

# The Texts of Literacy Instruction: Obstacles to or Opportunities for Educational Equity?

Literacy Research: Theory,  
Method, and Practice  
2017, Vol. 66, 117-134  
© The Author(s) 2017  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/2381336917718521  
journals.sagepub.com/home/lrx



Elfrieda H. Hiebert<sup>1</sup>

## Abstract

Texts are a central part of reading. Yet our understandings of appropriate text features and distributions of text diets at different points in students' reading development are limited. The thesis of the essay is that, if the trajectory of struggling readers is to change, attention is needed to the features of texts and students' text diets, especially those of students who attend schools in heavily impacted communities. Three issues are identified that require the attention of researchers to ensure appropriate texts and text diets for struggling readers: (a) texts from the earliest levels need to be meaningful, (b) at least part of struggling readers' text diets need to be with texts in which words with morphological and phonological consistency are repeated, and (c) amount of text read by struggling readers needs to be substantial for reading capacity to increase. For each issue, the manner in which current practices can contribute to potential obstacles for struggling readers is described. Next, research on alternative practices is presented that shows how shifts in texts and text diets can support higher reading proficiency. The essay ends with a description of a research agenda that uses digital resources to increase students' facility with vocabulary in complex texts.

## Keywords

text analysis, text features, text diets, struggling readers, core reading programs

---

<sup>1</sup> TextProject, Santa Cruz, CA, USA

## Corresponding Author:

Elfrieda H. Hiebert, TextProject, 106 Phelan Court, Santa Cruz, CA 95060, USA.  
Email: hiebert@textproject.org

Reading comprehension involves three elements: a reader, a text, and an activity or purpose for reading (RAND Reading Study Group, 2002). Of the three elements, text has been the one that has generated most debate among literacy researchers (e.g., Alvermann & Commeyras, 1994). Among the theorists calling for expanded views of texts have been Goodman and Goodman (1979), who noted that environmental print, such as stop signs, provided young children with early encounters with print. My first two published studies (Hiebert, 1978, 1981) drew on this perspective as I examined preschoolers' understanding of print in everyday situations. Even 3-year-olds showed an awareness of the functions of print, such as the labels on packages, in their environments (Hiebert, 1981). Masonheimer, Drum, and Ehri (1984) did a follow-up study in which they transposed environmental print (e.g., Coca-Cola on a stop sign). They concluded that young children couldn't recognize these words out of context.

It had never been my view that young children need to display independent word recognition of environmental print. Rather, I view environmental print as a way for young children to bring understandings of print to tasks of conventional literacy. Further, even as an early adopter of an alternative view of text, I recognized that individuals need to have proficiency with the texts of the workplace and community and that the design of texts is particularly critical if students who live in diverse economic, linguistic, and social communities are to attain high levels of literacy.

Despite the importance of text in reading, surprisingly few models of texts exist. One exception is the framework for beginning reading texts outlined by Mesmer, Cunningham, and Hiebert (2012). Mesmer et al. argued that the selection and creation of appropriate texts for beginning readers requires attention to programmatic, discourse, syntactic, and word-level features. In this article, I extend this work on model building to texts for struggling readers.

This article has three parts. First, I examine a current text solution that has become ubiquitous in programs for struggling readers since the launch of the Common Core State Standards (CCSS; National Governors Association Center for Best Practices [NGACBP] & Council of Chief State School Officers [CCSSO], 2010). Second, I present three principles of text features and use, describing a framework of text with fairly broad brushstrokes. Finally, I discuss my conclusions in relation to future research.

## **Illustration of a Current Text Solution**

The programs that I use to illustrate the vagaries of many text solutions consist of current events articles. Sets of expository articles are not unique. What is unique about these programs is the provision of texts on the same topic at different levels of the text complexity staircase identified by CCSS writers (NGACBP & CCSSO, 2010). The steps of the staircase of text complexity are depicted in Figure 1 in Lexiles, as recommended by CCSS writers in appendix A of the standards. As Figure 1



**Figure 1.** Common Core State Standards staircase of text complexity with average Lexile levels of leveled article texts.

demonstrates, the five options for texts of one of the current text solutions map onto the CCSS’s staircase of text complexity.

The idea that with a click or two each student can be provided with “just right” text is compelling. However, I could find no evidence in the archival literature that these programs positively support the reading capacity of struggling students. At present, there are two primary programs that I refer to as “leveled articles”: Newsela (2017) and Achieve3000 (2017). In lieu of empirical evidence, I examined the features of texts in one-leveled article program relative to findings of effective texts in the archival literature (Hiebert, 2016). My study consisted of 500-leveled articles, a text for each of the five levels for 100 topics from Newsela (2017). As shown in Table 1, number of words and mean sentence length increases as text levels increase. For the mean log word frequency (MLWF)—the Lexile Analyzer’s measure of semantics (Stenner, Burdick, Sanford, & Burdick, 2007)—the pattern is the opposite. The MLWF decreases with higher text levels. That is, lower level texts have higher MLWFs.

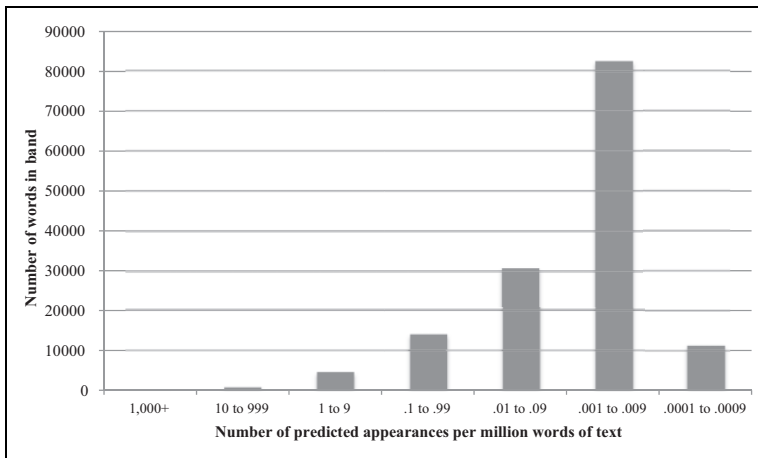
If struggling readers get the lowest level texts, they read texts that are 41% as long as those for proficient readers, sentences that are slightly less than half as long, and words with higher average frequency. Are these features validated by research? That is, do short texts with short sentences and many high-frequency words increase the capacity of struggling readers?

Research on the effects on comprehension of texts with short rather than more complex sentences is extensive and the findings consistent. For example, Pearson (1974) found that students had the same or better comprehension on complex than simple sentences. Pearson concluded that rewriting the ideas as two simpler sentences eliminates the causal relationship, placing a greater inferential load on readers.

By the mid-1980s, the findings on the negative effects of shortening sentences to obtain lower readability levels were substantial enough that the commission of reading’s report, *Becoming a Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1985), cautioned against overreliance on readability formulas for creating or selecting textbooks. Subsequent research has validated this conclusion. For example, Arya, Hiebert, and Pearson (2011) found that lower lexical complexity had a significant,

**Table 1.** Means (and Standard Deviations) for 100 Texts at Five Lexile Levels of a Leveled Article Program.

Publisher's Designated Lexile Level	Lexile	Mean Sentence Length	Mean Log Word Frequency	Total Words
600	597 (21.2)	9.23 (0.41)	3.56 (0.14)	413 (31.11)
850	800 (35.4)	11.75 (0.98)	3.50 (0.06)	612 (108.19)
1,040	994 (14.1)	14.63 (0.35)	3.42 (0.08)	750 (260.92)
1,210	1,150 (35.4)	17.48 (0.71)	3.35 (0.10)	827 (253.14)
Max	1,279 (155.6)	20.18 (3.14)	3.27 (0.11)	1,006 (362.75)

**Figure 2.** Number of words in frequency bands in the *Educator's Word Frequency Guide*.

positive effect on comprehension of two of the four topics but that syntactic changes, either harder or easier, did not influence comprehension of any topic.

The finding on lexical complexity is one of the most enduring in the field of literacy (Davis, 1942). The Lexile Framework, however, does not provide an in-depth measure of lexical complexity but, rather, uses an average of the frequencies of all words in a text as its semantic measure. An average of word frequency is a less than ideal measure of lexical complexity because of the skewed distribution of English words. Consider the distribution of words in Figure 2, which represents the 154,941 unique words derived from 17.25 million words from texts spanning Grades 1–13 within the *Educator's Word Frequency Guide* (Zeno, Ivens, Millard, & Duvvuri, 1995). Words that are predicted to account for 48% of total words in texts—107 words—account for less than 1% of the words in the database. The majority of words—89%—in the database appear less than once per million words. The disparate distribution of words means that the range of the word frequency average of texts is limited even when a logarithm is applied.

Not only is there no evidence that leveled articles positively influence struggling readers, but evidence from the archival literature indicates that features of these texts are contrary to those that support higher comprehension. But how did texts such as these become all the rage? One explanation is digital capacity. Texts can now be changed quickly with algorithms and a minimum of human input. This explanation is inadequate, however, because such capacity has been available for at least 20 years.

The justification for these programs, I argue, was the decision of CCSS writers (NGACBP & CCSSO, 2010) to identify explicit quantitative parameters for text levels within the staircase of text complexity. From the late 1980s, when states such as California (California English/Language Arts Committee, 1987) mandated that reading texts manipulated by readability formulas could not be purchased with state funds until 2010 when the CCSS was published, readability formulas have not been used to change or select texts in core reading programs (Hiebert, 2015). The CCSS writers went a step further than providing the staircase of text complexity based on the Lexile Framework when they provided an addendum with guidelines for additional quantitative text complexity systems—the only modification to the CCSS (Coleman & Pimentel, 2012). Such an action is more than tacit support for changing and assigning texts based, primarily, on the criterion of sentence length.

## **Rudiments of a Multidimensional and Adaptive Model**

This illustration demonstrates the urgency for attention to my thesis: For the majority of students to become highly literate, literacy scholars need to provide well-grounded principles that counter policies such as the staircase of text complexity. In this section, I identify three principles related to texts and reading experiences that aim to change the trajectories of struggling readers. The model is not comprehensive; rather, I have chosen three topics that require extensive and intensive attention from our community. Without such attention, educators are vulnerable to quick fixes that, at best, maintain the status quo and possibly create even more obstacles to educational opportunities for the students most in need.

### *Language Is a Meaning-Making Process*

The first principle of texts for struggling readers is that texts need to be meaningful. At first glance, this principle may seem so apparent as to be tautological. After all, reading, as a language process, focuses on meaning. Humans—young and old—consistently approach language interactions—oral and written—with the expectation that there will be meaning. When students with few prior academic literacy experiences are asked to learn to read with nonsensical texts, we should not be surprised if they fail to become proficient comprehenders. And nonsensical text is precisely what has formed the primary reading diets of a generation of American students, especially those who are perceived as at risk. Excerpts of texts called decodables appear in Table 2. In California's textbook adoption (California State

**Table 2.** Examples of Text Types.

Text Type	Excerpt
First-grade decodable <sup>a</sup>	Brad's ram spins and nabs his hat. Brad is mad. Brad nabs his hat. Snap! The hat!
Middle school decodable <sup>b</sup>	The con man said, "I can shave sheep. I can shape. I can shear" "But how is your rate at shearing?" The rancher asked
Meaning text <sup>c</sup>	"The sun is out. Do you have the seed?" asked Paul. "We need to plant it in the soil"

<sup>a</sup>Adams et al. (2000). <sup>b</sup>Englemann and Bruner (1995). <sup>c</sup>Hiebert, Brown, Taitague, Fisher, and Adler (2003).

Board of Education, 2006) during the No Child Left Behind Act (NCLB) period (and still in use in many classrooms across the country, especially for interventions), core reading programs needed to include a set of decodables at Grades K–3. Further, decodable texts continue to be mandated in programs for struggling readers and English learners (ELs), including Grades 4–8 (California State Board of Education, 2015).

The depth and quality of knowledge about the meanings of words influences students' word recognition, as Perfetti (2007) describes in his lexical quality hypothesis. Yet, in decodable texts, low-meaning words are numerous and rarely repeated by design to ensure that students are relying on their decoding skills and have not memorized words. Further, programs need to provide texts with practice on all 44 phonemes and their associated graphemes, of which there are approximately 250, to attain the decodable status (Stein, Johnson, & Gutlohn, 1999).

As an alternative to decodables, my colleagues and I (Hiebert, Brown, Taitague, Fisher, & Adler, 2003) generated a set of texts for first-grade ELs that was based, first and foremost, on the principle of meaningfulness. ELs might not know the English label for a concept, but particular concepts such as animals, school, playing, and food are salient in young children's lives. Such familiar but engaging concepts provided the basis for the "meaning texts," an example of which appears in Table 2.

From the target words for a topic, chosen words were ones with consistent, common vowel–consonant patterns or rimes and with substantial frequency in written English—for example, *dog* and *frog* but not *cog* or *bog*. The number of unique words in the meaning texts was 40% of the Open Court decodables (Adams et al., 2000). This figure indicates that meaning texts provided students with substantially more repetition of key words than decodable texts.

Hiebert and Fisher (2016) initiated an intervention with the meaning texts in a district with a high percentage of ELs. For 12 weeks, in three weekly half-hour sessions, the same teacher worked with two groups of three students from each of nine first-grade classrooms. The instructional routine was identical in both groups—students wrote extensively and engaged in word play. But 10 min of each half hour was devoted to reading texts, one group reading the Open Court decodables (Adams

et al., 2000) and the other the meaning texts (Hiebert et al., 2003). A third group remained in the classroom, where they received instruction in the district's core reading program, which had a second set of decodable texts (Cooper et al., 2003).

When the intervention began (second half of first grade), students had gone through one set of decodables in kindergarten and were more than halfway through a second set. Most, however, could recognize only a handful of words. Despite the equivalence of initial performances of the three groups, students in the meaning group outperformed those in the intervention decodable and classroom decodable groups in reading several texts at the end of the intervention. However, the meaning group did not perform better than the other groups on the decodable text in the assessment (Hiebert & Fisher, 2016). The meaning group also outperformed the other two groups on several word recognition tasks but not a decoding task of nonsense words. That is, when things didn't make sense, better reading skills didn't help.

### *A Modicum of Consistency in Linguistic Information Is Useful When Learning Language*

The meaning texts were filled with familiar concepts, but they also provided some consistency in phonics patterns. The second principle is that texts for struggling readers need to have a modicum of consistency in linguistic information. I underscore the word *modicum*. Learning language involves generalization, not having every single sound or word explicitly taught and practiced. The meaning texts, for example, emphasized approximately 30 phoneme-grapheme patterns, only a portion of the 250 in typical decodable programs (Stein et al., 1999).

Even representative knowledge about phonics can take readers only so far. To move beyond the novice stage in reading English, students need to be facile with the meaning parts of words or morphemes. To understand morphology's role in proficient reading requires an understanding of the lexicon or vocabulary of English. Estimates place the number of words in English at around 400,000 words (Nagy & Anderson, 1984, p. 313)—an impossible number of words to teach. When viewed from the perspective of morphology, however, the learning task becomes somewhat more manageable. Nagy and Anderson estimated that approximately 400,000 words could be parsed into 88,500 word families. An example of a morphological or word family is *adequate, adequately, and inadequate*.

Nagy and Anderson (1984) described an increase in the role of morphological families in the texts of the middle grades and beyond. The words at these grade levels would be expected to be those with lower frequencies in Figure 2. But an analysis of core reading programs showed that 52% of the unique words in Grade-1 texts and 65% in Grade-3 texts have multiple morphemes (Carlisle, Hiebert, & Kearns, 2015).

Colleagues and I (Hiebert, Goodwin, & Cervetti, 2017) set out to establish the presence of word families in the most frequent portion of the English lexicon—words with 10 or more predicted appearances in 1 million words of text. We identified 2,500 morphological families in this group. In a subsequent study (Hiebert, 2017),

**Table 3.** Staircase of Vocabulary Progression With Illustrative Texts.

Level	Core Vocabulary	Additional Linguistic Content <sup>a</sup>
A	300 Most frequent words	Words with regular short and long vowels
B	600 Most frequent words	Words with regular <i>r</i> -controlled vowels
C	1,000 Most frequent words	All monosyllabic words
D	1,000 Most frequent words	Two syllable with regular vowel patterns in both syllables
E	2,500 Most frequent words	All of above patterns
F	6,000 Most frequent words	All of above patterns

<sup>a</sup>Content from previous levels is included at each subsequent level.

I identified families among words with frequencies of one to nine predicted appearances per million words (see Figure 2). This group has approximately 1,500 word families. These two analyses indicate that there are approximately 4,000 word families in the portion of the lexicon with words of high to moderate frequency.

The next step in our research was to establish the presence of word families in the 200 exemplars of complex texts for different grade bands identified by CCSS writers. The 2,500 word families accounted for an average of 91.5% of the total words in the CCSS exemplars with percentages from 97% of total words in K–1 texts and 90% of total words in Grade 11 to college texts. The 1,500 families accounted for 5.1% of the remaining total words.

To date, my work has focused on creating texts that support struggling readers in getting facile with the core vocabulary. Core vocabulary refers to words with moderate to high frequency—the words with 10 or more predicted appearances in Figure 2. In designing “content texts” to support proficiency with core vocabulary, I drew heavily on research from the English as a Second Language (ESL) field, where a substantial amount of evidence shows that young adults who can read in their native language learn to read English well when instructional texts move through a progression of highly to moderately frequent words (Nation, 2015).

Building on this work, I generated a series of word zones around the 6,000 most frequent words—the corpus of words subsequently parsed into 2,500 word families by Hiebert et al. (2017). These word zones followed a staircase based on vocabulary rather than primarily sentence length, as in the CCSS’s staircase of text complexity. This staircase of core vocabulary (see Table 3) was the basis for creating texts in which at least 98% of the words come from the designated words at that level or lower levels (Hiebert, 2003).

Even with this goal of emphasizing the core vocabulary, a chief criterion of the content texts was on meaningfulness. Whereas current intervention texts for middle graders emphasize a con man shaving sheep, as illustrated in example in Table 2 (Englemann & Bruner, 1995), the content texts emphasized topics that are relevant to both narrative and expository texts—topics such as those included in Table 3. Many students have read these texts over the past 15 years and numerous investigators have studied their effects, concluding that consistent experiences with these texts result in



increased background knowledge, comprehension, and rate of reading (e.g., Trainin, Hayden, Wilson, & Erickson, 2016).

### *Struggling Readers Need to Read a Substantial Amount of Text to Increase Reading Capacity*

With any intervention including those with the content texts, the question can be raised as to whether instructional effects reflect more reading by students rather than the texts or strategies of the intervention. Studies have controlled for that possibility with the content texts but claims for the leveled article programs (e.g., Newsela) need to be viewed through this lens. Introducing even a little more reading into students' school days is likely to make a difference. After all, as Lewis and Samuels (2005) observed, in any human endeavor, even some degree of proficiency requires participating in the activity with some frequency.

The third principle has to do with text diets, which I define as the accumulation of texts that students read in schools. The gist of this principle is that struggling readers need to read a substantial amount of text to increase reading capacity. But the question of exactly how much reading is associated with particular levels of reading proficiency has not been established and infrequently investigated. A first step in answering this question is to examine present reading patterns—that is, how well are students reading? If students are reading at adequate levels, one assumption is that they are reading a sufficient amount.

The findings of the National Assessment of Educational Progress (NAEP) indicate that approximately 64% of a fourth-grade cohort is not attaining the proficient level (National Center for Education Statistics, 2015). Several features of the NAEP—a heavy dose of written responses, the aspirational nature of the text levels, and shifts in setting the proficient level—call for an examination of additional sources for insight into student proficiency.

One source of information comes from Spichtig et al. (2016) who compared data from two cohorts reading the same texts—one group in the late 1950s and the other in the early 2010s. High school performances in 2011 had a similar slope as that of high schoolers in the 1950s, but the lack of progress over the middle school years meant that high school graduates in the early 2010s ended their school careers with comprehension-based silent reading rates that were significantly lower than those of peers in an earlier era.

A study of the oral and silent reading of approximately 8,000 students from Grades 4–8 (Bielinski, Daniel, & Hiebert, 2015) showed that 90% or more of students read at 95% accuracy or higher—levels that many clinicians (e.g., Betts, 1946; Clay, 1991) have deemed satisfactory for comprehension. On a silent reading task, however, 60% of the sample failed to satisfactorily answer relatively easy comprehension questions (Bielinski et al., 2015). When a sample of the poor and moderate comprehenders was asked to read silently in a one-to-one context with an investigator, their comprehension rose significantly, but their reading rates decreased relative to those on the online,

independent assessment (Daniel, Hiebert, & Martin, 2016). Findings such as these suggest that the vast majority of students are proficient enough at word recognition but lack the silent reading skills to read on their own.

One explanation for these patterns of poor silent reading performances may lie in how much students read in school. From the early 1970s to early 1990s, observational studies of classroom reading showed approximately 14–15 min spent in daily reading in elementary classrooms (Fisher et al., 1980; Foertsch, 1992; Gambrell, Wilson, & Gantt, 1981). In the late 1990s (Donahue, Voelkel, Campbell, & Mazza, 1999), fourth graders reported reading an average of 10 or fewer pages daily or the equivalent of 8–12 min of daily reading in school. These results, which comprise the most recent, large-scale investigation of reading volume, suggest that the amount that students are reading in school, previously at fairly moderate levels, has possibly decreased.

This pattern of a potential decrease in reading time has occurred as the allocation to English language arts instruction has increased. Brenner, Hiebert, and Tompkins (2009) found that, in classrooms under NCLB mandates in one state, reading instructional time in third grade ranged from 90 to 120 min per day with an average of 94 min. The amount of time in which students had their eyes on text was quite consistent regardless of total classroom time: an average of 9.18 min or slightly less than 10% of the reading period.

A compelling finding in the Brenner et al. (2009) study was the mode of reading. Students spent half of their reading time in assisted reading, ostensibly following along as texts were read aloud by teacher or on an audio. This pattern of assisted reading seems to have extended across the grades. In a study of reading in high school English language arts, social studies, and science classes, Swanson et al. (2016) found that 80% of the reading occurred with someone reading aloud (teachers, peer, and audio) and students presumably reading along.

The research is not extensive, but those students who read more in their classrooms appear to do better, at least on the NAEP (Guthrie, Schafer, & Huang, 2001). Kuhn and Schwanenflugel (2009) found that an additional 1.8% of the school day spent reading distinguished classrooms with higher levels of student reading achievement than classrooms with lower levels.

As scholars have noted for decades, variables such as minutes or number of words are crude indices of students' reading experiences (Fisher et al., 1980). A study conducted by Rasinski, Samuels, Hiebert, Petscher, and Feller (2011) demonstrates the need to dig deeper in understanding reading volume. In that study, struggling middle and high schoolers only began to show improvements after they had spent a minimum of 20 hr in an online reading interventions. The study of reading volume requires attention to numerous variables including features of tasks and texts.

## **A View of Text: Conclusions and Questions**

I began with a description of a form of text that has swept through the educational marketplace to illustrate that many textual innovations are without roots in theory or

research. I next described two features of texts—meaningfulness and linguistic consistency—that are essential in the design and selection of texts for struggling readers. Before I address these two features in more depth, I review the conclusions regarding reading volume.

### *Reading Volume*

Data on student performances and classroom practices point to the need for more time spent reading in classrooms. In response to the question frequently asked by teachers as to how much reading is necessary, the answer is that we simply don't know how much of what kind of reading is needed at particular points in reading development. There are sufficient questions related to the study of texts and text diets for numerous scholars to engage in substantive research programs.

### *Meaningfulness and Linguistic Consistency*

In describing the issues of meaningfulness and linguistic consistency of texts, I presented two examples in which I have been a text developer and researcher. This involvement in text development should not be construed as a pursuit of the “holy grail of texts”—that is, the search to prove that one text type is superior to all other text types for readers of all proficiencies or for all purposes. The development of reading proficiency reflects participation with many text types. At the same time, I do believe that some texts better support some aspects of reading development for struggling readers than other text types. Our colleagues in ESL make intentional vocabulary a centerpiece of texts for young adults who can read in their native languages and are motivated to learn English (Nation, 2015). Yet, in reading instruction in the United States, we have dropped this critical feature of learning for struggling readers, even as our schools have become more linguistically, economically, and socially diverse.

We rightfully rejected the Dick-and-Jane genre in *Becoming a Nation of Readers* (Anderson et al., 1985), but the commission did not reject any control of vocabulary. The commission recommended that primary-level texts should meet the technical requirements for controlled vocabulary, while at the same time using language in artful ways. No one had addressed systematically what technical requirements and artful language meant when California's 1987 textbook adoption (California English/Language Arts Committee, 1987) mandated that only authentic texts (i.e., trade books) would be acceptable for purchase with state funds at all grade levels, including first grade. This policy was a vast misinterpretation of the report's conclusions. The number of unique words in beginning reading texts increased from around 5 new words per 100 in the mid-1980s program to 26 new words per 100 in the mid-1990s program of one core reading program (Hiebert, 2015). Such differences indicate a substantial decrease in repetition of vocabulary.

Even with policies for another type of texts—decodable ones—in the 1990s (e.g., Texas Education Agency, 1997), repetition of vocabulary was not addressed. Neither

has repetition of vocabulary been addressed in the guided reading books widely used in beginning reading instruction. A recent analysis shows that the numbers of unique words per 100 in guided reading and decodable beginning reading programs were 21 in the former and 22 in the latter (Murray, Munger, & Hiebert, 2014).

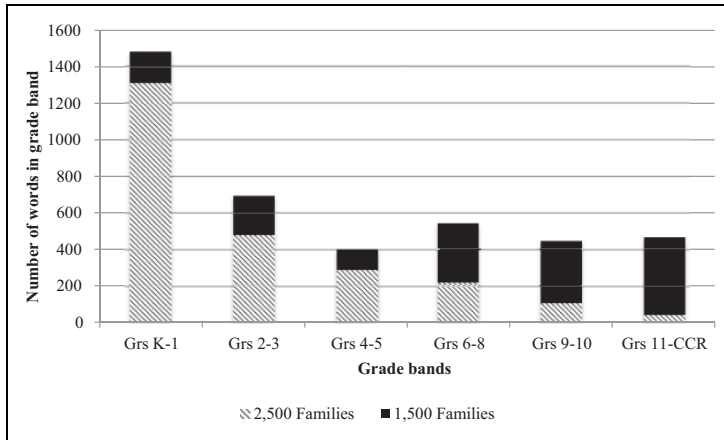
### *A Research Agenda on Parsimonious Vocabulary*

The answer is not to go back to the Dick-and-Jane genre of text but rather to address what it means to have a modicum of meaningful linguistic information in texts for struggling readers. The aim of Hiebert et al. (2017) research on a core vocabulary described earlier is to identify a “parsimonious vocabulary”—the smallest number of word families that account for the greatest proportion of words in text.

In pursuing research on parsimonious vocabulary, I’m particularly interested in the resources provided by digital capacity. Enormous numbers of texts can be analyzed to confirm the presence of the words at particular stages in students’ learning (see, e.g., Hiebert et al., 2017). Digitization also means that databases are available in which words have been tagged on features such as concreteness (Brysbart, Warriner, & Kuperman, 2014).

The prototype of a nonprofit group, Text Genome (2017), illustrates the manner in which digital resources can be used to create assessment and instruction around a parsimonious vocabulary. The 4,000 word families that were described earlier (Hiebert, 2017; Hiebert et al., 2017) were placed into the grade bands in which the lead word of a family first becomes prominent. Figure 3 depicts the distribution of the word families. The families from the 2,500 most frequent lead words are already prominent in the elementary grades but, during the middle and high school years, the 1,500 word families account for a critical portion of text.

The designation of word families to grade bands should not be viewed as a mandate for instruction of every word at a grade level. The aim of the Text Genome project (2017) is precisely the opposite. This project aims to identify the words within particular grade bands with which students are facile and those with which they could benefit from additional experience. The process begins by sorting words within grade bands. For example, the 450 words in the Grades 9–10 band are sorted into groups of 15–20 words based on 10 word features: frequency, frequency of morphological family, length, part of speech, polysemy, abstractness, age of acquisition, dispersion, semantic cluster, and identification as a general academic word. For example, the following words make up a band in the middle of Grades 9–10: *rejected*, *exceed*, *imported*, *reverse*, *qualified*, *guarantee*, *accumulated*, *conceived*, *installed*, *exert*, *inspired*, *inserted*, *imitate*, *translated*, and *anticipated*. These words all have at least one prominent meaning as a verb and are similar in the other nine features. Two words are chosen to represent a group—the words that have the mean for the majority of features. For this set of words, the target words are *installed* and *imitated*.



**Figure 3.** Distribution of 4,000 word families across grade bands.

Next, a database of texts is used to locate excerpts of the target words in texts that students are likely to encounter in high school, college, or careers, as illustrated in the following excerpt from Austen’s (1813/2008) *Pride and Prejudice*:

With a glance, she saw that he had lost none of his recent civility; and, to \_\_\_\_\_ his politeness, she began, as they met, to admire the beauty of the place; but she had not got beyond the words “delightful,” and “charming,” when some unlucky recollections obtruded, and she fancied that praise of Pemberley from her might be mischievously construed. (Chapter 43)

As well as the correct answer of imitate, accompanying choices might include *reject*, *accumulate*, and *inspire* (all from the target group). Students’ performances determine whether they receive instruction or move to the next level of assessment.

The Text Genome prototype illustrates how curriculum design and research can be conducted efficaciously. But numerous questions remain about the underlying model. For example, does knowledge of one word in a band (e.g., installed) mean that students know other words in the band (e.g., *guarantee*)? Pearson, Hiebert, and Kamil (2007) argued that students should know words with similar frequencies and features (e.g., part of speech and morphological families), if students have had particular levels of reading exposure. A preliminary study provides sufficient support for testing this hypothesis in further depth. From an analysis of 4.25 million responses from students in Grades 1–12 to 480 words, Hiebert and Castaneda (2017) found that a word’s frequency and the age of acquisition at which a word enters into students’ oral vocabularies predicted students’ vocabulary knowledge across all grades.

## Conclusions

In conclusion, our models of texts are few and have not been well articulated; however, as a community, we have a great deal of collective knowledge about texts. For struggling readers to have access to appropriate texts depends on the efforts of literacy researchers in integrating and communicating this knowledge to teachers.

This attention to text features and diets is not hypothetical nor simply a good idea. Many literacy scholars have commented on the need to emphasize critical literacies in instructional contexts in the digital and global age. I agree wholeheartedly with the call for a greater emphasis on critical literacies. Applying critical perspectives as listeners of news and opinions is an essential part of the experiences students require to become thoughtful citizens. But, at the same time, reflective and critical reading also requires students to read texts on their own without continually relying on teachers, audio readers, or peers to read texts for them.

Getting students onto the page with alacrity and proficiency is necessary for increasing educational equity and maintaining democratic ideals. It is not sufficient by any means, but it is necessary, if our literacy efforts are to support democratic ideals. I believe that ensuring teachers have the knowledge and will to support such proficiency is our primary obligation and contribution to a democratic society. One of our critical roles as reading researchers and teacher educators is to ensure that our teachers are knowledgeable and motivated to provide appropriate texts for all students, but particularly those struggling readers who depend on public schools to become highly literate.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## References

- Achieve3000. (2017). *Reach all learners in core English language arts classes*. Retrieved from <http://www.achieve3000.com/learning-solutions/>
- Adams, M. J., Bereiter, C., McKeough, A., Case, R., Roit, M., Hirschberg, J., . . . Treadway, G. H., Jr. (2000). *Open court reading*. Columbus, OH: SRA/McGraw-Hill.
- Alvermann, D. E., & Commeyras, M. (1994). Gender, text, and discussion: Expanding the possibilities. In R. Garner & P. A. Alexander (Eds.), *Beliefs about text and instruction with text* (pp. 183–199). Hillsdale, NJ: Lawrence Erlbaum Associates.
- 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*. Champaign, IL: The Center for the Study of Reading, National Institute of Education, National Academy of Education.

- Arya, D. J., Hiebert, E. H., & Pearson, P. D. (2011). The effects of syntactic and lexical complexity on the comprehension of elementary science texts. *International Electronic Journal of Elementary Education*, 4, 107–125.
- Austen, J. (1813/2008). *Pride and prejudice* (Project Gutenberg EBook). Retrieved from <http://www.gutenberg.org/files/1342/1342-0.txt>
- Betts, E. A. (1946). *Foundations of reading instruction, with emphasis on differentiated guidance*. New York, NY: American Book Company.
- Bielinski, J., Daniel, M., & Hiebert, E. H. (2015, February). *Patterns of silent reading fluency and accuracy: What they mean for instruction and intervention*. Symposium presented at the annual meeting of the National Association of School Psychologists, Orlando, FL. Retrieved from [https://www.academia.edu/11191138/Patterns\\_of\\_Silent\\_Reading\\_Fluency\\_and\\_Accuracy\\_What\\_They\\_Mean\\_for\\_Instruction\\_and\\_Intervention](https://www.academia.edu/11191138/Patterns_of_Silent_Reading_Fluency_and_Accuracy_What_They_Mean_for_Instruction_and_Intervention)
- Brenner, D., Hiebert, E. H., & Tompkins, R. (2009). How much and what are third graders reading? In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 118–140). New York, NY: Guilford.
- Bryson, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46, 904–911.
- California English/Language Arts Committee. (1987). *English-language arts framework for California public schools: Kindergarten through grade twelve*. Sacramento, CA: Department of Education.
- California State Board of Education. (2006). *Reading/language arts framework for California public schools: Kindergarten through grade twelve*. Sacramento, CA: Author. Retrieved from <http://edpolicyinca.org/sites/default/files/PUBLICATIONS/CDP/CAELAFrameworks.pdf>
- California State Board of Education. (2015). *2015 English language arts/English language development instructional materials adoption (K-8)*. Sacramento, CA: Author. Retrieved from [www.cde.ca.gov/ci/rl/im](http://www.cde.ca.gov/ci/rl/im)
- Carlisle, J. F., Hiebert, E. H., & Kearns, D. (2015, July). *Characteristics of complex words in early elementary texts*. Paper presented at the annual meeting of the Society of Scientific Study of Reading, The Big Island, Hawaii. Retrieved from [https://www.academia.edu/29968605/Characteristics\\_of\\_Complex\\_Words\\_in\\_Early\\_Elementary\\_Texts](https://www.academia.edu/29968605/Characteristics_of_Complex_Words_in_Early_Elementary_Texts)
- Clay, M. M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann Educational Books.
- Coleman, D., & Pimentel, S. (2012). *Revised publishers' criteria for the Common Core State Standards in English language arts and literacy, grades 3–12*. Retrieved from the Common Core Standards Initiative at [www.corestandards.org/assets/Publishers\\_Criteria\\_for\\_3-12.pdf](http://www.corestandards.org/assets/Publishers_Criteria_for_3-12.pdf)
- Cooper, J. D., Pikulski, J. J., Chard, D. J., Garcia, G., Goldenberg, C., Hunter, P., . . . Vogt, M. (2003). *Houghton Mifflin reading: The Nation's Choice California Phonics library* (Grade 1). Boston, MA: Houghton Mifflin.
- Daniel, M. H., Hiebert, E. H., & Martin, L. A. (2016, July). *Assessing silent reading fluency: The problem of poor comprehension*. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, Porto, Portugal. Retrieved from [https://www.academia.edu/29968580/Assessing\\_Silent\\_Reading\\_Fluency\\_The\\_problem\\_of\\_poor\\_comprehension](https://www.academia.edu/29968580/Assessing_Silent_Reading_Fluency_The_problem_of_poor_comprehension)

- Davis, F. B. (1942). Two new measures of reading ability. *Journal of Educational Psychology*, 33, 365–372.
- Donahue, P. L., Voelkel, K. E., Campbell, J. R., & Mazza, J. (1999). *NAEP 1998 reading: Report card for the nation*. Washington, DC: U.S. Department of Education.
- Englemann, S., & Bruner, E. C. (1995). *Reading mastery*. Columbus, OH: SRA/McGraw-Hill.
- Fisher, C. W., Berliner, D. C., Filby, N. N., Marliave, R., Cahen, L. S., & Dishaw, M. M. (1980). Teaching behaviors, academic learning time, and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), *Time to learn* (pp. 7–32). Washington, DC: U.S. Department of Education.
- Foertsch, M. A. (1992). *Reading in and out of school*. Darby, PA: Diane.
- Gambrell, L. B., Wilson, R. M., & Gantt, W. N. (1981). Classroom observations of task-attending behaviors of good and poor readers. *Journal of Educational Research*, 74, 400–404.
- Goodman, K. S., & Goodman, Y. M. (1979). Learning to read is natural. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading* (Vol. 1, pp. 137–154). Hillsdale, NJ: Lawrence Erlbaum.
- Guthrie, J. T., Schafer, W. D., & Huang, C. W. (2001). Benefits of opportunity to read and balanced instruction on the NAEP. *Journal of Educational Research*, 94, 145–162.
- Hiebert, E. H. (1978). Preschool children's understanding of written language. *Child Development*, 49, 1231–1248.
- Hiebert, E. H. (1981). Developmental patterns and interrelationships of preschool children's print awareness. *Reading Research Quarterly*, 16, 236–260.
- Hiebert, E. H. (2003). *Quick reads*. Parsippany, NJ: Pearson Learning Group.
- Hiebert, E. H. (2015). Changing readers, changing texts: Beginning reading texts from 1960 to 2010. *Journal of Education*, 195, 1–14.
- Hiebert, E. H. (2016). English learners and instructional texts: Evidence, current practice, and a call to action. In C. P. Proctor, A. Boardman, & E. H. Hiebert (Eds.), *Teaching emergent bilingual students: Flexible approaches in an era of new standards* (pp. 63–81). New York, NY: Guilford Press.
- Hiebert, E. H. (2017). *An analysis of morphological families within word zones six and seven*. Santa Cruz, CA: TextProject.
- Hiebert, E. H., Brown, Z. A., Taitague, C., Fisher, C. W., & Adler, M. A. (2003). Texts and English language learners: Scaffolding entrée to reading. In F. Boyd, C. Brock, & M. Rozendal (Eds.), *Multicultural and multilingual literacy and language practices* (pp. 50–73). New York, NY: Guilford.
- Hiebert, E. H., & Castaneda, R. (April, 2017). *An analysis of the features of words that influence vocabulary difficulty*. Paper presented at the annual meeting of the American Educational Research Association, San Antonio, TX.
- Hiebert, E. H., & Fisher, C. W. (2016). *A comparison of the effects of two phonetically regular text types on young English learners' literacy* (Reading Research Report 16.01). Santa Cruz, CA: TextProject.
- Hiebert, E. H., Goodwin, A. P., & Cervetti, G. (2017). Core vocabulary: Its morphological content and presence in exemplar texts. *Reading Research Quarterly*. doi:10.1002/rq.183



- Kuhn, M. R., & Schwanenflugel, P. J. (2009). Time, engagement, and support: Lessons from a 4-year fluency intervention. In E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 141–160). New York, NY: Guilford.
- Lewis, M., & Samuels, S. J. (2005). *Read more—read better? A meta-analysis of the literature on the relationship between exposure to reading and reading achievement*. Minneapolis: University of Minnesota.
- Masonheimer, P. E., Drum, P. A., & Ehri, L. C. (1984). Does environmental print identification lead children into word reading? *Journal of Literacy Research, 16*, 257–271.
- Mesmer, H. A., Cunningham, J. W., & Hiebert, E. H. (2012). Toward a theoretical model of primary-grade text complexity: Learning from the past, anticipating the future. *Reading Research Quarterly, 47*, 235–258.
- Murray, M., Munger, K. A., & Hiebert, E. H. (2014). An analysis of two reading intervention programs: How do the words, texts, and programs compare? *Elementary School Journal, 114*, 479–500.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*, 304–330. doi:10.2307/747823
- Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language, 27*, 136–145.
- National Center for Education Statistics. (2015). *The Nation's report card: 2015 mathematics and reading assessments* (NCES 2015-136). Washington, DC: Institute of Education Sciences, U.S. Department of Education.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects (including Appendices A & B)*. Washington, DC: Authors. Retrieved from [www.corestandards.org/assets/CCSSI\\_ELA%20Standards.pdf](http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf)
- Newsela. (2017). *Newsela learning and support*. Retrieved from <https://support.newsela.com/hc/en-us/articles/204041325-Why-should-I-sign-up-my-students-for-Newsela>
- Pearson, P. D. (1974). The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. *Reading Research Quarterly, 10*, 155–192.
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to know. *Reading Research Quarterly, 42*, 282–296.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357–383.
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Corporation.
- Rasinski, T., Samuels, S. J., Hiebert, E., Petscher, Y., & Feller, K. (2011). The relationship between a silent reading fluency instructional protocol on students' reading comprehension and achievement in an urban school setting. *Reading Psychology, 34*, 76–93.
- Spichtig, A. N., Hiebert, E. H., Vorstius, C., Pascoe, J. P., Pearson, P. D., & Radach, R. (2016). The decline of comprehension-based silent reading efficiency in the United States: A comparison of current data with performance in 1960. *Reading Research Quarterly, 51*, 239–259.

- Stein, M., Johnson, B., & Gutlohn, L. (1999). Analyzing beginning reading programs: The relationship between decoding instruction and text. *Journal for Special Educators, 20*, 275–287.
- Stenner, A. J., Burdick, H., Sanford, E. E., & Burdick, D. S. (2007). *The Lexile framework for reading* (Technical Report). Durham, NC: MetaMetrics.
- Swanson, E., Wanzek, J., McCulley, L., Stillman-Spisak, S., Vaughn, S., Simmons, D., . . . Hairrell, A. (2016). Literacy and text reading in middle and high school social studies and English language arts classrooms. *Reading & Writing Quarterly, 32*, 199–222.
- Texas Educational Agency. (1997). *Proclamation of the State Board of Education advertising for bids on textbooks*. Austin, TX: Author.
- Text Genome. (2017). *Our pedagogy and technology*. Retrieved from <http://www.textgenome.org>
- Trainin, G., Hayden, H. E., Wilson, K., & Erickson, J. (2016). Examining the impact of Quick-Reads technology and print formats on fluency, comprehension, and vocabulary development of elementary students. *Journal of Research on Educational Effectiveness, 9*, 93–116.
- Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science Associates.

### Author Biography

**Elfrieda H. Hiebert** is the president and CEO of TextProject, a non for profit that provides open-access resources for teachers. She is also a research associate at the University of California, Santa Cruz, CA.