The utility of phonic generalizations: Let's take another look at Clymer's conclusions

By: Francine P. Johnston

Johnston, F. R. (2001) The utility of phonic generalizations: Let's take another look at Clymer's conclusions. *The Reading Teacher.* 55; 132-143.

Abstract:

English orthography is not easily reduced to a few rules, but there are some general recommendations for teaching about vowels that can be helpful.

Article:

Theodore Clymer's analysis of the utility of 45 phonic generalizations (1963) has been hailed as a landmark study and been reprinted as a "classic" (1996). It continues to be frequently cited almost 40 years after it was originally published. His conclusion that only 18 out of 45 phonic generalizations met criteria of usefulness is more often than not used as evidence that phonics instruction itself is of limited usefulness. Depending on one's philosophical orientation Clymer's findings may be disappointing or vindicating, but in view of the current interest in phonics instruction it is worthwhile to take a closer look at the study.

Clymer collected 121 generalizations from the teacher manuals of four primary basal series used in the early 1960s. In a manner Clymer described as "somewhat arbitrary," he selected 45 generalizations that "were stated specifically enough so that it can be said to aid or hinder in the pronunciation of a particular word" (1963, p. 254). He combined the vocabulary introduced in the four basals with the Gates Reading Vocabulary for the Primary Grades to form a list of 2,600 words for analysis. Merriam-Webster 's New Collegiate Dictionary, 6th ed. was used as a guide to pronunciation. He compared the actual pronunciation of each word to the generalizations that should apply and calculated a percentage of agreement for each generalization by dividing the number of words actually conforming to the rule by the total number of words that could possibly conform. Using 75% as a reasonable level of utility, and eliminating any generalizations that did not apply to more than 20 words, he found that only 18 of the 45 generalizations met his criteria. He concluded, "many generalizations that are commonly taught are of limited value" (p. 255) but noted that "some of the generalizations might be useful if stated in different terms or restricted to certain types of words" (p. 258).

Clymer's original analysis needs to be reexamined, not only in terms of his findings and conclusions, but also in terms of what we currently know about how children learn to recognize words. After looking back at some little-known replications of Clymer's study, I will present a current analysis of vowel regularity done with slightly different parameters. The findings of this analysis will be compared to Clymer and other studies in order to demonstrate some of the issues that affect interpretation of this kind of analysis. The utility of phonic generalizations continues to be an issue as teachers make decisions about what to teach. Teachers will find some specific suggestions about the teaching of vowels scattered throughout this article, as well as some general recommendations at the end.

Replications of Clymer's work

After the appearance of Clymer's study in the early 1960s there was a flurry of replications. They appear to fall into two general categories: those that merely changed the collection of words for analysis and those that attempted to restate the generalizations or to call into question the kind of generalizations students should be taught.

Six studies replicated Clymer's analysis, using different sets of words but adhering closely to Clymer's procedures, including the use of Merriam-Webster 's New Collegiate Dictionary, 6th ed. (hereafter referred to as MW-6). Bally (1967) increased the word list to include all the words in eight basal series up through the sixth grade (5,773 words). Emans (1967) analyzed 1,944 words randomly sampled from The Teacher's Word Book of 30,000 Words (Thorndike & Lorge, 1944). Four dissertation studies (cited in a summary by Curry & Geis, 1976) analyzed words from math books (Ferguson, 1970), science books (Jernigan, 1969), social studies books (King, 1970), and spelling books (Davis, 1969). The utility of some generalizations varies depending on the source of words, sometimes climbing above the 75% criterion and sometimes falling below. Overall however, no matter the number of words or the source, the results still show that less than half of the generalizations, as stated, meet Clymer's criteria of 75% utility applied to at least 20 words across all types of text.

Several researchers questioned the generalizations themselves and made attempts to replace them with ones that would have higher utility. Gates (1983, 1986) analyzed 17,211 words taken from Hanna, Hanna, Hodges, and Rudorf (1966). She created new generalizations with high reliability, but unfortunately they include conditions and exceptions that make them unwieldy. For example, her generalization for the silent *e*, which is reliable at 93%, is restated as follows: When a word ends in a single vowel + consonant + *e*, the *e* is silent and the vowel is long, or has the short-*i* sound (except for words ending in *some, ove,* and *vowel + re*).

The "two vowels go walking" generalization is probably the most frequently cited example of how unreliable phonics can be. According to Clymer's analysis, this generalization is true only 45% of the time. Burmeister (1968) focused specifically upon this generalization in her analysis of words from Hanna et al. (1966), examining 26 vowel combinations separately. She concluded that the generalization does have limited usefulness when coupled with qualifications about when to apply it. She described several categories of vowel pairs that should be taught in different ways: (a) In some pairs the first vowel dominates as in *ai, ay, ea, ee, oa*, and *ow*. (b) In some pairs the vowels "cooperate" as in *au, aw, oi, oy,* and *oo*. (c) In some pairs a new sound is created as in *ei, ou, ey*, and *ew*. A few vowel pairs also separate into syllables as *ia* does in piano.

Caldwell, Roth, and Turner (1978) analyzed 18,000 words from Kucera and Francis (1967) to determine the consistency of 23 phonics generalizations much like those Clymer used. They determined percentages based on frequency of occurrence (tokens as opposed to merely types) in running text and found even lower percentages than Clymer for some generalizations such as the two-vowel (26%) and silent-*e* (39%) generalizations. They concluded, "a higher-order unit approach may be an efficient alternative for teaching the rules"(p. 96). Examples of these higher order units include *ine* or *ake*, combinations derived from a generalization such as the final-*e* rule.

Although Clymer is frequently cited, it is rare to see these follow-up studies cited. Both those that tend to validate Clymer's work and those that question his conclusions have failed to make the same dramatic impact as the original study.

The purpose of this study

Given the fact that many researchers have taken on the task of analyzing Clymer's generalizations, one might wonder why anyone would take it on again. There were two reasons initially: (a) A careful reading of Clymer's generalizations revealed some that seemed very badly stated and others listed exceptions that did not appear to apply. While some researchers had suggested alternative generalizations, none had really critiqued the ones listed by Clymer. (b) *The American Heritage Word Frequency Book* (Carroll, Davies, & Richman, 1971) provided a source of words that had not been analyzed before and seemed a valid representation of the kinds of words readers of today were most likely to encounter in the uncontrolled text used in many primary classrooms. Along the way, in the process of analyzing the words, I made some interesting discoveries and determined that there are better ways to state some of the generalizations. Some of these discoveries seemed to confirm alternative ways to conceive of phonics instruction that were worth sharing with a larger audience.

Table 1
Percentages of words conforming to the "two vowels go walking" generalization from earlier studies

Researcher	Data source	Percentage compliance
Clymer (1963)	Four primary basals, Gates	45
Baily (1967)	8 basals, Grades 1–6	34
Emans (1967)	Random sample of 30,000 words	18
Davis (1969)	6 spelling series	32
Jernigan (1969)	3 science series	34
Ferguson (1970)	3 math series	37
King (1970)	3 social studies series	38
Caldwell, Roth, & Turner (1978)	18,000 most frequent words by occurrence in running text	26

Not all of the 45 generalizations listed by Clymer are included in this investigation. My analysis was limited to a selection of vowel generalizations: vowel pairs, long vowels marked with silent *e*, and final *y*. Vowels pose the greatest difficulty for young readers (Calfee, 1998), and vowel generalizations may be the ones they need most. Consonants, on the other hand, are mastered more easily and are much more regular. Of the 45 generalizations investigated by Clymer, 10 involved consonants and all but one of those had a utility of 95% or better. Vowel generalizations are much less regular and more subject to interpretation.

Procedures

A database was created from the 3,000 most frequent words listed in *The American Heritage Word Frequency Book*. The words in this book are ordered by frequency and are based on a wide range of reading materials used in Grades 3 to 9, rather than upon the controlled vocabulary of basal texts used in many of the earlier studies. Each word was analyzed for vowel letters and combinations, and vowels were listed in fields. For example, *stone* has the *o-e* combination, while words like realize have both *ea* and *i-e*. A database application (Microsoft Works 4.0) was used to search for and sort the letter combinations, which were then categorized by sound. For example, words with the *ea* vowel pair were categorized as long *e* (*sneak*), short *e* (*bread*), long *a* (*great*), or miscellaneous (*create*). The sounds were categorized initially based on my own dialect and confirmed using *The American Heritage Dictionary of the English Language* (Morris, 1970) (hereafter referred to as AHD) whenever there was a question about how to categorize a vowel sound. A sample of words in each category was verified, and in some cases, such as words with *ue*, all words were checked against a dictionary.

Certain protocols need to be established when classifying words and the vowels within them. There are a number of issues Clymer did not comment on in his report so it is not certain how he dealt with them. Capitalized words, compounds, and inflected forms are listed as separate entries in *The American Heritage Word Frequency Book*, and how they are counted makes a difference when the total number of words in a category is small. I decided not to include capitalized words as separate instances but I did include compound words (*rainbow*) and inflected forms (*rains, raining*), as well as derived forms (*care, careful, carefully*). The vowels in unaccented syllables were subjected to the same analysis as accented syllables. Homographs were given the pronunciation suggested by the generalization (*read* with long *e*) unless it seemed likely that an alternative was more common (*live* with a short *i*). *W* and *v* are treated as vowels when preceded by a vowel (*e-w; ow, oy, ay*).

Findings

When two vowels go walking. The longevity of Clymer's work may be accounted for in part by his dramatic revelation regarding one particular generalization, the "two vowels go walking" rule. According to his analysis, the generalization that "When two vowels are side by side the long sound of the first one is heard and the second is usually silent," is true only 45% of the time. Subsequent follow-up studies by other researchers exploring different data sets of words found even lower percentages (see Table 1).

With such poor reliability it could be hard to argue that this generalization should be taught at all, yet when individual vowel pairs are examined a different picture emerges. The two vowels go walking generalization averages all vowel pairs into one percentage and this obscures the fact that some vowel pairs are highly regular, Vowel pairs are included in the generalization that never take the long sound (*ay, oi*, and *au* for example), lowering the percentage dramatically. However, these vowel pairs represent their respective sounds on a very consistent basis. Like Burmeister (1968), I found that vowel pairs could be more accurately described in sub-groups. See Table 2 for a complete listing. Each group will be described, and individual vowel pairs will be discussed to highlight some of the discoveries as well as problems encountered as I tried to make a complex orthography fit into narrow categories.

Some vowel pairs reliably represent one vowel. That is, the first one does the "talking." The most reliable pair is *ay* at 96.4%. Says was the only word in this data set that did not conform. Clymer listed *always* as an exception, but AHD shows it first with a long *a* and then with two other possible pronunciations. The 10 nonconforming words Clymer reported for *ay* are puzzling, but the use of MW-6 might account for them.

The *oa* combination is quite regular at 95%. In the AHD *roar* and *board* have a second pronunciation represented by the long-vowel sound. I included these words under long *o* because it seems likely that using the long-*a* sound would enable a reader to identify the word. *Broad* is an exception.

The *ee* pair is also highly regular at 95.9% regularity. *Been* is a familiar exception. According to AHD, words like *deer* and *engineer* do not strictly have the long sound due to the influence of *r*. However, like the long o in *oar*, the sound is very close to long *e* and would offer a close approximation. *Ai* takes the long-*a* sound 75% of the time. Eleven of the 19 words that did not conform had the *air* pattern (*pair*) which, like *oar* and *eer*, seems very close to the long-vowel sound. Nonconforming words include *again*, *said*, and the final unaccented syllable in words like *mountain* and *captain*. *Ey* takes the long-*e* sound 77% of the time but is not very common. *They* and *they're* are the familiar exceptions.

Some vowel pairs represent one sound quite reliably but that sound is not long. *Aw* (*saw*), *oy* (*boy*), and *oi* (*join*) are all reliable at 100%. *Au* is 78.9% reliable as in *cause*. The exceptions *laugh, laughs*, and *laughed* reduced the percentage noticeably in a set of only 19 words. Together these four patterns showed up in only 52 words in this data set, but they constitute two highly regular letter-to-sound correspondences.

Other vowel pairs can be considered variable but still regular. These are the ones like *ow, ew*, and *oo*, which alternate between two sounds very consistently. *Ow* represents the long-*o* sound heard in *snow* 68% of the time and the sound heard in *how* 31.9% of the time. Together that accounts for 99.9%. The only nonconforming word was *knowledge*. *Ew* alternates between the two long-*u* sounds; 88.3% take the sound heard in *blew* (/bloo/) while 18.7% take the sound heard in *view* (Nyoo/). Some words such as *new* are shown with both pronunciations in AHD. *Oo* represents the sound heard in *boot* 50% of the time and the sound heard in *hook* 40.4% of the time. The remaining 7.7% include four words with *oor* (including *door* and *floor*) and *blood*. *Ei* is rare but alternates between long *a* (50%) and long *e* (25%).

What remain are vowel combinations that present more of a challenge for readers who might encounter them in unknown words. Some are rare enough not to be of much concern, but others like *ea* and *ou* occur in many words. The *ea* pair represents a number of sounds, although a flexible strategy of using either long *e* as in *seat* (49.6%) or short *e* as in *head* (16.7%) can predict the sound 66.3% of the time. The *ear* pattern has several sounds, but the sound heard in *fear* is most common (14.3%). If the *ear* words could be combined with the long-*e* words, 63.4% of the *ea* words could be accounted for. *Great, bear, earn, really*, and *idea* are some of the words that do not conform to the three other categories.

The pair *ou* alternates between the sound in *out* (43.2%) and the short *u* in *touch* (17.8%), but that leaves 39% that vary between a number of sounds. Fourteen of those words have the *r*-influenced sound and pattern as in *your*. Other words include *through, thought, and although*. With *ie*, a flexible strategy can help the reader deal

with many unfamiliar words. The sound is long *e* (field) in 49% of the cases and long *i* (tied) in 27.2%. That leaves 23.6% that vary between a number of sounds including short *i* (*friend*) and *r*-influenced *e* (*fierce*). The *oe* pair is very rare and varies between three or more sounds.

Table 2
Vowel pairs in three studies showing number of words and regularity (sample words in parentheses)

	Clymer (1963) 2,600 primary words		Burmeister (1968) 17,310 core vocabulary words		Present study 3,000 most frequent words	
Pair	Number	Percentage	Number	Percentage	Number	Percentage
One sound						
ay	46	78 (play)	137	96.4 (gray)	38	96.4 (play)
oa	35	97 (boat)	138	93.5 (road)	16	95 (coat)
ee	87	98 (seem)	290	85.5 (sleet)	70	95.9 (feet)
ai	67	64 (nail)	309	74.4 (abstain)	76	75 (rain)
ey			69	8.7 (key)*	9	77 (monkey)
aw			77	100 (lawn)	11	100 (saw)
oy			50	98 (convoy)	8	100 (boy)
oi			102	98 (moist)	14	100 (join)
au			178	93.8 (auction)	19	78.9 (cause)
Two sounds alternate						
ow	85	59 (own)	250	50 (snow)	74	68 (snow)
				48.4 (how)		31.9 (how)
ew	26	35 (blew)	64	60.9 (blew)	16	88.3 (blew)
				34.4 (view)		18.7 (view)
oo			315	58.7 (boot)	52	50 (boot)
				36.2 (book)		40.4 (book)
ei			86	40 (eight)	12	50 (eight)
				25.6 (either)		25 (either)
Three or more sounds						
ea	162	64 (bead)	545	50.5 (seat)	177	49.6 (seat)
				25.7 (head)		16.7 (head)
				9 (fear)		14.3 (fear)
ou			815	**35 (out)	188	43.2 (out)
				3.7 (touch)		17.8 (touch)
				3.1 (your)		7 (your)
ie	47	17 (field)	156	35.9 (field)	55	49 (field)
				16.7 (tied)		27.2 (tied)
oe			22	59.1 (toes)	9	44.4 (toes)
				18.2 (shoe)		33.3 (shoe)
						22.2 (does)
Other						
ui	17	6 (suit)	34	47.1 (build)	17	52.9 (build)
				29.1 (fruit)		23.5 (fruit)
ia			5	***	24	54.2 (piano)
						45.8 (Asia)

* Using *Merriam-Webster's New Collegiate Dictionary, 6th ed.* as guide to pronunciation.
** Burmeister found that *ou* represented the schwa sound in 41.2% of the words she analyzed.
*** Burmeister evidently did not include words like *piano*, which split into syllables between the vowels.

Some pairs such as *ui* and *ia* are very rare. The *ui* pair represents either long *u* (*suit*) or short *i* (*build*) in only seven words, unless you consider words like *quiet* and *guide* as instances. In those words, the *u* should be associated with the preceding consonant rather than with the vowel. However, from the young reader's perspective the *ui* might be seen as a vowel pair and pose a problem. The *ia* pair either splits into two syllables (*piano* or *India*) or it represents the *schwa* sound in a final unaccented syllable (*Asia*, *official*). These are good examples of how complex a task it is to divide English orthography into well-defined categories.

Table 2 compares my analysis of individual vowel pairs with Clymer (when applicable) and Burmeister. Overall the percentages do not vary much except in a few instances. Using MW-6, Clymer and Burmeister could not categorize words such as *monkey* as having the long-*e* sound. The use of MVV-6 probably also accounts for the difference in the percentages for *ay*. Another discrepancy is found in the *ou* pattern. Burmeister found that the most common sound for *ou* was the *schwa* sound. This is probably due to the size of her database, which contained many words of two or more syllables in which the *ou* might be reduced to *schwa* (as in *dangerous*). It

is not clear how she categorized words with the *ia* pattern, but the small number (5) she reported suggests that she did not include all possible words.

Table 3
Patterns for final *e*

Vowel pattern	Number of words	Percentage of regularity	Other sounds
a-e	130	77.7 (cake)	8% have *are*, 13.8% are in final unaccented syllables
e-e	31	16.6 (these)	
i-e	128	74.2 (five, fire)	
o-e	89	58.4 (stove, more)	35.9% end in *ove* (move), *ome* (come), *one* (none)
u-e	26	76.9 (rule, refuse)	11.5% have *ure*

To summarize this section, five vowel pairs can be considered highly regular and worth teaching: *ai, ay, oa, ee*, and *ey*. Although the first vowel does the talking in these words, it would be setting a dangerous precedent to talk about all pairs in this fashion because there are only five pairs out of 19 that work that way with any regularity. Other vowel pairs do not have long sounds but are nevertheless highly regular and can be taught: *aw, oy, oi*, and *au*. The remaining vowel pairs have two or more alternative sounds and cannot be taught as easily. R-controlled vowels can in some cases be treated as though they are simply long, but should also be considered as variant sounds. Students will need a flexible strategy for these pairs, such as trying more than one sound and checking the results with their oral language and context.

The final-e generalization. Clymer reported that the generalization, "When there are two vowels and one of them is final *e*, the first vowel is long and the *e* is silent" is accurate only 63% of the time, falling short of his criteria. Like the two-vowels rule, this generalization too, can be broken down by specific vowels to see if there are some more regular than others. See Table 3. For my analysis the words were limited to those in which a single vowel is separated from a final *e* by one or two consonants. Note, however, that while some words with two consonants between the vowel and the final *e* conform to the generalization (*change, strange*), many words ending in *ce, ge, ve* do not (*edge, twelve, since, solve, dance*). In those words the final *e* marks the *c* or *g* "soft" or satisfies the generalization that words in English never end in final *v*. However, to include those exceptions in generalizations taught to young readers will burden them with unnecessary complexity, and so for now those words are treated as either conforming or not.

The *a-e* pattern represents the expected long sound 77.7% of the time in my database. Of the words that do not conform, 10 of them are spelled with *are* and have the same sound (*care, bare, stare*). Eighteen more of them have the *a-e* pattern in a final unaccented syllable that takes the short-*i* sound (*palace, message, distance*). Other exceptions include *have, dance, badge, are*, and *large*. The *e-e* combination is not common and represents a variety of sounds. Five of them, or 16.6%, have the long sound (*scene, these, Japanese, Chinese*, and *complete*). Five of them have a short-*e* sound (*edge, fence*), and the rest have various *r*-controlled sounds (*here, there, were*). The *i-e* pattern takes the expected long sound in 74.2% of the words. This includes words such as *fire* in which *r* follows the vowel since AHD represents the sound in the *ire* patterns as long. Exceptions include *live, give, since*, and two- or three-syllable words in which the final syllable is unaccented and takes the short-*i* sound (*opposite, favorite, determine*).

The *o-e* pattern takes the expected long sound in 58.4% of cases. These include words such as *more* in which the vowel is followed by *r*, as allowed by AHD. The words that do not conform fall into some interesting categories. Thirty-two of the 39 exceptions end in *ove, ome*, or *one*. Twenty-five of those words include or are derived from only four high-frequency words, *some, come, move*, and *one*. Aside from words with these patterns, there are only six words that do not conform (*purpose, Europe, solve, involve, lose, whose*).

The *u-e* pattern is complicated by the fact that it has two sounds that may be considered "long." *Rule* has one of

those sounds, /oo/, while *refuse* has the other, /yoo/. Dictionaries do not always agree which of these sounds goes with which words, and some words are listed with both (*tune* = /toon/ and /tyoon/). Of these types of words, 76.9% of the *u-e* words conform with one of these sounds. Nearly all of the nonconforming words have the *ure* pattern as in *sure, pure, pasture*, and *mixture*.

By breaking the broad generalization regarding silent e into specific vowels, three of the vowels were found to have a respectable degree of compliance as compared with the overall utility of 63% Clymer found. These are *a-e, i-e,* and *u-e* at 77.7%, 74.2%, and 76.9% respectively. If the rare combination of *e-e* at 16.6% were averaged with these others the percentage would be lower. *O-e* is an interesting case because, although it has a low percentage of utility (58.4%), it is due to a small set of high-frequency words. Vowels that precede *r* present an interesting problem as well. Sometimes they are designated as long in the AHD (*ire* and *ore*) and sometimes not (as in *are, are,* and *ere*).

It would appear that teachers can teach the final-*e* generalization with confidence, although encouraging a flexible strategy will be important. R-controlled vowels once more play havoc in some cases and may need special attention. In addition, the role of final *e* in words that end in *ge, ce,* and *ve* would be worth exploring at some point (Johnston, 1999). It hardly seems worthwhile to make much of an issue about the different sounds of long *u*. Although teachers should be aware of this subtle variation, it has been my experience that it does not cause any problems for children trying to read or spell such words.

Other discrepancies. In addition to the two generalizations governing long vowels, several more generalizations were surprisingly low in utility according to Clymer. Generalization 34 was initially very puzzling. It states, "When *y* or *ey* is seen in the last syllable that is not accented the long sound of *e* is heard." It listed *baby* as an exception, but based on my dialect that should have been long *e* and so should all the other words I found. Clymer had faithfully followed the pronunciations given in MW-6 and reported that he found zero words conforming and 157 exceptions. *Baby* adheres to the generalization according to AHD (as do *honey*, *lucky*, and *friendly*). Use of a different dictionary changes the utility of this generalization from 0% to 100%.

Clymer noted that some of the generalizations might be useful if stated in different terms or restricted to certain types of words. One of these is generalization 16, "When *y* is the final letter in a word it usually has a vowel sound." Clymer cited 32 exceptions to this for a utility of 84%, but exceptions included words such as *play* in which the final *y* is silent. I found that when the generalization was restated to say, "When *y* is the final letter in a word it *acts* as vowel," it was applicable 100%.

The problem of badly stated generalizations shows up in generalization 17, "When y is used as a vowel in words, it *sometimes* (emphasis mine) has the long sound of *i*." Clymer reported that this applied only 15% of the time and had 170 exceptions. A better rule can easily be created by saying, "When y is used as a vowel in words, it sometimes has the long-*e* sound." According to my analysis, out of 134 words, this is true 87% of the time and is therefore, by Clymer's criteria, a very useful rule. Most of the exceptions to this restated generalization are high-frequency one-syllable words that end in *y* (*fly*, *try*, *my*) or words derived from them such as *myself.* A more comprehensive generalization might state, "A final y takes the long-*e* sound in words of two syllables and the long-*i* sound in one-syllable words." This is true in 95% of the words analyzed. Exceptions include *July* and *supply* in which the accent falls on the second syllable.
Generalizations concerning the final y are surprisingly reliable when restated and reevaluated using a different dictionary. Teachers do need to listen for the regional variations in dialect that influence these sounds as well as all other vowel sounds. The purpose of phonics instruction is not to teach children the "correct" pronunciation but to help them map letters and letter combinations to their own pronunciation.

Discussion

Clymer's work is to be respected for focusing our attention on the shortcomings of broad generalizations. The analysis presented here regarding vowel pairs and final *e* does not contradict his findings, but instead shows that when broad generalizations with low percentages of utility are broken down into specific vowel combinations

there are, in many cases, high degrees of utility. My analysis corroborates the work of Burmeister (1968) and Caldwell, Roth, and Turner (1978) who also concluded that the study of specific vowel combinations was preferable to teaching broad generalizations. Some of these vowel combinations have more than one sound commonly associated with them but those should be considered alternatives rather than irregularities. Students need to develop flexible strategies or a "set for diversity" as suggested by Williams (1968) and Levin and Watson (1963). Their research determined that alternative sounds, such as the *ow* sounds in *how* and *grow*, could be taught simultaneously. Context can usually confirm which works in a word (The girl with the *bow* took a *bow*).

This analysis did reveal dramatically different results when the three generalizations governing final y were studied using a different dictionary or interpreted slightly differently. Clymer discussed the problem of using a dictionary that did not reflect his own dialect but did not question the authority of the dictionary. He did admit that some generalizations might be restated but his purpose was to test the generalizations as stated in teacher's guides, not to rewrite them. His remark about selecting generalizations that were stated specifically enough to aid or hinder pronunciation is curious. It suggests that Clymer was aware of selecting both good generalizations as well as poor ones, and we should therefore be careful about the conclusions drawn from his analysis. The inclusion of bad generalizations contributes to an overall reduction in the phonetic regularity readers are likely to attribute to English and to the faith teachers are likely to hold for the teaching of phonics.

Finding and establishing a level of utility is risky business, and Clymer recognized this. Percentages can change depending upon the manner in which words are analyzed. For example, the percentage for *ui* plunged when *gui* and *qui* words were included, but the *o-e* category improved when *ore* words were included as having a long *o*. The pair *au* would have had a higher percentage if the three derived forms of the irregular word *laugh* had not been included. Percentages can also change dramatically depending upon the dictionary used as a pronunciation guide. Percentages are offered here and compared with other studies, but using any numerical cutoff to determine instructional value is ill-advised. We simply should not conclude that vowel pairs are not worth teaching because they represent the long sound 45% of the time. That ignores the fact that thousands of words have vowel pairs that need to be used by readers.

Clymer's decision to ignore generalizations that did not apply to more than 20 words in the 2,600 he analyzed is also questionable. Even vowel pairs that occur only rarely in primary reading materials may be worth learning because they will continue to show up in higher level words. The *oy* pair is a good example. While it showed up in only eight words out of the 3,000 analyzed here, a search through the entire corpus of the AHD on CD-ROM turned up 312 words containing *oy*. All but *buoy* and *coyote* had the expected sound. Familiarity with the *oy* pattern will make it possible to read rare words like *gargoyle* and *clairvoyant*.

Many vowel generalizations would be of higher utility if dictionary pronunciations were not applied quite so strictly. Besides the fact that experts cannot agree and regional dialects will vary, there is the fact that context offers redundant sources of support for word recognition, especially for young readers. With this in mind I would argue that we should not be overly concerned about how strictly vowel sounds might be categorized by linguists and dictionaries. Words like *care, fair, fear,* and *cure,* which have *r*-controlled vowels and not long vowels according to linguists, seem close enough to long vowels for practical purposes. Any attempt at applying phonic generalizations is an approximation. When a young reader struggles to sound out an unfamiliar word encountered in print, phonemes are isolated, stretched out, and distorted. The resulting approximation combined with context clues is then likely to trigger recognition and a more accurate pronunciation of a word (Thompson, 1999; Venezky & Massaro, 1979).

Approximation and flexibility are two reasons why teachers should feel confident teaching some of the generalizations such as the final e signaling the long sound of the preceding vowel. Even when a consistent letter-to-sound match is lacking, a reader has limited options from which to choose. The *ea* pattern represents long *e,* long *a,* and short *e,* but never sounds associated with other vowels. Knowledge of letters and patterns and vowel sounds associated with them will help students get very close to a word's pronunciation, which is

then checked against context as well as their own oral language.

Clymer's work does not offer a useful guide to what needs to be taught. Vowels persist as the most difficult problem students face, yet only six generalizations involving vowels meet his criteria, and these address scattered issues. Simplistic broad generalizations do not capture the complexity of English orthography, yet when they are refined and stated in more specific ways there is the danger that they will become clumsy and complex. Gates's (1983) attempts at restating generalizations reassure us that English is more regular than Clymer's findings would predict, but they do not seem comprehensible for the young readers who might try to apply them. There is clearly a need to think beyond generalizations as we try to conceptualize the content that might be covered in a phonics program.

Alternatives to broad generalizations

The main issue addressed in Clymer's study is an important one we should heed. Broad phonic generalizations are not especially useful in very many cases. However, that should not be interpreted to mean that phonics instruction is not useful or that English orthography is too irregular to be the subject of study. Research in the years since Clymer's study offers ideas about alternatives to the teaching of generalizations.

Caldwell, Roth, and Turner (1978) proposed the study of higher order units such as *ope* or *op* as an alternative to complex rules about final *e* on the basis of their finding that teaching rules was effective only when the rules were simple. Recent research has established that word recognition depends more upon pattern recognition than abstract rules (Adams, 1990; Seidenburg & McClelland, 1989). As the reader sees patterns repeatedly (such as the *ai* in *mail*, *paint*, and *constrain*), the graphic letter pattern is stored in memory in association with sound. When a familiar pattern is detected in a word, that pattern evokes those stored associations. This enables a reader to decode a novel word such as *complaint*. Bussis, Chittenden, Amarel, and Klausner (1985) have described the brain as an exquisitely designed pattern detector, so it makes sense to focus more upon patterns than generalizations.

Patterns. What exactly is meant by the term *pattern*? Vowel patterns might be thought of at several levels of specificity. The lowest level, and most specific, are rimes that include a final consonant (*aid, ode*). The next level is limited to the vowels (*ai, ay*, or *oi*), and the highest level cuts across different vowels and can be represented as a sequence of consonants and vowels (CVVC or CVCe).

Wylie and Durell (1970) argued convincingly that phonograms (or rimes as we are apt to call them currently) are a powerful and effective way to teach letter-sound correspondences. Vowels tend to represent the same sound within a rime pattern and need not be identified as having long, short, or other vowel sounds. Research by Treiman (1985), Ehri and Robbins (1992), and Goswami and Bryant (1992) has established that even young readers can use rimes to figure out unfamiliar words (although they disagree about how soon readers can do this). However, there are hundreds of rimes. The rimes for the *ai* pattern alone include *aid, ail, aim, ain, air, ait*, and *aint*. There are simply too many rimes to study them all individually.

The study of a vowel pattern, such as the pair *ai*, is more encompassing because there are hundreds of words with the *ai* pair that work regularly, though not as regularly as the rimes. These are the patterns that have been analyzed in this study. I examined 24 vowel patterns in this study, although some *r*-controlled patterns constitute additional ones in some cases. The highest level of pattern is the most abstract (or general) and was the level of analysis used by Clymer. These patterns are described in broad generalizations and are not highly reliable. Still there are reasons why students should be able to recognize these abstract patterns. For example, a pattern such as CVC versus CV VC can cue a speller trying to decide whether to double the final consonant before adding *ing*, *er*, or *ed*.

Phonics instruction can address all of these levels at different times and for different reasons. Certainly it would make the teaching of phonics easier if all 19 vowel pairs could be reduced to a single, simply-stated generalization like "two vowels go walking," but this is not the case. However, we should not lose heart at the

prospect of teaching a dozen or so useful vowel patterns as an alternative.

Implications for teaching

The issue of how to teach these vowel patterns is beyond the scope of this article but a few comments seem in order. The teaching of phonics is broadly categorized as either synthetic, in which children are directly taught letter-sound correspondences and then how to blend them to generate a pronunciation, or analytic, in which children examine known words to discover patterns and regularities to extend to unknown words. In either case, teachers need to know what patterns have high utility (such as vowel pairs that represent one sound consistently and final *e*); as well as those that are variable and less reliable but still common (such as vowel pairs in which sounds alternate and *r*-controlled vowels). This latter group will be more challenging to teach but activities in which students categorize words by sound and patterns, as in word sorts (Bear, Invernizzi, Templeton, & Johnston, 2000; Bear & Templeton, 1998), hold promise for effective instruction within a comprehensive reading program that includes lots of reading and writing (Stahl, Duffy-Hester, & Stahl, 1998).

Word sorting activities encourage students to think flexibly about letter-sound correspondences and include the examination of exception words. Word sorts can also examine patterns at different levels of specificity. Word families might be sorted by rimes or phonograms (*cat, mat, sat, vs. bag, rag, tag*) in the early part of first grade (Johnston, 1999), while words might be sorted by different spellings of the long-*a* sound (*ai, ay, a-e*) in the latter part. In preparation for the study of two-syllable words and the addition of suffixes, students might sort words by CVC, CVCe, or CV patterns across the different vowels in second or third grade.

Another important issue to consider when planning phonics instruction is to determine whether particular students even need it. Thompson (1999) has argued that students learn phonics in two distinct ways: from instruction and from experiences with print. In the latter case students develop implicit unconscious knowledge about letter-sound relationships as they encounter words sharing similar features in their reading. Both sources of knowledge can develop concurrently and are available for students when they encounter an unfamiliar word. Some children are obviously better at implicit learning than others, and there are probably some phonics features that are easier to make sense of than others. A little phonics instruction may go a long way with some children, while others may need long-term systematic instruction to become independent readers. Teachers will need to carefully observe their students as they read and as they write to determine who needs what.

Phonics, patterns, and meaning

It is impossible to neatly categorize sounds and letter combinations in such a way that simple generalizations will work reliably. We have inherited a very complex orthography that cannot be reduced to a few simple rules of letter-sound correspondence easily taught by the end of second grade as recommended in *Becoming a Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1985). Instead, English must be examined across the grades as a complex system that is basically phonetic, but also relies upon patterns and meaning to provide an optimal system (Chomsky & Halle, 1968; Henderson & Templeton, 1986). Obviously this system can be mastered, but most students cannot master it quickly or easily, and some kind of instruction in phonics is important (Adams, 1990; Stahl, Duffy-Hester, & Stahl, 1998). hope this study, as well as review of research, provides some information and guidelines for teachers who may have different beliefs about the manner of teaching phonics as well as the extent of teaching required, but who share a common concern about accurately presenting children with information about English orthography, no matter how complex it might be.

References

Adams, M.J. (1990). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.
Anderson, R.C., Hiebert, E.H., Scott, J.A., & Wilkinson, I.A.G. (1985). Becoming a nation of readers: The report of the Commission on Reading. Washington. DC: National Institute of Education.
Bally, M.H. (1967). The utility of phonic generalizations in grades one through six. The Reading Teacher, 20, 413-418.
Bear, D., Invernizzi, M., Templeton. S., & Johnston. F. (2000). Words their way: Wordstudy for phonics, vocabulary, and spelling instruction, 2nd ed. Englewood Cliffs, NJ: Merrill.

Bear, D., & Templeton, S. (1998). Explorations in developmental spelling: Foundations for learning and teaching phonics, spelling, and vocabulary. The Reading Teacher, 52. 222-242.
Burmeister, L.E. (1968). Vowel pairs. The Reading Teacher. 21, 445-452.
Bussis, A., Chittenden, E., Amarel, A., & Klausner. E. (1985). Inquiry into meaning: An investigation of learning to read. Hillsdale. NJ: Erlbaum.
Caldwell, F.C., Roth, S.R., & Turner, R.R. (1978). A reconsideration of phonic generalizations. Journal of Reading Behavior, 10, 91-96.
Calfee. R. (1998). Phonics and phonemes: Learning to decode and spell in a literature based program. In L.C. Ehri & J. Metsala (Eds.), Word recognition in beginning literacy (pp. 315-340). Hillsdale, NJ: Erlbaum.
Carroll, J.B., Davies, P., & Richman, B. (1971). The American Heritage word frequency hook. New York: Houghton Mifflin.
Chomsky, N., & Halle. M. (1968). The sound pattern of English. New York: Harper & Row.
Clymer, T. (1963/1996). The utility of phonic generalizations in the primary grades. The Reading Teacher, 16/50, 252-258/182-185.
Curry, R.L., & Geis, L. (1976).A summary of studies on the usefulness of phonic generalizations. (ERIC Document Reproduction Service No. 132 560)
Davis, L.S. (1969). The applicability of phonic generalizations to selected spelling programs. Unpublished doctoral dissertation, University of Oklahoma, Norman.
Ehri. LC., & Robbins, C. (1992). Beginners need some decoding skill to read words by analogy. Reading Research Quarterly, 27, 13-26.
Emans, R. (1967). The usefulness of phonic generalizations above the primary level. The Reading Teacher, 20, 4 19-425.
Ferguson, L.H. (1970). The applicability of specific phonic generalizations to elementary mathematics textbooks. Unpublished doctoral dissertation, University of Oklahoma, Norman.
Gates. L. (1983). A face lift for the silent e. The Reading Teacher. 37.102-103.
Gates, L. (1986). The consonant generalizations revisited. Reading Horizons. 26, 232-236.
Goswami, U., & Bryant. P. (1992). Rhyme, analogy, and children's reading. In P. Gough, L.C. Ehri. & R. Treiman (Eds.), Reading acquisition (pp. 49-63). Hillsdale. NJ: Erlbaum.
Hanna. P.R., Hanna, J.S.. Hodges, R.E.. & RudorL E.H. (1966). Phoneme-grapheme correspondences as cues to spelling improvement. Washington, DC: U.S. Government Printing Office.
Henderson, E.H., & Templeton. S. (1986). A developmental perspective of formal spelling instruction through alphabet, pattern. and meaning. Elementary School Journal. 86, 305-316.
Jernigan, M.L. (1969). The utility of phonic generalizations to selected science series. Unpublished doctoral dissertation, University of Oklahoma, Norman.
Johnston, F. (1999). The timing and teaching of word families. The Reading Teacher. 53.64-75.
King. E.P. (1970). The applicability of phonic generalizations to social studies programs. Unpublished doctoral dissertation, University of Oklahoma, Norman.
Kucera, H.. & Francis, W.N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.
Levin, H., & Watson, J. (1963). The learning of variable grapheme-to-phoneme correspondence: variations in the initial consonant position. In A basic research program on reading, USOE Cooperative research project no. 639. Ithaca, NY: Cornell University.
Morris. W. (Ed.). (1970). The American heritage dictionary of the English language. New York: Houghton Mifflin.
Seidenburg. M.S., & McClelland, J.L. (1989). Sublexical structures in visual word recognition: Access units or orthographical redundancy. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 245-263). Hillsdale, NJ: Erlbaum.
Stahl. S.A.. Duffy-Hester, A.M., & Stahl. K.A.D. (1998). Everything you wanted to know about phonics (but were afraid to ask). Reading Research Quarterly, 33.338-355.
Thompson, G.B. (1999). The process of learning to identify words. In G.B. Thompson & T. Nicholson (Eds.), Learning to read: Beyond phonics and whole language (pp. 25-54). Newark. DE: International Reading Association.

Thorndike, EL., & Large, I. (1944). The teacher's hook of 30,000 words. New York; Teachers College Press.
Treiman, R. (1985). Onsets and rimes as units of spoken syllables: Evidence from children, Journal of Experimental Child Psychology. 39, 161-181.
Venezky, R.L., & Massaro, D.W. (1979). The role of orthographic regularity in word recognition. In L.B. Resnick & P.A.
Weaver (Eds.), Theory and practice of early reading, Vol 1. (pp. 85-107). Hillsdale, NJ: Erlbaum.
Williams, J.P. (1968). Successive vs, concurrent presentation of multiple grapheme-phoneme correspondences. Journal of Educational Psychology, 39.309-314.
Wylie, RE., & Durrell, D.D. (197Q). Teaching vowels through phonograms, Elementary English, 47. 787-791.