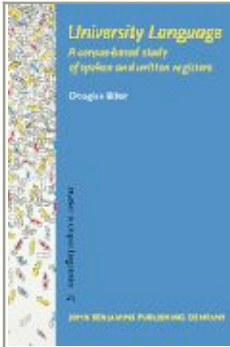


***University Language:  
A Corpus-based Study of Spoken and Written Registers***

<b>Author:</b>	Douglas Biber (2006)		
<b>Publisher:</b>	Amsterdam: John Benjamins		
<b>Pages</b>	<b>ISBN</b>	<b>Price</b>	
Pp. viii + 262	90-272-296-7 (paper)	£ 36.00 GBP; \$48.95 USD	

As linguists, we know that language matters. But to many people it might come as a surprise to realize that when students begin their university life, some of their main problems are linguistic ones. When students fail to communicate with professors in an appropriate register, or allow the relaxed style of a lecture to disrupt the formality of a term paper, they are falling into linguistic traps for which no one has prepared them. For non-native speakers, the difficulties may be particularly acute, but even for natives, mastering a range of new registers can pose serious difficulties. Yet this problem often passes unnoticed, perhaps because people are almost as oblivious to language as to the air they breathe or because tools have not been available to research the situation objectively. Even with the advent of corpus linguistics, most studies of academic language have focused on published texts, particularly research articles, and little is known about other ways language is used within the university setting, which might include textbooks, lectures, study groups, institutional publications and encounters with administrative and service staff.

*University Language* by Douglas Biber is an attempt to address the issue of language in the university across the board. The book has its background in the TOEFL 2000 Spoken and Written Academic Language (T2K-SWAL) Project sponsored by ETS and should therefore be of interest to teachers preparing students for TOEFL iBT, as well as to teachers on pre-sessional language courses and linguists interested in corpus research. The aim of the project, which defines the scope of this book as well, was to describe language use across a very wide range of university registers: spoken and written, formal and informal, embracing the major disciplines (humanities, natural and social sciences), the main academic levels (undergraduate and postgraduate), and the typical situations in which students might find themselves (lectures,

seminars, tutorials, interaction with administrative staff). Because of the deliberate focus on what students encounter, the design of the book intentionally excludes language that students produce independently: areas such as student presentations and term papers fall outside its scope. In Biber's own words, "the central goal of the book is simple: to provide a relatively comprehensive linguistic description of the range of university registers, surveying the distinctive linguistic characteristics of each register" (p. 22).

Since the usefulness of corpus-based studies depends heavily on the design of the original corpus, it is instructive first to glance briefly at the T2K-SWAL corpus before moving on to examine how the data are analyzed. The T2K-SWAL corpus consists of 2.7 million words captured at four US universities. Unusually for an academic corpus, almost 1.7 million of those words are from recorded spoken sources, while only 1 million originated in written material. Within these categories, most of the spoken data were obtained from class sessions (1.2 million words), while only 50,000 were from office hours. The author explains that this proportion reflects his assessment of the relative importance of these two types of academic encounter. Similarly, within the category of written sources, the vast majority of texts included were textbooks, whereas course packs and institutional publications were given a lower profile. Although this is fully consistent with the overall purpose of the corpus, that is, to reflect the language mix in the university in general, it does mean that the samples of certain types of language are rather small, which may render comparisons between samples hazardous on occasion.

The main challenge of corpus research is to select tools that will confirm intuitions or reveal unexpected patterns of language use. University Language relies heavily on the previous work carried out by the same author in collaboration with others for the *Longman Grammar of Spoken and Written English* (Biber et al., 1999) in which a comprehensive set of quantifiable features were measured across language samples from four different registers, one of which was defined as "academic prose" (a useful summary of the Longman project's findings on academic prose can be found on pp. 15-18 of *University Language*). This book builds on what has already been established about academic and other registers by taking many of the same measures (altogether, 129 features are investigated) and studying them across the university language situations outlined above. Thus we learn, for example, that despite a certain superficial similarity of pragmatic function, textbooks employ vastly more nouns and fewer verbs than are used in classroom teaching. Textbooks also contain more relative clauses but fewer complement and adverbial clauses than classroom teaching, and that the complement clauses used are less likely to be "that-" or "wh-" clauses. Classroom language on the other hand abounds in adverbs of certainty and likelihood, both of which are relatively uncommon in textbooks, and uses many more modal verbs of all kinds. These findings are perhaps most intriguing in that they situate classroom teaching firmly towards the "spoken" pole of the continuum from spoken to written, putting paid to any notion that a class is somehow equivalent to a textbook: the objective may be the same, but the way in which it is achieved is linguistically quite different.

One way this book represents an advance on earlier generations of corpus research can be found in the chapters it devotes to lexical bundles and multidimensional analysis. There is a growing awareness that bundles play a vital role as "building blocks of discourse," and this study illustrates the dramatic differences between the

bundle repertoires of different disciplines and registers. Few are encountered across the board ("at the end of," "as well as the"), while many are found as typical of specific areas ("an increase in the" for science and engineering, for example, or "by the fact that" in social sciences). Although there is obviously more work to be done in this area, this chapter affords a fascinating glimpse into the subliminal patterns that underpin specialized language.

Similarly, the chapter on multidimensional analysis constitutes a major step forward in that it attempts to create order out of the chaos of fragmentary data generated by corpus techniques. Biber applies factor analysis to 90 of the original 129 features, obtaining four different dimensions useful for characterizing different registers. Thus the register employed in study groups emerges as being near the oral pole of "oral vs. literate discourse," while it lies near the center of "procedural vs. content-focused discourse." This contrastive global approach opens up fruitful possibilities for future research, particularly in the scope it offers for pinpointing differences between disciplines or between apparently similar oral registers.

On the whole, this book is a masterly example of what can be done using specialized corpora. My only question, from the classroom floor, is how the very detailed knowledge gained by studies of this kind can be mobilized in students' language learning interests. At risk of violating the norms of descriptive corpus research, I would like to suggest that future research might examine how student language diverges from the target registers concerned, since this data could provide teachers with concrete evidence they can use to help students acquire broader and more appropriate linguistic resources.

## Reference

Biber, D., Johansson, G., Leech, Conrad, S. & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.

Ruth Breeze  
Institute of Modern Languages, University of Navarra, Spain  
<rbreeze@unav.es>

© Copyright rests with authors. Please cite TESL-EJ appropriately.