This is a version of an article that has been accepted for publication in *The Reading Teacher* (on November 1, 2019). Please note that this is neither the copy-edited nor published version of the article.

The Core Vocabulary:

The Foundation of Proficient Comprehension

Elfrieda H. Hiebert

TextProject

Abstract

A group of words, labeled the core vocabulary, can be expected to be prominent across all texts. Scholarship made possible by digital databases of words and new analytic systems has shown that approximately 2,500 morphological families account for the vast majority of words in texts—an average of 91.5% of all words in the CCSS exemplars from Grade 1 through College-and-Career Ready. The rare vocabulary of texts is so extensive and varied that, without a foundation in the core vocabulary, students are likely to struggle with the meanings of rare words. Words are in the core vocabulary because they represent critical concepts, are part of morphological or word families that share a root word, and are often versatile in their meaning and function. This article describes the semantic, morphological, and multiple-meaning knowledge represented by the 2,500 word families and ways in which teachers can foster these proficiencies.

The Core Vocabulary: The Foundation of Proficient Comprehension

Teaser Text: A core vocabulary of about 2,500 word families accounts for the majority of words in texts. This article describes three features of the word and world knowledge that underlie this core vocabulary.

Pause and Ponder

- To what degree does your school's ELA program emphasize selecting texts that promote knowledge of words and the world rather than assigned text complexity levels?
- How aware are your students that the number of rare words in new texts is small relative to the words they encounter frequently across texts?
- How often do students in your school talk about the richness of semantic and morphological families and multiple meanings associated with words?

The words our students know make a big difference in how well they comprehend texts. Consider this excerpt from a third-grade favorite, *The One and Only Ivan*: "Humans reek," Bob replies. "They just don't notice because they have incompetent noses." (Applegate, 1997, p. 201). If third graders cannot quickly recognize the meanings of moderately frequent words such as *replies* and *notice* and struggle with less frequent words such as *reek* and *incompetent*, they will have a hard time understanding Bob's critique of how bad humans are at smelling—and how bad they smell.

Words represent knowledge and knowledge about a text's topic strongly predicts students' comprehension of a text (Verhoeven, van Leeuwe, & Vermeer, 2011). When students' vocabularies are not large on entering school, their prospects for reading success depend on the richness of school experiences. As educators, however, we face a formidable task in providing rich vocabulary instruction. The challenge may seem to come from the vastly differing vocabularies of students on school entry, but the real obstacle is the tidal wave of words in the English dictionary. A comprehensive analysis of school texts identified over 150,000 different words in a sample of typical books from grades 1 through college (Zeno, Ivens, Millard, & Duvvuri, 1995). And English has even more words—as many as 300,000 more (Nagy &

Anderson, 1984; Stevenson, 2010). But even these numbers do not tell the whole story, since many unique words have multiple meanings. One example is *notice*, which has a meaning as a verb—as used by Bob in the previous excerpt—as well as two meanings as a noun (attention; notification).

The typical approach of teaching 6 to 8 vocabulary words from a specific story has not proven especially successful in extending students' vocabularies (see, e.g., Apthorp, 2006). Looking at the typical words selected for a week of instruction for third graders in a core reading program—*batted, buzzing, clattered, fetch, rattled, rough, slick, sniff,* and *thumped* (Baumann et al., 2014)—gives a clue as to why this approach has not been highly effective. All but one word (*rough)* are predicted to appear rarely in future texts (Zeno et al., 1995) and in conversations (Kuperman, Stadhagen-Gonzalez, & Brysbaert, 2012). Further, all words have easily recognizable synonyms (e.g., *fetch/get, clattered/crash*). After a week of instruction, students will not have added significantly to their vocabularies.

Which words should be taught, then? The same global-digital world that imposes higher demands for literacy offers evidence-based answers. Digital scanning of texts enables us to study millions of words from texts. Computers quickly provide information on word features—such as membership in morphological families and the age at which words are heard and used in students' oral language. Findings from these digital efforts offer insights about how to teach vocabulary strategically. This paper summarizes these insights and applies this evidence-based information to the design of instructional experiences in classrooms.

The Core Vocabulary: Why and How

The basis for selecting the 6-8 words for a week of vocabulary instruction in many reading programs has little documentation (see, e.g., Stallman et al., 1989). But over the past two

decades, several approaches for selecting words have been proposed as ways to bolster the rigor of vocabulary instruction.

Approaches for Selecting Vocabulary to Teach

The three-tier model (Beck, McKeown, & Kucan, 2013) has become the favored choice for vocabulary selection in pedagogy textbooks, core reading programs, and even the National Assessment of Educational Progress (National Center for Education Statistics, 2017). The instructional emphasis is on the middle tier—general academic words or synonyms of common words—rather than the everyday words of tier one or the technical words of tier three. Assigning words to tiers can be hard, however, even for experts. Of 13 words identified by writers of the Common Core State Standards (CCSS; NGA CBP & CCSSO, 2010) as exemplifying tier-two words for fourth through fifth graders, all are predicted to be in students' oral vocabularies by age eight (Kuperman et al., 2012)—words such as *early* and *pours*. At the same time, four of the five words that were identified as tier three are multiple-meaning words with a strong likelihood of appearing in both literary and informational texts (e.g., *crust, molten*). Evaluations of interventions with tier-two words have not shown consistent effects on either standardized comprehension (Elleman, Lindo, Morphy, & Compton, 2009; Wright & Cervetti, 2017) or vocabulary measures (Cervetti, Fitzgerald, Hiebert, & Hebert, 2019).

In Biemiller's (2010) "words worth teaching" approach, words come from an historical database (Dale & O'Rourke, 1981) that has recently been enhanced (Biemiller & Slonim, 2001). The methods of the original study remain vague, resulting in some unusual recommendations; for example, words considered worth teaching in the upper elementary grades come from both ends of the frequency spectrum—from *most* (very much) and *must* (have to) to *mulligan* (a

stew) and *muss* (a mess). While *mulligan* and *muss* are likely to appear fewer than once in every two million words of text, the words *most* and *must* are on most high-frequency word lists.

Nagy and Hiebert (2011) offer a third approach: a word selection framework based on research on word acquisition and distributions of words in texts. They recommend selecting vocabulary on the basis of answers to questions such as: (1) How often can a word be expected to appear in texts at specific grade levels? (2) How many close morphological relatives does a target word have? and (3) How likely is the word already known by students at a grade level, and, if unfamiliar, can the word be easily explained with already known concepts or experiences?

From this perspective, a word such as *incompetent* from *The One and Only Ivan* would be the focus of instruction, rather than *reek; incompetent* and members of its morphological family (e.g., words with the same root word, including *competence* and *competently*) are likely to appear over 125 times more frequently than *reek* and its family members in school texts (Zeno et al., 1995). *Reek* also can be explained easily with the synonym *smell*, a word that third graders typically know. In contrast, the synonyms for *incompetent* are themselves fairly complex words: *inept, inefficient,* and *inadequate*.

Identifying a Core Vocabulary

A group of researchers (Hiebert, Goodwin, & Cervetti, 2018) set out to determine if they could use the Nagy and Hiebert (2011) criteria to identify a "core vocabulary," a set of words that account for a substantial percentage of the words in school texts. Students who are automatic with the meanings of the majority of words in texts, these researchers reasoned, will have a strong foundation for text comprehension. The researchers targeted 90% of the words in text as the goal for a core vocabulary because scholars such as Clay (1991) have identified this level as sufficient for comprehension.

The core vocabulary was identified through three steps: (a) identifying words with moderate to high frequency in school texts, (b) sorting the identified words into morphological families, and (c) verifying that the core vocabulary makes up 90% or more of the words in school texts.

Identifying words with moderate to high frequency in school texts. Although several databases of word frequency (e.g., Davies, 2009) are based on as many as a half billion or more words of texts, an interest in the words of school texts led to the use of the *Educator's Frequency Word Book* (EWFG; Zeno et al., 1995). This list summarizes words in a sample of school texts across grades (1 through college) and content areas (literature and content areas). It also includes frequencies of words at different grade levels, which is critical in selecting words for instruction, as not all words appear with similar frequency across the grades. Take the words *government* and *big,* which are both among the 300 most frequent words. *Government* has few appearances in the primary grades, while *big* is frequent in the primary grades but not in middle to high school.

The distribution of all unique or different words from the *EWFG* across five categories of frequency appears in Table 1. To be included in the core vocabulary, words needed to appear with sufficient frequency in school texts to merit instructional attention. The number of repetitions associated with students' knowledge of a word's meaning has been reported to be about 10 (McKeown, Beck, Omanson, & Pople, 1985; Reitsma, 1988). Almost 6,000 words on the list met the criterion of 10 or more appearances per million words. When proper names were excluded, the group consisted of 5,586 words.

Sorting words into morphological families. The next step was to sort the words into morphological families—that is, groups of words sharing a root word (e.g., *help: helping, helpless, unhelpful*). Many beginning readers make the connection between words with the same

root word (e.g., *day, days*) (Anglin, 1983). Further, instruction can be successful for students who do not make these connections (Goodwin & Ahn, 2010), especially when morphological connections go beyond inflected endings to compound words (e.g., *birthday, daytime*) or affixes (e.g., *daily*). The 5,586 words were sorted into 2,451 (rounded up to 2,500) morphological families. With family members from the relatively rare group (see Table 1), the 2,500 lead words represent 11,298 words in all. That is, each lead word in the core vocabulary represents a family of approximately five words.

Verifying the presence of core vocabulary in texts. The prominence of the core vocabulary in texts was verified by an analysis of all 200 texts identified in Appendix B of the CCSS (NGA CBP & CCSSO, 2010) as exemplars of complex texts. The 2,500 word families averaged 91.5% of the total words in texts across six grade bands spanning from kindergarten through College-and-Career Ready (CCR). The percentage was highest in K-1 texts (97%) and decreased to 89% for Grade 11-CCR texts. A majority of the word families appear by the end of primary-level texts (1,787), another 511 become prominent in middle-grade through middle-school texts, and the remaining 153 families appear in high-school texts.

What Is Involved in Proficiency with the Core Vocabulary?

The core vocabulary is much more than a list of 2,500 word families to be memorized. Adeptness in recognizing the meanings represented in the core vocabulary includes: (a) developing connections across semantically related ideas, (b) generalizing root-word meanings across family members, and (c) recognizing that words can take on multiple meanings. The semantic, morphological, and multiple-meaning features of five words (all from the moderately frequent, concrete group in Table 1) are summarized in Table 2 and are used to illustrate the three types of connections.

Semantic connections

The meanings of words are connected to one another in semantic networks. As illustrated in Table 2, the meaning of *sick* is associated with words such as *disease* and *germs* and specific ailments (*cold*, *fever*). A framework that has proven useful in establishing concepts among words is the supercluster framework of Marzano and Marzano (1988), who set out to establish semantic relationships among 7,230 words from core reading programs. Words were organized around three hierarchical relationships: (a) 61 superclusters where words share a broad meaning (e.g., occupations, emotions, machines); (b) 430 clusters where words share a specific meaning (e.g., outdoor professions, one of 30 clusters within occupations); and (c) 1,500 miniclusters in which meanings are synonymous (e.g., hunter, one of 5 miniclusters within outdoor professions).

All 61 superclusters (Marzano & Marzano,1988) are present in the 2,500 word families. Some superclusters are heavily populated, such as animals and emotions. Other superclusters, such as chemicals and electricity/particles of matter, have few members, as is the case in the original classification. Even so, a broad and encompassing range of concepts is represented within the 2,500 word families. The size and diversity of semantic networks for the five focus words is evident in Table 2.

Information on when semantic relatives are predicted to appear in students' oral vocabularies is included in Table 2. The levels (i.e., Early, Middle, Late) at which words appear in students' oral language underscores the role that already known words can have in building understanding of new words. For example, students can bring their knowledge of a number of related concepts (e.g., *fair, wrong, right, crime)* when learning the word *trial* in the middle grades.

Morphological connections

Morphology has to do with shared meanings as a function of word parts. Morphemes, the smallest individual meaningful elements in a language, take several forms. The fundamental unit is the root word, which functions on its own (e.g., *invite*). The simplest form of bound morphemes, which are added to root words and do not function on their own, are inflected endings (e.g., *invited, inviting, invites*). Other examples of simple bound morphemes are possessives (e.g., *trial's*) and comparatives (e.g., *sicker, sickest*). In the next group of prefixes and suffixes, part of speech (e.g., *invitation*) and meaning (*disinvite*) can shift. Another morphological construction is the compound, where two (and sometimes more) root words are joined to form a new word. For example, *sick* forms compounds, both as a head word (*sickbed*) and as a base word (*carsick*).

The types of word families for the focus words in Table 2 vary from one another. These variations in morphological types are a function of linguistic origins. The first source of English—Anglo-Saxon, a Germanic language—is illustrated by the word *sick*. Like many words in this layer of English, the monosyllabic *sick* is in numerous compound words, many of which have idiosyncratic meanings. A *sickbed* is not a bed that is sick but a place where a sick person lies. In that compound words are extensive in English, applying the meanings of the words in compounds is fundamental to proficient reading of the core vocabulary.

The words *invite, current*, and *trial* come from the French layer of English, which originated in Latin. Words in this layer are frequently multisyllabic synonyms of Anglo-Saxon words, as illustrated by *current* (French) and *now* (Anglo-Saxon). Affixes are the primary way in which new meanings are generated in this layer of English.

In Table 2, *technology* represents a word with Greek roots, a third contributor to English. Similar to Anglo-Saxon words, a primary way of generating new words in Greek-origin words is

to create compounds. However, in Greek-origin words, the meanings are fairly straightforward. The meaning of *techno* (technical) stays the same in *technophobe* (a person who dislikes technology), while *logy/ology* (the study or science of a field) has the same meaning in *hydrology* (the study of water).

Knowledge of Multiple Meanings

Polysemy—from the Greek words *poly* (many) and *sema* (sign)—refers to the multiple meanings of words. Most of the lead words in the 2,500 word families have multiple meanings, as illustrated by the five focus words in Table 2. For some words, the meanings for a single word are not vastly different as in the case for the two meanings of *sick*. But some words, as illustrated by the word *current*, meanings can be quite different. The everyday meaning of a word such as *current*, which is almost always learned first, can interfere with students' learning of technical meanings (e.g., air, water, or electrical currents) (Cervetti, Hiebert, Pearson, & McClung, 2015). Awareness of polysemy is useful from the early stages of reading, since many polysemous words (e.g., *can, bill, sink*) often appear in beginning reading programs.

How Can Teachers Support Students' Facility with the Core Vocabulary?

The description of the semantic, morphological, and multiple-meaning connections within the core vocabulary shows why these words need to be taught in relation to one another and not as individual words. Before describing instructional ways to support these connections, two points merit attention.

First, not all words need to be taught. Table 1 shows examples of concrete words. In all, 32% of the core vocabulary words are highly concrete. Concrete words are typically learned easily, especially when words are nouns rather than verbs (McDonough, Song, Hirsch-Pasek,

Golinkoff, & Lannon, 2011), which is the case with most of the concrete words in the core vocabulary. Pictures can go a long way in supporting students' knowledge of concrete words.

Further, a significant group of words (about a quarter of the lead words) are likely to be in children's oral vocabularies when they enter school. Table 2 shows some of these words, which appear in the row for the Early Period. Young children may not use the words *invite* and *current* but they are likely to use words such as *ask* and *now.* These known words can be the anchor for acquiring new words. For example, a known lead word such as *sick* can support knowledge of potentially unknown words in the semantic family (e.g., *temperature, fever*) or morphological family (e.g., *sickness, sickly*).

Second, a focus on the core vocabulary does not mean that rare words are dismissed or disregarded. The core vocabulary can be the means for building a robust vocabulary of the words that occur less frequently in text. Returning to the excerpt from *The One and Only Ivan*, a word in the core vocabulary such as *smell* can be used to build knowledge of *reek* and other related words not in the text, including ones that are rare such as *olfactory* and *fetid*.

These two principles should be kept in mind when thinking about core vocabulary instruction: (a) not all words need to be taught and (b) known words can be the basis for expanding students' knowledge of rare words.

Extensive Reading

Just as with many other proficiencies in life, a person only gets good at reading by reading extensively. But not just any text will ensure facility with the core vocabulary. All texts are likely to have a high percentage of core vocabulary, but to get good at particular words within the core vocabulary requires repetition of those words. Remember that all of the words in the core vocabulary do not happen at once. The words that are new to the core vocabulary at a

grade band represent the growing edge of students' vocabulary learning. Such is the case with many of the words in the moderate frequency category in Table 1, which increase in prominence in middle-grade texts and beyond. Repeated encounters with these words in texts are critical, and such encounters are more likely to occur when instructional texts are organized around topics than simply around text complexity.

To illustrate the nature of opportunities with moderately frequent words when text sets are selected according to different criteria, 2,300 words were analyzed from two sets of texts. All texts were nonfiction and had the same guided reading levels (N through 0). The Text Complexity set came from a leveled program (Fountas & Pinnell, 2008) and covered various topics, including volcanoes, skateboarding, and unusual snakes. The second set of texts—labeled Topic/Text Complexity—came from several leveled text programs, but all were chosen for their content on growing giant fruits and vegetables.

A summary in Table 3 shows that the topically related texts had 50% more words that were repeated three times or more than the text set based on text complexity only. Quantity, however, is only one way in which the moderately frequent words differ in these two text types. Conceptual networks of repeated words also vary as a function of text type. In the texts organized by complexity only, two pairs of words represent semantic clusters. In the texts grouped by complexity and topic, five semantic clusters with two to eight members are evident. Topical sets of text give students the chance to generalize their knowledge of words across texts and, in the process, build a coherent knowledge base—the foundation of proficient comprehension.

Conversations

Conversations have been shown to be a powerful means of supporting students' awareness and knowledge of vocabulary (Cabell, Justice, McGinty, DeCoster, & Forston, 2015). The sophistication of the content of conversations changes over the grades, but two types of conversations support students' vocabulary development, whatever the grade level.

Conversations that highlight intriguing uses of words. One type of conversation foregrounds rich uses of words in high-quality literature. Since figurative language draws heavily on core vocabulary (Glucksberg, McGlone, Grodzinsky, & Amunts, 2001), students can develop a rich knowledge of metaphors and similes early in their school careers. Nancy, a second-grade teacher, has read several books to her students in which the full moon is described figuratively with words from the core vocabulary, examples of which appear in Table 4. Nancy demonstrates to her students how words they already use in reading and speaking can be used figuratively in their writing. Among the descriptions of the full moon given by Nancy's students were the following:

Jorge: giant baseball

Zari: pancake

Jin: marble

Maya: cotton ball with dirt on it

Leon: lightbulb

Bella: grumpy man

Emma: fortune teller's ball

Conversations on the presence of core vocabulary in texts. The second type of conversation addresses the ratio of core to rare words in texts. In particular, awareness of the prominence of core vocabulary in virtually all texts can be useful when students confront the

summative assessments of their district or state. Malcolm, a third-grade teacher, uses a passage illustrative of those on summative assessments to demonstrate to his students that they have the proficiency to successfully read the texts on assessments:

> Some teachers say that children's personal toys can create conflicts on the playground and in the classroom. Some students do not want to share their toys, which can cause **frustration** and **envy** in their classmates. Or, if a toy is misplaced or broken, the owner of the toy might blame other students for the loss or damage. (Hiebert, 2013)

Malcolm asks students to use their knowledge of the core vocabulary, context, and word features to give meaning to the rare words in the text (in boldface in the example above). Through conversations such as this one, Malcolm's students have learned that a small number of rare words will always be present in texts, but that their core vocabulary proficiency enables them to understand the meaning of rare words and to comprehend the questions that follow the passage.

Mini-Lessons

A project in Bernice's fourth-grade classroom illustrates how mini-lessons on the core vocabulary involve students as partners in identifying multiple meanings of core vocabulary words in texts. Multiple meanings of words can be especially confusing for students when words are common in conversations but have technical, specific meanings in content areas (Cervetti et al., 2015). Bernice has posted a set of such words: *channel, cycle, force, power,* and *concentration.* Among the text excerpts that Bernice's students identified in their reading for the word *concentration* are the following.

> *Literature:* I worked in deep **concentration** as did the other kids....(*Soto, 1996)*

> *Social Studies:* The men who wrote the Constitution wanted to guard against the **concentration** of power in any part of the government.... (*History.com, 2018)*
>
> *Science:* We get it [our drinking water] from one of the many freshwater sources that have a lower **concentration** of salt and other dissolved solids than seawater. (*Bix, 1995*)

In mini-lessons centered on the text excerpts, students read one another's contributions and compare the uses of the same word. Mini-lessons end with Bernice asking students to write a summary of what they have learned about words in their notebooks. An example of a student's summary is the following: "Often, the same word can have very different meanings and uses in texts."

Conclusion

Words are central to acquiring, remembering, and using knowledge. The digital revolution has led to new understandings about the words in school texts. One of the insights from this work has been the role of a relatively small part of the English lexicon—a group of 2,500 word families—that accounts for a sizable portion of texts at all levels. A solid foundation in this core vocabulary is built on knowledge of underlying systems and features of words, not simply memorizing the meanings of individual words. Deep knowledge of the words in the core vocabulary comes from guided conversations and mini-lessons where shared and unique features are discussed, as well from as extensive reading. Such experiences are critical for all students if they are to acquire the vocabulary foundations that underlie the literacy proficiency required for the digital-global age.

References

Anglin, J.M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development*, pp. i-186.

Apthorp, H. S. (2006). Effects of a supplemental vocabulary program in third-grade reading/language arts. *The Journal of Educational Research*, *100*(2), 67-79.

Baumann, J., Chard, D.J., Cooks, J., Cooper, J.D., Gersten, R., Lipson, M….(2014). *Journeys: Common Core*. Boston, MA: Houghton Mifflin Harcourt.

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. Guilford Press.

Biemiller, A. (2010). *Words worth teaching: Closing the vocabulary gap*. Columbus, OH: McGraw-Hill SRA.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, *93*(3), 498-520.

Cabell, S. Q., Justice, L. M., McGinty, A. S., DeCoster, J., & Forston, L. D. (2015). Teacher–child conversations in preschool classrooms: Contributions to children's vocabulary development. *Early Childhood Research Quarterly*, *30*, 80-92.

Cervetti, G.N., Hiebert, E.H., Pearson, P.D., & McClung, N. (2015). Factors that influence the difficulty of science words. *Journal of Literacy Research, 47*(2), 153-185.

Cervetti, G.N., Fitzgerald, M., Hiebert, E.H., & Hebert, M. (July 18, 2019). The influence of vocabulary instruction on word knowledge and word solving. Poster presented at the annual conference of the Society for the Scientific Study of Reading, Toronto, ON. Retrieved from:

https://www.triplesr.org/sites/default/files/uploads/conference_program_sssr_2019_toronto.july.15th.pdf

Clay, M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.

Dale, E., & O'Rourke, J. (1981). *The living word vocabulary: A national vocabulary inventory*. Chicago, IL: World Book-Childcraft.

Davies, M. (2009). Corpus of contemporary American English: Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, *14*(2), 159-190.

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44.

Fountas, I. C., & Pinnell, G. S. (2008). *Leveled literacy intervention*. Portsmouth, NH: Heinemann.

Glucksberg, S., McGlone, M. S., Grodzinsky, Y., & Amunts, K. (2001). *Understanding figurative language: From metaphor to idioms* (No. 36). Oxford University Press on Demand.

Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of dyslexia*, *60*(2), 183-208.

Hiebert, E.H., Goodwin, A.P., & Cervetti, G.N. (2018). Core vocabulary: Its morphological content and presence in exemplar texts. *Reading Research Quarterly*, 53*(1),* 29-49.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4) 978–990

Marzano, R.J., & Marzano, J.S. (1988). *A cluster approach to elementary vocabulary instruction*. Newark, DE: International Reading Association.

McDonough, C., Song, L., Hirsch-Pasek, K. Golinkoff, R.M., & Lannon, R. (2011). An image is worth a thousand words: Why nouns tend to dominate verbs in early word learning. *D, Developmental science*, *14*(2), 181-189.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly*, 20(5), 522-535.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, *19*(3), 304-330.

Nagy, W.E., & Hiebert, E.H., (2011). Toward a theory of word selection. In M.L. Kamil, P.D. Pearson, E.B. Moje, & P.P. Afflerbach (Eds.), *Handbook of Reading Research* (Vol. 4; pp. 388-404). NY: Longman.

National Center for Education Statistics (2017). *National Assessment of Educational Progress (NAEP): 1992-2017 Reading Assessments.* Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from https://www.nationsreportcard.gov/reading_2017/

National Governors Association Center for Best Practices (NGA CBP) & Council of Chief State School Officers (CCSSO) (2010). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects*; Appendices A & B. Washington, DC: Author. Available online: www.corestandards.org/assets/.

Reitsma, P. (1988). Reading practice for beginners: Effects of guided reading, reading-while-listening, and independent reading with computer-based speech feedback. *Reading Research Quarterly, 23*(2), 219-235.

Stallman, A. C., Commeyras, M., Kerr, B., Reimer, K., Jimenez, R., Hartman, D. K., & Pearson, P. D. (1989). Are "new" words really new?. *Literacy Research and Instruction*, *29*(2), 12-29.

Stevenson, A. (2010). *Oxford Dictionary of English* (3rd Ed.). DOI: 10.1093/acref/9780199571123.001.0001 http://www.oxfordreference.com

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*(1), 8-25.

Wright, T. S., & Cervetti, G. N. (2017). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, *52*(2), 203-226.

Zeno, S.M., Ivens, S.H., Millard, R.T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, MA: Touchstone Applied Science Associates Inc.

Children and Young Adult Texts

Applegate, K. (2012). *The one and only Ivan*. New York, NY: HarperCollins. Author (2011)

Bix, C.O. (1995). *Water, water everywhere*. San Francisco, CA: Sierra Club Books for Children.

Fletcher, R. (2003). *Hello, harvest* moon. New York, NY: Clarion Books.

Henke, K. (2004). *Kitten's first full moon.* New York, NY: HarperCollins.

Hiebert, E.H. (2013). "Bringing toys to school," *Talking Points for Kids.* Santa Cruz, CA: TextProject.org. Retrieved from: http://textproject.org/classroom-materials/students/talking-points-for-kids/bringing-toys-to-school/

History.com (August 21, 2018). *Checks and balances*. Retrieved from: https://www.history.com/topics/us-government/checks-and-balances

Soto, G. (1996. *Living up the street*. New York, NY: Laurel Leaf Library.

Stevenson, R.L.. & Pearson, T.C. (2006). *The moon*. New York: Farrar, Straus, & Giroux.

Yolen, J. (1987). *Owl moon*. New York: Philomel Books.

Take Action!

1. Use sets of topically connected texts that are available for free download at readworks.org. Have students read the texts over a week, identifying the shared vocabulary across the texts on a classroom word wall.

2. Select a topic for a specific period of the school year (e.g., a month, a quarter) and cluster available texts (i.e., leveled texts, trade books, selections in core reading anthologies, magazine articles) in your classroom around the topic. Encourage students both individually and as a group to identify the critical vocabulary for a topic.

3. Put a typical text from frequently used materials in your classroom on a white board. Lead a conversation with students about which words are rare and how known words in the text can be useful in comprehending the meaning of unknown words.

More to Explore

For more information on how to develop clusters of vocabulary surrounding topics, morphological awareness, and multiple-meanings of words:

Cervetti, G. N., & Hiebert, E. H. (2019). Knowledge at the center of English Language Arts instruction. *The Reading Teacher*, *72*(4), 499-507.

Goodwin, A. P., & Perkins, J. (2015). Word detectives: Morphological instruction that supports academic language. *The Reading Teacher*, *68*(7), 510-523.

Greenwood, S. C., & Flanigan, K. (2007). Overlapping vocabulary and comprehension: Context clues complement semantic gradients. *The Reading Teacher*, *61*(3), 249-254.

Table 1. Distribution of Words and Word Families in Written English

Word Category	**Predicted appear-ances per million words of text**	**Percentage Accounted for in CCSS texts**	**Number of Words**	**Number of Morph-ological Families**	**Example of Lead Word**		
					Highly Abstract (1-1.5 on 5-point scale)	**Average (3 on 5-point scale)**	**Highly Concrete (5 on 5-point scale)**
Highly Frequent	68,006 – 300	65.2	310	240	would, if, though, because, as	example, said, last, several, look	water, house, feet, face, sun
Frequent	299 - 100	13	620	380	perhaps, actually, suppose, instead, rather	return, force, meet, spent, grow	trees, horse, fish, birds, bed
Moderately Frequent	99 - 10	14.7	4,656	1,832	despite, seldom, normal, merely, ideal	current, invite, sick, technology, trial	turtle, tomatoes, stairs, sand, pillow
Relatively Rare	9-1	4.9	13,882	N/A	whatsoever infinitely ultimately intangible, thereof	trance, pry, jumble, gossip, accumulate	walrus, vase, umbrella, tulip, tractor
Very Rare	<1	2.2	135,473	N/A	abject jeopardy awry pathogenic deviate	inflammable martyr juncture, psychedelic, relinquished	oriole, smock, tattoo, rhubarb, thistle

Table 2. Illustrations Of Semantic, Morphological, and Multiple-Meaning Connections Of Moderately Frequent Words In The Core Vocabulary[1]

Category	**Period or Type**	**sick (Early[2])**	**invite (Middle[2])**	**current (Middle)**	**technology (Late[2])**	**trial (Late)**
Semantic	Early	care doctor fever Miller s hospital ill nurse pain sore medicine fever	ask welcome	now present happening latest		fair wrong right
	Middle	patient wound disease operation poorly cancer medical temperature heal unwell infection treatment	offer suggest request attract	live popular actual modern existing	camera equipment platform	crime guilty defend law court judge
	Late	bacteria	propose		terminal computer machinery mechanics	ruling sentence
Morphological[3]	Inflected endings & comparatives	sicken sicker sickest	invites invited inviting	currents	technologies	trials
	Affixes	sickly sicklier sickish sickness	invitingly inviter invitee invitingness invitation invitatory disinvite uninvited reinvite	currently currentness currency	technological technologic technologist technologize	mistrial pretrial posttrial retrial

	Compound words	heartsick homesick sickroom sickbed seasick airsick carsick sick-list sick leave sick pay				
Multiple Meanings		adjective: affected by physical or mental illness adjective: feeling nauseous and wanting to vomit	verb: make a formal or polite request of someone verb: to elicit a particular response or to tempt someone to do something noun: an invitation	adjective: existing or being used now noun: continuous movement of water or air in a particular direction noun: a flow of electricity; the rate of flow of electric charge noun: the general course of events or opinion	noun: machines and equipment developed from the application of scientific knowledge noun: the branch of knowledge dealing with engineering or applied science	noun: a legal process in which a court of law examines a case to decide if someone is guilty of a crime noun: a test of performance of someone or something verb: test to assess its performance

[1]Semantic and morphological relatives of the focus word are based on the first meaning of the word (final row of this table).
[2]Early, Middle, & Late refer to bands of age at which the word is predicted to be in students' oral language: Early = age 5.9 or earlier; Middle = Ages 6 through 8.9; Later = Ages 9 and beyond
[3]The morphological relatives are in the EWFG, although not all are in the highly or moderately frequent groups.

Table 3. Moderately Frequent Words With Repetitions (3 times plus)[1] in Semantically Connected Groups in Two Text Sets

Text Complexity	Topic + Text Complexity
1) journey 4[2]	1) flower 6
train 5	garden 6
explorers 5	harvest 3
	nutty 3
2) protect 4	orange 7
safety 5	seeds 3
	vegetable 3
chemicals 3	vines 9
competitions 5	
created 3	2) healthy 3
double 4	nutrients 4
melted 4	vitamin 3
meter 3	
mount 4	3) competition 3
music 6	contests 3
palace 3	winners 9
threw 3	
trick 5	4) transport 3
volcano 8	truck 4
	5) scales 3
	weighed 7
	carve 4
	chains 10
	circle 3
	create 4
	football 3
	native 3
	patch 6
	perfect 3
	smooth 3
	sweet 3

[1]Proper names and words with 4 or fewer letters excluded

[2]Number of times word is repeated in the sample of 2,300 words.

Table 4. Examples of Figurative Language in the Core Vocabulary

Book (Author)	**Figurative Language**[1]
Hello, Harvest Moon (Fletcher, 2017)	"It comes up round, ripe, and huge over autumn fields of corn and wheat. Hello, harvest moon. *With silent slippers it climbs the night stairs*," (p. 4)
Owl Moon (J. Yolen, 1987)	"The moon made his face *into a silver mask."* (p. 10).
Kitten's First Full Moon (Henke, 2010)	"It was Kitten's first full moon. When she saw it, she thought, 'There's *a little bowl of milk in the sky.'* " (p. 1)
"The Moon" (Stevenson & Pearson/1885/2006)	"The moon has *a face like the clock in the hall;*"(p. 3)

[1]Figurative language is in italics.