Syntax 2: prenominal adjective order

Wulff (2003) is a corpus-based study on the factors governing the order of two prenominal adjectives in English. She starts out from the observation that, for example, *the big red ball* is preferred over *the red big ball* and reviews as well as tests a variety of variables that have been claimed to influence prenominal adjectives on the basis of corpus data. One of the many variables that has been claimed to be relevant is the length of the two adjectives. More specifically, it was claimed that, just as in many other cases, short elements precede long elements. In this section, we will test this claim, operationalizing lengths of adjectives as lengths in letters. Again, this is a phenomenon that cries out for corpus-linguistic analysis especially since even the lengths of the adjectives – at least when operationalized as suggested – can be automated.

Assignment 1

On the basis of the above characterization, formulate the alternative hypotheses and null hypotheses in text form and in statistical form.

When you are done, load the file `<C:/_qclwr/_scripts/syntax_2-assignment1_adj-adj.r>` and compare your solution with it.

Assignment 2

Write a script that has the following characteristics and performs the following operations:

(i) The script prompts the user to choose corpus files to search (`<C:/_qclwr/_inputfiles/corp_bnc_sgml *.txt>`, files 1 to 4).

(ii) The script loads the files, retrieves only those lines tagged as sentences and removes unwanted annotation.

(iii) The script finds all occurrences of two words that are tagged (either with simple tags or with some part of a portmanteau tag) as adjectives ("AJ."; cf. the Appendix) followed by a word tagged as a noun (use "<w·N..>") in the files in question.

(iv) The script retrieves for the adjectives from each pair (a) the adjectives, (b) their positions in a pair, and (c) their lengths (I leave it up to you how many data structures etc. you use for that).

When you are done, load `<C:/_qclwr/_scripts/syntax_2-assignment2_adj-adj.r>` and compare your solution with them.

Assignment 3

Explore the data statistically in R such that you

(i) represent the distribution of adjective lengths in the two positions (first vs. second) graphically;

(ii) analyze the distribution statistically and interpret the findings.
When you are done, load `<C:/qclwr/scripts/syntax_2-assignment3_adj-adj.r>` as well as
`<C:/qclwr/outputfiles/syntax_2_assignment3-plot1.png>`,
`<C:/qclwr/outputfiles/syntax_2_assignment3-plot2.png>`,
`<C:/qclwr/outputfiles/syntax_2_assignment3-plot3.png>`, and
`<C:/qclwr/outputfiles/syntax_2_assignment3-plot123.png>` and compare your solution with them.