SECTION C: *Excerpt from:* ***Advancing Our Students' Language and Literacy: The Challenge of Complex Texts***

Marilyn Jager Adams is a research professor in the Cognitive, Linguistic, and Psychological Sciences Department of Brown University and former chief scientist for Soliloquy Learning Inc. She is the author of numerous scholarly papers and several books, including two landmark works: Beginning to Read: Thinking and Learning about Print *and* Phonemic Awareness in Young Children*. This article is adapted with permission of the Guilford Press from "The Challenge of Advanced Texts: The Interdependence of Reading and Learning," which Adams wrote for* Reading More, Reading Better, *edited by Elfrieda H. Hiebert, copyright 2009 by Guilford Press.*

Insights from a Computer Model of Vocabulary Acquisition

For another way to think about vocabulary acquisition, let's con- sider an intriguing computer model called Latent Semantic Analysis (LSA) that was developed by Tom Landauer and his colleagues.[31] The core mechanism underlying the LSA model is "associative learning." When a text is input into the LSA model, the computer builds an association between each individual word of the text and the total set of words—that is, the context—in which the word has appeared. Where a word shows up in multiple contexts, the strength of the association between the word and each of the separate contexts is weakened through competition. Where a word arises repeatedly in one particular context, the association between the two is strengthened.

Importantly, the associations between words and contexts in the LSA model are bidirectional. That is, there are links from each word to each of its contexts and also from each context to all of its words. As a result, the full complex of knowledge that is called forth as each word is "read" extends through its contexts to other words, and through those words to other contexts and words. Thus, as the model "reads" the next word of the text and the next and the next, activation spreads to other, related complexes of knowledge, which may well include clusters that have never before been directly represented by any combination of words and contexts the model has ever "read" before.

Moreover, because the model's knowledge is represented relationally, the addition or modification of any one connection impacts many others, pulling some closer together, pushing some further apart, and otherwise altering the strengths and patterns of connections among words and contexts. Through this dynamic, reading causes the connections that collectively capture LSA's knowledge of words to grow, shrink, and shift continuously, continually, and always in relation to one another.

In short, the model's response to any text it "reads" extends well beyond what is denoted by the specific words of the text. Further, the richness of the model's representation of any text that it "reads" depends on how much it already knows. Just as with people,[32] the larger its starting vocabulary and the more it has read before, the more it will learn and understand from the next text.

In comparing LSA's word-learning to that of schoolchildren, the researchers began by "training" it with a set of texts judged comparable to the lifelong learning of a typical seventh-grader. The researchers then gave the model new texts to "read" and measured its vocabulary growth. The results showed that the likelihood that the computer gained adequate understanding of new words it encountered in these new texts was 0.05—just exactly the same as researchers have found for schoolchildren.[33]

But the results showed something else, too. It turned out that, with each new reading, the model effectively increased its under- standing not just of words that were in the text but also of words that were *not* in the text. Indeed, measured in terms of total vocabulary gain, the amount the model learned about words that did *not* appear in a given reading was three times as much as what it learned about words that *were* in the reading.

"What?" we cry, "How can that be? How can reading a text produce increases in knowledge of words that it does not even contain? That is not credible! It makes no sense!" But wait. If we were talking about knowledge rather than words, then it would make lots of sense. Every concept—simple or complex, concrete or abstract—is learned in terms of its similarities, differences, and relationships with other concepts with which we are familiar. As a simplistic example, when we read about tigers, then, by dint of both similarities and contrasts, we learn more about all sorts of cats and, further, about every subtopic mentioned along the way. The more deeply we read about tigers, the more nuanced and complex these concepts and their interrelations become.

As explained earlier, it was to be expected that LSA's full response to any new text would spread beyond the content of the text itself. The unexpected discovery was that this dynamic would impact the model's understanding of individual words. Given that words are really nothing more than labels for interrelated bundles of knowledge, perhaps this should not have been surprising.

In the study that modeled a seventh-grader, the researchers were able to gauge LSA's overall vocabulary growth by computationally examining changes in the representation of every word to which it had ever been exposed. Yet here is a mull-worthy correlate: unavoidably, the bundles of concepts and relations that emerged or were strengthened through LSA's reading experience included many that pertained to words that the model had never seen before. An analogous effect might explain why researchers have found time and again that the strength of students' vocabulary predicts the likelihood that they will learn new words from context,[34] the probability that they will correctly infer a new word's meaning from context,[35] and both the amount and nature of their reasoning when they are asked to explain how they do so.[36] Even when students are *told* the meaning of a new word, their prior vocabulary strength predicts the likelihood that they will retain it.[37] (These are known as "Matthew effects," referring to the notion that the rich get richer and the poor get poorer.) As the reader's linguistic and conceptual knowledge grows in rich- ness and complexity, it will increasingly support the meanings of many new words and the representation of many new spheres of knowledge.

Cognitive psychologists broadly agree that the meaning of any word consists of bundles of features and associations that are the cumulative product of the reader's experience with both the word in context and the concepts to which it refers. What is unique about the LSA model is its demonstration that this structure and dynamic can so richly and powerfully evolve through accrued experience with the various contexts in which words do and do not occur—that is, sheerly through reading.

Another way to state the larger point here is that words are not just words. They are the nexus—the interface—between communication and thought. When we read, it is through words that we build, refine, and modify our knowledge. What makes vocabulary valuable and important is not the words themselves so much as the understandings they afford. The reason we need to know the meanings of words is that they point to the knowledge from which we are to construct, interpret, and reflect on the meaning of the text. A core implication of the LSA model is that students' knowledge of words grows less through any process of inferring their meanings, one by one, based on the sentences in which they arise, than as a product of learning more generally about the contexts in which they arise and of understanding the concepts and relationships to which they refer.

Knowledge, Cognitive Strategies, and Inferences

If reading results in so rich a network of knowledge through nothing more than overlaps and contrasts in associations, then shouldn't students learn far more efficiently, given active, incisive inference and comprehension strategies? Research indicates that such strategies can be taught and suggests that doing so may improve comprehension.[38] However, inference and comprehension strategies seem to do little to compensate for weak domain knowledge.[39] Instead, research repeatedly shows prior domain knowledge to be a far stronger predictor of students' ability to comprehend or to learn from advanced texts.[40] Of course, students' comprehension and learning is also influenced by their reading skills (such as decoding and fluency). But even the advantage of strong reading skills turns out to be greatest for students with strong domain knowledge.[41]

Again, such findings should not be surprising. Cognitive research affirms that there are two modes of reasoning.[42] The first, most common mode is knowledge-based. This sort of reasoning is rapid, extensive, and automatic. This is the sort of reasoning that ensures, for example, that we properly understand the meaning of *fan* depending on whether the text is about a *soccer fan*, a *ceiling fan*, or a *peacock's fan*. This is the sort of reasoning that computer models such as LSA statistically emulate.

The second mode of reasoning is conscious and rule-based. Such logical, analytic thought also warrants instructional attention in our schools, as it is our means of deliberately evaluating and vetting our thoughts for bias, happenstance, and inconsistencies. However, no reasoning strategy, however well-structured, can rival the speed, power, or clarity of knowledge-driven under- standing;[43] nor can it compensate for an absence of sufficient information.

There may one day be modes and methods of information delivery that are as efficient and powerful as text, but for now there is no contest. To grow, our students must read lots. More specifically, they must read lots of "complex" texts—texts that offer them new language, new knowledge, and new modes of thought. Beyond the basics, as E. D. Hirsch, Jr., the founder of Core Knowledge, has so forcefully argued, the reading deficit is integrally tied to a knowledge deficit.[44]