# Corpus-Assisted Discourse Studies (CADS)

## Video 4 transcript

Full resource, see: <https://www.ncrm.ac.uk/resources/online/all/?id=20855>

Hi. Welcome to the final video in this online tutorial. In this video I’m going to talk about a recent project I’ve been involved in with colleagues Sally Hunt and Sarah Mukherjee, where we’re looking at young adult fiction and in particular comparing the writing of a group of female authors with that of male authors.

In this project we aim to compare the writing in order to see what differences exist in the fictional worlds created by each author group. We thought this is an important area because young people often gather their views of the world and have insights into the world through this reflective space created by YA fiction. As well as a corpus analysis, we’re also interested in what the young readers themselves think about our findings, so we went into a secondary school and talked to students in focus groups.

The corpus comprises 50 best-selling YA books, which were best selling in a five-year period 2017 to 2022. Here are the top 10 of those 50 best-selling books. You might recognise some of them from TV adaptations or movie tie-ins.

First of all in our analysis we used Wmatrix software to pull out key semantic domains in each author corpus, and this is quite early work, so findings are still being explored. We used part of the female author corpus, the larger of the two, and we couldn’t upload it all into Wmatrix, so these are partial findings. So, we compared part of the female author corpus with all of the male author corpus. And these are the key semantic domains. So, if you look in the second column on the left under domain, these are the semantic areas which were key in the female author corpus when compare to the male-authored one. And the top one, the domain unimportant is comprised of lexis or words such as regardless, minor, meaningless, which on their own may not be a keyword, but taken together as a category they are a key category or domain.

One thing we found of interest and which I’m going to focus on here is the category in the male author corpus of warfare, defence and the army. So, to explore this further, here are a few of the words that occur in the category of warfare in the male-authored corpus. And if you look through this, some of them are literal usage, but some of them appear to be metaphorical, for example, take a shot on me , I’m a ticking time bomb, shoot that dream in the foot. And we thought this was an interesting area to explore further. Maybe we could do that by taking a lot of concordance lines and categorising them, looking at the proportion of actual literal references and metaphorical ones to see what proportion of the language is figurative. So, that’s further work we can do. And you can see here how sematic domain analysis provides a sort of way in or entry point to exploring these corpora further.

Alongside this corpus analysis we went to one secondary school, we plan to do more, and we talked to students in six different focus groups, and we asked them… first of all we showed them our findings and we asked them does it surprise you that female and male authors wrist about different things? I’ll give you a moment to read some of the things they said. So, these are just a few students, but it’s interesting that they had these responses. We felt the students were quite insightful and helped to confirm our findings and to give us further areas to explore, and that work is ongoing. Thank you.

National Centre for Research Methods (NCRM)  
Social Sciences  
Murray Building (Bldg 58)  
University of Southampton  
Southampton SO17 1BJ  
United Kingdom

**Web** www.ncrm.ac.uk   
**Email** info@ncrm.ac.uk  
**Tel** +44 23 8059 4539  
**Twitter** @NCRMUK