Aspects of Text Complexity
Vocabulary Research Base
David Liben
March 6, 2010

Introduction

This research base supports the March public release of the Common Core Standards vocabulary section. As that draft moves into circulation, it will be helpful to insure that this research base reflects the best current thinking about how students learn words and networks of words.

In particular, we need to know about research and work we may be overlooking.

Widely accepted findings regarding vocabulary in general:

- Vocabulary has been empirically connected to reading comprehension since 1925 (Whipple 1925).
- The most succinct summary comes from the National Reading Panel's review of hundreds of studies; “Vocabulary instruction leads to gains in comprehension” (NRP 2000), see also Reading Next (Snow 2002).
- Vocabulary needs to be taught and developed from text, as the frequency of rare words in even educated adult conversation is 17.3 per 1000 words, and the ratio in children’s books is 30.9 per 1000. (Hiebert 2009).
- Informational (non-narrative) text contains more rare words than narrative, yet constitutes only 7-15% of instructional material in elementary and middle school (Hoffman et al 1994, Moss and Newton 2002, Yopp and Yopp 2006).

The reason undergirding this last bullet is that informational text contains large quantities of Tier Three words (Beck, McKeown, and Kucan, 2002), which almost by definition are less common.

Accepted by most researchers:

- Vocabulary deficiency is one of the primary causes of the achievement gap (Becker 1977, Baumann & Kameenui, 1991, Stanovich 1986).
- Vocabulary instruction is neither frequent nor systematic in most schools (Durkin 1979, Scott and Nagey 1997, Biemiller 2001).

Awareness of this as a problem is now becoming widespread, which may explain the urgent call for vocabulary to be included on its own in the Common Core Standards.

What is new but becoming more widely accepted:


What we know about successful vocabulary instruction:

• It combines direct systematic instruction of words as well as teaching students how best to learn word meanings from the context of their reading (Snow 2007, Blachowicz et al 2006, NRP 2000, Stahl 1999, Moats 2005 and numerous others).
• The question of what constitutes academic vocabulary does not get the attention it deserves. Not only tier three words, “academic vocabulary” includes a type of tier two words more likely to appear in academic texts (e.g. relative, vary, formulate, specificity, accumulate) and cause more problems than tier three words. Tier three words are introduced by the teacher, often explicitly defined in the text, included in a glossary, and repeated a number of times in each chapter. None of these supports are provided for tier two words.
• Students need to think actively about words; how and why specific words are used, how changing one word can change the meaning or tone of a text, how one word can have varied but related meanings depending on context and why another word might be more appropriate.
• The meaning of words varies with the context, thus the more varied the context provided to teach the meaning of a word the more effective (Texas was admitted to the union, he admitted his errors, admission was too expensive…). Both differences and relations need to be explored.
• Dictionaries alone don’t work (Graves 2006 provides complete explanation, also Beck et al 2002, Stahl 1999).

These findings make effective teaching of vocabulary from context difficult. An ELA teacher choosing to teach vocabulary using the literature she is working with (as opposed to a workbook or a vocabulary program) would need to do the following:

• Find words that most of the students are not likely to know.
• From this group choose the words essential to understanding the text.
• Other words, not essential to understanding the specific text, may be important words for students to know. These include words that carry a lot of weight and words that are likely to appear in academic or literary texts students will need to read in future grades (Hiebert 2009). From these, words that are part of a semantic group (e.g. social, socially, socialite, society) should be given priority (Hiebert 2009).

Once this list is assembled, the words need to be taught. If following the accepted research noted above, instruction would require the following:

• User friendly definitions, though not the bulk of the lessons, are provided.
• Words are taught and reviewed in varied contexts: the more the better.
• Activities that require thinking about the word are used instead of rote memorization.
• The figure for the number of repetitions needed before a word is learned varies depending on which study is consulted. These numbers include students’ previous exposure to the word. The lowest number of repetitions we have found is six. Assuming that students may have averaged one previous exposure to each new word, five additional repetitions seems a reasonable minimum. But to solidify learning, many students will need more repetitions and exposures.
• Distributed practice has long been shown to be superior to more concentrated practice so all of the above would need to be paced over a week or more.
• Finally, if we accept the work of Beck, Perfetti and McKeown (1982) and Stahl and Nagy (2006), morphology and etymology should be folded into instruction.

Very few teachers are able to do this systematically, and past elementary school, the number decreases. High school and middle school Language Arts teachers are required to teach literature, writing, and grammar. Where is the time for this needed attention to vocabulary?
Teachers need tools – not programs – that make this feasible, as addressed very briefly at the end of this document. Additionally, standards pointing to the right priorities and toward the right practices will help teachers focus their energies for maximal student benefit.

Until fairly recently, a major debate centered on direct instruction of vocabulary versus learning from context through wide reading. Advocates for context argued that the number of words students need to learn would be impossible to teach directly within the school day/week/year, and thus the emphasis should be on wide reading. (This was the origin of the “25 book standard” prominent in the mid 90’s and still present in many states’ standards.) Direct instruction advocates argued that word awareness should be cultivated through systematic study of morphology and word lists.

The debate ended with a body of work (Nagy et al 1985, Cunnigham and Stanovich 1988, Baker et al 1995, Beck et al 2002, Biemiller 2004, Moats 2005)) showing that not all words are created equal. Some words and classes of words appear far more frequently in student texts; more important, learning these particular words makes students more likely to learn words independently from context. This effect is strengthened if instructional methods are research-based (Nagy and Hiebert 2007, Moats 2005, Beck et al 2002, Snow 2007). Thus teaching words directly enhances independent learning from context, and both are necessary.

That being said, another implication for vocabulary acquisition is that student texts should increase steadily in complexity over time; eighth graders might expand their vocabulary somewhat by reading a 5th grade text, but far more growth will occur with texts written at an 8th grade or higher level of complexity.

Nietzsche wasn’t thinking of modern American education when he wrote of eternal recurrence, though he could have been - vocabulary’s importance just may be making a comeback.
Bibliography for Vocabulary Piece


Blachowicz et al. (2006, October-December). Vocabulary: Questions from the classroom. *Reading Research Quarterly* Vol. 41, No. 4. **City of publication?**


Nagy, W., Anderson, R. C., & Herman, P. (1986). The influence of word and text properties on learning from context (No. 369). Is this an article within a book or journal? Name of book or journal? City of publication?


National Reading Panel’s review of hundreds of studies. Vocabulary instruction leads to gains in comprehension. Rockville, Maryland: National Reading Panel. of Reading. Date of publication? Where does “of Reading” go or should I just delete it?

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. Scientific Studies. 11(4), 357-383. Should Scientific Studies have of Reading added to it (see next entry)? City of publication?


