***First Pass Modeling of the ANES Text Data***

This initial analysis considers the office recognition questions in the ANES in which respondents were asked to identify four people from the political sphere: Gordon Brown, Dick Cheney, Nancy Polosi , and John Roberts. The analysis shown here uses only the data from the interviewer transcript and the scoring codes associated with the interview transcript.

As a first ‘reality check’, the following four histograms enumerate the word tokens in the ‘somewhat parsed’ interviewer transcript. All have a highly skewed, geometric shape.



graph the number of word tokens on the number of assigned codes. Each point represents a respondent.

