Assignment Guide for Mini-Project 2
Introduction to Methods in Corpus Linguistics

Purpose of project

To demonstrate functional knowledge of common analyses in corpus linguistics including:
a) Conducting a lexical bundle analysis
b) Conducting a collocation analysis

Procedure:
Step 1: Corpus description

You can use any corpus (other than the Brown corpus). You will report the following information about the corpus you are using

Be sure to report:
a) The language use domain that your corpus represents
b) The number of words in your corpus
c) The number of documents included in your corpus
d) How you collected the data
e) Benefits and limitations of your corpus

Step 2: Generate a list of clusters/bundles/n-grams

Using AntConc, generate of list of clusters/bundles/*n*-grams. Be sure to report what your frequency, range, and size thresholds were. Include a table with the twenty most frequent items, including the following information:
a) Rank
b) Normed frequency
c) Normed range (percent of texts in which the item occurs)

Step 3: Conduct a use analysis of a cluster/bundles/n-gram

Choose a relatively frequent item from your list and report on its use. In particular, report the following information:
a) Meaning(s) and/or function(s) of the bundle
b) Frequency of each meaning/function
c) Frequency by register (if applicable)

Step 4: Generate a collocates list

Using AntConc, generate a collocate list for a frequent noun or verb using both MI and T. Be sure to set the minimum frequency to at least 5 and to use a search span of 5L to 5R. In two

separate tables (one for MI and one for T), report the top ten collocates for the noun or verb you chose. Be sure to report the following:

a) Total frequency
b) Frequency to the left
c) Frequency to the right
d) Statistic (MI score or T score)

Step 5: Conduct a use analysis of two collocates

Choose two collocates from your list and report on its use. In particular, report the following information:

a) Syntactic/grammatical relationship between the node word and its collocate
b) Typical meaning(s) and/or function(s) of the collocation

Mini-Project 2

Introduction
<insert introduction here – provide an overview of the project and indicate why the analysis you are conducting might be interesting/useful>

Corpus Description
<Describe characteristics of your corpus here. Be sure to outline the target language use domain and the purpose of the corpus.>

Bundle Analysis Lists
<Include a prose description of your criteria for selecting bundles. Include your lists here. Additionally, submit a supplementary file with the complete list when you submit your report.>

Bundle Use Analysis
<Report the findings of the use analysis here.>

Collocation Lists
<Include a prose description of your criteria for selecting collocations. Include your two lists of twenty items here. Additionally, submit a supplementary file with the complete list when you submit your report.>

Collocation Use Analysis
<Report the findings of the use analysis here.>

Conclusion
<summarize your results here>