Burrows's Delta verstehen

Stilometrische Verfahren der Autorschaftsattribution haben eine lange Tradition in den digitalen Geisteswissenschaften: Mit der Analyse der Federalist Papers durch Mosteller und Wallace (1963) konnten schon Anfang der 1960er Jahre Erfolge verzeichnet werden. Überblicksbeiträge von Patrick Juola (2006) und Efstathios Stamatatos (2009) belegen die Vielfältigkeit der Bestrebungen, stilometrische Verfahren für die Autorschaftsattribution einzusetzen und weiterzuentwickeln. Ein jüngerer Meilenstein der stilometrischen Autorschaftsattribution ist ohne Zweifel das von John Burrows (2002) vorgeschlagene "Delta"-Maß zur Bestimmung der stilistischen Ähnlichkeit zwischen Texten. Die beeindruckend gute Performance von Delta in verschiedenen Sprachen und Gattungen sollte allerdings nicht darüber hinwegtäuschen, dass die theoretischen Hintergründe weitgehend unverstanden geblieben sind (Argamon 2008). Anders ausgedrückt: Wir wissen, dass Delta funktioniert, aber nicht, warum es funktioniert.

In diesem Vortrag stellen wir den aktuellen Stand der Forschung in der stilometrischen Autorschaftsattribution mit Delta und seinen Varianten vor und berichten neue Beobachtungen und Erkenntnisse aus eigenen Untersuchungen.

Stefan Evert, Thomas Proisl

FAU Erlangen-Nürnberg

Professur für Korpuslinguistik