

## Workshop Report: Corpus Approaches to the Language of Literature

Authors: Ylva Berglund and Martin Wynne

### INTRODUCTION

The *Corpus Approaches to the Language of Literature* workshop took place at Oxford University Computing Services on 17-18th May 2006. The event was the first in the series of advanced workshops funded by the AHRC ICT Methods Networks, and it gathered more than 20 participants from different geographical areas, research backgrounds, and subject fields to attend a series of presentations and practical sessions over a two-day period.

The event aimed to disseminate advanced methods in linguistic analysis using linguistic corpora to researchers in literary studies.

The workshop built on networks and discussions at a workshop on Corpus Approaches to Literature held at the Corpus Linguistics 2005 conference in Birmingham and recent Poetics and Linguistics Association (PALA) conferences. These discussions resulted in a clear feeling that while there was a recognition of the potential usefulness of corpora, there are practical barriers to progress. It was decided that it would be useful to run an event which would (a) disseminate examples of exemplary work in the field, and (b) introduce in a practical way literary scholars to the techniques and methods of corpus linguistics. The workshop was supported by the Poetics and Linguistics Association, and access to their email list and website proved to be a very useful way to publicise the event to a key constituency. The workshop was the founding event for a PALA special interest group on corpus stylistics, and a short report was published in the *Parlance*, the PALA newsletter.

The workshop looked at practical ways to exploit the potential for more widespread use of corpora to study literature. Work in stylistics relies on the evidence of the language of literature. Corpus linguistics is also an empirical approach to linguistic description, relying on the evidence of language usage as collected and analysed in corpora. As linguists and stylisticians become more aware of the possibilities offered by corpus resources and techniques, then a useful exchange of ideas and methods can be facilitated.

The workshop was an opportunity to disseminate and discuss examples of successful research which has shed new light on literary texts through the techniques of corpus linguistics. Furthermore, it pointed to ways forward in demonstrating the resources and techniques necessary for such work in the future. Participants were armed with arguments, language resources, tools and methods to take back to their departments to train colleagues and to use in their research and teaching.

Discussion addressed the following topics:

- (i) the study of literary effects (or 'deviations') in texts by using the evidence of language norms in a reference corpus, including the use of collocations, colligations and semantic prosody;
- (ii) creativity in language, as identified or analysed with reference to corpus evidence;
- (iii) corpus annotation and analysis as a means of conducting a thorough and exhaustive analysis of linguistic features in literary texts;
- (iv) theoretical and practical problems with the use of corpora in literary study;
- (v) resources and techniques for the study of literature using corpora.

The techniques which were explored are also of use in teaching literature and linguistics. It was therefore possible to work with the HEA centre for English to publicise the event. HEA English expressed strong support for the venture, and a representative was present at the event.

Participants were asked to provide a short statement indicating the area of work relevant to the workshop in which they were involved, and explaining what they hoped to get out of the workshop. This information is provided as an appendix below.

## REPORT ON THE PROCEEDINGS

The event started with a short introduction to the area from Martin Wynne (AHDS Literature, Languages and Linguistics), where it was suggested that this is a relatively new field, a field where corpus linguistic methods meet and mingle with literary analysis and stylistics. The introduction was followed by a practical session where the participants had a chance to explore some key tools and techniques as well as the resources made available.

Jonathan Culpepper (Lancaster University) then described how techniques developed in corpus linguistics can be used to produce a new kind of dictionary based on usage. With illustrations from a number of case studies, he showed how he used familiar notions in corpus linguistics, such as collocation, cluster (multiword unit), keyword and grammatical and semantic annotation to examine the language of Shakespeare.

Michaela Mahlberg (University of Liverpool) further developed the notion of 'corpus stylistics' as a meeting between corpus linguistics and literary stylistics, stressing that it is not simply the application of corpus linguistic methodology to the study of style. With illustrations from her studies of the language of Dickens', she showed that if we use innovative categories to describe linguistic norms, deviations from these norms will shed new light on the way in which we analyse style in literary texts.

The third presentation was by Bill Louw (University of Zimbabwe). In his talk on 'Collocations, corpora and criticism' he provided novel illustrations and inspirational examples of how to look at collocations when examining literary works. He suggested that 'collocation has begun to offer proof of its ability to produce tangible results which exceed the results of close reading and far outstrip approaches fettered by grammar and syntax alone'.

The practical, hands-on sessions were an important part of the workshop. Each presentation was followed by a practical session where the workshop participants were given an opportunity to explore the methods used by each presenter. The sessions were structured to allow a natural progression, from simpler to more complex methods and they were followed by general discussions where problems and success stories were shared. To allow not only the participants to benefit from these, the workshop material (abstracts of presentation, exercises and practical guides) will be made available on the workshop webpage (<http://www.ahds.ac.uk/litlangling/events/approaches/home.htm>) and via the Methods Network webpages.

## OUTCOMES

The workshop was viewed as a success by all organisers and participants, and excellent feedback was received, both through the feedback forms completed by participants, and the many votes of thanks and encouragement received verbally and by email, which still continue to arrive.

This event led directly to a one-day pre-conference workshop at the PALA conference in Joensuu in Finland, at which more PALA members and other international scholars were able to participate.

The exercises and presentations are currently online and freely available. Work is ongoing to improve the accessibility and sustainability of the online resource.

The workshop was successful in linking and 'joining up' various services supporting academic work in the UK, via the involvement of AHDS, Methods Network and HEA. The workshop was also successful in raising the profile of these services and in building links between organisations and communities in linguistics, literature, stylistics and humanities computing. The workshop itself only lasted two days, but the aspiration is that it will be the beginning of future fruitful collaboration and exchange of ideas. Among the possible initiatives that were suggested was a special group for new post-graduates in the field, and the development of some materials to exploit the pedagogical potential of some existing resources. Participants who were interested continuing discussion of the issues raised in the course were invited to join the 'corpus-style' email list (<http://www.jiscmail.ac.uk/lists/corpus-style.html>).