
What the Quasi-Regular Orthography of English Means for Bringing Students to Proficient Reading

Elfrieda H. Hiebert
TextProject

ABSTRACT

English orthography, described as quasi-regular, exhibits variability in the correspondences between letters and sounds, particularly with vowels. Proficiency in reading demands automaticity in connecting letters and sounds, necessitating systematic phonics instruction. However, the complexity of English orthography and its morphology means that becoming proficient in reading takes time and requires substantial exposure to text. Success in reading English requires a set for variability, where readers apply phonetic knowledge flexibly to decode unfamiliar words. While evidence supports instruction in letter-sound correspondences, questions persist regarding the extent of content coverage and the efficacy of decodable texts.

Elfrieda H. Hiebert is president and CEO of TextProject, an educational nonprofit, Santa Cruz, CA, USA

hiebert@textproject.org
<https://orcid.org/0000-0002-2864-7549>



In 1919, William S. Gray published "Principles of Method in Teaching Reading, as Derived from Scientific Investigation." Among Gray's 500 publications over the next 40 years, the words *research* and *investigation* would figure prominently. Gray was intent on translating his research into practice, including creating a reading assessment that remains in use (Gray Oral Reading Test). He also founded the International Reading Association with the goal of giving practitioners access to scientific information.

This glimpse into history is a reminder that efforts to ensure that reading instruction builds on scientific evidence have a long history. One of the first federal investments into educational research was to gather evidence on the best methods for teaching beginning readers (Bond & Dykstra, 1967). In subsequent decades, several national commissions revisited and updated research on beginning reading (Anderson et al., 1985; National Reading Panel, 2000; Snow et al., 1998).

The current science of reading movement has the same objective of increasing evidence-based reading instruction as previous initiatives, but it is distinct in one important way: Educators have no nationally vetted report they can refer to. What is deemed scientific can be based on newspaper articles, blogs, and social media posts. This is not to negate the presence of reputable sources within publications on the science of reading (see, e.g., special issues of *Reading Research Quarterly*, 2021). However, pressing questions remain about who is giving studies the imprimatur of science and whether the representation of research is comprehensive in the many online posts.

The number of studies on reading processes and instruction is gargantuan. To illustrate, I extracted data on scholarly journals with "reading" in the title from a database of referred journals (Scimago Journal & Country Rank). As shown in Table 1, approximately 400 articles were published in 2022 in the six journals that met my criterion. In addition to the backlog in these journals (approximately 15,000 articles), research on reading processes and instruction is published in approximately another 100 journals, among them the highly rated *Journal of Educational Psychology* and *American Educational Research Journal*.

I could list additional areas of research that are germane to understanding processes of reading and instruction, but that is not the purpose of this essay. My intent is to describe a body of work that rarely enters discussions of reading instruction: linguistic research on the nature of the English lexicon and its underlying systems. The field of linguistics is as complex and diverse as reading research; in this context, I focus on the orthographic system of English. Attention to the uniqueness of the English orthographic system is required for success in efforts to transform reading instruction. The remainder of this essay includes (a) a description of the English orthographic system, (b) conclusions from research that

support specific instructional practices, and (c) critical issues related to instruction that have yet to be addressed.

Table 1: *Number of Papers Published in 2022 and Potential Publications for Six Journals*

	Papers: 2022 (#)	Years Published (#)	Potential Papers (#)
<i>Journal of Research in Reading</i>	30	45	1,350
<i>Reading and Writing</i>	170	34	5,780
<i>Reading and Writing Quarterly</i>	51	31	1,581
<i>Reading Psychology</i>	29	44	1,276
<i>Reading Research Quarterly</i>	67	58	3,886
<i>Scientific Studies of Reading</i>	45	19	855

English Orthography and What It Means for Proficient Reading and Its Instruction

Overview of English Orthography

English is an alphabetic language, which means the letters of its alphabet represent the sounds in spoken language. Some languages have highly regular associations between the letters of their alphabets and the sounds. English does not. Linguists have described the English orthography (i.e., spelling system) as quasi-regular (Share, 2021). This means that, although English writing is always alphabetic, the relationships between the 26 letters of the English alphabet and the 44 sounds of spoken language can vary.

Relations between the 21 letters and 24 phonemes that are associated with consonants are consistent (Moats & Brady, 2000). It is with vowels that the variability in letter-sound correspondences is substantial. In Fry's (2004) analysis of 17,310 words, the 20 vowel phonemes had correspondences with 127 letters or letter combinations. Table 2 illustrates the variability in relations between vowel phonemes and graphemes with the phoneme /i/, which sounds like the name for the letter e.

In looking at the frequencies of associations in Table 2, remember that Fry (2004) did

not differentiate between appearances in monosyllabic and multisyllabic words. Each syllable in a multisyllabic word has a vowel, and multisyllabic words outnumber monosyllabic words in English (Vousden, 2008), meaning there is a heavy representation of multisyllabic words in the total numbers in the data in Table 2. Multisyllabic words, especially ones formed with the addition of suffixes and prefixes, become exceedingly more common as texts become more complex.

What the Features of English Orthography Mean for Learning and Instruction

Within the literature on reading, researchers have addressed critical questions. For some of the questions, strong evidence supports practices and conclusions. In this section I summarize four central findings.

Proficiency in reading an alphabetic language, even with a quasi-regular orthography, will not occur without automaticity in connecting letters and sounds. As far back as the First-Grade Studies in the 1960s (Bond & Dykstra, 1967), research has been unequivocal that automatically connecting letters and sounds is fundamental to proficient reading. In a national report I was part of, Anderson et al. (1985) concluded that “children who are taught phonics get off to a better start in learning to read than children who are not taught phonics” (p. 37). A meta-analysis conducted by members of the National Reading Panel (2000) concluded that systematic phonics instruction enhances children’s success in learning to read.

Success in reading requires that readers develop a set for variability early on. In describing the role of systematic phonics instruction, the National Reading Panel (2000) noted that “programs that focus too much on the teaching of letter–sound relations and not enough on putting them to use are unlikely to be very effective” (p. 2–96). A reason for this caution is that although there are many words in English with consistent and common vowel patterns, variability in vowel patterns is substantial among frequently used words in writing. Of the 10 most-frequent words in written English (Zeno et al., 1995), half have irregular letter–sound correspondences for vowels: *the, of, to, a, was*. Early on, successful readers of English require a set for variability—a recognition that they need to test whether their decoding effort has produced the right word (Steady et al., 2019).

Becoming proficient in a quasi-regular orthography takes time. In a study of reading development in European Union countries, Seymour et al. (2003) concluded that children reading in a deep orthography—such as English—took twice as long to learn to read as those learning to read in a shallow orthography. The reason was neither the quality of

instruction nor when children entered school; instead, it was a function of the complexity of syllables and the orthographic relationships.

Further, becoming adept at recognizing English words involves morphological knowledge. English is a phonomorphemic orthographic system (Nagy et al., 2012), which means that spellings and pronunciations of words are influenced by morphological units in words. Such shifts are evident in families of words such as *pronounce* and *pronunciation* and *nation* and *national*. Guiding students to recognize the root words in morphological families and anticipate that most words are members of morphological families will greatly benefit their reading development (Lee et al., 2023). Attention to the complexity of the English orthography and morphology must extend beyond the primary grades.

Becoming automatic with the orthography and morphology of English requires substantial exposure to text. The frequency with which words appear in language and the volume of text read by students work together to ensure successful learning of both rule-governed words and their irregular neighbors. In the volume *Language at the Speed of Sight*, Seidenberg (2017) never talked about decodable texts based on the LTTM model. But he did describe the need for exposure to substantial amounts of text to develop reading proficiency in English, stating: “Readers become orthographic experts by absorbing a lot of data, which is one reason why the sheer amount and variety of texts that children read is important. . . . Major statistical patterns emerge as the child encounters a larger sample of words. . . . We don’t study orthographic patterns in order to be able to read; we gain orthographic expertise by reading” (Seidenberg, 2017, p. 92).

Summary

The English quasi-regular spelling system poses challenges for proficient reading. However, research has pointed to properties of instruction and learners’ processes that support proficient reading. Mastery in a deep orthography such as English requires time and extensive exposure to text as well as a set for variability. Morphological complexity further complicates word recognition, especially with multisyllabic and rare words. Ultimately, navigating the intricacies of English orthography and morphology demands comprehensive instructional approaches across students’ school experiences.

Where Evidence Is Not as Clear

Research is clear that instruction in letter-sound correspondences is essential to establish for students to become proficient readers. However, in national reports and in

numerous studies, questions are raised about what we do not know. Two questions that are often raised pertain to essential aspects of instruction and learners' experiences as readers.

How many letter–sound correspondences need to be taught? The description of the quasi-orthographic nature of English illustrated the variability in letter–sound correspondences. In the current science of reading, indicators point to a view of comprehensive content coverage of letter–sound correspondences (Moats & Tolman, 2019), rather than a focus on letter–sound correspondences that appear consistently in many frequent words. One publisher, in promoting the alignment of its beginning reading program with the science of reading, advertises, “Instruction guides you in explicitly teaching the 150 spellings for the 44 sounds of English, with an intentional progression and review of skills to set your students up for success” (Amplify CKLA, 2024).

As is evident in Table 2, the variation in numbers of words that share a vowel pattern is considerable. For example, vowel digraphs such as the “ee” and “ea” appear in a relatively large number of words, many among the 500 most-frequent words in written language (e.g., *green*, *tree*, *street*). Other patterns such as the “ey” in *money* appear in few words.

Table 2: Graphemes Associated With the Phoneme /i/

Grapheme	Example	Number of Appearances in Words ¹
e	me	1765
y	very	1801
ee	keep	249
ea	eat	245
e-e	these	62
ie	field	62
i-e	police	44
ey	key	40
i	ski	38
ea-e	peace	30
ie-e	piece	23
ei	sheik	16
ei-e	seize	6
eo	people	2

¹Based on database of 17,310 words (adapted from Fry, 2004)

What is the evidence for comprehensive content coverage of letter–sound correspondences? The Commission on Reading (Anderson et al., 1985) concluded that “phonics instruction should aim to teach only the most important and regular of letter-to-sound relationships because this is the sort of instruction that will most directly lay bare the alphabetic principle” (pp. 37–38). Fifteen years later, the National Reading Panel (NRP; 2000) gave a similar response: “Specifically, how many letter–sound relations should be taught, and how many different ways of using these relations to read and write words should be practiced for the benefits of phonics to be maximized? These questions remain for future research” (p. 2-98).

Has literature emerged over the almost 25 years since the NRP (2000) that answers these questions? I have searched the literature assiduously and have raised these questions at research conferences (e.g., Hiebert, 2023). To date, I have been unable to locate research that would justify a response other than the recommendation in *Becoming a Nation of Readers* to emphasize the most important and regular letter-sound relationships and the questions about letter–sound correspondence content coverage that the NRP raised.

What types of texts are most supportive for reading acquisition? Advocates of the science of reading frequently describe the critical role of decodable texts for beginning readers (The Reading League, 2020). At present, the model that is the basis for defining decodable text is Lesson-To-Text Match (LTTM; Stein et al., 1999). In the LTTM model, texts are deemed decodable to the degree that all letter–sound correspondences in words of a text have been covered in prior lessons of the program. For example, if a teacher’s guide has provided lessons on a/[æ], i/[ɪ], m/[m], n/[n], s/[s] and the words *on* and *a* as sight words, the following text is judged to be decodable: “Sam sat on a mat. A man sat on a mat. Nan sat on a mat.”

In California, the LTTM model has been part of the mandates for textbook adoptions since the early 2000s. In 2002, the California State Board of Education stated that “materials designated by the publisher as decodable must have at least 75% of the words comprised solely of previously taught sound–spelling correspondences, and from 15% to 20% of the words comprised of previously taught high frequency words and story words.” (pp. 4-5). The grade-one component of a program needed to include two decodable books per sound–spelling with a minimum of 8,000 words; at grade-two, the minimum number of words was 9,000. In the 2015 textbook requirements (California State Board of Education, 2015), 9,000 words of decodable text were stipulated in programs for interventions in Grades 4 to 8.

Despite the prominence of the LTTM model over the past 25 years, evidence for its

efficacy is sparse. Cheatham and Allor (2012) found a single study of substantial duration using the LTTM model—one conducted by Jenkins et al. (2004) who compared small-group instruction, where students received the same phonics lessons but applied their knowledge in either LTTM-based texts or leveled texts. Both intervention groups performed significantly better than comparison students on decoding and comprehension, but there were no significant effects reported between the two groups in the experimental treatments.

These findings should not be interpreted to mean that the words in beginning texts do not matter. Research has shown the critical role of texts that give students opportunities to apply the phonics they are being taught (Ehri et al., 2007; Juel & Roper/Schneider, 1985; Menon & Hiebert, 2005). Unlike the LTTM model that classifies words as decodable once the letter–sound correspondences within them have been taught, the texts in these studies were chosen to follow systematic phonics curricula that focus on consistent letter-sound correspondences in numerous frequent words.

At the same time, research on leveled texts is sparse. Studies have consistently shown that one variable—total number of words in texts—accounts for the most variance in assignment of a text's level, rather than word frequency or letter–sound correspondences (Cunningham et al., 2005; Hiebert & Tortorelli, 2022). To date, no researchers have compared the efficacy of leveled texts with another text type when all other dimensions of instruction have been held constant.

In a meta-analysis, Pugh et al. (2003) reported that the combination of decodable texts and texts based on multiple criteria such as word frequency, familiarity, and decodability may be more supportive than either text type on its own. Neither text type differed in effects on student outcomes but interventions that included reading of both decodable and texts chosen for multiple criteria had greater effects than interventions using only one of the text types.

Summary

Both the scope of the curriculum (i.e., the number of letter-sound correspondences that are taught) and the text diets of beginning readers are critical aspects of phonics instruction. Scientists have provided responses to numerous questions about phonics instruction; these two critical aspects of reading instruction illustrate that important questions have yet to be answered.

Conclusion

In this essay I have focused on scientific evidence related to word recognition. This does not mean that word recognition is the only proficiency that contributes to comprehension. If reading is not used to learn and think, the effects of an advantageous start wane. Numerous bodies of research on reading development and its instruction that have not been discussed in this article, including background knowledge and engagement and motivation. However, if we are to ensure successful initiation to reading, students' ability to decode letter-sound correspondences in unknown words needs to be automatic. For such automaticity to occur requires that students be exposed to massive quantities of text. Daily and frequent experiences with texts are essential if students are to become automatic in recognizing letter-sound correspondences and developing a set for variability. To achieve the large amounts of texts required for this automaticity to occur, the decodable texts that are prominent in beginning reading classrooms can be augmented with leveled texts and trade books that have been organized around letter-sound correspondences and topics (Hiebert, 2024). Texts that are clustered around patterns in words and topics assist students in building the bodies of knowledge that underlie proficient comprehension, while developing automaticity in decoding.

As is often the case in education, where choices of practices have consequences for students, educators are bombarded with claims and counterclaims. It is important for educators to remain aware of what research does and does not reveal and to seek support from professional organizations as they engage their students in practices that support the high levels of literacy required in the digital global age.

References

- Amplify CKLA (2024). Developing foundational skills with explicit, systematic instruction. Retrieved from <https://amplify.com/programs/amplify-core-knowledge-language-arts/>
- Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. (1985). *Becoming a nation of readers: The report of the commission on reading*. National Academy of Education.
- Bond, G. L., & Dykstra, R. (1967). The cooperative research program in first-grade reading instruction. *Reading Research Quarterly*, 2(1), 5-142.
- California State Board of Education (2002). 2002 Reading/Language Arts/English Language Development (RLA/ELD Adoption (Final Report). California Department of Education. Retrieved from: <http://eds-courses.ucsd.edu/tep129/ELArpt.pdf>
- California State of Board of Education (2015). *2015 English Language Arts/English Language Development Instructional Materials Adoption (K-8)*. California department of Education. Retrieved from: <http://www.cde.ca.gov/ci/rl/im/>
- Cheatham, J. P., & Allor, J. H. (2012). The influence of decodability in early reading text on reading achievement: A review of the evidence. *Reading and Writing*, 25(9), 2223–2246.
- Cunningham, J. W., Spadorcia, S. A., Erickson, K. A., Koppenhaver, D. A., Sturm, J. M., & Yoder, D. E. (2005). Investigating the instructional supportiveness of leveled texts. *Reading Research Quarterly*, 40, 410–427.
- Ehri, L. C., Dreyer, L. G., Flugman, B., & Gross, A. (2007). Reading rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Educational Research Journal*, 44(2), 414-448.
- Fry, E. (2004). Phonics: A large phoneme-grapheme frequency count revised. *Journal of Literacy Research*, 36(1), 85-98
- Gray, W. S. (1919). Principles of method in teaching reading, as derived from scientific investigation. *Teachers College Record*, 20(7), 26-51.
- Hiebert, E. H. Enhancing opportunities for decoding and knowledge building through beginning texts. *The Reading Teacher*. Retrieved from <https://textproject.org/wp-content/uploads/2024/03/2024TRT-Hiebert-Enhancing-Opportunities-for-Decoding-and-Knowledge.pdf>
- Hiebert, E.H. (2023). *The research we need: Decodable texts*. Presentation at the annual meeting of the Literacy Research Association as part of the Reading Hall of Fame symposium on Reading Research in the 21st Century. Retrieved from: <https://independent.academia.edu/ElfriedaHiebert/Conference-Presentations>
- Hiebert, E. H., & Tortorelli, L. S. (2022). The role of word-, sentence-, and text-level variables in predicting guided reading levels of kindergarten and first-grade texts. *The Elementary School Journal*, 122(4), 557-590.
- Jenkins, J. R., Peyton, J. A., Sanders, E. A., & Vadasy, P. F. (2004). Effects of reading decodable texts in supplemental first-grade tutoring. *Scientific Studies of Reading*, 8(1), 53–85.

- Juel, C., & Roper-Schneider, D. (1985). The influence of basal readers on first grade reading. *Reading Research Quarterly*, 20(2), 134–152. <https://doi.org/10.2307/747751>
- Lee, J. W., Wolters, A., & Grace Kim, Y. S. (2023). The relations of morphological awareness with language and literacy skills vary depending on orthographic depth and nature of morphological awareness. *Review of Educational Research*, 93(4), 528–558.
- Menon, S., & Hiebert, E. H. (2005). A comparison of first graders' reading with little books or literature based basal anthologies. *Reading Research Quarterly*, 40(1), 12–38.
- Moats, L. C., & Brady, S. (2000). *Speech to print: Language essentials for teachers*. Paul H. Brookes Pub.
- Moats, L.C., & Tolman, C.A. (2019). *LETRS* (3rd Ed.). Voyager Sopris Learning.
- Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle-school students. *Journal of Educational Psychology*, 98(1), 134–147.
- National Reading Panel (2000), *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. National Institute of Child Health and Human Development.
- Pugh, A., Kearns, D. M., & Hiebert, E. H. (2023). Text types and their relation to efficacy in beginning reading interventions. *Reading Research Quarterly*, 58(4), 710–732.
- Reading League, The (2020, May). *Decodable text sources*. <https://www.thereadingleague.org/wp-content/uploads/2020/05/FINAL-Decodables-Update-May-2020.pdf>
- Seidenberg, M. (2017). *Language at the speed of sight: How we read, why so many can't, and what can be done about it*. Basic Books.
- Seymour, P. H., Aro, M., Erskine, J. M., & Collaboration with COST Action A8 Network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174.
- Share, D. L. (2021). Is the science of reading just the science of reading English?. *Reading Research Quarterly*, 56, S391–S402.
- Snow, C. E., Griffin, P., & Burns, M. S. (Eds.). (1998). *Preventing reading difficulties in young children*. National Research Council.
- Steady, L. M., Wade-Woolley, L., Rueckl, J. G., Pugh, K. R., Elliott, J. D., & Compton, D. L. (2019). The role of set for variability in irregular word reading: Word and child predictors in typically developing readers and students at-risk for reading disabilities. *Scientific Studies of Reading*, 23(6), 523–532.
- Stein, M., Johnson, B., & Gotlohn, L. (1999). Analyzing beginning reading programs: The relationship between decoding instruction and text. *Remedial and Special Education*, 20, 257–287.
- Vousden, J. I. (2008). Units of English spelling to sound mapping: A rational approach to reading instruction. *Applied Cognitive Psychology*, 22(2), 247–272.
- Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Touchstone Applied Science Associates, Inc.