Academic Vocabulary: Analyzing College Texts and Creating Post-Secondary Vocabulary Lists

**Problem Background and Overview of the Literature**

Recent federal reports indicate that vocabulary knowledge is one of the five essential components of reading comprehension (e.g., RAND Reading Study Group, 2002). The more words students know, the easier it is for them to comprehend academic texts and lectures (Francis & Simpson, 2009). Some studies estimate that students need to know 90% to 98% of word families in a text for optimal comprehension (Hu & Nation, 2000; Schmitt, 2008) and that students who know less than 80% suffer poor comprehension (Bonk, 2000). However, discussing word families may underestimate the difficult task of vocabulary learning. Each word family may include several word forms, which are not always readily transferrable (Schmitt, 2008). For example, a student may know the word *colleague* but not be able to comprehend the word *collegial*. Thus, an emphasis on learning word families alone is limited.

Students must learn a large number of words to be able to be proficient in English; however, the goal of vocabulary instruction in middle and high schools has increasingly been on learning content-specific terms (e.g., *meiosis* and *mitosis* in a biology text) (Harmon, Hedrick, & Wood, 2005), often at the expense of general academic vocabulary. Thus, many students enter college lacking a vocabulary developed enough for comprehending academic texts. To succeed in college, students need to develop an extensive academic vocabulary (Francis & Simpson, 2009).

An academic word list created by Coxhead (2000) was created using texts intended for an international audience, but only 2% of the texts used in this study were produced for use in the United States. Additionally, the 570 words in Coxhead’s list are often more basic than many students need at the college level. For example the words *role, item, image, odd* are part of the list, but more advanced words commonly used in text such as *ostensibly, multitude, plethora* are not. Thus, this research will extend Coxhead’s approach by generating a more advanced list of academic words currently in use in college textbooks.

# Purpose

The purpose of this study is to examine the frequency of words used in today’s college textbooks and to create an advanced word list of general (non-content-specific) college-level vocabulary commonly used in the texts that college students are required to read.

# Project Design

Ten college texts from each of ten disciplines (100 textbooks total) will be analyzed. For the purposes of this study, the disciplines selected are those that a majority of students will enroll within their first two years of college. They are comprised of psychology, history, biology, philosophy, sociology, political science, economics, communications, art history, and physical science. Although mathematics and literature are courses that students take in the first two years, they are not included in this list. Literature vocabulary can vary widely and mathematics texts, although filled with content-specific terms such as *addends* and *denominator* do not have enough running text for the purposes of this study. Pearson Education Higher Education division has agreed to supply the electronic PDF copies of the texts that will be used in this study.

## Data collection.

Data will be collected in two stages. First, 100 textbooks will be analyzed for word frequency using textual analysis software that will allow the words to be tagged by specific textbook and content area. A potential word list of 750–1000 words will be developed. The goal is to identify the academic vocabulary that college students need to know to be able to comprehend college-level texts (e.g., *myriad, heuristic, palpable*). These words will be used to create an online assessment for students to help identify commonly used academic words that are not known by today’s undergraduate students. Participants will be drawn from freshman and sophomore level courses in the ten disciplines used in this study. The goal is for 200 students to complete the online vocabulary assessment. Each participant will respond to items over a subset of 50 words. First, students will rate their knowledge of the word on a Likert-type scale ranging from 1 (I have never seen this word before) to 7 (I know this word and am able to use it in a sentence). Second, students will respond to a multiple-choice question about each word’s definition.

## Data analysis.

First, 100 textbooks will be analyzed for word frequencies. PDF’s of the textbooks will be run through textual analysis software that will allow each word to be tagged by text and discipline. Second, the word lists will be analyzed to create the potential final list of words to be tested. This list will be generated by eliminating Dolch’s (1948) and Fry’s (1980) lists of high-frequency sight words (e.g., *tree, boy, the*), words known by a majority of students (to be determined by pilot testing word lists, but may include words such as *team, background, negative*) and content-specific words (e.g., *eidetic, expressionism, geomorphology*). Third, data from the online assessment will be analyzed in several ways. The data will be examined for words students claim to know, but do not; those that they claim not to know but answer correctly; words that students say they do not know and do not respond correctly; words students say they know and respond correctly. A matrix of words will be created to generate the final list of academic words.

# Importance and Relation to Applicant’s Field of Research

The idea for this study came directly from my work with the Emerging Stars students this summer. All of the students enrolled in my class (RDG 1300) were also enrolled in a Philosophy course. I asked them to jot down any unfamiliar words in their philosophy text as a way to introduce the concept of vocabulary learning in college and was stunned when the class of 14 students submitted over 65 discrete words from one night’s worth of textbook reading. My extant research has focused primarily on the study strategies students need in their transition from high school to college learning; this project will extend my work on helping students succeed in college by addressing the critical need for vocabulary development. Analyzing the texts and creating the word lists will be the beginning of an intensive line of research. In addition to publication of the results of this study, I anticipate creating an interactive web-based vocabulary website as well as a vocabulary textbook.

# Budget

Faculty Salary: $6,000 1/12 release time Summer Session I/10

Capital: $700 Textual analysis software

Programming: $1300 computer programming

Total: $8000

# Budget Justification

* The budget for this project is anticipated to be $8,000.
* $6,000 used for summer session I/10 release to allow time to analyze the PDF’s of the 100 textbooks and create the vocabulary lists and questions for 750-1000 words to use with students in Fall 2010.
* $700 used for textual analysis software to analyze the textbooks
* $1,300 used for hiring a computer programmer to create an online web site for the word list vocabulary testing.

# References

Bonk, W.J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening, 14,* 14–31.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34, 213–238.*

Dolch, E.W. (1948). *Problems in reading.* Champaign, IL: Garrad Press.

Francis, M.A., & Simpson, M.L. (2009). Vocabulary development. In R.F. Flippo & D.C. Caverly (Eds.), *Handbook of college reading and study strategy research* (2nd Ed.). New York: Routledge, 97–120.

Fry, E. (1980). The new instant word list. *Reading Teacher, 34,* 284–289.

Harmon, J.M., Hedrick, W.B., & Wood, K.D. (2005). Research on vocabulary instruction in the content areas: Implications for struggling readers. *Reading and Writing Quarterly, 21,* 261–280.

Hu, M., & Nation, I.S.P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language, 23,* 403–430.

RAND Reading Study Group. (2002). *Reading for understanding: Toward a research and development program in reading comprehension.* Prepared for the Office of Educational Research and Improvement (OERI), U.S. Department of Education. Santa Monica, CA: RAND Education.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12,* 329–363.

Vita of Scholarly and Creative Activities

**University Experience**

2009-present Texas State University-San Marcos Associate Professor

2006-2009 University of Georgia Associate Professor

## Textbooks.

Holschuh, J.P., & Nist, S.L. (in press). *Effective college learning* (2nd Ed.). New York: Longman/ Dorling Kindersley.

Nist, S.L., & Holschuh, J.P. (2009). *College success strategies* (3rd Ed.) (Penguin Academic Series). New York: Longman.

Nist, S.L., & Holschuh, J.P. (2000). *Active learning: Strategies for college success.* Needham Heights, MA: Allyn & Bacon.

## Chapters in books.

Holschuh, J.P., & Aultman, L. (2009) Comprehension development. In R.F. Flippo &   
D.C. Caverly (Eds.), *Handbook of College Reading and Study Strategy Research (2nd Ed.).* New York: Routledge.

Holschuh, J.P., Hynd, C., & Oldfather, P. (2000). Theory into practice: Current motivation theories and the RWCT project. In D. Klooster, J. Steele, & P. Bloem (Eds.), *Ideas without boundaries: Educational reform through reading and writing for critical thinking* (pp. 108–128). Newark, DE: International Reading Association.

## Creative books.

Nist, S.L., & Holschuh, J.P. (2007). College rules! How to study, survive, and succeed in college (2nd Ed.). Berkeley, CA: Ten Speed Press.

## Articles.

Holschuh, J.P. (2006). Assessing beliefs: The epistemological scenario. *Academic Exchange Quarterly, 10,* 172–176.

Nist, S.L., & Holschuh, J.P. (2005). Practical applications of the research on epistemological beliefs. *Journal of College Reading and Learning, 35*(2), 84–92.

Hynd-Shanahan, C.R., Holschuh, J.P., & Hubbard, B.P. (2004). Thinking like a historian: College students’ reading of multiple historical documents. *Journal of Literacy Research, 4,* 238–250.

Holschuh, J.P., Nist, S.L., & Olejnik, S. (2001). Attributions to failure: The effects of effort, ability, and learning strategy use on perceptions of future goals and emotional responses. *Reading Psychology, 22,* 153–173*.*

Holschuh, J.P. (2000) Do as I say, not as I do: High, average, and low performing students’ strategy use in biology. *Journal of College Reading and Learning, 31,* 94–107.

Hynd, C., Holschuh, J.L., & Nist, S.L. (2000). Learning science: What motivates students to learn complex scientific information*. Reading and Writing Quarterly, 16,* 23–57.

## Papers presented at professional meetings.

Cornelius, T., Owen-De Schryver, J., & Holschuh, J.P. (2008, March). *The Effectiveness of Parallel Note-taking on Undergraduate Learning*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Holschuh, J.P., & Nist, S.L. (2007, March). *Debunking Conventional Wisdom: What Works for Learning to Learn.* Paper presented at the annual meeting of the National Association for Developmental Educators, Nashville, TN.

Holschuh, J.P., & Aultman, L.P. (2007, March). *Why do Students Think That They Don’t Need to Read Their Textbooks.* Paper presented at the annual meeting of the National Association for Developmental Educators, Nashville, TN.

Holschuh, J.P. (2006, August). *Teaching Students to Evaluate Internet Information by Thinking Like a Historian.* Paper presented at the International Reading Association World Congress, Budapest, Hungary.

Holschuh, J.P., & Domizi, D.P. (October, 2004). *Teaching Students to Evaluate Internet Information by Thinking Like a Historian.* Paper presented at the annual meeting of the College Reading Association, Delray Beach, FL.

## Funded internal grants and contracts.

President’s venture fund, University of Georgia (2008). $4,500 to conduct research entitled “My Undergraduate Semester: Examining Academic Tasks in Today’s Undergraduate Courses.”

International travel grant, University of Georgia (2004). $2,000 grant funded by the office of the senior vice president for academic affairs and provost to present seminars on teaching and learning in Tallinn, Estonia.

President’s venture fund, University of Georgia (2004). $500 grant for materials and expenses for teaching and learning seminars in Tallinn, Estonia.