

### *Comparing two factor analyses of 19th-century English*

This study compares the results of two factor analyses carried out on a 1-million-word collection of 19th-century texts, the CONCE corpus (A Corpus of Nineteenth-Century English, see Kytö et al. 2000). The CONCE corpus is a multi-register corpus consisting of Drama, Trials, Personal letters, History, Debates, Science, and Fiction. In Biber's multi-dimensional framework, factor analysis is utilized to uncover co-occurrence patterns among linguistic features. These patterns are grouped into dimensions to allow comparisons between registers. The first analysis used a subset of Biber's (1988) features (as reported in Geisler 2004, see summary in Table 1 below), while the second will use features presented in Biber (2003).

Table 1. Functional interpretation of CONCE dimensions in Geisler (2004).  
 (highly marked registers in parenthesis)

Dimension 1	Lexical elaboration (History/Science/Debates)	Personal / interpersonal texts (Drama/Trials/Personal letters)
Dimension 2	Reduced packaging of content (Trials/Drama)	Lexical density (History/Science)
Dimension 3	Affective/emotive focus (Science/Personal letters) cf. Biber 1995: 169; Academic hedging	Non-affective/non-emotive
Dimension 4	Persuasion / argumentation (Personal letters/Debates)	Non-persuasion/non-argumentative (Trials)
Dimension 5	Narrative texts (Trials/Fiction)	Non-narrative texts (Science)

In Biber's (2003) study, the set of grammatical features was extended and refined. In particular, *that*-complement clauses and *to*-infinitives were analysed according to the type of controlling matrix predicate, and adverbs were subdivided into semantic classes. Additional features included frequently occurring lexical bundles and various semantic classes of verbs. The second multi-dimensional analysis is expected to identify new patterns of co-occurring features, and thus yield a fuller picture of nineteenth-century English.

Since the texts in CONCE are divided into three subperiods, it is also possible to trace developments among the registers across the century.

#### References

- Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. 1995. *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.
- Biber, D. 2003. Variation among university spoken and written registers: a new multi-dimensional analysis. In Pepi Leistyna & Charles F. Meyer (eds.), *Corpus analysis: Language structure and language use*. 47-67. Amsterdam: Rodopi.
- Geisler, C. 2004. *A multi-dimensional analysis of 19th-century English*. Paper presented at The Fifth North American Symposium on Corpus Linguistics, Montclair State University, May 21-23, 2004.
- Kytö, M., J. Rudanko and E. Smitterberg. 2000. Building a bridge between the present and the past: A corpus of 19th-century English. *ICAME Journal* 24: 85-97.