



# CENTRAL INSTITUTE OF INDIAN LANGUAGES

## 53<sup>rd</sup> Foundation Day Lecture Series

### Lecture 17

Monday, September 27, 2021 at 3:30 pm IST

Keywords, Clustering and Discourse Analysis -  
the case of Islam in the UK Press



#### Tony McEnery

Tony McEnery is Distinguished Professor in the Department of English Language and Linguistics at Lancaster University, United Kingdom and Changjiang Chair at Xi'an Jiaotong University, China. He has published widely on corpus linguistics and is the author of *Corpus Linguistics: Method, Theory and Practice* (with Andrew Hardie, Cambridge University Press, 2011). He has created many corpora in his career, including the EMILLE corpus of South Asian languages, created in part in collaboration with the Central Institute of Indian Languages.

This talk introduces a new method for grouping keywords and examines the extent to which it also allows analysts to explore the interaction of discourse and subregister.

Keywords have become a widely used technique in applied linguistics, in particular in discourse analysis, where it permits the investigation of large bodies of corpus data. In these studies, keywords provide a meaningful way of down-sampling and gisting discourses. This talk will introduce a new technique which permits for a slightly more systematic way of grouping keywords into the main themes that seem to dominate a discourse. This is done using the multivariate statistical technique, Multiple Correspondence Analysis, to reveal dimensions of keywords which co-occur across the texts of a corpus. These dimensions are then interpreted in terms of the discourses to which they contribute within the data, thus forming the basis of a corpus-assisted discourse analysis.

The approach is demonstrated through analysis of the discourses that are used to represent Muslims and Islam in a corpus of UK national newspaper articles published on these topics spanning 2010–2019. The approach reveals an interaction between discourse and subregister, hence this article argues for the need for (corpus-assisted) discourse analysts to account for subregister as a level of meaningful variation when analysing press discourse.

Join us via Google Meet at <https://meet.google.com/eup-czkk-eos>

LIVE Streaming on



YouTube

<https://www.youtube.com/c/CIILMysore1>