Text Complexity and English Learners— Building Vocabulary

Elfrieda H. Hiebert
TextProject & University of California, Santa Cruz

*The Common Core State Standards (CCSS) are changing curriculum planning and classroom instruction in many ways. One significant change involves the difficulty levels of text. In the past, standards documents have referred to proficiency with grade-level texts. However, **grade level** was not defined. The CCSS represents a departure from this practice. Standard 10 of the CCSS specifically calls for increasing levels of text complexity across the grades to ensure students' proficiency with the texts of college and career. This standard affects all students, but it represents a special challenge to English Learners. Many educators ask what increases in text complexity mean for English Learners, many of whom struggle with their current texts.*

For more information about Text Matters, visit **www.textproject.org/text-matters**

What, then, is text complexity, and how can English Learners achieve success with this standard? First, an understanding of what makes texts complex is in order. Archaic language, lengthy sentences, new topics, unusual writing styles, unique text structures—these features and many others affect the complexity, and hence the comprehensibility, of text. However, the foremost challenge for English Learners is a text's vocabulary (Pasquarella, Gottardo, & Grant, 2012). The syntax of a new language does present obstacles to comprehension; however, vocabulary is the most significant hurdle when learning a second language.

Two issues of *Text Matters* are devoted to the topic of complex text and English learners. This first issue describes support for English Learners in developing strategies and knowledge about the vocabulary of complex texts. The second issue of *Text Matters* presents guidelines for selecting appropriate texts that move English Learners up the staircase of text complexity. Both topics depend on teachers' understanding of how English words work.

The Distribution of English Words

A small group of words—4,000 simple word families (e.g., *help*, *helps*, *helping*, *helped*, *helper*)—accounts for about 90% of the words in most texts. This vocabulary forms the *core* of any text, even complex ones. In the exemplars of complex texts listed in Appendix B of the CCSS, the core vocabulary accounts for 93% of the Grades 2–3 exemplars, 92% of those in Grades 4–5, 90% of those in Grades 6–8 and in Grades 9–10, and 88% of those in Grade 11–College and Career Ready.

This *Text Matters* focuses on the *extended* vocabulary—the 300,000 or more words that account for approximately 10% of the words in texts. [Readers who are interested in learning more about the core vocabulary can read Hiebert (2012, 2013).] Unlike the words in the core vocabulary, many words in the extended vocabulary appear less than once per every million words of text. Consequently, they are often described as *rare*. When rare words do appear in text, they often are essential to the content and quality of texts. In narrative texts, words in the extended vocabulary often describe the traits of characters and the nuances of plots. In informational texts, they convey specialized terms in chemistry, entomology, and many other topics.

The percentage of rare words in a text can vary considerably, even among complex texts (see Table 1). An increase of only one or two percent in rare vocabulary can make texts considerably more complex. When viewed from the vantage point of a thousand-word text, a rate of 8% means the text has about 80 rare words, while a rate of 10% means that a text has about 100 rare words. An additional two rare words in every 100 words can increase the challenge of a text.

The next *Text Matters* gives guidelines on appropriate rates of rare words for English Learners at different developmental levels. In this issue, the focus is on the extended vocabulary and ways teachers can support English Learners in understanding this vocabulary.

Systematic Vocabulary Instruction

To help English Learners build strong vocabularies, teachers need

Table 1

Vocabulary Profiles of CCSS Selected Exemplars From Grades 6–8

Content Area	Text	Core Vocabulary (%)	Extended Vocabulary (%)
Technology/ Mathematics	*Geeks*	83	17
	Math Trek	88	12
Social Studies	*A Night to Remember*	92.5	7.5
	Narrative of the Life of Frederick Douglass	93	7
Literature	*Adventures of Tom Sawyer*	90	10
	Dark is Rising	95	5

Table 2

Categories Within Extended (Rare) Vocabulary of Two Exemplar Texts

	A Night to Remember	*Geeks*
Total Rare Words	180 (11 words per 100)	157 (17 words per 100)
Single Syllable	26%	48%
Multisyllabic Words: Proper Names	23%	8%
Multisyllabic: Picturable	26%	16%
Multisyllabic: Compound Words	16%	16%
Multisyllabic: Remaining	8% of rare words (1 word per 100 of entire text)	12% of rare words (2 words per 100 of entire text)

to focus on general principles and strategies of word learning. They also need to conduct short lessons and discussions about the vocabularies of specific texts.

This *Text Matters* focuses on the rare vocabulary of literary texts, not that of content-area texts. Content-area standards are explicit about the topics and the concepts underlying those topics (Marzano, 2004). In that concepts are represented by vocabulary, the critical words in a physics unit are clear within standards and curricula (e.g., *magnetic attraction*, *repel*, *polarity*). This vocabulary becomes part and parcel of activities and discussion. Words such as *magnetic attraction*, for example, are used repeatedly as students engage in inquiry with magnets.

Such clarity is not evident in English/Language Arts standards where literary texts are the focus (Marzano, 2004). Literary texts often present rare words that are unique to a particular story. Each text has its own rare words. Thus, students cannot become proficient in the meaning of these words through repetition.

These rare words, however, represent specific elements of stories. An author of a literary text chooses a word intentionally from the extended vocabulary to communicate an action, a social relationship, the feature of a place or event, and the feelings and attitudes of characters. In *Geeks,* for example, Katz (2000) could have used numerous words to describe the condition of the furniture within the apartment of his protagonists. However, by describing the beanbag chair as *moldering,* readers get a clear idea of the condition of the apartment. Just as in mathematics where lessons build understanding for future problem solving, considerable time needs to be spent in developing the linguistic foundation of English vocabulary for the future reading of complex literary texts.

This pattern of single-appearing, rare vocabulary does not only appear in narratives, though. It also appears in magazine articles on topics of science, history, and civics to describe traits, features, interactions, and contexts. This style also extends to full-length texts with a literary stance about technology

and science (e.g., *Geeks*) and history and political science (e.g., *A Night to Remember*). English Learners need to become adept with such vocabulary for a variety of reasons, including the heavy presence of literary texts on assessments.

Instruction of General Principles and Strategies

Analyses of two Grade 6–8 texts from the CCSS's Appendix B exemplar list—*Geeks* and *A Night to Remember* (*NTR*)—demonstrate the two types of vocabulary instruction needed for proficient reading of literary texts: (a) general principles/strategies and (b) lessons with specific texts. Both texts typify the literary texts offered as CCSS exemplars. Both have higher levels of extended vocabulary than other Grade 6–8 exemplars (see Table 1), which is why these texts were selected for illustration in this *Text Matters*. Analyses of these two texts show that, even in these vocabulary-dense texts, most words fall into particular groups—groups that share underlying features.

Table 2 shows the types of words in these two texts. Although there

are numerous monosyllabic words, students typically recognize these less-complex words more readily than they do multisyllabic words. Without instruction on multisyllabic words, however, students can develop dysfunctional word-recognition strategies. This is why, beginning in the late primary grades, multisyllabic words should receive the lion's share of vocabulary instruction.

The multisyllabic words in texts such as *Geeks* or *NTR* are grist for lessons on four types of words in literary texts. These types of vocabulary contribute to the meaning of the text, but they are not necessarily complex in content. Students who are not prepared to deal with this vocabulary will find literary texts difficult. Following are the four types of words in literary texts and strategies for helping students understand them.

Proper names. Stories and magazine articles are typically replete with proper names, many of which are difficult to pronounce (e.g., *Boise* in *Geeks*). Students need to learn that capitalized words within sentences often are proper names and that accurate pronunciation of these words is not a priority.

Picturable words. Research has shown that concrete words that can be represented in pictures are learned more easily than abstract words (Strain, Patterson, & Seidenberg, 2002). Using pictures to create a context for a new concrete word (e.g., *smelting*) or to support English Learners in relating a known concept with the English label (e.g., *necklace*) are especially effective ways to support the vocabulary development of English Learners.

Pictures that illustrate these words will support students' recognition much more effectively than extended discussions. Sources for images (with certain copyright restrictions but free and downloadable) include [Flickr](#) and [Wikimedia Commons](#).

Compound words. Compounding of two root words is a primary way in which many new words are added to English. Some compound words in *Geeks* illustrate how new words are generated, especially with inventions in fields such as digital technology: *software*, *motherboard*, *network*, *upgrade*, and *playlist*.

Compound words typically have a connection to the root words within them, but they often have idiomatic meanings. As a consequence, compound words with the same headword (e.g., *up* in *upgrade*, *uproar*, *uptown*, *uptight*, *upkeep*) cannot be taught in the same way as words from the same morphological family (e.g., *suspicion*, *suspiciously*, *unsuspicious*). The upside of compound words is that most headwords (and often the second word as well) belong to the core vocabulary. Once students learn to use the headword to predict meanings of compound words, their word-recognition vocabularies expand considerably.

Morphological families. Becoming facile with inflected endings and affixes is also a critical part of preparing students for reading complex text. Among the 300,000 words of the extended vocabulary, most belong to morphological word families with an average of approximately four members. Lessons on the relationships among members of a morphological family are essential to developing the expecta-

tion that words are connected to one another structurally (e.g., *formality*, *formal*, *formalize*, *informal*, *informally*). Such lessons provide students with opportunities to use words in meaningful ways, not simply to memorize the meanings of suffixes and prefixes.

Vocabulary Related to Specific Texts

Even when students have been taught strategies to recognize a high percentage of the words in complex texts, a group of words remains to be learned (see Table 2). These words need individual study. Lessons on the vocabulary of specific texts have two dimensions: (a) an overview of the task and (b) instruction on specific words.

An overview of the task. The CCSS refers to the scaffolding of complex texts for challenged readers, but the forms of this scaffolding are not described. Frequently, scaffolding has been interpreted as reading a text for students or leading students through a guided reading of the text. However, students need to take responsibility for reading, including initial reads of texts, if they are to improve their comprehension. But teachers also need to give students a realistic view of the challenges of texts by identifying core (in black) and extended (in gray) vocabulary in samples of text, such as the following:

> He wasn't just a kid at a computer, but something more, something new, an **impresario** and an Information Age CEO, **transfixed** and concentrated, almost part of the machinery, conducting the **digi-**

tal **ensemble** that controlled his life. (Katz, 2000, p. 19)

Four days before, she had playfully **teased** him for putting a **life belt** in her **stateroom**, if the ship was meant to be so **unsinkable**. At the time he had laughed and assured her it was a **formality** … she would never have to wear it. (Lord, 1955, p. 22)

One of several text analysis schemes can be used to distinguish between core and extended vocabulary (e.g., Laurence Anthony's [AntWordProfiler](http://) software).

Next, teachers should address the words for which previously taught strategies should be applicable: proper names, picturable words, compound words, and morphological families. Teachers can't review all of the words in these categories, but they can give students examples of words of different types within the text.

Instruction of specific words in extended clusters. The remaining words become the grist for instruction. In *NTR*, these remaining words account for approximately one rare multisyllabic word per 100 words of text, including words such as *adamant*, *formality*, *solicitous*, and *suspiciously*. In *Geeks*, the number of rare multisyllabic words per 100 is two, including words such as *alumnus*, *ensemble*, *impresario*, *transfixed*, and *contemplated*.

A handful of the most critical words—those that are fundamental to the meaning of the text—can be introduced before students read the text. For example, the unsinkable reputation of the Titanic led to particular stances on the part of passengers (e.g., *bewildered*, *protested*, *suspicious*) as well as on the part of the crew (e.g., *solicitous*, *reassuring*, *adamant*).

Short lessons on critical vocabulary should also follow the initial reading of a text. In addition, discussions of critical vocabulary are an essential part of the close reading of text. For example, in the segment from *Geeks* above, the author's use of the phrase "conducting the digital ensemble" merits discussion, as do words such as *dispensable*. Would the use of *disposable*, *superfluous*, *unnecessary*, or *useless* have served the same function as *dispensable*?

A final post-reading vocabulary activity asks students to record critical words and their semantic connections (e.g., the above-mentioned synonyms of *dispensable*) and morphological derivatives. For English Learners, such records are important as references for writing and as records of what they have learned.

Summary

A rich vocabulary and strategies that permit students to read texts with new words are essential to comprehending complex text. For English Learners, a rich vocabulary and strong strategies result from intentional instruction on the part of their teachers. This intentional instruction is not a one-shot occurrence but rather a sustained effort that focuses on categories of words (e.g., compound words, picturable words) and also on words within specific texts, especially words which are part of extended networks of words.

References

Common Core State Standards Initiative (2010). *Common Core State Standards for English language arts & literacy in history/social studies, science, and technical subjects.* Washington, DC: CCSSO & National Governors Association.

Hiebert, E.H. (2013). Core vocabulary and the challenge of complex text. In S. Neuman & L. Gambrell (Eds.), *Reading Research in the Age of the Common Core State Standards.* Newark, DE: IRA. [Pre-publication version of the chapter is available at http://textproject.org/library/articles/core-vocabulary-and-the-challenge-of-complex-text/]

Hiebert, E.H. (2012). Core vocabulary: The foundation for successful reading of complex text, *Text Matters* 1.2. Santa Cruz, CA: TextProject. Retrieved from http://textproject.org/professional-development/text-matters/core-vocabulary/

Katz, J. (2000). *Geeks: How Two Lost Boys Rode the Internet out of Idaho.* New York, NY: Broadway Books.

Lord, W. (1955). *A night to remember.* New York, NY: Bantam Books.

Marzano, R.J. (2004). *Building background knowledge for academic achievement: Research on what works in schools.* Alexandria, VA: ASCD.

Strain, E., Patterson, K., & Seidenberg, M.S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 207–215.