

# Tackling Meaning and Aboutness

## Relevant URLs:

English get started guide:

[http://www.lexically.net/wordsmith/step\\_by\\_step\\_guide/index.html](http://www.lexically.net/wordsmith/step_by_step_guide/index.html)

British National Corpus and Guardian newspaper wordlists:

<http://www.lexically.net/wordsmith/index.html>

Shakespeare corpus:

<http://www.lexically.net/wordsmith/index.html>

## Exercise

1. Download the Shakespeare corpus and the 2 big wordlists to your **c:\temp** drive.
2. Extract each into some sensible place (for example the 2 wordlists would be best in **c:\temp\wsmith5\wordlist**, wouldn't they, and maybe the corpus could go into **c:\temp\wsmith5\texts**).
3. Make a wordlist of all the Shakespeare plays. Only the plays, not the character files! Save it as **all\_shakespeare\_plays.lst**.
4. Make a batch of wordlists of the plays: one wordlist per play.
5. By now you should have 38 wordlists altogether. (They might be 1 as a normal wordlist (**all\_shakespeare\_plays.lst**) plus 37 zipped up into one big **.zip** file. If you have the 37 zipped, then please extract them so you have individual wordlists.)
6. Choose 1 play and look at the frequency ordered wordlist, skim down to the first lexical word (as opposed to function word).
7. Check out the wordlist of all the plays **all\_shakespeare\_plays.lst** and look down to find the first lexical word there.

Findings

.....

.....

.....

8. Start the KeyWords tool and make a keywords list of the play you chose in step 6. Use **all\_shakespeare\_plays.lst** as a reference corpus. Examine the key words. With a neighbour, try to decide whether you feel these KWs really reflect the play. Are there any particular types of words appearing as “key”?
9. Are the key words that WordSmith found, really “key” in your opinion as a human reader? Which would you quibble with? Are there any you would never normally consider key but once you see them seem to be key? Which words would you have identified as key but WordSmith didn’t?

<p>Findings</p> <p>.....</p> <p>.....</p> <p>.....</p>
--

10. Now make 2 more KW lists of the same play using the Guardian then BNC as reference corpora.
11. Compare the 3 KW lists. Can you get them lined up nicely on the screen so you can see all 3 at once?

<p>Findings</p> <p>.....</p> <p>.....</p> <p>.....</p>
--

12. Go back to using **all\_shakespeare\_plays.lst** as a reference corpus. Adjust the KeyWords settings so the p value is smaller and the minimum frequency is higher. What effect does this have on new keywords computations?
13. Now let’s try to find the KW clusters of Shakespeare. Make an index of all the Shakespeare plays, compute 2-4 word clusters with a minimum frequency of 2. Save it.
14. Now do the same with the play you chose in step 6. And finally use the KeyWords tool to compute the KW clusters. You may find altering the p value to 0.001 or even 0.01 is helpful.

<p>Findings</p> <p>.....</p> <p>.....</p> <p>.....</p>
--