**British National Corpus Baby Word Count**

**CMIST**

**University of Manchester**

**John Mcloughlin**

**September 2016**

Extract from corpus via xsltproc

Split into 10 files for loading into NVIVO. NVIVO can’t load full corpus. Have to split.

Load ten files into NVIVO.

NVIVO cannot export more than 65536 rows to an Excel file. In this case the output is smaller than the Excel limit, so an output Excel file can be created from the NVIVO word count query.

To create bar charts, use grep ( Linux/ Mac ) or FINDSTR ( Windows ) to extract the required word from the CSV file. EG

grep teach BNC\_Baby\_word\_count.csv > teach.csv

This will extract values for all words related to the stem ‘teach’ and create a new CSV file, teach.csv. Load this into Excel and create a bar chart. Save as BNC\_Baby\_teach\_bar.xsls