Core Vocabulary and The Challenge of Complex Text

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

TextProject Article Series
July 2014

TextProject, Inc.
SANTA CRUZ, CALIFORNIA

This paper is an earlier version of: E.H. Hiebert (2013). Core vocabulary and the challenge of complex text. In S. Neuman & L. Gambrell (Eds.), *Quality Reading Instruction in the Age of Common Core Standards* (pp. 149-161). Newark, DE: IRA.

Core Vocabulary and the Challenge of Complex Text

Standard 10 is the feature that distinguishes the Common Core State Standards (CCSS; CCSS Initiative, 2010) from previous standards documents. In the past, a proficiency such as comparing and contrasting two grade-level texts might be given as a fifth-grade standard but what was meant by grade-level text was never defined. Not only is there a standard devoted specifically to students' ability to read increasingly more complex text but explicit guidelines are given in the form of quantitative indices in Appendix A and in illustrations of exemplar texts in Appendix B. An emphasis on increasing capacity with complex text makes perfect sense and we can only ask why it has been ignored for the past several decades.

But the topic has been ignored. Consequently, when the Common Core proposed a three-part system for establishing text complexity, there were few systems for evaluating two of the three evaluation components: qualitative and reader-task dimensions of texts. By contrast, there is a long history of quantitative systems that typically use some measure of sentence length and vocabulary difficulty (Klare, 1984). The Common Core writers drew on a current digital text difficulty system—the Lexile Framework—to propose a staircase of text complexity that begins with grade 2 and extends to college and career readiness (CCR). The staircase was designed to ensure that students' reading proficiencies increase across the school years to a level of CCR text by high school graduation.

The CCSS writers based this staircase on the assumption that there has been a dip in text complexity across all grades over the past 50 years. Data confirm this assumption in the middle-

and high-school grades (Hayes, Wolfer, & Wolfe, 1996; Williamson, 2008) but, as Hiebert and Mesmer (in press) have illustrated, the statement does not apply to texts in the primary grades. Indeed, analyses such as that of Foorman, Francis, Davidson, Harm, and Griffin (2004) have shown that the texts of the primary grades have accelerated in difficulty since *Becoming A Nation of Readers* (Anderson, Hiebert, Scott, & Wilkinson, 1985) called for an end to vocabulary control in reading textbooks.

Many educators question how their students can read even more complex text when a sizable portion of a grade-four cohort struggles with current texts which now fall below grade-level expectations (National Center for Education Statistics, 2011). The perspective developed in this chapter is that current performances can partly be explained by the small amounts of reading most American students do in school. Further, when students get "easier" texts, these texts are often simply shorter in length, not substantially different in the load of vocabulary. Finally, when texts are perceived as too hard, teachers frequently read texts aloud. When this occurs, students have even fewer opportunities to develop independence and proficiency in reading. The task confronting teachers is to support their students in reading more of the texts that are currently available in their classrooms, many of which are sufficiently complex for many students.

The aim of this chapter is to provide teachers with an understanding of why and how movement up the Common Core's staircase of text complexity begins with a focus on the current texts of the elementary school. To accomplish this aim, the chapter addresses three topics: (a) the distribution of vocabulary in texts, (b) students' performances with the core vocabulary, and (c) scaffolds that support students' reading of current texts.

Understanding How Vocabulary Functions In Texts

This section begins the distribution of vocabulary across texts in general and then demonstrates the consistency of the distribution with specific sets of texts.

The Distribution Of Vocabulary Within Texts: The 90-10 Phenomenon

Figure 1 shows the distribution of words in a sample of K-college schoolbooks (Zeno, Ivens, Millard, & Duvvuri, 1995), representing content areas taught in school. Approximately 930 words in this sample accounted for two-thirds of all words in texts. Adding another 4,700 words brings the total to 80% of total words. The remaining words occur much less frequently with those that occur less than once per million forming the largest group.

I have examined how many words in the most frequent 5,860 words share the same root word. Word families were defined as the root word and related words with simple derivational endings (e.g., *er, est, ly*), inflected endings, and possessives. Examples of simple word families are *help, helps, helping, helped,* and *helper* and *locate, locating, located, locates*. This process produced 4,000 word families among the 5,860 words (Hiebert, 2012). The label "simple" is used to distinguish this process from complex word families. For example, the complex word family for *help* would include *helpful* and *helpless*, while the family for *locate* would include *location* and *relocate/relocation.*

The next step was to examine the words in word zones five and six (see Figure 1) to identify members of the 4,000 simple word families. Adding rare members to the database brings the percentage of total words accounted for by the simple 4,000 word families to approximately 90% of most texts. Because of its central role in text, this set of 4,000 simple word families can be considered a "core vocabulary" (Hiebert, 2012).

The forms of words in these families may be simple but the words themselves are not necessarily simple to learn. The most frequent words (e.g., *the, a, of*) often have variant vowel

patterns and are quite abstract in meaning. But the core vocabulary includes many words beyond the high-utility words. Words such as *current* and *light* illustrate that many words in the 4,000 simple word families are frequent because they have multiple uses. They take on different parts of speech (e.g., *light* from the sun (noun), a *light* color (adjective)). Some have technical meanings (e.g., *current* in electricity) as well as everyday meanings (e.g., *current* fashion). Many are also used in compound words where the meanings are often idiosyncratic (e.g., *light-headed, light-weight*) and idiomatic phrases (e.g., *give the green light, see the light of day*).

Demonstrations Of The 90-10 Pattern In Authentic Texts

Verification that the 4,000 simple word families consistently account for the majority of words in texts comes from two analyses. The first considered at least 1,000 words (or all words in shorter texts such as *The Gettysburg Address)* from all of the text exemplars in Appendix B of the Common Core. The information in Table 1 shows that, while the percentage of core vocabulary is somewhat lower in high school texts and for informational texts than for elementary and narrative texts, the vast majority of the vocabulary of all sets of exemplar texts is accounted for by the core vocabulary.

The second analysis considered the distributions of core vocabulary *across* texts offered for students of different proficiency levels in one grade level of a core-reading program (Beck et al., 2008). This program, similar to most core reading programs has an anthology with primary and secondary selections for a week and five guided reading texts designated for readers of different proficiencies (advanced, on-level, below-level, English learner, and intervention). The data in Table 2 represent the average percentage for the texts offered for six weeks of instruction.

With one exception, the core vocabulary for texts accounts for 92-94% of all words. The only text that falls outside this range is the text for below-level students. At 89%, much of the

vocabulary continues to be from the core but approximately 11% of the words are rare. Many, although not all, of the rare words in the below-level texts aim to support decoding skills with monosyllabic words (e.g., *plank, daze*) or two-syllable word patterns (e.g., *hinder*, *indoor*). The percentage of rare words is high and the rare words are often not repeated, leading to the conclusion that the automaticity of below-level students with the core vocabulary will likely not increase if these texts are their primary learning experience. Texts intended for other levels may support their automaticity to a greater extent.

Students' Performances With The Core Vocabulary

The question of how students are doing with the core vocabulary is a challenging one to answer. Since the late 1980s when state policies demanded an end to controlled texts, even state assessments stopped using texts controlled for core vocabulary. There is an exception to this pattern—the DIBELS oral reading fluency (ORF) passages. DIBELS developers used the Spache readability formula to establish the difficulty of the ORF passages (Good & Kaminski, 2002). In developing a readability formula, Spache (1953) identified specific sets of high frequency words for particular grade levels and measured a text's readability against those words. As a result, the core vocabulary accounts for a high percentage of words in exit-level ORF passages: from 97% at the end of grade 2 to 92% at the end of grade six.

The use of the Spache in validating ORF passage difficulty means that DIBELS norms can be used to get a sense of how students do with core vocabulary. The text below fits the parameters of the end-of-year DIBELS ORF passage at grade 2.

> Once upon a time a woman was frying some pancakes. As she
> turned a cake in the pan, she said to her little boy, "If you were a
> little older, I would send you to the sawmill with some of these

> cakes for your father's/[10] dinner. But as it is, he must wait till supper for them. ” “Oh, do let me take them,” said the/[20] boy, whose name was Karl. “Do let me go. ” And he begged and/[30] begged, till at last his mother selected the brownest/[40] and crispest pancakes. She put them on a/[50] plate with a napkin over them and bade her son take them to the mill. (Lindsay, 1965)

The superscript following a slash represents the number of words read by a decile group. Not every student within a decile group would be expected to have read every single word up to that point but to have read 65 words (20th percentile) or 95 words (50th percentile) students needed to read a significant number of words, many of which are in the core vocabulary. For many students, this reading is often slow which means that their comprehension is compromised.

In sum, the vast majority of American students have fundamental recognition of the core vocabulary by the end of second grade and certainly by the end of third and fourth grades. As Stanovich (1986) suggested, a lack of automaticity means that students are less likely to read which decreases reading proficiency even more. To stop this cycle of the poorer getting poorer depends on instructional scaffolds that are carefully and intentionally enacted.

Scaffolds To Support Struggling Readers

The standards for the grade-bands of the Common Core suggest that there are ways to scaffold students in reading harder text:

> 10. By the end of the year, read and comprehend literature, including stories, dramas, and poetry, in the grades 4-5 text complexity band proficiently ***with scaffolding as needed at the high end of the range***

> 10. By the end of the year, read, and comprehend literature, including stories, dramas, and poetry, at the high end of the grades 4-5 text complexity band ***independently and proficiently.*** (italics and boldface added).

Scaffold, as defined in a standard dictionary, refers to a temporary structure that workers use to reach high places as they work on buildings. The parts of the standards that have been highlighted suggest that proficient reading can be supported in ways that support students' independent reading in subsequent grades.

Before attending to three scaffolds that can support this movement up the staircase of text complexity, the efficacy of one frequently used scaffold in classrooms merits attention: teachers' read-alouds of instructional text. In many classrooms, teacher read-alouds are no longer scaffolds but permanent parts of the instructional landscape. This practice of teacher read-alouds of instructional texts is distinct from teachers' read-alouds of high-quality texts, which has been found to enhance students' vocabulary and comprehension (Greene & Lynch-Brown, 2002). At the present time, however, there is no evidence that shows that teachers' read alouds of instructional/learning texts lead to increases in students' ability to read more fluently and proficiently in independent contexts.

Teacher read-alouds with student read-alongs of short but focused parts of text may be one of the tools in a teacher's toolkit, but it cannot be the primary one. There is research support for three alternate types of scaffolds.

Responsibility

Almost 30 years ago, Pearson and Gallagher (1983) described the critical role of gradual release of responsibility for text guidance from teacher to students. Pearson and Gallagher's (1983) cycle can also be applied to tasks with texts. Through intentional instructional steps, a

teacher moves students from depending on the teacher for supervision and guidance in a reading task to independent responsibility for the reading task. The goal of reading instruction is to ensure that students read texts independently, not for students to expect their teachers or audio devices will do the reading for them when texts are challenging.

Teachers need to peruse texts to answer the question, "Is this a text that this group/class of students should be able to read?" If the answer is yes, the teacher next looks at the features that might need an overview prior to students' reading. The prior reading support should not exceed the time-spent reading but an anticipatory set for reading can be created with a few, well-chosen comments. Next, the teacher might provide an overview such as the following:

> "One of your goals as a third-grader is to be able to read texts that have words that you might not know. When you read texts where you learn new words and ideas, you are growing as a reader. This text might look like it is hard and it may even be on the first read. But I've studied the text and you know most of the words. We are going to do our word warm-up where we identify new words. I'll also be available to talk with you about words you don't know after you've read the first three pages by yourself. What is important is for you to read to learn the message of the text. After you've read the first three pages by yourself, we'll talk about new words and also about what the text tells you about what school was like in a time long ago. "

The DIBELS data suggest that American students, even those in the bottom quartile, can recognize many words. To develop a sense of agency and resilience as readers, students need to take responsibility for appropriate texts. The texts in Table 2 average 7% rare words. Research is unclear as to precise percentages of rare words students can read in scaffolded and unscaffolded

settings. However, as the description of the next scaffold—vocabulary—indicates, part of instructional support comes from teachers' review of potentially challenging vocabulary with students.

Vocabulary

There are numerous reasons why vocabulary receives substantial attention in a Common Core classroom. First, while not the only dimension that makes texts complex, vocabulary knowledge is closely tied to comprehension (Just & Carpenter, 1987). Unlike syntax where research has been ambiguous about how it can be strengthened through instruction, there is a rich research literature on ways of supporting students' vocabulary. There are two levels to this instruction: (a) strengthening vocabulary knowledge overall and (b) addressing the vocabulary of specific texts.

Teaching students about morphological word families is one way to expand students' vocabularies (Carlisle, 2010). Students' vocabularies can also be strengthened by understanding distinctions between unique words (i.e., those not in the 4,000 simple word families) in narrative and informational texts (Hiebert & Cervetti, 2012). The words in narrative texts typically fit into clusters with similar meanings (e.g., *lackadaisical, apathetic*), while the unique vocabulary in informational text belong to clusters where words have distinctive but conceptually interrelated meanings (e.g., *acidic, abrasive, alkaline* are all properties of substances).

Students also need to be guided in the specific words, which may be unknown in instructional texts. Figure 2 is a summary of the infrequent words from a week's worth of guided reading texts (included in the summary in Table 2). Very short lessons can be conducted on particular groups of words (e.g., dealing with proper names). As important, students need to be

aware that the number of new words is finite and that they can figure out almost all but a small percentage of the new words with their already-acquired decoding and root word strategies.

Volume

In a study of the amount of time that students spent reading in schools part of the No Child Left Behind initiative, Brenner, Hiebert, and Tompkins (2009) found that the amount of time devoted to reading *instruction* doubled in most classrooms. The time spent in reading *practice* increased by 15%. Much of the increased instructional time was devoted to teachers' talking about reading. True, some of this talk was discussion about what had been read but students read for less than 20% of the reading period.

The amount that students read in their classrooms needs to be a structural change in American classrooms. Viewed from that perspective, the component is not a temporary structure. However, scaffolds are needed to initiate substantial increases in independent reading time. Increasing the amount of independent reading that students are doing does not happen in one flail swoop. The increase of reading at any point in time needs to occur intentionally and be consistently monitored.

The information in Table 2 shows that the distinction between many instructional texts is in the length of texts, not the percentage of core vocabulary. For below-level students, 400 fewer words in a weekly text compared the text read by higher performing peers may not seem great but over a school year that difference is substantial. Almost 35 years ago, Allington (1977) asked the question "if they don't read much how they ever gonna get good?" Struggling readers, in particular, need to read substantially more text in school. Since the guided reading texts in a unit typically cluster around a theme, the below-level students can read several texts (e.g., below-

level, intervention, EL). By asking below-level students to read three texts rather than one, their opportunities to read over the school year are more than double.

Some may argue that repeated reading of single texts will accomplish the same goal. However, in the landmark study of fluency conducted by Kuhn and Schwanenflugel (2009), it was wide reading, which was defined as reading several texts, not simply repeated reading of single texts that proved critical in increasing students' fluency. To ensure that students "get good," students need to read widely and deeply. Wide reading means reading more rather than fewer texts. Deep reading means that students reread portions of these texts for connections and extensions of their knowledge.

Summary

A frequent response of teachers to hearing the message of "complex text" is that this is impossible in their classrooms since their students are far below grade level. The content of this chapter shows that, across all grade levels, the core vocabulary accounts for the majority of the words in texts. By the end of second grade, almost all students can recognize the core vocabulary but a significant portion of a grade cohort are not very automatic with this vocabulary. Students need to read widely and deeply to become automatic with this vocabulary. When teachers over-use the scaffold of reading texts aloud for students, students' automaticity with the core vocabulary is not aided. Teachers' toolboxes of appropriate scaffolds need to expand to conversations where students take responsibility for the texts themselves, where strategies for unique vocabulary are developed, and where the amount that students read during school time increases substantially. Just like knowledge is the center of the Common Core for students, so too teachers' knowledge about how language works can go a long way to providing the foundation for practices that increase students' capacity with complex text.

Discussion Questions

1. Vocabulary is a large component of reading comprehension. Given vocabulary's importance within texts, explain vocabulary distribution and the 90-10 phenomenon. Why is this concept vital in students' comprehension of complex texts?
2. Hiebert spends much time discussing the importance of teachers' read-alouds of high-quality texts. Why should teachers read aloud high-quality texts rather than instructional texts?
3. How do the length of the text selected to read and reading time impact students' success with texts? How can you scaffold your students to spend more time reading in your classroom?

Try This!

Try This!

Before reading a chapter about microorganisms in a life sciences class, have students brainstorm and then use list-group-label to further enhance their knowledge of key ideas. First, identify a key concept that reflects one of the main topics to be studied in the text. Then, have students work in small groups to generate a list of words related to the concept in 60 seconds. List words from each group on the board, then have the class form learning teams to group the words into logical arrangements. Then, have the teams label each arrangement. Finally, ask the students to make predictions about what is to be studied.

Try This!

Use a modified Cloze passage to reinforce technical vocabulary. Choose a 200- to 500-word text segment that represents one of the most important parts of the reading assignment. Students can supply the missing words either before or after reading the entire assignment. Use discussion to build meaning for key terms and to raise expectations for the entire assignment if students work on the cloze activity before reading. If you assign the cloze passage after reading, the passage will reinforce concepts attained through reading.

References

Allington, R. L. (1977). If they don't read much, how they ever gonna get good? *Journal of Reading, 21,* 57-61.

Anderson, R. C., Hiebert, E. H. , Scott, J. A. , & Wilkinson, I. A. G. (1985). *Becoming a Nation of Readers: The Report of the Commission on Reading.* Champaign, IL: The Center for the Study of Reading, National Institute of Education, National Academy of Education.

Beck, I. L., Farr, R. C., Strickland, D. S., Ada, A. F., Hudson, R. F., McKeown, M. G., Scarcella, R. C., & Washington, J. A. (2008). *Storytown*. Orlando, FL: Harcourt School Publishers.

Brenner, D., Hiebert, E. H., & Tompkins, R. , (2009). How much and what are third graders reading? E. H. Hiebert (Ed.), *Reading more, reading better* (pp. 118-140). New York, NY: Guilford.

Carlisle, J. F. (2010). An integrative review of the effects of instruction in morphological awareness on literacy achievement. *Reading Research Quarterly*, *45*(4), 464-487.

Common Core State Standards Initiative. (2010). *Common Core State Standards for English Language Arts & Literacy in History/Social Studies, Science, and Technical Subjects*. Washington, DC: CCSSO & National Governors Association. Common_Core_State_ Standards_Supporting_Districts_and_Teachers_with_Text_Complexity. html

Foorman, B. R., Francis, D. J., Davidson, K. C., Harm, M. W., & Griffin, J. (2004). Variability in text features in six grade 1 basal reading programs. *Scientific Studies of Reading*, *8* (2), 167-197.

Good, R. H., & Kaminski, R. A. (2002). *DIBELS Oral Reading Fluency Passages for First through Third Grade*s (Technical Report No. 10). Eugene, OR: University of Oregon.

Good, R. H., Wallin, J. U., Simmons, D. C., Kaem'enui, E. J., & Kaminski, R. A. (2002).

System-wide percentile norms for DIBELS benchmark assessment (Technical Report No. 9). Eugene, OR: University of Oregon.

Greene, E. B., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology, 94*, 465-473.

Hayes, D. P., Wolfer, L. T., & Wolfe, M. F. (1996). Schoolbook simplification and its relation to the decline in SAT-Verbal Scores. *American Educational Research Journal*, *33*, 489-508.

Hiebert, E. H. (2012). Core vocabulary: The foundation for successful reading of complex text (Text Matters 12. 2). Santa Cruz, CA: TextProject. Retrieved from: *http://textproject. org/teachers/text-matters/core-vocabulary/*

Hiebert, E. H., & Cervetti, G. N. (2012). What Differences in Narrative and Informational Texts Mean for the Learning and Instruction of Vocabulary. In J. Baumann and E. Kame'enui (Eds.), *Vocabulary Instruction: Research to Practice (2nd Ed.)*. (pp. 322-344). New York, NY: Guilford Publications.

Hiebert, E. H., & Mesmer, H. A. (in press). Upping the ante of text complexity in the Common Core State Standards: Examining its potential impact on young readers. *Educational Researcher*.

Klare, G. (1984). Readability. In P. D. Pearson, R. Barr, M. L. Kamil, & P. Mosenthal (Eds.), Handbook of reading research (pp. 681-744). New York, NY: Longman.

Kuhn, M. R., & Schwanenflugel, P. J. (Eds.). (2008). Fluency in the classroom. New York, NY: Guilford Press.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston, MA: Allyn & Bacon.

Lindsay, M. (1965). The plate of pancakes. In H. M. Robinson, M. Monroe, A. S. Artley, C. S. Huck, & W. A. Jenkins (Eds.), *More roads to follow*. Chicago, IL: Scott, Foresman and Company.

National Center for Education Statistics (2011). *The Nation's Report Card: Reading 2*011 (NCES 2011–457). Institute of Education Sciences, U. S. Department of Education, Washington, DC.

Pearson, P. D., & Gallagher, M. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, *8,* 317-345.

Spache, G. (1953). A new readability formula for primary-grade reading materials. *The Elementary School Journal, 53,* 410–413.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.

Williamson, G. L. (2008). A text readability continuum for postsecondary readiness. *Journal of Advanced Academics*, *19*, 602-632.

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

Table 1. *Core Vocabulary Distributions for CCSS Exemplars (Appendix B)*

Grade	Narrative	Informational
2-3	. 93	. 92
4-5	. 92	. 91
6-8	. 93	. 87
9-10	. 89	. 91
11-CCR	. 89	. 87

Table 2. *Texts Offered for Different Readers: Six Units/Weeks of Third Grade*

	% Core Vocabulary	WORD COUNT
ADVANCED	93	1010
ON-LEVEL	94	891
BELOW	89	541
EL	94	479
INTERVENTION	93	287
ANTHOLOGY (Primary)	92	1049
ANTHOLOGY (Secondary)	94	153

Figure 1. Distribution of Words in Written English

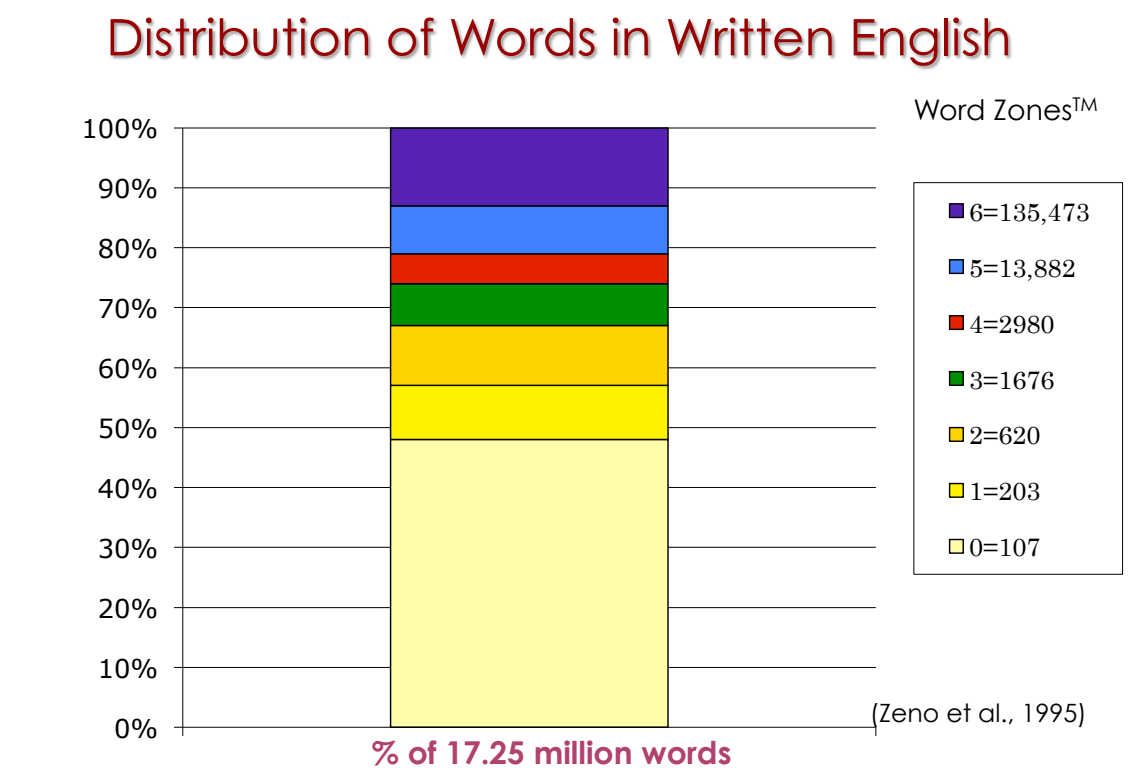

Figure 2. Attending to Infrequent Words & Strategies for Addressing These Words