

## WSWG 28

11 October 2014

Campus Center, University of Massachusetts at Amherst

### *Stylometrics and the Question of Interpolation in Paul*



#### *Explaining BIRD*

BIRD (the Brooks Index of Rhetorical Difference) is a set of measures of stylistic similarity and difference. They are based on the Significance (S) of the variations from normal usage of 14 function words (which reflect the continuity rather than the content of a message). S is calculated so as to be self-interpreting: an S value of 1.00 indicates that there is 99% assurance that the departure from general-frequency expectation is significant; that is, not a result of random variation. This is the statistical standard for situations when accuracy is especially important. From the S values for the 14 words, the following measures are calculated, for the text or texts under study:

**N** or Neutrality, the degree to which the 14 words together depart from general-frequency expectation. Minus S values may be canceled out by plus S values, leaving a value of 0.00 (perfect conformity with general frequency). Higher N values indicate increasing departure from that norm, in one direction or the other. Plus values indicate more than average use of connectives, implying a logical or otherwise closely organized style. Minus values imply greater reliance on simple statements, aphorisms, quotations, or the like.

**V** or Volatility, the total amount of variability, whether or not plus values are compensated by minus values. High V numbers imply tension, strong engagement with the subject, or other varieties of emotional disturbance; low ones, the opposite. Both N and V are measures of the style of *single* texts or text segments.

**D** or Difference, a measure of the fit between *two* texts. Low values imply that, whether neutral or not in general terms, their profiles tend to coincide, a zig in one text being matched by a zig in the other. High values suggest that, however similar their N or V readings may be, the two texts reach that result by different routes.

D results are not to be taken as direct indications of same or different authorship. It has been found that the range of variation *within* an author can exceed the range of variation *between* authors. Paul, with whom this conference is especially concerned, is a highly variable author, sometimes ingratiating, sometimes expounding a position, sometimes defending his Apostolic credentials, sometimes accusing his betrayers. These states of mind tend to produce different expression, and thus to yield different D numbers. Nor it is impossible that Paul's editors, or those who later wrote under Paul's name, could imitate his style sufficiently to pass these tests.

This is not to say that the results have no value for authorship questions, only that they must be weighed understandingly by those who interpret them. In general, based on the work of calibration so far, it appears that D values below 0.5 are consistent with a hypothesis of same authorship, those between 0.5 and 1.0 are ambiguous, and those above 1.0 are more supportive of a different-author than of a same-author hypothesis.

The terms of any analysis are set by a human investigator, and the results of the analysis are referred to human investigators for interpretation and decision, in the light of other available information. This is as it should be. There are no magic formulas in this area. On reflection, it would be a strange world if there were.