Productive Vocabulary: A Look at the Effect of Task and Language in a Corpus of Academic Student Writing

Randi REPPEN
Department of English
Northern Arizona University
Flagstaff, Arizona, U.S.
Randi.Reppen@nau.edu

There is no dispute that the role of vocabulary is central to second language success. As researchers, we also know that there are different ways measuring vocabulary knowledge (Nation, 2001; Wesche & Paribakht, 1996). This presentation will focus on productive vocabulary use in a corpus of writing collected from first year students in required university writing course. The over 500 million word corpus used in this study is composed of five academic writing tasks produced by students from three different first languages (Arabic, Asian (Chinese and a few Korean) and English).

The study explores vocabulary at several levels including traditional vocabulary measures (e.g., lexical diversity, General Service List (GSL) and Academic Word List (AWL coverage)), but also includes a careful examination of vocabulary variability across different task types (e.g., argumentative, analytical, reflective tasks) and across the three different first language groups. In addition to traditional measures, several innovative measures will be explored such as, the use of multi-word versus single word verbs and the use of four word n-grams. The four word n-grams will be used to identify patterns of formulaic sequences across these three language groups as they encounter different academic writing tasks. It is expected that certain task types will elicit certain types of formulaic sequences, for example, a task involving different points of view or comparisons will most likely have the four word n-gram *on the other hand.* The variability of these formulaic expressions that serve as sign posts will be explored across tasks. In addition to comparing the similarities and differences due to the effect of task, the use four word n-grams produced by the writers from different first languages groups will also be explored. The examination of multi-word versus single word verbs will be used to identify more oral language (multi-word verbs) from more written forms (single word verbs). In addition to these indicators of oral and literate forms, linguistic features of development proposed in Biber, Gray & Poonpon (2010) that can be explored from lexico-grammatical level will be included to compare not only across language groups, but also within language groups across task types.

By using these diverse vocabulary measures across a corpus of writing that is carefully controlled for task, and that represents three different first languages, it is expected that a clearer picture of both first language influences and the effect of task will emerge. These results will yield insights as to what is known about the productive vocabulary in academic writing tasks of students from these three different first languages in a first year writing course. Implications for instruction will also be discussed.

References

Biber, D. Gray, B. & Poonpon, K. (2010). 'Should we use characteristics of conversation to measure grammatical complexity in L2 writing development?' *TESOL Quarterly* 45: 5-35.

Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Wesche, M. & Paribakht, T. (1996). 'Assessing L2 vocabulary knowledge: Depth versus breadth.' *Canadian Modern Language Review* 53: 13-40.