Visualising gendered language in Shakespeare and modern texts
By Gordon Rugg

Background: Patterns of use in gendered language can provide many insights into a text and its world. This article looks at gendered language in Shakespeare, and shows how the same approach can be used to check for gender bias in texts.

This article was originally posted on July 13, 2012

To what extent does the language of Shakespeare's plays indicate a male-dominated world? One way to see is by looking at the distribution of gendered words within the texts.

The figure below shows the location of the words *he, him, his, she* and *her* in *Midsummer Night's Dream*. Each of the tiny rectangles represents a word in the text; the coloured words represent the keywords, and the blank rectangles represent the other words. This representation ignores linebreaks in the original text. The images are roughly equivalent to a miniaturised image of the text laid out as a scroll, with the keywords marked with coloured highlighter.

In each pair of images below, the nominative forms of the keywords are in red, and the other forms such as accusatives are in green, to show whether one gender appears more often in an active role.

(Apologies to any readers who are red/green colour blind; the Search Visualiser software itself takes account of colour blindness in its options, but shrinking the images down to fit into blog format loses the contrast.)

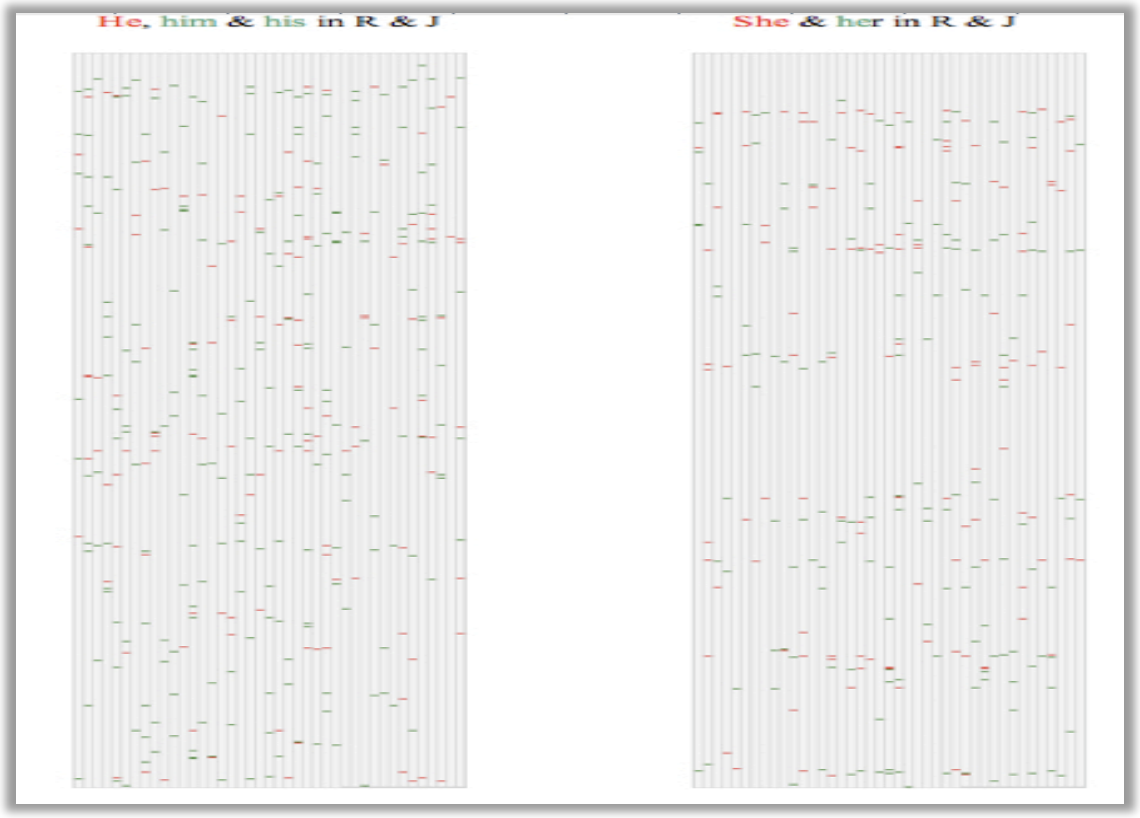
The column on the left shows the distribution of the words *he/him/his* in *Midsummer Night's Dream*. The column on the right shows the distribution of the words *she/her* in the same play.

Gendered language in Midsummer Night's Dream.



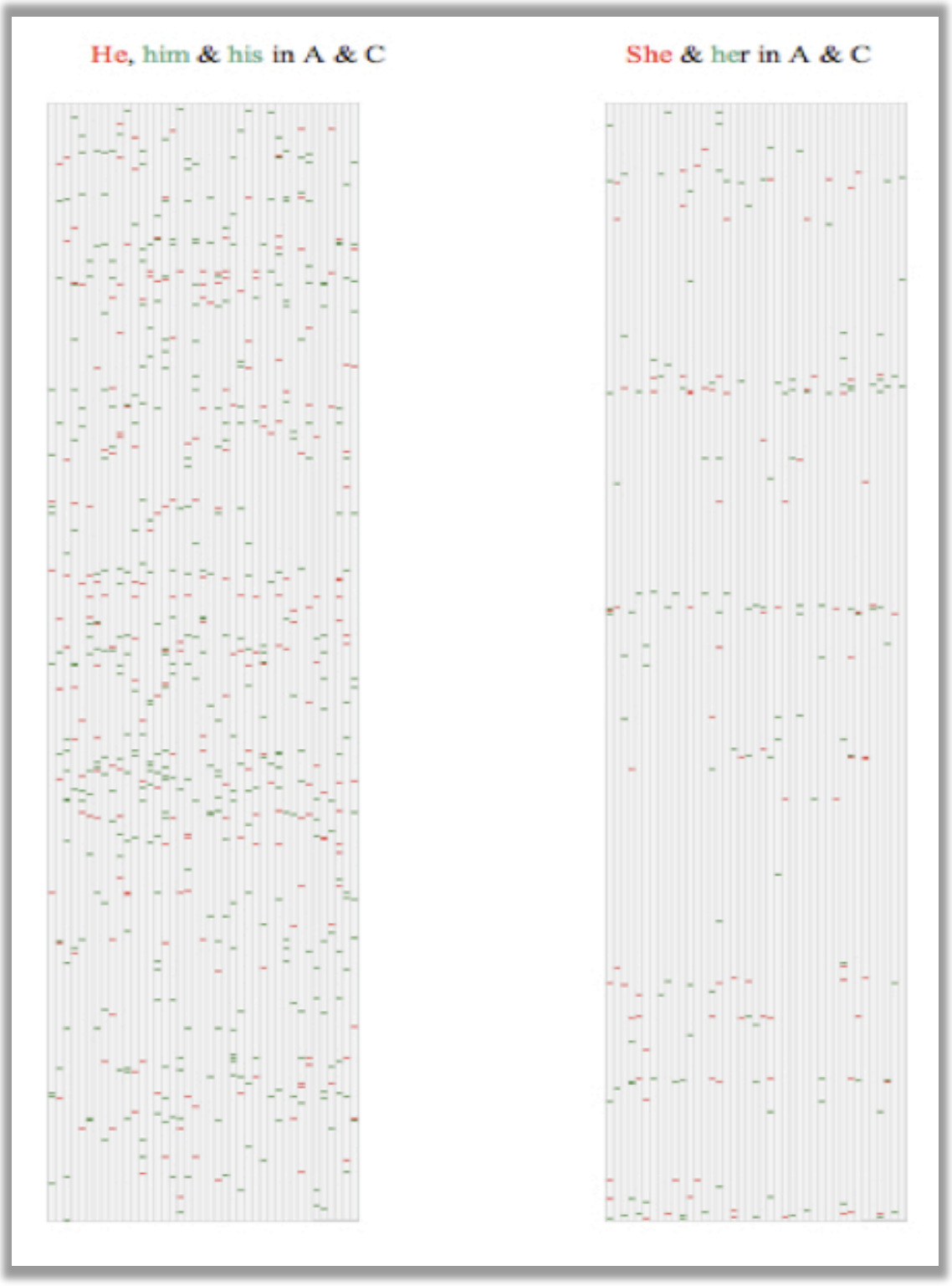
There are more male pronouns, but the difference is not huge; both male and female pronouns occur frequently throughout the play.

Here, for comparison, is the corresponding figure for *Romeo and Juliet*.



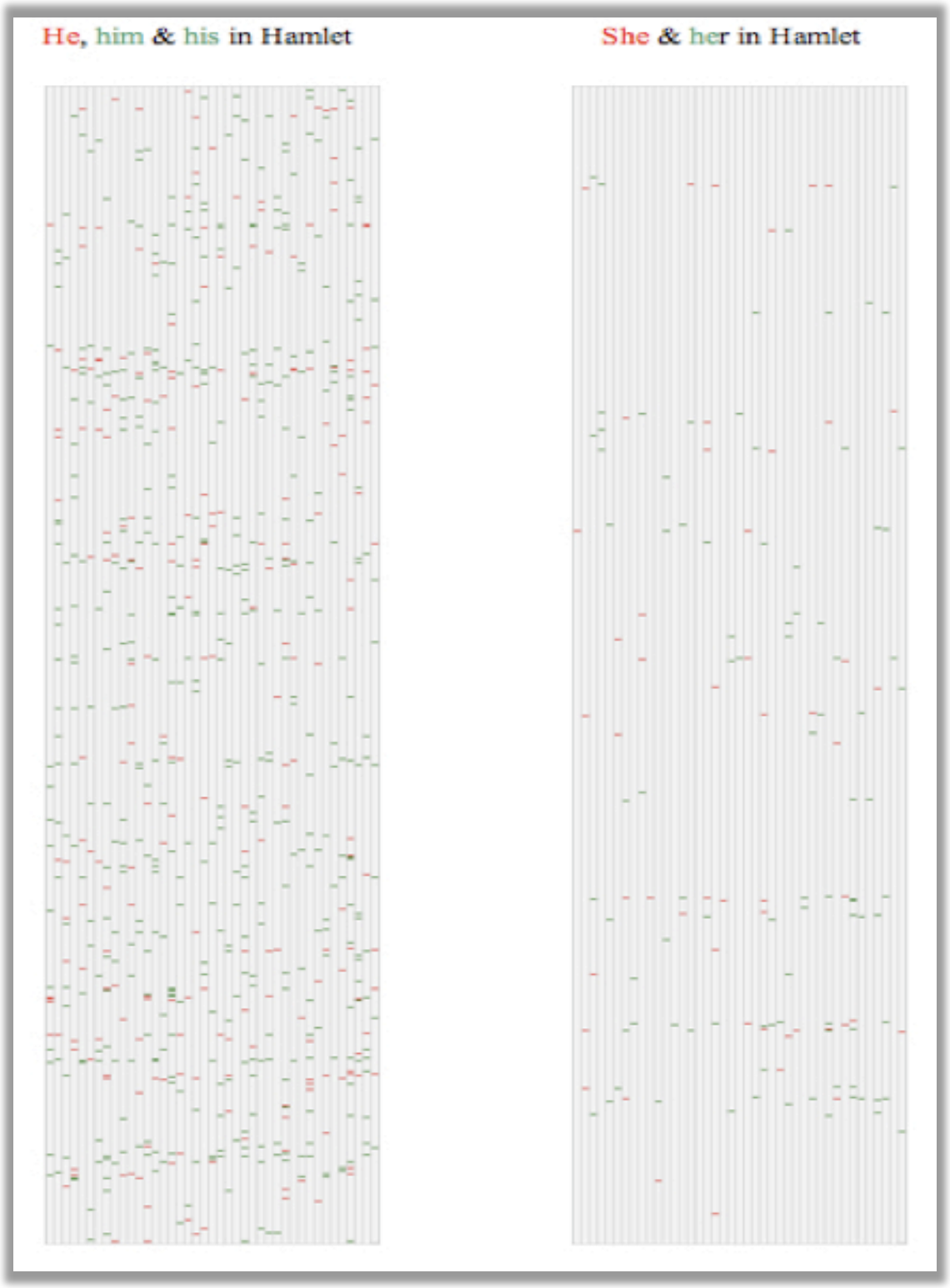
Again, there are more male pronouns, but the difference is more marked, and there's one section of the play which contains no female pronouns. This isn't a simple artefact of the plot device of Juliet's death. A comparison with another Shakespeare love story makes the point clearly.

The figure below shows the distribution of male and female pronouns in *Antony and Cleopatra*.



In *Antony and Cleopatra*, there are many more male pronouns than female, and there are several stretches of the play which contain no female pronouns. *Antony and Cleopatra* is about power politics, as well as a love story. The next two figures show results from two Shakespeare tragedies about power politics, both with strong female characters, but where love is not a central theme.

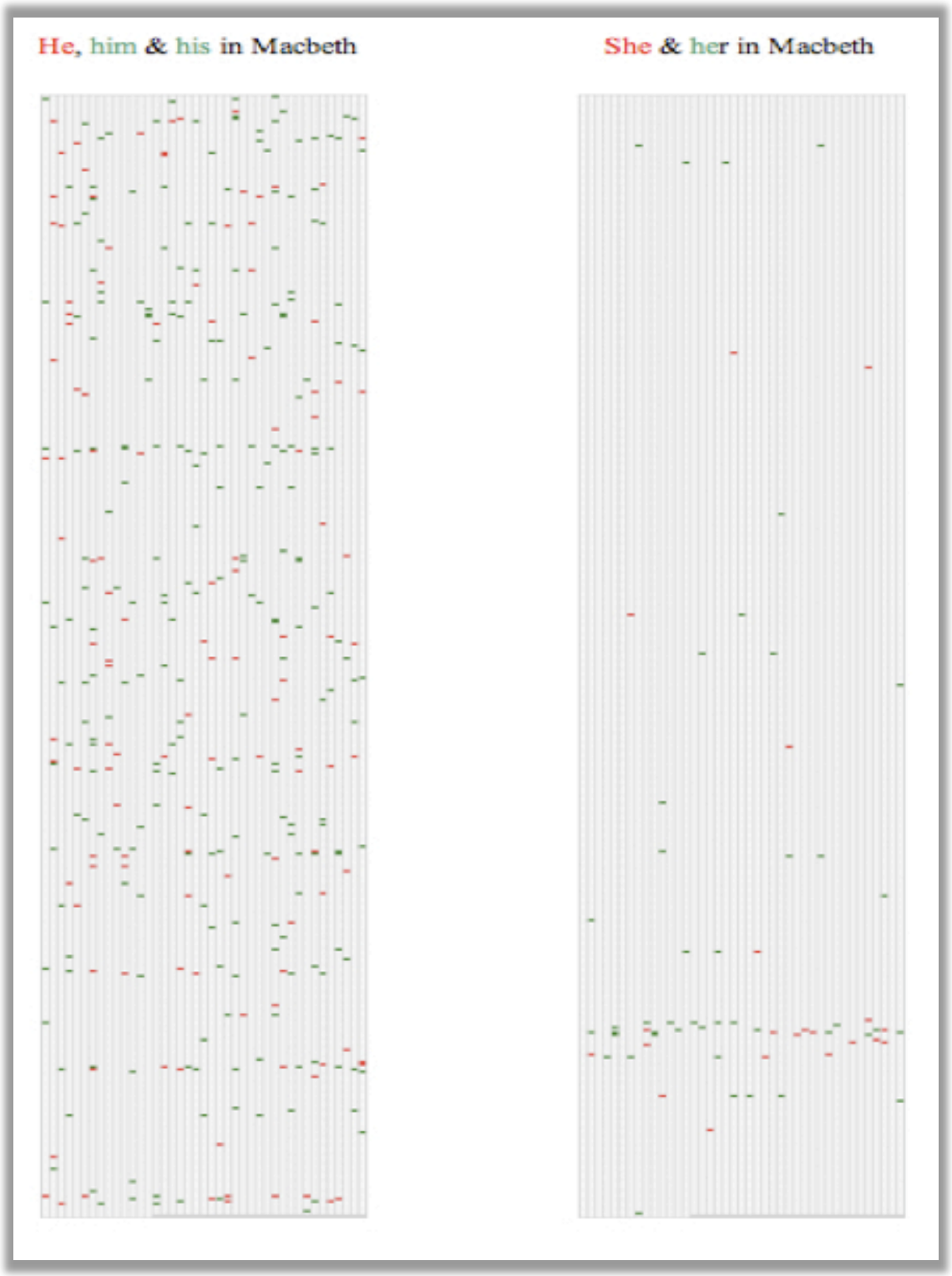
The following figure shows the results from *Hamlet*.



This play shows the same trend as *Antony and Cleopatra*, but taken further. Another play goes further still.

The next figure shows results from *Macbeth*.

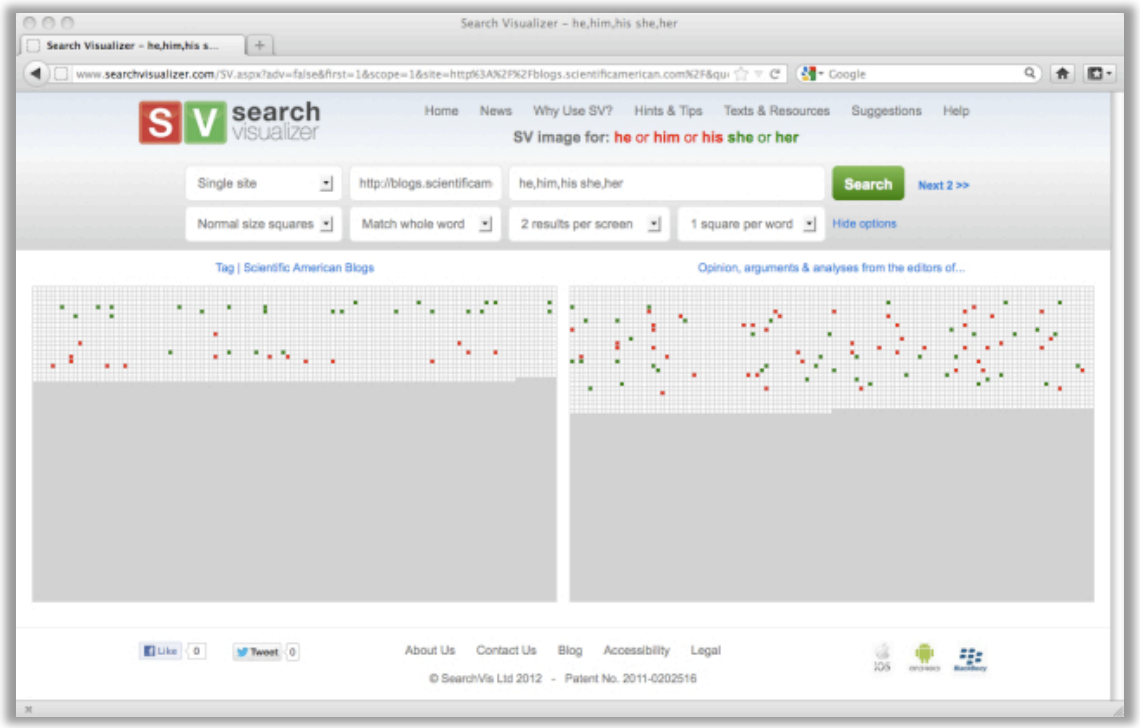
Gendered language in Macbeth.



Although Lady Macbeth is one of the most famous female characters in drama, the play contains hardly any female pronouns, and most of those are clustered toward the end.

How do the findings above compare with modern patterns of gendered language?

Here, as a closing comparison, is a screenshot of the first two records that came up when I searched the Scientific American blog site (http://blogs.scientificamerican.com/) using the "single site" option, for the same gendered words. This time, *he/him/his* appear in red, and *she/her* appear in green.



In terms of gendered language, these are much more balanced than any of the Shakespeare texts.

They're also an interesting comparison because they show two different ways of reaching the same end point. One has a clearly defined structure, with a stratum of *she/her* and then a stratum of *he/him/his* in a way that implies a deliberate attempt to achieve balance. The other has the two sets of terms intermingled throughout the text. Not all blogs achieve this level of balance, but these two articles show that it can happen, and show ways that it can be done.

Conclusion

Displaying where key terms occur within a text can give swift and powerful new insights into a text and into a body of texts.

This approach has the advantage of showing relative frequencies etc swiftly and easily. It also has the advantage of showing distributions of words within a text, making it easier for the researcher to examine how vocabulary reflects thematic structure.

We're working on other analyses of texts, ranging from great literature to popular fiction and blogs. We'd be glad to hear from anyone working in this area who has suggestions, questions or experiences involving insights that this approach can offer.

Further information and technical notes

The Search Visualiser is available for online use, free, at:
www.searchvisualiser.com

The "classic texts" section of the site contains the five Shakespeare plays above, plus other texts, and articles about using the Search Visualiser.

The searches above all used the "whole word" option, as opposed to the "partial match" option, to avoid false positives from words such as "the" which contain "he".