**Guidelines for writing a term paper**

**General comments**

Your paper should demonstrate the following:

1) you can use R to extract and analyze corpus data. Your paper should contain the data in the form of tables and graphs. You can add the R code in an appendix;

2) you can interpret the results and go back to the original data. Each type of analysis should be accompanied by an interpretation and linguistic examples from the corpus. It is important to strike a balance between the quantitative observations and their linguistic interpretation;

3) you can express your thoughts clearly and use the terminology correctly.

Note: corpus (sg.) – corpora (pl.)!!!

Lemma (sg.) – lemmas or lemmata (pl).

**Evaluation criteria**

**There is no minimum or maximum number of pages.**

1. A well-formulated, reasonable research question: **10** **points** maximum

2. Corpus (attached to the file), large and clean (**motivate your choice!**): **20** **points** maximum

3. Analysis of frequencies (lemmas or wordforms): **10** **points** maximum

4. Analysis of frequencies by different parts of speech: **10** **points** maximum

5. Visualizations of frequencies: **10** **points** maximum

6. Keyword analysis: **15** **points** maximum

7. Clustering of texts and interpretation of common and distinctive words/lemmas: **15** **points**

8. Ngram analysis (2-, 3- and 4-grams): **15** **points** maximum

9. Analysis of collocations: **10 points** maximum

10. A convincing theoretical generalization of the results: **10** **points** maximum

11. Critical analysis of parsing mistakes and their correction and re-analysis: **10 points** maximum

Points and grades:

95-100 points 1

90-94 points 1.3

85-89 points 1.7

80-84 points 2

75-79 points 2.3

70-74 points 2.7

65-69 points 3

60-64 points 3.3

55-59 points 3.7

50-54 points 4

less than 50 points 5 (FAIL)

**Submission procedure**

Submission can occur only electronically. Send your paper to [natalevs@gmail.com](mailto:natalevs@gmail.com)

Either attach the corpus as a separate file, or make one archive together with the term paper. If the corpus is very large, please compress it first.