Journal of English for Academic Purposes 18 (2015) 25–39

Contents lists available at ScienceDirect

Journal of English for Academic Purposes

journal homepage: www.elsevier.com/locate/jeap

Linguistic dimensions of impromptu test essays compared with successful student disciplinary writing: Effects of language background, topic, and L2 proficiency

Sara Cushing Weigle*, Eric Friginal

Georgia State University, USA

ARTICLE INFO

Article history:
Available online 6 May 2015

Keywords:
Multi-dimensional analysis
Writing assessment
Validation

ABSTRACT

One important validity question with regard to writing assessment is the degree to which performance on a timed writing test can predict performance on future academic writing. Recent developments in corpus linguistics have allowed scholars to describe in detail the linguistic features of a variety of academic texts, including genres of disciplinary writing and writing on essay tests, which can aid in answering this question. The purpose of this paper is to compare the linguistic features of test essays written by native and non-native speakers with a comparison corpus of successful student writing across a range of disciplines using Biber's (1988) multidimensional analysis framework. Essays written on two different test prompts were analyzed along dimensions of successful student writing revealed by an analysis of the Michigan Corpus of Upper-level Student Writing (MICUSP) conducted by Hardy and Römer (2013). Results demonstrated that test essays differed in significant ways from disciplinary writing, particularly in the natural and health sciences. Furthermore, language background (native vs. non-native), prompt, and language proficiency (i.e., essay scores) were systematically related to scores on all four dimensions. Implications for pedagogy and language assessment are discussed.

1. Introduction

In recent years, many applied linguists concerned with second language (L2) writing instruction and assessment have turned to corpus linguistics as a means for analyzing and understanding the nature of academic texts. By studying large electronic collections of texts (e.g., published articles or successful student writing in specific disciplines), which can be automatically parsed and tagged for linguistic features such as parts of speech, clause types, or semantic categories of words, researchers have been able to describe genres of academic writing (e.g., Biber, 2006; Biber, Conrad, & Cortes, 2003; Mudraya, 2006; Robinson, Stoller, Constanza-Robinson, & Jones, 2008) and also to assess the quality and/or describe the linguistic characteristics of writing by students enrolled in universities (e.g., Hinkel, 2002; Horst, Cobb, & Nicolae, 2005; Hyland, 2004; Lee & Swales, 2006; Swales, 1990). Such studies are important for informing writing instruction, as they provide empirical evidence of the kinds of language structures students need to master to be successful writer in different disciplines, and of the differences between learner texts and professional or successful student texts.

Corpus linguistic studies are also important for language assessment, as they can provide information that can be useful in both test development and test validation. A convenient distinction here can be made between reference corpora and learner

* Corresponding author.
E-mail address: sweigle@gsu.edu (S.C. Weigle).

http://dx.doi.org/10.1016/j.jeap.2015.03.006
1475-1585/

corpora. By reference corpus we mean a corpus of authentic language that represents a specific domain of language use. For example, the 2.7 million word TOEFL 2000 Spoken and Written Academic Language (T2K-SWAL) Corpus (Biber et al., 2004) was commissioned by Educational Testing Service to inform the development of what eventually became the TOEFL Internet-Based Test (TOEFL iBT). A reference corpus such as the T2K-SWAL can be beneficial in describing what Bachman (1990) calls the target language use (TLU) domain: in this case, the literacy and language demands of North American university studies. Reference corpora are also useful in test item writing, to ensure that test items are representative of authentic language in important ways.

Another important type of corpus for language testing is a learner corpus: that is, a collection of texts produced by language learners. Corpora of this type are useful for describing the linguistic features of texts produced by language learners at different levels of proficiency, learners from different first language backgrounds and so on. In the test development and validation process, learner corpora are useful for investigating whether open-ended items (speaking or writing) elicit the language forms and functions of interest to the test developers, and whether language varies with proficiency in line with the construct being tested. Another potentially important application of corpus analysis is in investigating differences across writing prompts to explore task effects on scores.

It is also of interest for language test validation to compare the linguistic features of test taker responses to those of native speakers, particularly in the case where non-native speakers as a group are required to take a test that is not required of native speakers. The native/non-native distinction is of course an oversimplification (see Hall & Navarro, 2011) but the fact remains that non-native speakers are often required to demonstrate language proficiency in ways that native speakers are not.

Finally, one potentially useful application of corpus linguistics to writing assessment is to compare learner language produced during a writing test with a reference corpus of successful student writing. While it has long been acknowledged that timed writing tests are insufficient as a measure of writing ability in both first and second language settings (e.g., Behizadeh, 2014; Camp, 1993; Hamp-Lyons & Condon, 2000; Weigle, 2002), the degree to which essays written in response to typical prompts can predict performance in the domain of academic writing is an important validity question. Looking at the linguistic features of a learner corpus (in this case, a corpus of tests essays) in comparison to those of a reference corpus (in this case, successful university student writing) can provide important evidence for the validity of the essay test.

In this paper we present a corpus-based analysis of timed test essays written by native and non-native speakers and investigate the linguistic features of these test essays with reference to those of a large corpus of successful student writing across a wide variety of disciplines. We look at three research questions:

1. How similar are the linguistic features of successful timed impromptu test essays written by native and non-native speakers of English to those of successful student writing across a range of disciplines?
2. Do first language status (native vs. non-native) and essay topic influence linguistic features of test essays?
3. For non-native speakers, is there a relationship between essay scores and linguistic features?

In answering these questions, we use a multi-dimensional (MD) approach to the study of text features (Biber, 1988; 1995), described below. In the following section, we introduce this approach and then describe the two corpora used in the study (a reference corpus and a learner corpus).

2. Background and context of the study

Many corpus-based studies of academic writing have looked at predictive or correlational data showing the relationship between individual textual features and quality of writing scores given by instructors or raters. A substantial number of studies have identified linguistic features (e.g., subordination, prepositions, linking adverbials, etc.) that are predictive of scores given by instructors/raters and features that distinguish differences between students' disciplines (e.g., Römer & Wulff, 2010; Tarone, Dwyer, Gillette, & Icke, 1998) and various demographic factors (e.g., language proficiency levels, graduate vs. undergraduate) (e.g., Ferris, 1994; Grant & Ginther, 2000; Hinkel, 2002).

The identification of linguistic features found to be statistically significant indicators of writing quality has interested researchers because of its obvious pedagogical import (Grant & Ginther, 2000). However, correlational and predictive indicators on the relationship between linguistic distributions and quality of writing have been somewhat inconclusive (Jarvis, Grant, Bikowski, & Ferris, 2003). It is clear that contextual and disciplinary differences, including the roles of topics and prompts, contribute to variations in academic texts (Beers & Nagy, 2009; Norris & Ortega, 2009). Thus, the assumption of a linear relationship between linguistic distributions and assessment scores or grades on one type of writing may be limited when applied to other types of writing.

Rather than looking at individual features, then, our study takes a multidimensional (MD) approach to the study of text features (Biber, 1988, 1995). The MD approach is based on the assumption that differences in registers are associated with "patterns of co-occurring lexico-grammatical features" (Halliday, 1988, p. 162). The term "register" in this context has the specific meaning of a specific language use situation, which can be defined broadly, as in speech as opposed to writing, or more narrowly, e.g., academic writing, blogs, office hour conversations. For example, conversation tends to include more pronouns and shorter sentences than academic prose; if one looked only at the use of pronouns in these two registers one would miss out on the other linguistic features that tend to co-occur with pronouns. These systematic differences occur because individuals make lexical and grammatical choices appropriate to the register in which they are speaking and writing (Biber & Conrad, 2001; Hymes, 1984). Importantly, these co-occurrences are not one-dimensional.