VOCABULARY ASSESSMENT METHODS

Sulaymonova Sojida
Norpulatova Shaxnoza

After a nearly 15-year absence from center stage, vocabulary has returned to a promi-nent place in discussions of reading, and it is alive and well in reading instructionand reading research. We have no doubt that the renaissance is due, at least in part,to the salutary findings about vocabulary in the report of the National ReadingPanel (NRP; National Institute of Child Health and Human Development[NICHD], 2000) and, even more important, the use of the NRP findings to shapepolicy and practice via the Reading First component of No Child Left Behind(2002). We regard these developments as positive, for we think there is good reasonto teach vocabulary more aggressively and even better reason to study its relation tocomprehension more carefully. However, if we are going to teach it more effectivelyand if we are going to better understand how it is implicated in reading compre-hension, we must first address the vexing question of how we assess vocabularyknowledge and, even more challenging, vocabulary growth. In this essay, we arguethat vocabulary assessment is grossly undernourished, both in its theoretical andpractical aspects—that it has been driven by tradition, convenience, psychometricstandards, and a quest for economy of effort rather than a clear conceptualizationof its nature and relation to other aspects of reading expertise, most notably com-prehension. We hope that our essay will serve as one small step in providing thenourishment it needs.There is no doubt that vocabulary is closely tied to comprehension (Davis,1942; Just & Carpenter, 1987; Whipple, 1925)—in study after study, vocabularyknowledge predicts comprehension performance consistently with positive correla-282

Theory and Research Into Practice283tions typically between .6 and .8. But a correlation isnot an explanation of a conceptual relation betweenfactors. Anderson and Freebody (1985) understoodthis complexity well when they put forward threehypotheses to explain the ubiquitous finding of ahigh correlation between comprehension and vocab-ulary. The instrumentalist hypothesis argues thatlearning the words causes comprehension. The verbalaptitude hypothesis suggests that general verbal abili-ty is the root cause of both vocabulary and compre-hension performance. The knowledge hypothesisargues that both vocabulary and comprehension re-sult from increases in knowledge.More to the point, it is one thing to demon-strate a correlation and quite another to demonstratea causal relation between vocabulary instruction orlearning and comprehension. In that regard, it isworth noting the conclusions of the subgroup forvocabulary of the NRP (NICHD, 2000), whichdocument a consistent and robust relation betweenlearning vocabulary in specific texts and performanceon experimenter-designed

comprehension measuresderived from those same texts. By contrast, thegroup found only two studies showing that vocabu-lary instruction transferred beyond text-specific increases in vocabulary to far transfer measures, such asnorm-referenced comprehension reading tests. Aquestion of interest raised by the NRP report iswhether its conclusions are generalizable or are theartifact of some special characteristic of the ways inwhich the outcomes were measured in the studiesthey examined. Even though experimentally documented ef-fects of vocabulary instruction on measures of gen-eral reading comprehension are weak, at least asindexed by effects on standardized measures, vocabulary instruction has returned to a place of promi-nence in the reading curriculum; vocabulary serves acore role in commercial reading programs and inother curricular areas such as science, history, or for-eign language. Its ubiquity and gravity are capturedby the complaint, at least of science educators, thatthe bulk of text-centered science instruction is learn-ing the meanings of hundreds of new scientific termsrather than experiencing the intellectual rush ofhands-on inquiry (Armstrong & Collier, 1990). There are at least three plausible explanationsfor the weak empirical link between vocabulary in-struction and some transfer measures of readingcomprehension. The first position is that there is noactual link between the two: that a vocabulary mythhas clouded our reasoning and our pedagogy for cen-turies and that learning words does not cause com-prehension. The second is that vocabularyinstruction does not promote far transfer—that is, itis conceptually incapable of moving beyond the textsto which it is tied. Hence it shows up in local butnot global indicators of text understanding. Thethird position, and the one we take up in this essay,is that our measures of vocabulary are inadequate tothe challenge of documenting the relationship be-tween word learning and global measures of comprehension. That is, it might be that our instruction isimproving vocabulary learning, which might lead toimprovements in general comprehension, but the in-struments we use to measure vocabulary are so insen-sitive that they prevent us from documenting therelationship. In particular, the fact that standardizedassessments do not often include types of text thatare found in textbooks is an example of this potentialmasking of effects. The National Assessment ofEducational Progress (NAEP) 2009 framework hasaddressed this issue by dividing what have tradition-ally been labeled expository texts into more explicitand descriptive subcategories (National AssessmentGoverning Board [NAGB], 2005). Exposition hasbeen separated from, for example, literary nonfictionin recognition of the fact that these different genreshave, at the very least, different vocabulary loads.We don’t want to dismiss the first two posi-tions out of hand, but we want to press the measure-ment question so that it can be ruled in or out as themost plausible explanation for the paucity of docu-mented transfer effects. We will never know untiland unless we have developed and tested vocabularymeasures that are as conceptually rich as the phenomenon (vocabulary knowledge) they are intendedto measure.We begin by defining vocabulary and offeringa short historical account of vocabulary assessment.Then we

examine the literature—research, commonpractices, and theoretical analyses—on vocabularyassessment to answer three questions:1. What do vocabulary assessments (both pastand current) measure? 2. What could vocabulary assessments measure? 3. What research will we have to conduct overthe next decade in order to develop and validatemeasures that will serve us in our quest to improveboth vocabulary research and, ultimately, vocabulary instruction?

Perspectives

In line with the literature, several arguments must be taken into account when designing new high-quality vocabulary tools for assessment or instruction using digital technologies. In the context of vocabulary assessment, digital tools could provide accurate data collection and analysis of children's responses as well as response time and correct responses, an increased validity and reliability of assessment, and attractive support to enhance children's engagement in the task. These features are germane to the development of an accurate and detailed pro le of children both in terms of vocabulary and reading skills. Moreover, future assessment tools should target adaptive assessment based on, for example, item response theory. In the context of vocabulary instruction, computer-based tools should provide an interactive and multi-modal environment to favor motivation and engagement, systematic feedback appropriate to learners' characteristics such as age or initial vocabulary level to favor autonomy and active learning, and differentiated instruction depending on learners' initial vocabulary level enabling children to progress at their own pace without group or teacher pressure. More generally, the effectiveness of digital tools on learning requires more in-depth collaboration between researchers (cognitive aspects), teachers (pedagogical aspects) and developers (ergonomic aspects) to promote wider use of their potentialities and benefits.