



GENERATIVE VOCABULARY INSTRUCTION

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

P. David Pearson
University of California, Berkeley

It is hard to put an exact number on the number of words in the English language, but there is agreement that English has more words than most languages.¹ In everyday conversations, people, even highly educated adults, use only a small portion of the words available to them. Print is a different matter. Many more words from the English lexicon are used in written language. Even so, a very small group of words continues to account for the majority of the words in texts, but authors also use many rare words to define, describe, elaborate, or add nuance to their ideas. One of the signatures of complex texts is the presence of rare vocabulary.

Not all of the rare words that students will encounter in complex texts, whether in school or in their careers, can be taught. There are simply too many words in written English. To successfully understand complex texts, students need to be able to generate the meanings of new words, based on their knowledge about how words work in English.

Background on the Rare Vocabulary of Complex Texts

The most obvious feature of a complex text is the presence of rare words. The rare words typically account for no more than 10% of all of the words in a text but these words introduce and extend concepts. By using *shriek* rather than *say*, in *Charlotte's Web*, author E. B. White conveys Fern's outrage at her father's plans for the runt of the litter. By describing the tarantula found in a pile of leaves as a *Theraphosa* in *The Tarantula Scientist*, author Sy Montgomery is revealing that not all tarantulas are alike.

Words such as *shriek* and *Theraphosa* are sprinkled among the core vocabulary of complex texts. Core vocabulary consists of a relatively small group of words that does the heavy lifting in English. Around 4,000 word families account for an average of 90% of the words in the majority of texts. In English, the more common a word, the greater the likelihood that it will have multiple meanings; contrast a word like *run*, with 645 meanings documented in the Oxford English Dictionary² with a word like *permute*, which has a single meaning. This means that the core vocabulary of English—words such as *say* and *leaves*—will often have multiple meanings and can even function as different parts of speech. With their multiple meanings and functions, words in the core vocabulary are prominent in both narrative and informational texts. While these words may inhibit comprehension and should be defined at point of use, they don't require the same intensive instruction as rare words.

Unlike core vocabulary, rare vocabulary (i.e., the other 10% of words) comes from an enormous bank of words—at least 85,000 unique families of words. A word family is a group of words connected with inflected endings and affixes, such as *certify*, *certified*, *certifying*, *certifies*, *certification*, and *certificate*. Not all of these word families can be taught to children. If schools started with five words for every school day in kindergarten, students would be 100 years old before all of the words had been covered.

A critical group of words can be taught, but both the choice of words and the nature of instruction need to be **generative**, if students are to be prepared to unlock the meanings of the many rare words they will encounter in complex texts. *Generative* refers to the ability to apply knowledge of how words work when encountering new words. Generative instruction aims to make visible to students critical features and functions of words and connections among words. This knowledge is intended to support students in generating meanings of unknown words in texts. A student who recognizes that *spectacles* must have something to do with sight because the word has *spect* in it is using a generative strategy to unlock the new word. Or a student who infers that *glamorous* must have something to do with beauty because it is grouped with *attractive*, *beautiful*, and *pretty* is also using a generative strategy.

Currently, in English Language Arts (ELA) instruction, five to six words are chosen from a text and these words are taught in a variety of contexts over the course of a week. Consider the 10 Benchmark Vocabulary words for the first two chapters of *Charlotte's Web* found in Pearson's *ReadyGEN* program (Grade 2, Unit 1) and identified using a generative approach: *delightfully*, *injustice*, *increased*, *blissful*, *untimely*, *runty*, *adoring*, *specimen*, *distribution*, *selecting*. These words are often described as "Tier 2" words. Tier 2 words are referred to in the Common Core State Standards as *general academic words*. These are words that are more common in written texts than in speech, including informational, technical, and literary texts. Tier 2 words provide writers with subtle ways to express simple concepts, such as *blissful* rather than *happy*. The generative approach incorporates the Tier 2 approach by identifying a set of words to teach intensively with a text or set of texts. In addition, the generative approach extends the Tier 2 words in two ways.

First, words are chosen for the richness of their morphological family, their centrality to academic texts, and their role in a particular text. Beyond the core vocabulary, most new words that students encounter are members of families of at least four words. By emphasizing the morphological connections, and the Spanish cognates, when appropriate, students begin to generate an understanding of how the English language works. Thus, when the word *injustice* is taught in *Charlotte's Web*, reference can be made to other members of the family to which this word belongs (e.g., *justice*, *just*, *justly*, *unjust*, *unjustly*, *justification*).

Second, generative words are taught as networks of ideas rather than as single words. Currently, in American core reading programs, each of these focus vocabulary words would be taught as a separate word with a cycle of activities devoted to each word. These activities include using “friendly” definitions, discussing appropriate uses of a word, and writing sentences with a word. The words are not taught or practiced in relation to a central theme or to one another. Fern’s *blissful* state, for example, is not discussed in relation to *delightfully*; *runty* is not discussed in relation to the pig being a poor *specimen* of the breed.

The vocabulary map in Figure 1 shows that many of the same Tier 2 words are also selected for instruction within a generative approach (although there is not a perfect match). Where the generative approach differs from the Tier 2 approach is in how the words in the generative approach are taught as networks of ideas.

By teaching words in clusters of ideas, students are learning many more words. In stories, some of the words may be new to students but many of the concepts such as shouting or talking loudly (represented in this text by *shrieked*, *yelled*, and *cried*) are not new to students. Lengthy discussions or hands-on activities are not needed to understand the concept of shouting or talking loudly. What students need to learn is how a variety of words can be used to describe different degrees (*tepid*, *warm*, *hot*, *piping hot*, *steaming*, *sizzling*) or nuances (*pretty*, *glamorous*, *stunning*). These words carry different connotations when we encounter them in a text.

By making the connections among words rather than teaching them individually, which is the case with present Tier 2 instruction in American core reading programs, students are gaining the vocabulary awareness that allows them to approach unknown words with confidence and knowledge. Therein lies the essence of the generative approach to vocabulary instruction and learning.

How the Three Tiers of Words Fits into Generative Vocabulary Instruction

For the most part, Tier 1 words are core words, those words that children encounter and use mainly in conversation. These words also comprise 90% of the words they encounter in written text, but not in everyday conversations or in media such as TV shows. Tier 2 and 3 words are the other 10%, the rare words that students usually encounter in complex written text but not in everyday conversations or in media such as TV shows. Categorizing words into tiers is helpful for teachers when analyzing text complexity; it can provide important information about the academic complexity of a given text and potential entry points for instruction. Putting words into tiers can be challenging because the placement of a word in a tier can change depending on the text type and context for the word. For example, in an informational text such as *The Tarantula Scientist*, the words *barbed* and *venom* are Tier 3 words (i.e., key content words) with precise meanings that accurately describe the spider's defense mechanisms. When encountered in a narrative text, these same words take on Tier 2 significance: "the old man's *barbed* tongue and *venomous* laugh frightened the children."

The generative approach to vocabulary instruction builds on the tier concept by providing teachers with a method for identifying the most important Tier 2 and 3 words to teach in a text. These are words that contain both the content and conceptual knowledge needed to understand the text at hand and that provide opportunities for teaching how words and language function within different text types, subjects, and contexts.

Illustration of a Generative Stance in Action

Consider a text identified for a fifth-grade assessment³ about an architect who lives in and designs tree houses. Of the 460 words in this text, 7% of the words are rare. As is the case with topics about new inventions or innovations, many of the rare words are compound words (e.g., *crabapple*, *ziplines*, *hardwoods*, *overlook*). It is highly unlikely that fifth graders will have been taught all, or even any, of these words before they encounter them on the assessment. But when instructional words have been selected for their generative nature and instruction has supported a generative stance toward vocabulary, students will not be surprised when they encounter compound words in a text on a specialized subject. For example, when the text *The Birchbark House* (Pearson's *ReadyGEN* Grade 4 Unit 2) was taught earlier in the curriculum, students were taught about how people label newly encountered or discovered phenomena by putting together existing words. Words such as *birchbark* and *namesake* from the first chapters of that text illustrate the kinds of words used for developing a generative stance toward vocabulary.

Another rare word in the fifth-grade assessment text is *arborist*. The word is defined in the context of the text. In a generative approach, students are supported in applying what they have been taught to new words in context. Spanish speakers who have been taught using a generative approach will recognize that this word is a cognate for *arból*, while students who speak native languages other than Spanish will recognize this word from the study of Arbor Day and the connection of that celebration to trees.

Picking Generative Vocabulary in Narrative and Informational Texts

Since narrative and informational texts serve different functions, the rare vocabularies in these genres are also unique. It is important that teachers consider these different functions when choosing words to teach using a generative approach. Through instruction, students will gain an understanding of how authors of narrative and informational texts use words differently to achieve their purposes. But, as well as attending to the differences between the unique words of narrative and informational texts, it is also important to keep in mind a feature that both types of words share—membership in morphological families.

Morphological Families: The Foundation of a Generative Approach

The foundation of a successful generative stance is in knowing about word families or the morphological systems of English. The phonology-orthography of English is the foundation of decoding. Knowledge of the morphological systems of English is fundamental to reading complex texts. Morphology is as critical to comprehension and learning from complex text as “breaking the code” is to reading acquisition.

Uniqueness of Rare Vocabulary in Narrative Texts

In narrative texts, rare words represent new labels for representing known concepts. Middle-grade students may not immediately recognize the meaning of the word *nonplussed* but most know what it means to be confused. Rare words in narratives have connections to what students know in two important ways. First, the underlying meanings of the rare words—*shrieked*, *swooped*, *runty*, or *blissful*—are often known to students. That is, most students know about yelling, jumping or leaping, tiny, and happy, but they may not know the less-used synonyms for these ideas when they encounter them in complex narratives.

Second, the rare words serve consistent functions in narratives. The characters of stories, whatever their form (e.g., animals, machines, or people), have recognizable personality traits and experience emotions

with which students are acquainted. Most stories also share a familiar plot arc: rising action, climax, and resolution. Plots themselves often fall into well-worn patterns that readers might recognize, such as a hero's journey with a goal and obstacles or a mystery to be solved.

In generative instruction, students are taught about the ways in which authors use rich and diverse words to describe these components of stories. Most of the rare words in stories are used for a handful of functions including communication, traits of characters and settings, emotions, action and motion, and general and proper names of characters and places. There are a few others, but these are the typical categories from which authors draw on words to create characters, setting, and plot. Learning how authors vary their use of words for these functions is a critical part of the generative stance of successful readers.

Table 1 illustrates common words and Tier 2 words for the five prominent semantic categories in narrative texts. Part of a generative stance toward vocabulary is understanding that, as readers, students know the underlying concepts of most of the new words they encounter in narrative texts. What they need to learn are the many synonyms for familiar ideas that authors use in complex texts.

Uniqueness of Rare Vocabulary in Informational Texts

In informational texts, rare words typically represent complex concepts that are new to students and require factual information or a related system of concepts to understand. Understanding *electrochemical energy*, for example, requires that students understand terms such as *chemical energy*, *conversion*, and *electric energy*. Unlike the words in narratives, elementary students usually don't know the underlying concepts of rare words in informational texts. Furthermore, the rare words of informational texts are part of a shared network of words, where the ideas are interconnected but not synonyms for one another. *Conversion* is connected to an understanding of *electrochemical energy*, but it is not a synonym for that term. The critical words within the conceptual map for *The Tarantula Scientist* in Figure 2 illustrate these networks. *Spinnerets* and *exoskeleton* are part of a network of words having to do with the spider's body parts. *Barbed hairs*, *venom*, and *fangs* are part of a network of words having to do with a spider's defense mechanisms. Teaching the words within the concept network develops students' understanding of the individual ideas and the larger concepts.

Teaching Vocabulary in a Generative Fashion

Generative instruction applies to the rare words of both narrative and informational texts. There are some differences in the conceptual networks of the two types of words, but that does not change the instructional approach. Generative instruction uncovers or makes visible the patterns within and across words.

There are two primary instructional routes with generative vocabulary, both of which are critical to developing the vocabulary awareness that is necessary to be successful with complex texts: in-depth instruction of generative words and words in context.

Generative Vocabulary

Words chosen to teach in depth as part of generative vocabulary are simultaneously text-critical and language-critical. First and foremost, the words are critical to comprehending and building knowledge from a text (text-critical). At the same time, the vocabulary has been chosen to build students' knowledge about how words work (language-critical).

The criteria for choosing text-critical vocabulary are the points of departure. For example, in *The Tarantula Scientist*, the *defense* mechanisms of tarantulas is an important concept in the first chapters. In *Charlotte's Web*, Fern's belief in the *injustice* of killing the runt of the litter is the point of initiation for the story of Wilbur and Charlotte.

Injustice and *defense* mechanisms are central to the content of the texts and, as such, require discussion and understanding. At the same time, these words are highly generative both in morphology and semantics. *Injustice* and *defense* are both part of rich and extended morphological families (e.g., *justice*, *just*; *defensive*, *self-defense*). Even if all of the morphological family members are not covered in a lesson, knowledge of one morphological family member is often sufficient for students to use their knowledge in figuring out unknown words. In both cases, these words have transparent Spanish cognates (*justicia*, *defensa*).

Both words are also part of rich semantic networks. Again, all of the synonyms or related concepts for these words do not have to be covered in a lesson in order for students to deepen their knowledge of words that are critical to academic and literary texts.

Words in Context

The total percentage of rare words in narrative and informational texts may be the same, but the number of *different* words, or synonyms, within the entire group of rare vocabulary is often greater in narrative than in informational texts. Authors such as Louise Erdrich or Madeleine L'Engle use a palette of vocabulary to convey a sense of place or setting rather than repeating the same word. A character may experience a setting as *eerie*, *disturbing*, *spooky*, or even *hair-raising*. It is unlikely, however, that any of these words will be used more than once in the same chapter or section of a narrative text. By contrast, Joy Hakim in *The History of US*, in a chapter about the Plains Indians, uses words such as *plains* and *tepees* repeatedly. She is not attempting to add nuance to words that are central to her description of Plains Indians. So she doesn't substitute *grasslands*, *flatlands*, *meadow*, or even *prairie* for *plains*, nor does she use *bivouac*, *lean-to*, or *shanty* for *tepee*. In informational text, the percentage of rare words may be the same as in narrative texts, but the number of different rare words is typically less. New rare words in informational texts tend to get used many times. This is because they represent a big idea in those texts.

Narrative texts, on the other hand, have a large number of rare words that appear a single time. Even in a generative approach, where the critical number of words emphasized within a text is greater than the typical 5-8 words of core reading lessons, all of the rare or critical words cannot be taught. Furthermore, all of these words should *not* be taught. If students believe that they need exposure to every word in a text to read it meaningfully, they will have trouble when they get to new texts, such as the texts in assessments.

An important part of the generative instructional strategy is to ensure that students learn how, when, and why to apply these generative learning tools so consistently that it becomes a habit in reading. For this reason, the words-in-context strategy should be a central part of generative vocabulary instruction. (In Pearson's *ReadyGEN*, words addressed in context are called By-the-Way words.) In a narrative text, in particular, the words-in-context strategy is important because of the presence of many rare words—most appearing only once. But the strategy is also useful with informational texts.

Teachers use the words-in-context strategy to demonstrate application of word knowledge, which includes the knowledge of word features and knowledge of the two kinds of families, morphological and semantic, to which words belong. Although the word features may be ones that have already been taught, students will still need to use the words-in-context strategy to unpack the meaning of the text. The words-in-context strategy gives students an opportunity to use their knowledge about how words work to generate the meanings of new words. As new words are encountered in context, students get to share their understandings of unknown words as well as the strategies they used to develop their interpretations. Teachers get valuable information about their students' capacity for applying word knowledge; thus they can offer them quick feedback to help students refine and expand their knowledge of how words work.

In conclusion, a generative vocabulary stance is both a way to teach and a way to learn. For teachers, a generative stance means making critical features and functions of words visible to students. This knowledge of how words work allows students to generate the meanings of the new words they encounter in complex texts.

Category	Common Words*	Tier-2 Words
Communication	call, shout	exclaim, roar, shriek, holler, screech, yell, bellow
Traits	kind, nice	sympathetic, compassionate, benevolent, generous, humane, charitable, considerate
Emotions	mad, angry	furious, livid, irate, infuriated, outraged
Movement/Action	go, leave	depart, exit, proceed, advance, escape, decamp, journey, bolt,
General names of characters	salesman, saleswoman	peddler, merchant, trader, seller, vendor, clerk, shopkeeper, storekeeper

* Common words are often Tier 1 words.

Figure 1.

Word Map for Chapters 1-2 of *Charlotte's Web*

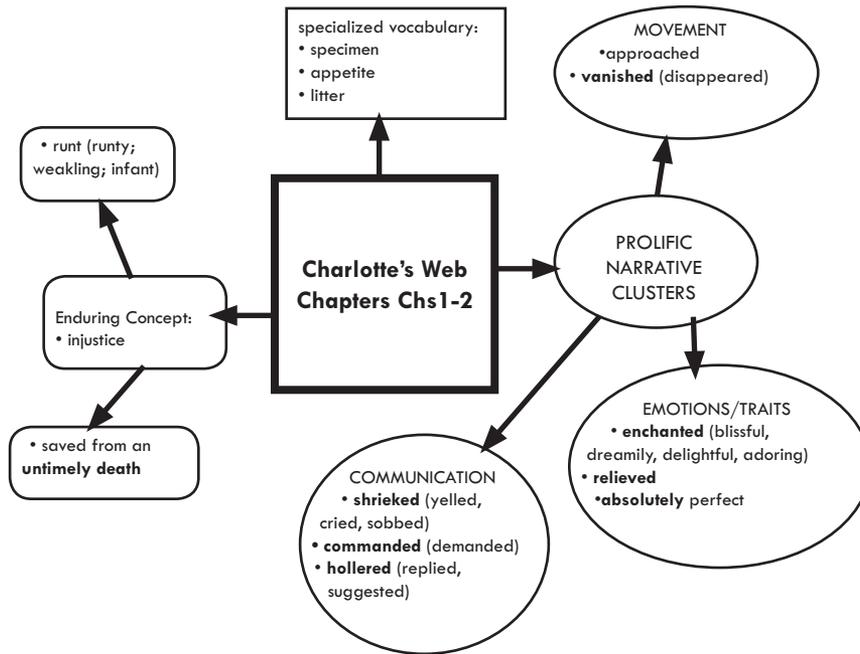
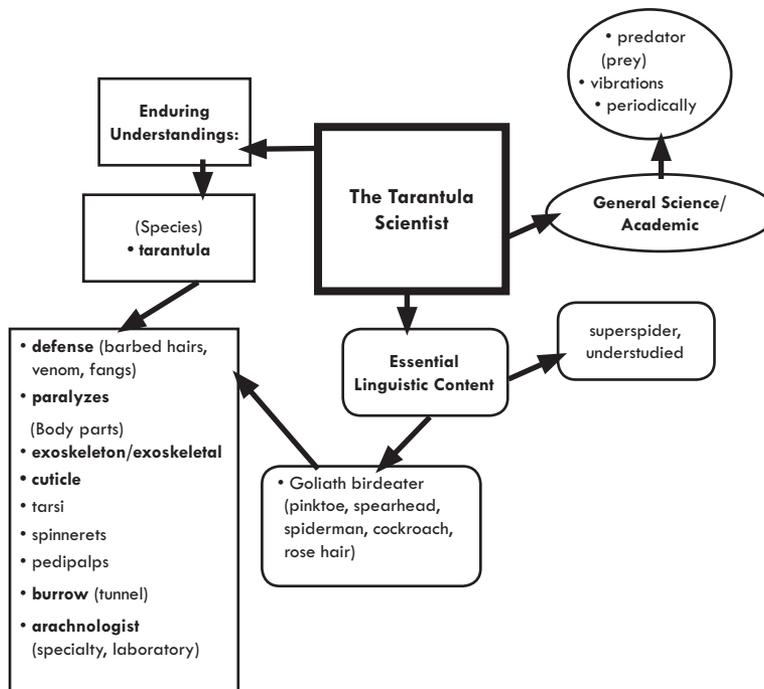


Figure 2.

Word Map for Chapters 1-2 of *The Tarantula Scientist*



Citations

¹Leech, G., Rayson, P., & Wilson, A. (2001). Word frequencies in written and spoken English based on The British National Corpus. London: Longman.

²Simpson, J., & Weiner, E. (2009). Oxford English Dictionary. New York, NY: Oxford University Press.

³http://www.parcconline.org/sites/parcc/files/PARCC_SampleItems_ELA-Literacy_Grade5Passage1_081513_Final_0.pdf



GENERATIVE VOCABULARY INSTRUCTION

ISBN-13: 978-0-328-81186-1
ISBN-10: 0-328-81186-6

