Sounds/Grapheme</td>
<td>2.3</td>
<td>2.8</td>
<td>2.5</td>
</tr>
<tr>
<td>Number of Graphemes with One Sound Value</td>
<td>46</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>Average Number of Words Accounted for by &quot;Consistent&quot; Graphemes</td>
<td>6</td>
<td>11</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1 indicates that the average number of sounds represented by each grapheme is about 2.5. The average is a misleading figure, however:
one must note that, while about half of the graphemes have one sound value, these graphemes account for very few occurrences in words. Removal of these "high utility" but seldom-used graphemes greatly increases the number of sounds represented by the remaining 50% of the graphemes. For example, even the single vowels a, e, i, o, and u represent 7, 6, 4, 7, and 6 different sounds respectively.

Many basal reading programs teach that single vowels between consonants have their "short" sound. Clymer (6) found this true 57% of the time; Hillerich (14), 56.5%. More specifically, linguistic programs begin by trying (inductively) to teach that a = /a/ ("A fat cat sat on a mat"). This specific generalization is true only 40 to 54% of the time. (Perhaps developers of the linguistic programs were looking at the other side of the coin: for spelling, /a/ = a 96 to 98% of the time.)

To make matters worse, 25% of all vowel sounds represented on a page are the schwa sound. Any one of 26 different graphemes -- one-third of all the graphemes reported -- may be used to represent that sound. What can be taught about such a high-frequency sound represented by such a variety of graphemes?

As if the observed lack of consistency were not enough, adult analysts can't even agree on classifications. Hanna fails to recognize existence of the vowel grapheme io (action), whereas Johnson's (18) printout from Venezky's study indicated this as the most frequently occurring vowel cluster, usually representing /ə/. Hanna got more "consistency" by counting the ti in words such as action as /š/ and the o as /ə/. Dewey (7) developed a unique practice of giving half credit to both consonant and vowel.
Debates on such items could go on indefinitely. In sign, is \textit{ig} the spelling of /I/ (Hillerich) or is \textit{gn} the spelling of /n/ (Hanna)? In diamond, is \textit{ia} the spelling of /I/ (Hillerich, Dewey) or does \textit{i} represent /I/ and \textit{a} /e/ (Hanna's "spelling pronunciation")?

Rather than focus on such disagreements, teachers should look at the practical implications. For example, how does one pronounce the nonce word \textit{wriat}? If the reader uses the Dewey and Hillerich classification, it is "write;" the Hanna classification leads to "riot." Which is correct can only be determined if the word is already known -- or, of course, if one is using context and also has the word in his listening/speaking vocabulary. In the latter case, however, the consonants alone would have been as helpful.

As a result of the many analyses which have been made of the language, this writer finds that rules relating to vowels seem to fit neatly into two categories: (a) rules that apply to many words but which are not true very often, or (b) rules which are true most of the time but which apply to very few words. Is either kind of rule worth teaching?

\textbf{Effectiveness of Teaching Vowel Generalizations to Children}

This author is still too much of a practitioner to accept arm-chair analyses of language as a final answer when it comes to children and learning. What are the effects of teaching vowel generalizations to primary children?

There is considerable research dealing with the generalization approach in spelling and also with intensive versus delayed phonics in reading. Unfortunately, however, specific research relating to the effectiveness of teaching vowel generalizations in reading is limited. Most effort has been devoted to analyses of the language itself.
First of all, since there is at least agreement that the major vowel graphemes each represent more than one sound, basic studies by both Levin (19) and Williams (21) point a consistent direction. These researchers investigated the relative merits of successive versus concurrent presentation of the multiple correspondences between grapheme and phoneme. For example, when $ai$ may represent /e/ (said) or /a/ (paid), is it better to teach one correspondence at a time (successive) or both at once (concurrent)? Both investigators found the concurrent approach best; children taught each generalization separately seemed to develop a "mind set" for consistency and became more easily confused when faced with the "irregularity" of the natural language.

Hillerich (15) studied the effectiveness of teaching vowel generalizations to first grade pupils. He compared all first graders ($N=742$) in two comparable school districts in terms of reading achievement at the end of grade one. In the one school district, first graders were taught vowel generalizations as part of the reading program; in the other school district, they were not. Results at the end of the year indicated that those pupils who had not been taught vowel generalizations scored significantly higher in reading achievement. Most important, the entire difference in reading achievement of the two groups was reflected in the subtest of comprehension, suggesting that excessive attention to the vowels led to over-analysis of words rather than to concern for meaning.

In a follow-up study, Hillerich (16) investigated the effectiveness of teaching about vowels in second grade. In this study of six classrooms, two classes of second graders were taught vowel generalizations, two were taught about vowels only at the hearing level, and two classes were not taught anything about vowels during the entire second-grade
year. At the end of the year the reading achievement test indicated that those who had been taught only at the hearing level -- to listen for and to recognize the various vowel sounds -- scored significantly higher than the next highest group, which was the group that had not been taught anything about vowels. Significantly lowest in reading achievement was the group that had been taught vowel generalizations.

This latter study not only points up the weakness of teaching vowel generalizations, but underscores the importance of an auditory emphasis. In working at the hearing level, second graders were not taught a letter-sound generalization; they were taught to identify a particular vowel sound under discussion and then to see how many other words they could find with that same vowel sound, to list those words and to classify them by spelling pattern. Such experience is truly inductive; it is discovery without the usual rigged sample, since it releases children to the entire world of language. Further, this emphasis on listening and classifying the sounds is in keeping with the findings of Bateman (2) and others: an auditory emphasis is most effective in beginning reading, regardless of preferred learning modality.

Rather than add to the many research studies on the effectiveness of "phonics" in reading, this author would suggest the need for more controlled research on specifics, research to clarify which elements of phonics contribute to reading success. Further research in this area is certainly needed to support or reject the findings proposed here. Meanwhile, just because "we have always taught vowel generalizations" is poor justification for continuing such instruction. From the view of research, as opposed to tradition, it seems that the burden of proof ought to be with those who seek to continue teaching vowel generalizations.
In conclusion, the writer finds that, from the standpoint of analysis of language, vowel generalizations have little validity; from the standpoint of instruction with primary children, they have little effectiveness.

On the other hand, the vowel symbols and sounds do exist in the language and cannot be ignored. Once children are underway in reading, opportunity to explore the vagaries of vowels can be helpful. However, the vowels should not be taken too seriously -- and here's the rub! Children have difficulty with vowel rules and are often given extra practice sheets on vowels. Such activities require time for filling in blanks -- time which could be devoted to reading or to having fun exploring the language. Most teachers recognize that, once started, children also learn to read by reading.
References


