The Collocates of Women in Liberal Versus Politico-Islamic Discourse

Author: Emad Soliman Mohamed
Institution: Carnegie Mellon University in Qatar, Doha, Qatar
Email: emohamed@qatar.cmu.edu

The status of women in the Middle East has always been a controversial issue, but there does not seem to have been quantitative studies of how the political and religious attitudes reflect on the presentation of women in the Arabic media.

The current study examines the way women are presented in two websites of general circulation with the purpose of unearthing the linguistic connections of "woman" in liberal versus politico-Islamic discourse.

The study uses computational and corpus linguistics techniques to investigate a corpus of 40 million words extracted from the political Islamic website http://www.ikhwanonline.com and the liberal website www.almasry-alyoum.com in search for the collocates of the word "woman" in all its possible morphological representations. Due to the morphological richness of Arabic, the words were first stemmed, i.e. all the clitics were removed from the words, before all words within the range of 5 words before and after "woman" were subjected to collocational analysis using a combination of pointwise mutual information and raw frequency counts as a measure of association. The top 100 stems in each corpus source were selected for comparison.

The results of the study give a strong support to the conclusion that liberal and Islamic websites treat women differently. While the liberals stress political participation, rights, and work equality, the Islamists stress home, marriage, pregnancy and children. While the results are not new in the sense that this may sound like common knowledge, the study provides quantitative analysis and proof of the results.

We have provided quantitative analysis of large amounts of data in the study of a linguistic and humanities filed, an area that is still lacking in Arabic linguistic research. We hope to augment this with qualitative research in the future over the same data sets.