READING RESEARCH REPORT

#11.02
June 2011

Growing Capacity with the Vocabulary of English Language Arts Programs: Vocabulary Megaclusters

Elfrieda H. Hiebert
TextProject, Inc.

TextProject

textproject.org

RRR-11.02 V.1.01 JUNE 2011

Growing Capacity with the Vocabulary of English Language Arts Programs: Vocabulary Megaclusters

Elfrieda H. Hiebert
TextProject, Inc.

Abstract

The typical approach to teaching vocabulary in English/Language Arts programs has been to focus on six to eight words per text. Even though these words may add meaning to a particular story, the target words are often rare and their generalizability is limited. The Vocabulary Megaclusters provides a framework for selecting and teaching words according to their shared meaning and function in stories. Within each of the 13 Vocabulary Megaclusters, words are clustered according to conceptual connections such as emotions and traits of characters or the actions and motions in which characters engage. This article develops the rationale for the Vocabulary Megaclusters framework and illustrates how this approach provides students with a metacognitive stance that supports them in developing a rich literary vocabulary.

READING RESEARCH REPORT #11.02
June 2011

TextProject, Inc.
SANTA CRUZ, CALIFORNIA

Contents

Growing Capacity with the Vocabulary of English Language Arts Programs: Vocabulary Megaclusters

parlor	rumbling	massive	forecasts	dangle
draft	tropical	coward	inland	pollinate
frost	biologist	gleamed	expected	wondrous
terror	bluff	chorus	shatter	fragrant
fascinated	lagoon	shimmering	destruction	pollen
quaint		brilliant	surge	canopy
timid				dappled
etched				slithered

The words in the lists above are those identified for instruction within a unit of a widely used core reading program. For a five-week period in the school lives of fourth-graders, this group of 33 words will provide the focus of vocabulary instruction. Are these the most critical words in the text? Is the approximately 15 minutes of time devoted to each of these words over a five-week period the best use of the scarce instructional time in schools, especially for the students who depend on schools to overcome a huge gap in knowledge and vocabulary?

Over the past decade, I have raised questions about the typical words chosen for instruction within English Language Arts (ELA) programs, where the majority of elementary schoolchildren's formal vocabulary guidance occurs (Hiebert, 2005; Nagy & Hiebert, 2010). My primary concern has been that the processes for selecting these words lack any apparent theoretical framework. The words are picked on a story-by-story basis, not on larger units of themes or semantic or morphological relatedness. Even within a story, such as Van Allsburg's (1986) *The Stranger* (from which the first list of eight words come), reasons for the choices of particular words are not clear. In content areas, certain words are employed because of their contributions to a theme. For example, if the words of focus in a science text are ***precipitation*** and ***condensation,*** the reader is able to anticipate the content. Experts often claim that it is impossible to identify a set of words that coalesce within a narrative (Snow, Griffin, & Burns, 2005).

On closer examination, however, words exist within *The Stranger* (and the other texts represented above) that are more semantically related than ***parlor***

and ***draft*** and that would lead to a richer interaction with the text. Three of the words from the list do share a potential for connections to one another and to the theme of the text: ***terror, timid,*** and ***fascinated.*** All three describe aspects of the mysterious stranger who enters the life of the Bailey family: ***terror*** (the stranger's initial response when hit by the truck), ***timid*** (when meeting the other members of the Bailey family), and ***fascinated*** (his response to seeing steam rising off of food). The text contains other words that further describe responses of the stranger and that are likely to be less familiar but more relevant for vocabulary learning than words such as ***parlor*** and ***draft.*** The Baileys wonder if the stranger is a ***hermit.*** The stranger ***shyly tagged along.*** The weather becomes very ***peculiar.*** The stranger is ***hypnotized*** by a flock of geese heading south. His hand is ***trembling*** as he holds a green leaf from a tree and he becomes ***upset*** about the tree's leaves. Connections could be drawn among these words, as well as with some of the focus vocabulary (***terror, timid, fascinated, quaint***). By contrast, a concentration on ***parlor*** (a tangential reference to a place in the Bailey home) is likely to divert attention from the magnetic pull of the stranger to nature and to his peculiar traits.

A framework for categorizing the vocabularies of narratives is available, but it has been understudied by researchers and underused in pedagogical projects. This framework is the semantic cluster approach (Marzano & Marzano, 1988). I have revisited the underlying constructs of this approach and have refined it as the Vocabulary Megacluster approach. This article has three purposes: (a) an examination of why the Vocabulary Megacluster approach is essential for the development of the vocabulary of narratives, (b) an overview of the Vocabulary Megaclusters, and (c) an application of the Vocabulary Megacluster approach to the texts that were the source for the vocabulary that introduced this chapter.

Rationale for the Vocabulary Megacluster Approach

Relatively few words in English account for a majority of the total words that are read in text. In English, approximately 100 words account for almost 50% of the total words in text, and approximately 5,500 words account for 80% of the total words (Zeno, Ivens, Millard, & Duvvuri, 1995). Approximately 750,000 words (Leech, Rayson, & Wilson, 2001) account for the remaining 20% of the words in English texts. Most of the words within this last group appear less than once per one million (or even ten million) words.

The common view is that words in informational texts dominate this last group of rare words, not the words in narrative texts. While it is true that the vocabulary of informational texts challenges students, this vocabulary is challenging because it is conceptually complex, not necessarily because the words are rare. Narrative texts are actually more likely than informational texts to have a higher percentage of rare words. One of the reasons for this higher percentage is that the number of ***different*** rare words is higher in a narrative text. A rare word in an informational text will be repeated, as is the case in an article

on thermal energy with vocabulary such as ***convection*** and ***radiation.*** In a narrative about a stranger coming into a community, the author will likely use a variety of words to convey the reticence of the character (e.g., ***terror, fright, tremble***), rather than repeating the same word over and over. As a result, narrative texts are likely to have more unique rare words.

These features of the vocabulary of narrative texts require an instructional stance that recognizes these differences in kind and number of unique words. I bring to bear three sources of evidence to illustrate the distinctive nature of the vocabulary of narrative text from that of informational text: (a) a comparison of the words highlighted within standards documents for ELA and a content area (science), (b) the results of a study that compared the vocabulary identified for instruction in ELA and science programs, and (c) an analysis of the focus words that introduce this paper as well as of all the words in the texts from which these words came.

Words in standards documents

Evidence for a lack of theoretical or thematic purpose in the selection of words for ELA programs comes from a summary of the vocabulary in standards documents (Marzano, 2004). A list with ten terms from science and ELA vocabulary lists for four different grade bands appears in Table 1. Even with an alphabetic listing rather than a thematic one, it is evident that particular themes underlie the science vocabulary, such as weather in grades K–2 (e.g., ***weather pattern, precipitation, thermometer, weather conditions***). This vocabulary would be expected to appear in texts or materials that students read and use for inquiry.

The ELA vocabulary is quite different. Vocabulary is represented that fits into particular groups that cut across grade levels, for instance, parts of speech (e.g., ***adjective, common noun, relative pronoun***). This vocabulary is likely to be part

TABLE 1

Ten Terms in ELA and Science Standards Documents at Four Grade Bands

	ELA	Science
Level 1 (K–2)	alphabet, back cover, consonant blend, folktale, long vowel, number word, purpose, sight word, textbook, vowel combination	air, daily weather pattern, energy, insect, mixture, precipitation, salt water, states of matter, thermometer, weather conditions
Level 2 (3–5)	adjective, common noun, contraction, essay, inference, motive, object, regular verb, tone, word choice	acceleration, conductivity, electrical current, friction, light emission, mass, omnivore, pollution, reproduction, volcanic eruption
Level 3 (6–8)	adverb phrase, business letter, comparative adjective, dialect, figure of speech, historical fiction, jargon, metaphor, relative pronoun, verb phrase	asteroid, chemical element, eclipse, fungus, hydrosphere, lithosphere, muscular system, radiation, sunlight reflection, vertebrate
Level 4 (9–12)	acronym, censorship, denotative meaning, feature article, logographic system, mythology, past perfect verb tense, reflexive pronoun, structural analysis, visual text	biotic components of ecosystems, catalyst, electric potential, genetic mutation, meiosis, ohm, particle emission, radioactive dating, semiconductor, torque, weight of subatomic particles

of teachers' lessons or workbook exercises, but not in the narratives that comprise the core reading programs commonly used in ELA instruction. It is highly unlikely, for example, to find a narrative that uses any of the words that are listed as the ELA vocabulary for grades 3–5 such as ***contraction*** or ***inference.***

A comparison of vocabulary in ELA and content-area texts

A comparison of the words identified for instruction within the ELA and science textbooks, fourth-grade programs, of the same publishers also illustrates the qualities of literary vocabulary (Hiebert & Cervetti, 2011). For the ELA program, publishers had identified 209 words, seven from each selection, for the focus of instruction and assessment. The 207 focus words in the science program were distributed across 19 lessons, each with an average of 11 words. Hiebert and Cervetti established six features of the 416 words: a) length of words; (b) predicted frequency per one million words of text (Zeno et al., 1995); (c) morphological frequency: the predicted frequency per one million words of text of the words transparently related to the focus word (Zeno et al., 1995); (d) familiarity (Biemiller, 2008; Dale & O'Rourke, 1976); (e) dispersion, which indicates how widely a word appears in different subject areas (Zeno et al., 1995); and (f) conceptual complexity (Hiebert & Cervetti, 2011).

The features of words in narrative and informational texts were statistically different on all measures except for the frequency of morphological families of words and the dispersion index. On three of the remaining four features—length, familiarity, and conceptual complexity—the focus ELA words had averages that classed them as "easier" than the focus science words. On the fourth feature, frequency, the ELA words were deemed harder than the science words, with an average frequency of 14 occurrences per one million words of texts for the former and 39 for the latter. The target ELA words, then, were somewhat shorter, more familiar, and less conceptually complex than the target science vocabulary, but they were less frequent.

An analysis of all unique words in an ELA program

A view of the vocabulary demands of narrative texts comes from an analysis that identified the entire pool of words from which the focus words at the beginning of this essay came. The analysis began with all 6,410 words in the five texts of the focus unit (Afflerbach et al., 2007). Of these words, 1,204 were unique or distinctive words.

The frequency of these words was established through the WordZones Profiler (WZP; Hiebert, 2011) that draws on the Zeno et al. (1995) database. The eight word zones within the WZP differ according to the frequency with which words in a zone are predicted to occur in one million words of text. A small group of words (930) make up zones 0–2. These words occur at least 100 or more times per one million words of text. Approximately 4,900 words are in zones 3 and 4 where words are predicted to appear with moderate frequency (from 10 to 99 times per one million words). The approximately 8,240 words

TABLE 2

WordZone Distribution of Entire, Unique, and Target Vocabulary: ELA Program

WordZones	Total Words (n=6420)	Unique Words (n=1240)		33 Target Vocabulary Words	
		proportion	repetitions (X)	proportion	repetitions (X)
0–2	.81	.50	8.3	0	NA
3	.06	.15	2.1	.09	2.3
4	.09	.15	1.9	.27	1.4
5	.02	.13	2.0	.42	2.1
6	.01	.03	1.3	.15	1
7	.02	.04	2.6	.06	1

that make up zone five are relatively rare (2 to 9 appearances per one million words). A group of approximately 5,650 words occur about once per one million words and make up zone six. The remaining words fall in zone seven and occur less than once per one million (approximately 135,475 words in the Zeno et al. analysis of a 17.25-million-word corpus).

The distribution of the unique words in the five texts into the word zones appears in Table 2. It is the words of zones 5 and 6 that vocabulary instruction will likely emphasize since students are unlikely to have encountered these words previously in text. Words in zones 0–4 are ones which fourth-graders should know since these are words that are used frequently in their texts. Of the words chosen for instruction, approximately 63% were rare words. Most of these words appear once in the entire unit. Another 24% appeared 2–3 times. The remaining 10% appeared 4 times or more. Only 2% of this group appeared the requisite 10 times or more. The 33 words that introduced this essay represent approximately 5% of the rare/moderate words in the texts.

What can be concluded from this analysis is that there are numerous single-appearing words in narrative texts that appear infrequently in written language as a whole. It would be impossible for teachers to cover all of the words in lessons, even for a single text. Further, since the words can be exceedingly diverse in meaning, the instructional approach and task would be arduous. Finally, since almost all of these rare words appear infrequently in the text and are unlikely to reappear again in the texts that students are reading in other content areas, the longevity of students' learning—even of words that are taught intensively but appear a single time in a text—is uncertain.

A Description of the Vocabulary Megacluster Approach

Narratives have particular features, most notably a setting, problem, goal, action, outcome, resolution, and theme (Stein & Glenn, 1979; Whaley, 1981). The structure of narratives, as Bruner argued (1990), mirrors the way in which hu-

man beings describe their life experiences. Narratives are familiar structures and, typically, accessible to students. Within the instructional applications of story structure to the interpretation of texts, as Duke and Pearson (2002) showed, the structural aspects of text organization are emphasized rather than the concepts represented by these features of narrative. Students are taught to identify the particular components rather than to identify evidence within the text. The elements of stories were taught as structures rather than as concepts represented by ideas or words.

In a conceptual, rather than structural, approach to narratives, students learn to expect that characters will be involved in actions to deal with problems or conflicts. Authors use a variety of words to label the characters, the actions, the events, and the particular contexts in which these events occur. It is likely impossible to predict the words that an author will use in a narrative. For example, it is rare that a writer of a mystery will use the words ***suspense*** or even ***mystery.*** The words used by an author to describe a particular trait may vary from sentence to sentence but there are likely to be words that describe the traits and emotions of characters. These words can be clustered into categories that share particular meanings. Students can be taught to anticipate that authors will use words to describe the various components of the narrative. It is in this anticipation—or a meta-linguistic awareness—that vocabulary development can occur.

The basis for a categorization scheme for the primary concepts of narratives and expository content can be found in a 1998 publication by Marzano and Marzano. They presented a categorization of 7,230 words taken from a number of sources common in elementary school texts (e.g., Carroll, Davies, & Richman, 1971; Dahl, 1979; Harris & Jacobson, 1972). Marzano and Marzano ordered these words into 61 superclusters that were semantically related.

One of the most prolific superclusters that they identified was occupations of people. All 364 words that they assigned to that supercluster had to do with occupations of people, but those words were further categorized into clusters of words even more closely aligned. The supercluster of occupations contained 30 clusters, each pertaining to a different type of job such as people in sports, entertainers, and royalty/statesmen. Words within clusters were further organized into miniclusters. For example, within the entertainers cluster, miniclusters include actress, clown, and entertainer. Each of the miniclusters has at least a handful of words, sometimes more. The minicluster of clown consists of: clown, barker, magician, comic, and juggler. In all, Marzano and Marzano identified 430 clusters within the 61 superclusters and, within the clusters, 1,500 miniclusters where words have the strongest semantic ties.

The Marzano and Marzano (1988) clusters were published just when attention in reading education moved to the "whole text." Extensions of and experimentation with the clusters have been limited. These clusters, however, provide a means for much-needed support for the selection and instruction of vocabulary that Nagy and Hiebert (2010) have described. In particular, the cluster

approach may provide considerable guidance to publishers and curriculum developers regarding the many unique words in narrative texts that lack the thematic cohesiveness typically present in content-area texts and instruction.

For the cluster system to be useful to educators in selecting words for instruction, however, the system needs refinement. For one, the system needs to be able to integrate words from additional sources, such as the trade books that have become the basis for core reading programs. For such expansion, the superclusters themselves need to be defined and understood as conceptual sources of content. Another aspect of the superclusters that makes their use less than conceptual is the system's organization. Marzano and Marzano presented the superclusters in order of size (e.g., occupations first, types of motion next, and so on). The number of clusters—61—is also unwieldy.

To enable teachers and publishers to select the words to teach more efficiently, I have reconfigured the superclusters into 13 megaclusters, each of which represents a "big idea" about the content of texts. The development of the 13 megaclusters involved two steps: (a) eliminating and collapsing the superclusters and (b) identifying megaclusters from among the superclusters.

Eliminating and collapsing of superclusters

A first step was to eliminate a group of superclusters devoted to grammar: Pronouns, Contractions, and Auxiliary/helping verbs. The reason for this elimination is that the focus of the Vocabulary Megaclusters is on the conceptual content of words and the grammatical functions do not serve that purpose.

Subsequent changes to the remaining 58 superclusters are presented in Table 3. These changes involved collapsing several superclusters into related superclusters: (a) Health/disease was integrated into Human body, (b) an overall supercluster entitled Action was created from Helpful/destructive actions, Touching/grabbing actions, and Actions involving the legs, and (c) Noises/sounds and Facial expressions/actions were added to Communication, leaving 53 superclusters.

Identification of Vocabulary Megaclusters

The remaining 53 superclusters were examined with the aim of emphasizing particular components of narrative and expository texts. Many different perspectives could be brought to bear on the designations—just as is true with the original designations by Marzano and Marzano (1988). The final set of 13 megaclusters was derived from the primary components of narrative texts (Stein & Glenn, 1979; Whaley, 1981).

It should be noted that the integrity of the superclusters has been retained. Within the database, vocabulary can still be viewed in relation to superclusters, clusters, and miniclusters where the greatest similarity exists. For example, while Occupations, Types of people, and Types of groups form the megacluster of Characters, the data on the supercluster of Occupations continue to be

distinguished from the superclusters of Types of people and Types of groups. Researchers and educators working with the database can continue to identify the words in the closest possible grouping.

Expanding and validating the database

The database, initiated with the original words designated by superclusters, clusters, and miniclusters, has been extended to approximately 8,500 words. New words being added to the vocabulary corpus undergo a vetting process. First, the synonyms of a word are identified and a rater experienced in the semantic clustering procedures identifies a match to a particular minicluster, cluster, supercluster, and megacluster. Periodically, a second rater, also experienced in semantic clustering, independently categorizes the words that have been added to the database. When disagreements between raters arise, the nature of the disagreements and their resolution are recorded. After every 350 new additions to the database, a third experienced rater examines 20% of the additions as well as a randomly selected set of words that were part of the original database to ensure the fidelity of categorizations.

TABLE 3

Contents of Vocabulary Megaclusters and Adapted Superclusters

Vocabulary Megacluster	Superclusters (in Original Marzano & Marzano, 1988)	Changes from Original Superclusters
I. EMOTIONS & ATTITUDES	Feelings/emotion; Attitudinals	
II. COMMUNICATION	Communication; Mental actions; Senses/perceptions	(Communication subsumes Facial expressions/ actions & Noises/sounds)
III. TRAITS OF CHARACTERS	Nonemotional traits; Physical traits of people	
IV. SOCIAL RELATIONSHIPS	Ownership/possession; Popularity/knownness; Life/survival; Conformity/complexity	
V. CHARACTERS	Occupations; Types of people; Types of groups	
VI. ACTION & MOTION	Action; Motion	(Action subsumes Touching/grabbing, Actions involving legs, Helpful/destructive actions)
VII. HUMAN BODY	Human body; Clothing	(Human body subsumes Health/disease)
VIII. FEATURES OF EVENTS/THINGS/ PEOPLE	Value/correctness; Similarity/dissimilarity; Cleanliness/uncleanliness; Difficulty/danger; Causality	
IX. PLACES/EVENTS	Places where people live; Dwellings/shelter; Rooms/furnishings; Events	
X. PHYSICAL ATTRIBUTES OF THINGS/EVENT/EXPERIENCE	Size/quantity; Time; Location/direction; Shapes/dimensions; Texture/durability; Color	
XI. NATURAL ENVIRONMENT	Animals; Foods; Water/liquids; Land/terrain; Vegetation; Soil/metal/rock; Light; Weather; Mathematics; Temperature/fire; Chemicals; Electricity	
XII. MACHINES	Machines/engines/tools; Transportation; Materials	
XIII. SOCIAL SYSTEMS	Literature/writing; Money/finance; Sports/recreation; Language; Entertainment/arts	

TABLE 4

Distribution of Megaclusters Across Rare Words of an ELA Unit

Megacluster	Examples	Narrative
• Communication	summoned, shrieked	.11
• Emotions & Attitudes	anticipation, expected	.04
• Traits of Characters	daring, dignified	.02
• Social Relationships	peculiar, free	.03
• Characters (Occupations, People, Groups)	duke, magician	.06
• Action & Motion	dangled, swatted	.15
• Comparatives/Values	identical; useless	.04
• Body & Health	muscles, vaccine	.03
• Places/Dwellings	homeland; mansion	.04
• Physical Attributes	massive	.10
• Nature	precipitation, sedimentary	.26
• Machines	pulley, vehicle	.07
• Social Systems	sculptures, payroll	.05

To determine how well the Vocabulary Megaclusters accounted for the newly added vocabulary, words that had been identified as rare in Table 2 (i.e., the 149 that appeared in WordZones 5 and 6 in the ELA unit) were examined in relation to the Vocabulary Megaclusters database. The summary of this classification is given in Table 4. A Vocabulary Megacluster that is also common to informational texts—Nature—had the largest corpus of words. Since two of the five texts in the unit are magazine articles that have both narrative and informational elements, the appearance of words having to do with nature is understandable. As would be expected of narrative text, the Vocabulary Megaclusters of Communication and Action were also heavily represented.

Application of the Vocabulary Megaclusters

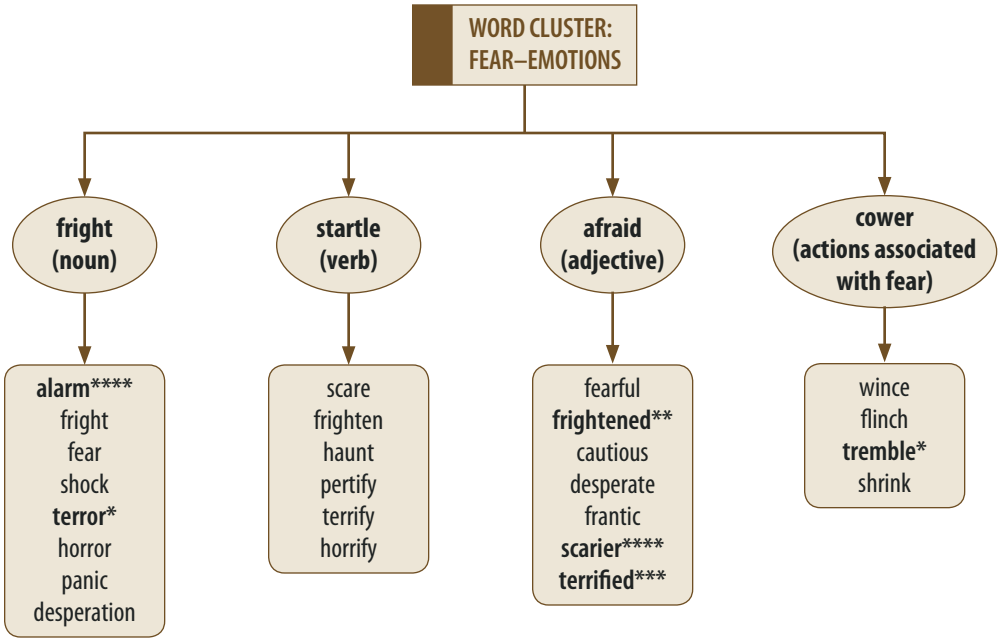
How might the information about Vocabulary Megaclusters be used by teachers? I offer two potential routes for a "conceptual" approach to narrative: (a) shared Vocabulary Megaclusters across a set of texts and (b) unique Vocabulary Megaclusters as a function of the author's use of language in a specific text.

Shared Vocabulary Megaclusters

Vocabulary that is typically critical in a narrative has to do with the ways in which characters communicate and with the characters' emotions and attitudes. Vocabulary associated with these critical aspects of narratives could be developed with particular clusterings of texts and also could be a focus of instruction in different grades. Emotions of fear, joy, and anger, for example,

FIGURE 1

Example of Vocabulary from All Texts Within a Unit Organized Around a Single Megacluster



* *The Stranger* (Van Allsburg, 1986)
** *Adelina's Whales* (Sobol, 2003)
*** *How Night Came from the Sea* (Gerson, 1994)
**** *Eyes of the Storm* (Kramer, 1997)

would be represented by larger and richer vocabularies through the grades, with the vocabulary of narratives read in earlier grades becoming the foundation for expanding and enriching the vocabulary related to a concept in later grades.

In the particular set of texts that formed the focus unit, the emotion of fear was present in four of the five texts. In *Adelina's Whales* (Sobol, 2003), Adelina's grandfather is frightened. In *The Stranger* (Van Allsburg, 1986), the stranger is filled with terror, and in *How Night Came From the Sea* (Gerson, 1994), the servants are terrified. In *Eyes of the Storm* (Kramer, 1997), Warren (the storm chaser) describes how the situation is getting scarier and scarier. The words used in the texts are the basis for a semantic map that appears in Figure 1. The number of synonyms and semantically related words for this concept is enormous. A recommendation that Nagy and Hiebert (2010) have made is for teachers to guide students' attention to a core set of words and then, gradually, to examine the semantic map. If too many words are introduced at once, students may have difficulty establishing the nuances of meanings.

Unique Vocabulary Megaclusters of individual texts

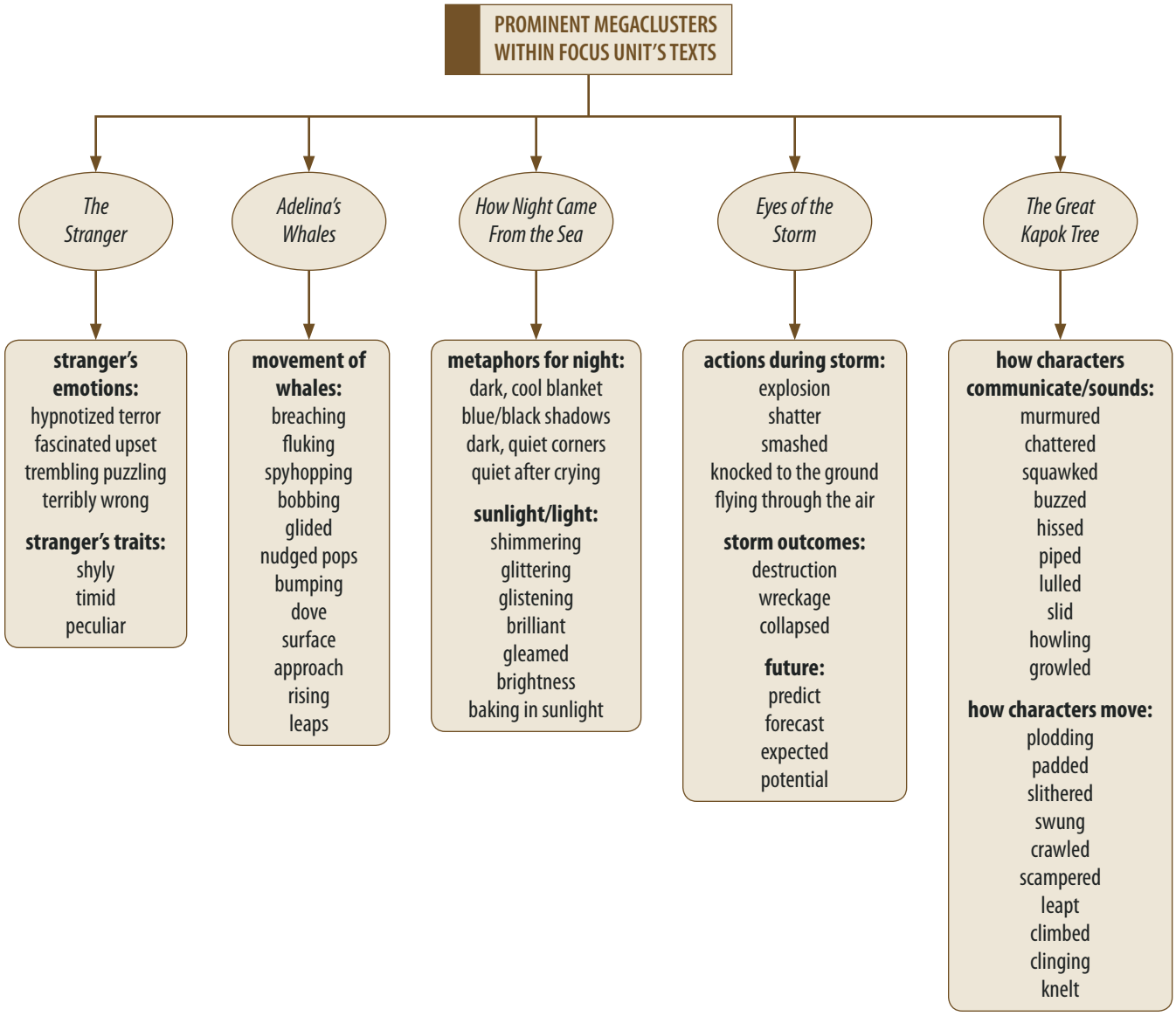
By examining the unique megaclusters represented in each of the five texts in this unit, I was able to identify possible emphases for the instruction of vo-

cabulary in each text. The uniqueness of each of the texts is represented in the semantic map in Figure 2.

In the two narratives—*How Night Came From the Sea* (Gerson, 1994) and *The Great Kapok Tree* (Cherry, 1990)—language is used richly and uniquely. *How Night Came From the Sea* has an array of adjectives that are used to describe the brightness of the light, which is new and jarring for the unnamed woman in the story. For the night—which represents the woman's previous experiences—the author uses numerous metaphors. A teacher could initiate an interesting conversation as to whether there are descriptions for darkness of the same variety as those for brightness, or whether metaphors are typically used to describe darkness and night.

FIGURE 2

Example of Selection of Megaclusters Based on Individual Texts Within a Unit



The two magazine articles—*Adelina's Whales* and *Eyes of the Storm*—use different words from those in the narratives. But there are, in addition, significant differences in the ways that the authors employ language. For example, the vocabulary of *Adelina's Whales* is fairly straightforward, as might be expected in a magazine article, with one exception: the actions of the whales (e.g., ***fluking, spyhopping, breaching***). The author's choice of these less common words allows for a discussion of how compound words are created (such as ***spyhopping***) and the nature of old Anglo-Saxon words (***fluke, breach***).

One could even imagine a set of texts chosen because they illustrate ways in which authors use language to express various contexts. London's (1906/2010) *White Fang* and Paulsen's (1987) *Hatchet* could be compared for their representations of adventure and danger. Erdich's (2002) *The Birchbark Tree* could be discussed with O'Dell's (1960) *Island of the Blue Dolphins* to understand how different authors might communicate despair, hope, and human ingenuity.

The current instructional approach of focusing on six to eight disparate words over a week does little to develop a strategic stance on the part of students. If students are to develop a deep understanding of vocabulary in literary texts, instruction needs to uncover the underlying uses of language in narratives. The Vocabulary Megaclusters provide a framework for teachers, publishers, and curriculum developers to select vocabulary and design instruction around critical concepts within narratives. By focusing on principles of language rather than only on the individual word, students gain the generative stance that is needed to engage in lifelong expansion of vocabulary.

References

Afflerbach, P., Blachowicz, C.L.Z., Boyd, C.D., Cheyney, W., Juel, C., Kame'enui, E.J., Leu, D.J., Paratore, J.R., Pearson, P.D., Sebasta, S.L., Simmons, D., Vaughn, S., Watts-Taffe, S., & Wixson, K.K. (2007). *Reading Street.* Glenview, IL: Scott Foresman.

Biemiller, A. (2008). *Words worth teaching.* Columbus, OH: SRA/McGraw-Hill.

Bruner, J. (1990). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Carroll, J., Davies, P., & Richman, B. (1971). *The American heritage word frequency book.* Boston, MA: Houghton Mifflin.

Cherry, L. (1990). *The Great Kapok Tree: A Tale of the Amazon Rain Forest.* New York, NY: Harcourt, Inc.

Dahl, H. (1979). *Word frequencies of spoken American English.* Essex, CT: Verbatim.

Dale, D., & O'Rourke, J. (1976). *The living word vocabulary.* Elgin, IL: Field Enterprises Educational Corporation.

Duke, N.K., & & Pearson, P.D. (2002). Effective practice for developing reading comprehension. In A.E. Farstrup & S.J. Samuels (Eds.), *What research has to say about reading instruction* (3rd Ed., pp. 205–260). Newark, DE: International Reading Association.

Erdich, L. (2002). *The Birchbark House.* New York, NY: Hyperion Press.

Gerson, M.J. 91994). *How night came from the sea.* New York, NY: Little Brown & Co.

Harris, A., & Jacobson, M. (1972). *Basic elementary reading vocabularies.* New York, NY: Macmillan.

Hiebert, E.H. (2011). *WordZones Profiler.* Santa Cruz, CA: TextProject.

Hiebert, E.H. (2005). In pursuit of an effective, efficient vocabulary curriculum for the elementary grades. In E.H. Hiebert & M. Kamil (Eds.), *The teaching and learning of vocabulary: Bringing scientific research to practice* (pp. 243–263). Mahwah, NJ: LEA.

Hiebert, E.H., & Cervetti, G.N. (2011). *What differences in narrative and informational texts mean for the learning and instruction of vocabulary* (Reading Research Report 11.01). Santa Cruz, CA: TextProject.

Kramer, S. (1997). *Eye of the storm.* New York, NY: Penguin Young Readers Group.

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

London, J. (1906/2010). *White Fang.* New York, NY: Simon & Brown.

Marzano, R. J. (2004). *Building background knowledge for academic achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.

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

Nagy, W.E., Hiebert, E.H., (2010). 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). New York, NY: Longman.

O'Dell, S. (1960). *Island of the blue dolphins.* Boston, MA: Houghton Mifflin.

Paulsen, G. (1987). *Hatchet.* New York, NY: Simon & Schuster Children's Publishing.

Snow, C.E., Griffin, P., & Burns, M.S. (Eds.) (2005). *Knowledge to support the teaching of reading: Preparing teachers for a changing world.* San Francisco, CA: Jossey-Bass.

Sobol, R. (2003). *Adelina's whales.* New York, NY: Dutton Children's Books.

Stein, N. L., & Glenn, C.G. (1979). *An analysis of story comprehension in elementary school children. In R.O. Freedle (Ed.), Advances in discourse processes* (Vol. 2, pp. 53–120). Norwood, NJ: Ablex.

Van Allsburg, C. (1986). *The stranger.* Boston, MA: Houghton Mifflin Books for Children.

Whaley, J.F. (1981). Readers' expectations for story structure. *Reading Research Quarterly, 17*(1), 90–114.

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

TextProject, Inc. is a non-profit public benefit corporation. Its aim is to bring beginning and struggling readers (of any age) to high levels of literacy through a variety of strategies and tools, particularly the texts used for reading instruction.

Find out more at **textproject.org**