

Improving Reading Across Subject Areas With Word Generation

Joshua F. Lawrence, Claire White, and Catherine E. Snow, Harvard Graduate School of Education

Introduction

In 2008, approximately 10.8 million children ages 5–17 in the United States spoke a language other than English in the home (Aud et al., 2010). While most language minority students receive all of their instruction in English, 3.8 million students received English language learner services during the 2003-2004 school year (Capps, Fix, Murray, Ost, Passel, & Herwanto, 2005). Compared with their native English-speaking peers, language minority students on average have lower reading performance in English (August & Shanahan, 2006). While numerous factors account for this gap, researchers have pointed to differences in word knowledge as part of the explanation. Language minority students have both less depth (Verhallen & Schoonen, 1993) and less breadth of vocabulary. In order to read with clear comprehension, students also need to understand the words they read, construct an interpretive cognitive model of what the author is trying to say, and have the requisite background knowledge to categorize, interpret, and remember what an author is saying in relation to established facts or a field of understanding (such as a content area subject). Although a deficit in any of these areas may prevent an adolescent reader from comprehending grade-level texts, deficits in vocabulary knowledge (and in the world knowledge indexed by vocabulary knowledge) may be the most widely shared problem among struggling adolescent readers.

One way that vocabulary supports reading comprehension is through reading subskills. There is substantial evidence that phonological, orthographic, and semantic processing of words are interrelated, such that vocabulary knowledge predicts rates of word reading (Nation & Snowling, 2004), and students with better semantic abili-

ties also have advantages in orthographic identification tasks (Yang & Perfetti, 2006). Vocabulary knowledge also plays an important role in students' higher order comprehension abilities. Established as well as current models of reading comprehension argue that word meaning and form selection are critical to creating a situation model from text and for integrating new knowledge from the text with prior background knowledge (Kintsch, 1986; Perfetti, Landi, & Oakhill, 2005). Recent reviews of research on adolescent literacy demonstrate that these higher order processes are exactly where most struggling adolescent readers break down (Kamil, 2003; Rand Reading Study Group, 2002); thus it is not surprising that vocabulary scores show increasingly strong correlations with reading comprehension scores as students move from primary to middle and secondary grades (Snow, Porche, Tabors, & Harris, 2007).

A challenging subdomain of vocabulary knowledge acquisition and instruction is all-purpose academic vocabulary, a segment of the lexicon that becomes particularly relevant to comprehension in adolescence. All-purpose academic vocabulary is a category with somewhat fuzzy boundaries, but prototypical members are words used for making fine distinctions in referring to communicative intents (e.g., *affirm*, *confirm*), argumentation (e.g., *evidence*, *conclusion*, *warrant*), abstract entities (e.g., *theory*, *factor*, *process*), and categories (e.g., *vehicle*, *utensil*, *artifact*). All-purpose academic words are used across content areas, occur frequently in glossaries where content-area words are defined, and receive little explicit instructional attention precisely because they are not seen as the responsibility of any content-area teacher. Yet control over this segment of the lexicon is crucial to comprehending and producing academic language.

This is the vocabulary domain on which the research program described in this brief focuses.

By the time normally developing children enter middle school, most will have mastered thousands of words for oral use, but comprehension of the rich language of text requires an understanding of more and different words (Nation, 2006). In middle school, students begin to take core subject area classes and are expected to read and understand expository texts with increasingly difficult vocabulary demands (Gardner, 2004). Clearly, exposure to new words in texts is one of the primary vehicles for word learning (Nagy & Anderson, 1984; Nagy & Herman, 1987). However, there are differences in students' abilities to learn new words incidentally while reading; these differences relate to their concurrent vocabulary levels (McKeown, 1985) and to their comprehension levels (Swanborn & de Glopper, 2002). Without instruction and support, independent reading is unlikely to improve word-learning outcomes for students of low socioeconomic status, although highly skilled readers may benefit (Lawrence, 2009).

Given the evidence that reading comprehension supports vocabulary development and that vocabulary development supports reading comprehension, we can describe the relationship between these two processes as one of reciprocal causation. It has been widely noted that less able students are likely to fall farther and farther behind if they struggle with learning processes linked by reciprocal causation (Stanovich, 1986). Fortunately, there is evidence that vocabulary instruction can have an important and lasting impact on student word learning (Beck, Perfetti, & McKeown, 1982; Carlo et al., 2004). There is reason to think, then, that a robust vocabulary intervention that targets academic language may improve vocabulary and reading comprehension in the short run while also supporting the struggling reader's facility at learning new words independently. The research project described here presents findings from an unmatched quasi-experiment of the Word Generation Program, an intervention firmly grounded in what is currently known about effective practice, while also casting light on how enhanced vocabulary levels relate to improved reading comprehension. While findings from a quasi-experiment are not firm grounds for causal inference, the data here are suggestive and form the basis for our ongoing randomized trial. The actual program itself can be downloaded for free at www.serpoinstitute.org/word-generation. More information about how the program was created and how words were selected is also avail-

able on the Web site and in published studies (Lawrence, White, & Snow, 2010; Snow, Lawrence, & White, 2009).

Program Implementation

The Word Generation materials define a list of key elements that are used to organize instruction. Those elements include the following:

- **Monday launch:** Reading a paragraph aloud with students that introduces a civic dilemma, modeling comprehension processes and word inferencing during reading, guiding discussion through comprehension questions, highlighting focus words, eliciting student opinions on the controversy of the week. This is usually done in English class.
- **Math activity:** Recurrently using target vocabulary in all-purpose and (if applicable) math-specific ways, engaging students in discussion of math problems, reminding students of controversy and soliciting their thinking about it, revoicing student comments to model clarity and target word use.
- **Science activity:** Recurrently using target vocabulary in all-purpose and (if applicable) science-specific ways, linking topic of the week to science content, reminding students of controversy and soliciting their thinking about it, revoicing student comments to model clarity and target word use.
- **Social studies activity:** Recurrently using target vocabulary in all-purpose and (if applicable) social studies specific ways, structuring a debate format, giving all students a chance to participate in debate, revoicing student comments to model clarity and target word use.
- **Friday writing:** Reviewing the controversy and reading the prompt aloud, reminding students to reread the paragraph or their notes to make good arguments for their point of view, having the target words posted or written on the board, ensuring quiet and order so students can write uninterrupted.

Quasi-Experiment of Word Generation in Partnership With Boston Public Schools

In 2007, our research team began a quasi-experimental study in which academic word learning by students in five schools implementing the Word Generation Program was compared to academic word-learning by students in three schools within the same system that did not choose to implement the program. Because the imple-

menting schools were those that volunteered for the program, selection effects must be taken into account in interpreting the findings. The results presented here have been described in greater depth elsewhere (Snow, Lawrence, & White, 2009).

Participants and Setting

Schools

At the start of the study, the average scores of the intervention schools on the state accountability assessment were lower than those of the comparison schools (mean of 56% failing in the treatment schools compared with 45% failing in the comparison schools). This is not surprising; the participating schools volunteered to do so, and those with lower scores were more likely to show an interest. We do not have detailed information about the vocabulary instruction in the comparison schools. Through limited observation and interviews we know that there is discipline-specific vocabulary being taught in each of the comparison schools as required by that state's curriculum framework. Furthermore, in one school, a long-time literacy coach had coordinated vocabulary instruction to some extent through a school-wide word-of-the-week effort. That being said, none of the comparison schools was using a commercial vocabulary program, nor were any heavily invested in a school-wide approach to vocabulary instruction.

Students

All students in the treatment schools received the intervention; both pre- and posttest data were available on 697 sixth-, seventh-, and eighth-grade students (349 girls and 348 boys) in the five treatment schools and 319 students (162 girls and 157 boys) across the three comparison schools. Of these, 438 students were classified as language minority (i.e., students whose parents reported preferring to receive materials in a language other than English); 287 in treatment schools and 151 in comparison schools. The vast majority of students in both treatment and comparison schools were low-income.

Research Design

Data Collection and Analysis

The efficacy of the intervention was assessed using a 48-item multiple choice test that sampled words from throughout the year. A high proportion of students failed to complete the vocabulary assessments in the time available. Because items at the end of the assessment had particularly low rates of completion, we dropped the last four items from our analysis of both pre- and posttest.

The reliability of the test with the 40 items that remained was acceptable (Cronbach's $\alpha = .876$).

This instrument was administered to students in all of the treatment schools in October 2007, before the introduction of Word Generation materials. Because of difficulty recruiting the comparison schools, the pretest was not administered at these schools until January. The posttest (identical to the pretest except for the order of items) was administered in all of the schools in late May. Because of the unfortunate disparity in interval between pre- and posttesting in the two groups of schools, we analyzed words learned per month as well as total words learned.

In addition to this curriculum-based assessment, we had access to most students' spring 2008 scores on the Massachusetts Comprehensive Assessment System (MCAS) English Language Arts. Additionally, we had Group Reading and Diagnostic Evaluation (GRADE; Williams, 2000) for both spring and fall for a selection of students in comparison ($n = 133$) and treatment ($n = 256$) schools. These scores were provided by the district for all the students for whom data were available. The decision to administer the GRADE was made at the school and classroom level. Thus, while these data are far from complete, we have no reason to think that there was a particular sampling bias across the schools.

Findings

Descriptive statistics show that students in the Word Generation Program learned approximately the number of words that differentiated eighth from sixth graders on the pretest. In other words, participation in 20–22 weeks of the curriculum was equivalent to two years of incidental learning. Unfortunately, the relative improvements in the Word Generation schools will be exaggerated by the differences in timing of the pretest. In order to account for the differences in test administration times, the pre-to-post improvement in all schools was divided by the number of months between the pre- and posttest administrations: the average improvement per months in the treatment schools was greater than that in the comparison schools. The average effect size of program participation on the researcher-developed vocabulary assessment was 0.49 (controlling for the improvement attained in the comparison schools).

Regression analysis was used to determine if participation in Word Generation predicted improved vocabulary outcomes, controlling for the pretest. Boys learned more words than girls ($\beta = -0.052, p < .007$)

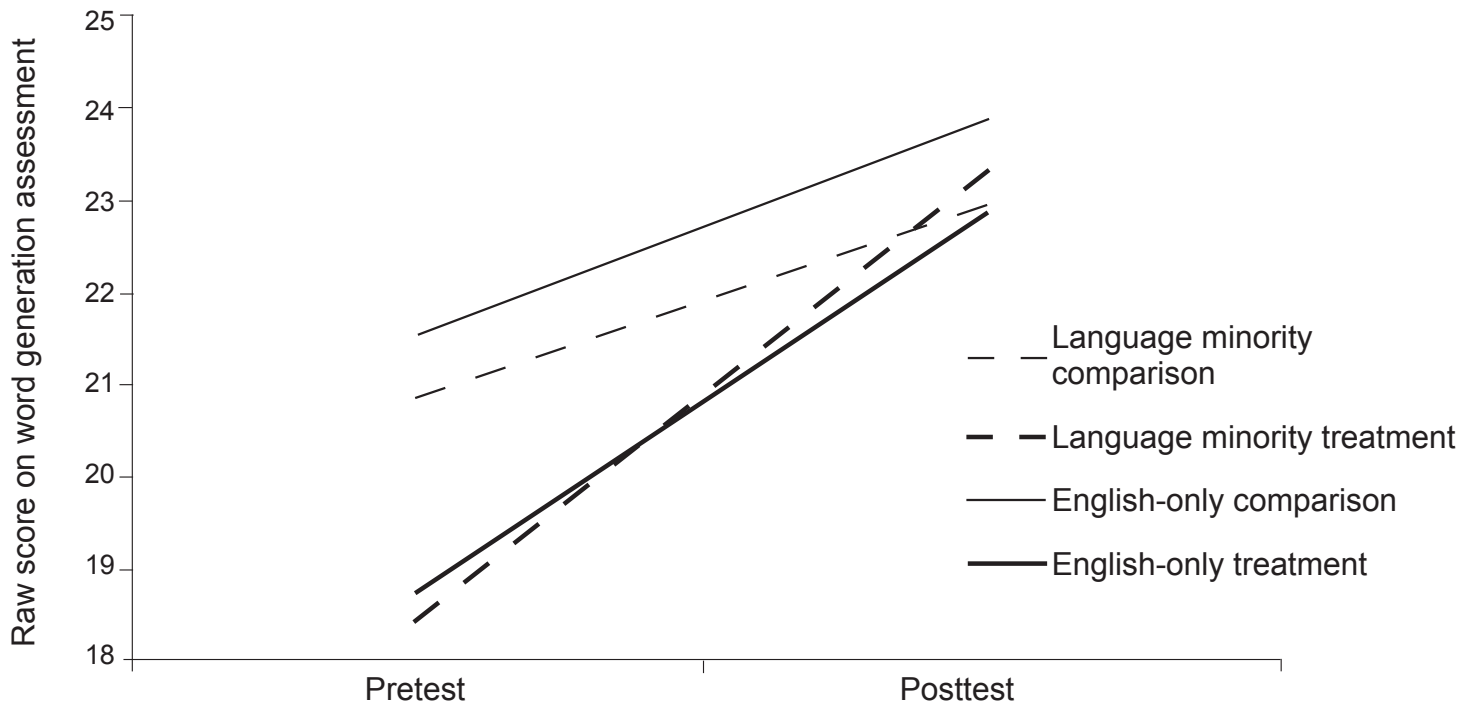


Figure 1. Vocabulary improvement for English-only and language minority students in Word Generation and comparison schools.

and participants in the program learned more words than nonparticipants ($\beta = 0.166, p < .001$). Language status (language minority versus English only) was not a significant independent predictor of word learning, but language minority students learned words at a relatively faster rate than English-only students in treatment schools, but not comparison schools (language status interacted with treatment at the margin of significance, $p = .055$); including the interaction improved the overall model. Interestingly, student pretest vocabulary did not interact with treatment in predicting posttest scores. It was decided to split the data set to investigate the home language variable more closely. The first set of regressions used pretests and gender to predict posttest scores in the comparison schools ($r^2 = .62$) and Word Generation schools ($r^2 = .64$). In Word Generation schools language minority status predicted improved vocabulary ($\beta = -0.053, p = .022$), but it was not a significant predictor in comparison schools. These results are demonstrated in Figure 1: In the comparison schools (the light lines) English-only students improved more than language minority students, in the treatment schools (bold lines) language-minority students improved more than English-only students.

In order to determine whether participation in Word Generation had any relationship to performance on the

MCAS, a regression model was fit with MCAS scores in April 2008 as the outcome; gender, treatment status, pretest, and posttest scores were used as predictors. We added an interaction term to see if posttest scores on the curriculum-based assessment interacted with treatment in predicting MCAS scores (controlling for pretest scores). The interaction term was significant ($\beta = .21, p = .01$) and its inclusion improved the model. In other words, students who benefited most from participation in Word Generation had higher MCAS scores than students with similarly improved vocabularies acquired without Word Generation exposure.

We further explored the interaction between treatment and vocabulary improvement by splitting the data and refitting the models to data from the treatment and comparison schools separately. The fitted model for comparison school data did not predict MCAS achievement ($R^2 = .41$) as well as the fitted model for the treatment school data ($R^2 = .49$). In the Word Generation schools student vocabulary posttest scores ($\beta = 0.527, p < .001$) were much stronger predictors of MCAS achievement than pretest scores were ($\beta = 0.201, p < .001$), perhaps because the posttest scores captured not only target vocabulary knowledge at the end of the year, but also level of student participation in the Word Generation program. These conclusions maintained even when we

used available baseline GRADE data as a covariate in our models.

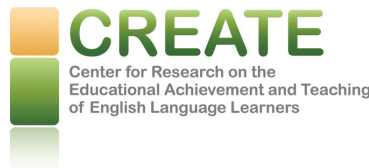
Longitudinal Follow-Up on Quasi-Experiment

The goal of Word Generation is to improve vocabulary so that it results in improved reading comprehension; clearly, short-term vocabulary learning will not generate long-term comprehension improvement. Despite the evidence of vocabulary gains for all Word Generation participants on average, and in particular for language minority participants, we did not know whether these students maintained vocabulary knowledge after summer vacation and through the following school year. We conducted a follow-up longitudinal study to examine the effects of Word Generation on the learning, maintenance, and consolidation of academic vocabulary for students from English-speaking homes, proficient English speakers from language minority homes, and limited English proficient students. Using individual growth modeling, we found that students receiving Word Generation improved on average on target words during the instructional period. We confirmed that there was an interaction between instruction and language status such that English-proficient students from language minority homes improved more than English-proficient students from English-speaking homes. We administered follow-up assessments in the fall after the instructional period ended and the spring of the following year to determine how well students maintained and consolidated target academic words. Students who participated in the intervention maintained their relative improvements at both follow-up assessments (Lawrence, Capotosto, Branum-Martin, White, & Snow, 2010). We thus have reason to expect that these students will display improved reading comprehension and enhanced academic learning. A randomized experimental study of Word Generation now underway will enable us to test this expectation more rigorously.

References

- Aud, S., Hussar, W., Planty, M., Snyder, T., Bianco, K., Fox, M., Frohlich, L., Kemp, J., Drake, L. (2010). *The Condition of Education 2010* (NCES 2010-028). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.
- Beck, I., Perfetti, C., & McKeown, M. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, 74(4), 506-521.
- Capps, R., Fix, M. E., Murray, J., Ost, J., Passel, J. S., & Herwanto Hernandez, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Washington, DC: The Urban Institute.
- Carlo, M., August, D., McLaughlin, B., Snow, C., Dressler, C., Lippman, D., ... White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, 39(2), 188-215.
- Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics*, 25(1), 1-37.
- Kamil, M. L. (2003). *Adolescents and literacy: Reading for the 21st century*. Washington, DC: Alliance for Excellent Education.
- Kintsch, W. (1986). Learning from text. *Cognition and Instruction*, 3(2), 87-108.
- Lawrence, J. (2009). Summer reading: Predicting adolescent word learning from aptitude, time spent reading, and text type. *Reading Psychology*, 30(5), 445-465.
- Lawrence, J., Capotosto, L., Branum-Martin, L., White, C., & Snow, C. (2010). *Language proficiency, home-language status, and English vocabulary development: A longitudinal follow-up of the Word Generation program*. Unpublished manuscript, Harvard University
- Lawrence, J., White, C., & Snow, C. (2010). The words students need. *Educational Leadership*, 68(2), 22-26.
- McKeown, M. (1985). The acquisition of word meaning from context by children of high and low ability. *Reading Research Quarterly*, 20(4), 482-496.
- Nagy, W., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19(3), 304-330.
- Nagy, W., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (Vol. 19-35). Hillsdale, NJ: Lawrence Erlbaum.
- Nation, I. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language*

- Review/*La Revue canadienne des langues vivantes*, 63(1), 59–82.
- Nation, K., & Snowling, M. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, 27(4), 342–356.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Malden, MA: Blackwell Publishing.
- Planty, M., Hussar, W., Snyder, T., Kena, G., KewalRaman, A., Kemp, J., ... Dinkes, R. (2009). *The condition of education 2009. (NCES 2009-081)*. Washington, DC: National Center for Education Statistics, Institute of Educational Sciences, U.S. Department of Education.
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.
- Snow, C., Lawrence, J., & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness*, 2(4), 325–344.
- Snow, C. E., Porche, M. V., Tabors, P., & Harris, S. (2007). *Is literacy enough? Pathways to academic success for adolescents*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360–407.
- Swanborn, M., & de Glosper, K. (2002). Impact of reading purpose on incidental word learning from context. *Language Learning*, 52(1), 95–117.
- Verhallen, M., & Schoonen, R. (1993). Vocabulary knowledge of monolingual and bilingual children. *Applied Linguistics*, 14, 344–363.
- Williams, K. T. (2000). *Group reading assessment and diagnostic evaluation*. Circle Pines, MN: American Guidance Service.
- Yang, C.-L., & Perfetti, C. (2006, July). *Reading skill and the acquisition of high quality representations of new words*. Paper presented at the Society of the Scientific Study of Reading Annual Meeting, Vancouver, Canada.



CREATE, the Center for Research on the Educational Achievement and Teaching of English Language Learners, is conducting a program of research designed to address the critical challenge of improving educational outcomes for English language learners in the middle grades by

- Enhancing the empirical research base for readers in Grades 4–8
- Using both narrative and expository text to develop and test effective interventions that promote content knowledge and language and literacy development
- Investigating the features of instruction and text modifications that facilitate learning for English learners (e.g., traditional instruction vs. ESL-enhanced instruction, teacher-guided instruction vs. group work, traditional text vs. modified text)
- Designing, testing, and delivering professional development that ensures that teachers implement effective classroom practices to help English learners achieve high standards

For more information, visit the CREATE Web site

www.cal.org/create

The Center for Research on the Educational Achievement and Teaching of English Language Learners (CREATE) conducts a program of research designed to address specific challenges in the education of English language learners in Grades 4-8. CREATE is a partnership of researchers from six institutions:

- Texas Institute for Measurement, Evaluation, and Statistics, University of Houston
- Harvard University
- California State University, Long Beach
- University of California-Berkeley
- Center for Applied Linguistics
- Vaughn Gross Center, University of Texas at Austin

This brief was produced by CREATE with funding from the U.S. Department of Education (ED), Institute of Education Sciences, under Contract No. ED-R305A05056. The opinions expressed in this report do not necessarily reflect the positions or policies of ED. This document is in the public domain and may be reproduced without permission.