The Vocabulary Filter Process

Elfrieda H. Hiebert

TextProject & University of California, Santa Cruz

In every 100 words of text, two to three words are likely rare—a word that is not frequent in written language. What this pattern of rare words means is that texts with 700 to 1,000 words can have numerous rare words and a chapter book might have several hundred rare words. Which rare words should be the focus of valuable instructional time? The Vocabulary Filter process provides a set of six questions for teachers to ask in choosing the words to teach. Further, teaching students about the different types of words represented by each filter can support independent vocabulary recognition proficiency.

How can we as educators be strategic in the words that we teach students? In particular, how do we do that in light of the enormous lexicon or dictionary of English—approximately 600,000 words (Oxford English Dictionary[1])? A first understanding for teachers and students has to do with the core vocabulary—the approximately 2,500 morphological families that account for an average of 91.5% of the total words in texts from kindergarten through college[2].

1   Steven, A. (Ed.)(2010. Oxford Dictionary of English (3rd Ed.). Online version: DOI: 10.1093/acref/9780199571123.001.0001

2   Hiebert, E. H., Goodwin, A. P., &

Knowing that many words will be known or have known root words can give students confidence. This knowledge can also encourage them to use the context of known words to figure out unknown words.

Even with an increased sense of confidence, however, there will be rare words in any text. In general, approximately 2 to 3 words of every 100 words in children's book will be rare[3]. Rare words are ones that are predicted to occur less than once in every million words of text. Most students are unlikely to have read more than one million words until at the end of the middle grades. That means that these rare words can be expected to be first-time encounters for many students. It's impossible to teach all of the rare words in any text. Nor should that be the aim of teachers. The goal of vocabulary should be to focus on words that ensure that students develop strategic knowledge that they can apply in figuring out word meaning in independent reading. To achieve this goal, however, depends on guidelines for teachers in knowing which words to highlight in their instruction.

The Vocabulary Filter[4] provides a system for teachers to apply in selecting the words worth teaching. It consists of six filters, each of which focuses on a different type of rare word in text: (a) words for familiar concepts, (b) concrete words, (c) semantic families of words, (d) words that share root words, (e) multiple-meaning words, and (f) words with highly complex meanings.

The *One and Only Ivan* (Ivan; Applegate, 2012),[5] a popular book in upper-elementary and middle-school classrooms, will be used to demonstrate The Vocabulary Filter process. In a book of approximately 27,000 total words, approximately 7% of the words that students will encounter are rare. But, since each rare word is repeated an average of 4 times, the number of unique rare words is lower.

Before applying The Vocabulary Filter to *Ivan,* a group of words that fall into the rare category needs to be addressed—proper names. Proper names are not like other groups of words in that they usually do not have a clear definition and can be pronounced in unusual ways. But proper names can add complexity to texts, especially when names are unusual. In the case of Ivan, 20 proper names account for 47% of all the rare words. That is, when proper names are taken out of the mix, the number of rare words in Ivan decreases to about 4% (with most of these words repeated an average of two times). A discussion on the characters in a text can be useful in setting the stage for a book and can waylay potential challenges some students may have in dealing with unusual names.

FILTER 1: WORDS WITH FAMILIAR CONCEPTS

Numerous words in narrative texts represent concepts with which students are familiar in oral language but are ones that students simply have not encountered previously in written language. Consider, for example, the word *swagger.* Most students know the term "show off." The word swagger will not require an extended discussion but simply a mention or comment about the meaning of the word should do the trick.

When an author's style includes many rare words with familiar meanings that are similar in function in a text, a short lesson may be justified. *Swagger* isn't the only rare word that Applegate uses to describe the motions of Ivan and his friends in the arcade. Many of the words are rare but represent familiar concepts to students—words such as *scampers, scuffling, hobbles, nuzzles,* and *topples.* A lesson that focuses on Applegate's use of these words could be useful in understanding the story and also in the role of inventive words in good writing (including students' own writing).

In studying an author's use of rare vocabulary in a text, teachers should be aware that some compound words fall into a unique group of familiar concepts but rare words. The last word or base word in a compound word is almost always a word that occurs with considerable frequency in written language. Most of the descriptive or first words in the compound are

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

3   Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of 'motherese'?. *Journal of Child Language, 15*(2), 395-410.

4   http://textproject.org/assets/library/powerpoints/HiebertIRA2011Vocabulary-Filters.pdf

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

usually quite frequent as well. For example, students know the words *arrow* and *root* and, while most may never have seen an *arrowroot,* the meaning can be inferred. When compound words are not literal meanings of the two component words, some students may benefit from a discussion of the word's meaning. For example, students know *lift* and *fork* but the word *forklift* in *Ivan* is not a dinner fork that is lifted while eating. In Ivan, unusual compound words are few. In some texts, such as *The Birchbark House,[6]* unusual compound words are many and a short lesson on this unique aspect of the book might support the vocabulary recognition of some students. Keeping an eye out for unique uses of vocabulary in texts can support students in attending to author's craft and applying that knowledge to their own writing.

FILTER 2: CONCRETE WORDS

The second filter addresses the question: Would the presentation of a picture convey the concept of the word more readily than an extended discussion or definition? A critical group of words in Ivan occur frequently in the text and are easily picturable: *gorilla, knuckle, silverback, chimpanzees,* and *arcade.*

6   Erdrich, L. (1999). *The Birchbark House.* New York, NY: Hyperion Books for children.

The latter is especially important because the arcade provides the setting where Ivan spends the majority of the book (and 27 years of his life). *Silverback* may be easy to decode but seeing a picture will help understand the compound word (and, in this case, can be seen on the cover of the book). *Knuckle* occurs only a couple of times in the text but it is important because Ivan remonstrates about the challenge of "knuckle walking."

The aim here is not to encourage a strategy of "look at the picture to figure out the word.," which is frequently promoted in the primary grades. Teaching young readers to rely on pictures can create problems in reading development. Further, in extended books such as Ivan, pictures are few. If the concept is an important one to the story and it is highly concrete, a picture can be worth many words of explanation and definition. This technique is an especially effective one with English learners who may have a concept in their native languages on which they can draw to understand the English word.

FILTER 3: WORDS THAT SHARE SEMANTIC FAMILIES

Compelling stories contain many rare words that describe characters' movements, traits, emotions, and ways of communicating. An author of a compelling story does not repeat the same word over and over again—even if the word is an intriguing one, such as *temperamental* or *undaunted.* The variety of word choices adds quality to a narrative. For example, Ivan's descriptions of humans as clumsy and his sister as nimble clarify his views of humans in a way that is not conveyed in a sentence: "Humans do not move very well."

There are particular categories that typically are the focus of these different descriptive words--emotions and traits, ways of communicating, and forms of movement. Authors use a variety of intriguing words in specific and nuanced ways, much like artists use a palette of colors. Many of these evocative words can be expected

Figure 1
Examples of Semantic Networks (Filter 3)
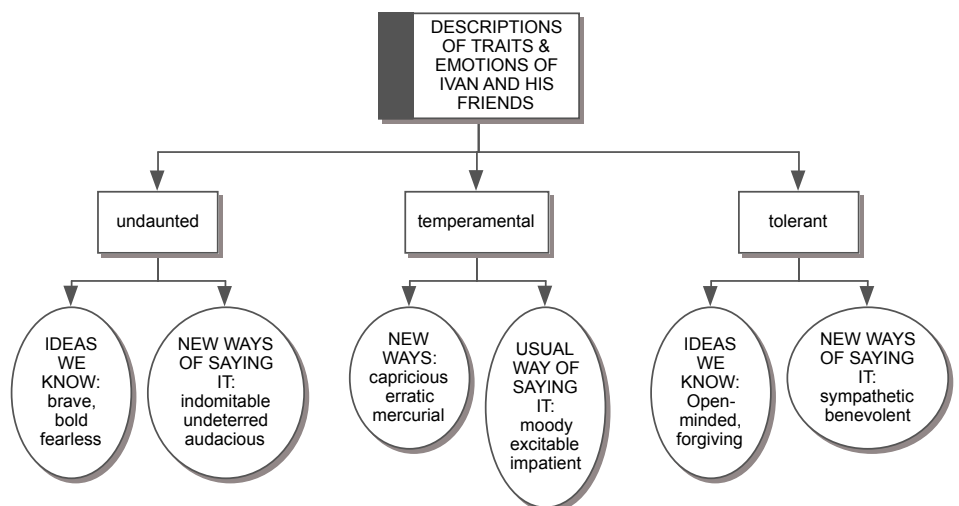
to occur infrequently in written language but these words are part of shared networks of words.

Figure 1 depicts the networks around several of the words that, Applegate, the author of *Ivan,* uses to describe the traits and emotions Ivan and his menagerie in the arcade. Some of these words are positive, some are negative. But each word—*undaunted, temperamental,* and *tolerant*—is part of a rich network of words with connected meanings. The manner in which authors use words in expressive and poignant ways merits instructional attention. In conducting lessons on these words, teachers want to move beyond the single word that the author uses to illustrate the different hues or nuances of words in the network. For example, Bob (Ivan's friend at the arcade, a dog) describes himself as undaunted. More common words for this trait are brave, bold, and fearless. But there are other words that convey the sense of strength and courage conveyed by undaunted: indomitable, undeterred, steadfast, and intrepid. By discussing and learning a network of words, students learn about the richness of language available to them as readers and writers.

FILTER 4: WORDS THAT SHARE ROOTS

The fourth filter relates to teaching students about the way in which meaning-unit or morphemes are shared across families of words. Remember that a large portion of rare words have root words that students have already encountered.

Table 1
Example of a Morphological Family (Filter 4)

Forms of the Root Word (Syllabic Structure)	Text Example	Meaning of the Affix
hesitant (hes•i•tant)	Part of me was hesitant to answer but something in her voice told me that I didn't need to be afraid.	-ant: the capacity of
unhesitant (un•hes•i•tant)	"He's always been the best man I know," was the **unhesitant** response.	un-: lacking, opposite of
hesitation (hes•i•ta•tion)	"I have a favor to ask," she said slowly, with a hint of hesitation in her voice.	-tion: The act of hesitating

A recent analysis showed that 40% of rare words had root words that students had already encountered.[7] Frequent demonstrations of how words in morphological families retain the meaning of their shared root word but also take on different functions and meanings when affixes are added are important to giving students the prowess they need to be proficient in reading complex texts.

The word *hesitating* illustrates how a word from a text can be used to demonstrate the richness of morphological families. Beginning with the use of the word in the text is critical in contextualizing the instruction.

"Oh," Ruby says. "Oh. Mack." She puts her trunk between the bars. "Do you think " She hesitates. "Do you think Mack is mad because I hurt him today?"

Students can infer from this that *hesitates* has to do with waiting

7   Hiebert, E.H., & Pugh, A. (July 19, 2018). *An Examination of Rare Words in Texts Across Grades and Genres.* Paper presented at the annual meeting of the Society for the Scientific Study of Reading, Brighton, UK.

or being uncertain. The inflected endings can be introduced (*hesitated, hesitating, hesitates*) but, more importantly, is the presentation of family members with affixes. The three most common entries in the Oxford Dictionary of the morphological family members of *hesitate* with affixes appear in Table 1. All possible members do not have to be the focus of a lesson, especially when some family members (e.g., *hesitance, hesitatious, hesitative*) are exceptionally rare or even archaic. The examples in Table 1 will be sufficient to illustrate how a suffix typically changes the meaning but not the function of a word, while unhesitant illustrates the effect of a prefix in changing the meaning.

As is evident in Table 1, adding affixes (suffixes, prefixes) invariably means that words add syllables. It is often with multisyllabic words that some students lose confidence in reading and begin reading in less than sustained ways. Many students have not been taught to read multisyllabic words. For that reason, an important part of teaching students about the

addition of affixes and inflected endings involves demonstrations of the pronunciation of words. The aim is not to require students to divide words into syllables on their own (a challenging and often tedious task for many students) but to ensure that they are guided in the pronunciation of multisyllabic words.

Another aspect of selecting words to demonstrate morphological families, especially in the middle elementary grades and beyond has to do with Spanish cognates. Many of multisyllabic words in the literary and academic layer of English share a linguistic history with Spanish. Two of the multisyllabic words in *Ivan, contemplate* and *confident,* have close Spanish cognates (*contemplar, confidente*). In the case of hesitate, however, the typical Spanish word is not a cognate—*vacilar.* Bringing this word into the discussion, however, can be illuminating since it is a cognate for vacillate, a word that is part of the semantic family of hesitate. Drawing students' attention to the Spanish connection is not simply for the benefit of native Spanish speakers (which is considerable) but it is also to support native speakers of English and of other languages to attend to the ways in which complex words change function and meaning.

FILTER 5: MULTIPLE-MEANING WORDS

The next step in the filter process is aimed at teaching students to approach texts with the expectation that many words in English take on multiple and unique meanings, depending on the context of the text. Some words can shift dramatically in their meanings (e.g., *arms, subjects, pupils*). Many common words, in particular, are used for multiple purposes, some of which are quite divergent from one another. The word *face,* for example, has two distinct meanings (front of the head; to confront something) and is used in idioms (e.g., "egg on his face"). Especially for students who are struggling with literacy, the variability of meanings of words in English can present a challenge.

All of the words in a unit of text are previewed to determine which words might be used to illustrate the feature of multiple meanings, especially ones that can create an obstacle for students' comprehension. Two potentially confusing words with multiple meanings in *Ivan* are *domain* and *juvenile.* Ivan consistently refers to the place where he lives as his domain. That in itself makes this word worth highlighting in class discussions because, at least in students' encounters with the word in social studies, domain typically describe a territory controlled by a ruler or government. The author's use of the word *domain* conveys Ivan's sense of control and agency—even in dire circumstances—which would not be the case if he referred to his lodging as a cell or a cage.

Additionally, however, students can expect to encounter domain in other contexts. In computing, for example, domain refers to a subset of the Internet with addresses sharing a common suffix. Domain can also be used to refer to specific activities or areas such as the domain of science or of medicine.

The word *juvenile* is not used as frequently as *domain* in *Ivan* and, when it is used in the book, only one of its meanings is used. *Juvenile* is a word that merits attention in scientists use the word to refer to the young of a species (which is how it is used in *Ivan*). But it can also be used in a pejorative sense to describe someone's behavior as childish. The word *juvenile* illustrates how words can be used in academic texts but also can be expected to be heard in conversations and in literary dialogues or essays.

FILTER 6: WORDS WITH COMPLEX MEANINGS

As previous filters have shown, many words may be new to students but the underlying concepts will not be. But some words represent concepts that are new to students. Such words are typically more prevalent in informational than in narrative texts but even in narrative texts there can be words that represent complex and abstract concepts that can challenge students. Many schoolchildren will have encountered the word *habitat* in discussions of the environment but it illustrates a word that represents a complex idea in *Ivan.*

The source for words with complex meanings in lessons surrounding the reading and instruction of *Ivan* may lie in informational articles about the history of the real Ivan and about gorillas in their native habitats. A beautiful tribute to Ivan is available

on the Atlanta Zoo website[8] (where Ivan spent the last 18 years of his life. Two concepts are developed in this description of Ivan that, while not stated explicitly in Applegate's book, merit attention: *legacy* and *credibility.* Ivan left a legacy—the recognition that wild animals cannot be locked in small spaces. The Atlanta Zoo also had credibility because it had been successful in creating naturalistic habitats for gorillas.

When concepts are important and complex, the Frayer method[9] can be a good way to support students in developing understanding.   Fundamentally, it involves a 2 x 2 table as shown in Figure 2.. What's important to remember about the Frayer method is its function in conversations around critical concepts. *Credibility* and *legacy* represent complex ideas for students. The examples, non-examples, and characteristics of the word *credibility* in Figure 2 include ideas that merit discussion and explanation. The Frayer method does not need to be applied to all words by any stretch of the imagination but words such as *credibility* and *legacy,* where finding antonyms and non-examples is challenging, are exactly the kinds of words for which the Frayer method is appropriate.

Figure 2
Frayer chart for the word credibility (Filter 6)

Definition	Characteristics
The quality of being trusted, believed in	Truthful Dependable Honest predictable
Examples of credibility	**Non-Examples**

• Atlantic Zoo:
  A zoo has taken good care of a wild gorilla.
• Julia
  Julia has taught Ivan to paint.
  Julia has talked with Ivan while she waits for her father.
	• Chicken Little sho says that the sky is falling
	• Stange things that happen on April Fool's Day.

SUMMARY OF THE FILTERS

Each of the filters represents a fundamental stance toward words in English: (a) words for familiar concepts, (b) words that represent concrete and picturable ideas, (c) words within the same semantic family, (d) words that share root words, (e) words with multiple meanings, and (f) words that represent highly complex concepts. Schools can never be responsible for teaching all of the words in English. However, students can be taught stances that can aid them in learning new words. The Vocabulary Filter provides a means for educators to choose words in a way that increases students' word knowledge and also strategies for knowing how words work. TM

References

Apex Learning (2011). *English III: American literature.* Seattle, WA: Author.

Bergman, L., & Pearson, P.D. (2008). *Into the soil.* Berkeley, CA: Lawrence Hall of Science.

Chall, J.S. (1983). *Stages of reading development.* New York, NY: McGraw-Hill Book Co.

Common Core State Standards Initiative (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects.* Washington, DC: National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Dale, E., & Chall, J.S. (1948). A formula for predicting readability and instructions. *Educational Research Bulletin, 27,* Jan. & Feb., 1948, pp. 11–20, 28, 37–54.

Dolch, E.W. (1948). *Problems in reading.* Champaign, IL: Garrard Press.

Harris, A., & Sipay, E.R. (1990). *How to increase reading ability: A guide to developmental and remedial methods.*

Hiebert, E.H. & Folkins, A.L. (2012). Big seeds, little seeds. In *BegininngReads.* Santa Cruz, CA: TextProject. Retrieved from http://textproject.org/products/beginningreads/

McKeown, M., 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.

National Center for Education Statistics (2009). *The Nation's Report Card: Reading 2009* (NCES 2010–458). Institute of Education Sciences, U.S. Department of Education, Washington, D.C.

Simpson, J., & Weiner, E. (2009). *Oxford English Dictionary.* New York, NY: Oxford University Press.

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

---

8   Meet Ivan. https://zooatlanta.org/animal-legend/ivan/
9   Frayer, D. A., Fredrick, W. C., & Klausmeier, H. J. (1969). A schema for testing the level of concept mastery (working paper No. 16). *Madison, WI: Wisconsin Research and Development Center for Cognitive Learning.*