

## Recommended Readings in Stylometry

v 0.3, April 12, 2017

The following list of recommended readings in Stylometry contains four types of entries: papers that have proven to be milestones in the history of stylometry; survey papers, overviews and introductions that help newcomers get a grip of the history and structure of the field; important papers on more specific methodological aspects; and applications of stylometry to problems of authorship or other issues that demonstrate best practices or showcase the usefulness of stylometry in various areas.

This list is being published by the 'Stylometry Group' at <https://github.com/christofs/stylometry-bibliography>. See also our online, sortable, searchable, machine-readable, and evolving *Stylometry Bibliography* on Zotero: [https://www.zotero.org/groups/stylometry\\_bibliography/items](https://www.zotero.org/groups/stylometry_bibliography/items). Contact: Christof Schöch, [c.schoech@gmail.com](mailto:c.schoech@gmail.com).

---

### ***Milestones in the history of stylometry***

Lutoslawski, Wincenty. "Chapter III: The Style of Plato." In *The Origin and Growth of Plato's Logic: With an Account of Plato's Style and of the Chronology of His Writings*, 64–193. London: Longmans, Green and Company, 1897. <https://archive.org/details/originandgrowthoolutogoog>.

Mosteller, Frederick, and David L. Wallace. "Inference in an Authorship Problem." *Journal of the American Statistical Association* 58, no. 302 (1963): 275–309. <http://www.jstor.org/stable/2283270>.

Burrows, John. "'Delta': A Measure of Stylistic Difference and a Guide to Likely Authorship." *Literary and Linguistic Computing* 17, no. 3 (2002): 267–87. <https://doi.org/10.1093/lc/17.3.267>.

### ***Survey papers and introductions***

Chaski, Carole E. "Who Wrote It?: Steps Toward a Science of Authorship Identification." *National Institute of Justice Journal*, no. 233 (1997): 15–24. <https://www.ncjrs.gov/App/publications/abstract.aspx?ID=184604>.

Craig, Hugh. "Stylistic Analysis and Authorship Studies." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 273–88. Oxford: Blackwell, 2004. <http://www.digitalhumanities.org/companion/>.

Delcourt, Christian. "Stylometry." *Revue belge de philologie et d'histoire. Belgisch tijdschrift voor philologie en geschiedenis* 80, no. 3 (2002): 979–1002. [http://www.persee.fr/doc/rbph\\_0035-0818\\_2002\\_num\\_80\\_3\\_4651?q=stylometry%20delcourt](http://www.persee.fr/doc/rbph_0035-0818_2002_num_80_3_4651?q=stylometry%20delcourt).

- Stamatatos, Efstathios. "A Survey of Modern Authorship Attribution Methods." *Journal of the Association for Information Science and Technology* 60, no. 3 (2009): 538–556. <https://doi.org/10.1002/asi.v60:3>.
- Holmes, David I. "Authorship Attribution." *Computers and the Humanities* 28, no. 2 (1994): 87–106. <https://www.jstor.org/stable/30200315>.
- Holmes, David I. "The Evolution of Stylometry in Humanities Scholarship." *Literary and Linguistic Computing* 13, no. 3 (1998): 111–17. <https://doi.org/10.1093/lc/13.3.111>.
- Hoover, David L. "Quantitative Analysis and Literary Studies." In *A Companion to Digital Literary Studies*, 517–33. Oxford: Blackwell, 2008. <http://www.digitalhumanities.org/companionDLS/>.
- Rudman, Joseph. "The State of Authorship Attribution Studies: Some Problems and Solutions." *Computers and the Humanities* 31, no. 4 (1998): 351–65. <http://www.jstor.org/stable/30200436>.
- Juola, Patrick. "Authorship Attribution." *Foundations and Trends in Information Retrieval* 1, no. 3 (2006): 233–334. <http://www.mathcs.duq.edu/~juola/papers.d/fnt-aa.pdf>.
- Koppel, Moshe, Jonathan Schler, and Shlomo Argamon. "Computational Methods in Authorship Attribution." *Journal of the American Society for Information Science and Technology* 60, no. 1 (2008): 9–26. <http://onlinelibrary.wiley.com/doi/10.1002/asi.20961/full>.
- Love, Harold. *Attributing Authorship : An Introduction*. Cambridge: Cambridge University Press, 2002.
- Milic, Louis T. "Progress in Stylistics: Theory, Statistics, Computers." *Computers and the Humanities* 25, no. 6 (1991): 393–400. <https://www.jstor.org/stable/30208122>.
- Stamou, Constantina. "Stylochronometry: Stylistic Development, Sequence of Composition, and Relative Dating." *Literary and Linguistic Computing* 23, no. 2 (2008): 181–99. <https://doi.org/10.1093/lc/fqm029>.

### ***Papers on specific methodological aspects***

- Argamon, Shlomo. "Interpreting Burrows' Delta: Geometric and Probabilistic Foundations." *Literary and Linguistic Computing* 23, no. 2 (2008): 131–47. <https://doi.org/10.1093/lc/fqn003>.
- Baayen, H., H. van Halteren, and F. Tweedie. "Outside the Cave of Shadows: Using Syntactic Annotation to Enhance Authorship Attribution." *Literary and Linguistic Computing* 11, no. 3 (1996): 121–32. <https://doi.org/10.1093/lc/11.3.121>.
- Biber, Douglas. "The Multi-Dimensional Approach to Linguistic Analyses of Genre Variation: An Overview of Methodology and Findings." *Computers and the Humanities* 26, no. 5–6 (1993): 331–45. <http://www.jstor.org/stable/30204629>.
- Brunet, Étienne. "Peut-on mesurer la distance entre deux textes ?" *Corpus*, no. 2 (2003). <http://corpus.revues.org/30>.
- Eder, Maciej. "Visualization in Stylometry: Cluster Analysis Using Networks." *Digital Scholarship in the Humanities* 32, no. 1 (2017): 50–64. <https://doi.org/10.1093/lc/fqv061>.
- Grieve, Jack W. "Quantitative Authorship Attribution: An Evaluation of Techniques." *Literary and Linguistic Computing* 22, no. 3 (July 2007): 251–270. <https://doi.org/10.1093/lc/fqm020>.

- Rybicki, Jan, and Maciej Eder. "Deeper Delta across Genres and Languages: Do We Really Need the Most Frequent Words?" *Literary and Linguistic Computing* 26, no. 3 (2011): 315–21. <https://doi.org/10.1093/lc/fqr031>.
- Smith, Peter W. H., and W. Aldridge. "Improving Authorship Attribution: Optimizing Burrows' Delta Method." *Journal of Quantitative Linguistics* 18, no. 1 (2011): 63–88. <https://doi.org/10.1080/09296174.2011.533591>.
- Stamatatos, Efstathios, N. Fakotakis, and G. Kokkinakis. "Automatic Text Categorization in Terms of Genre and Author." *Computational Linguistics* 26, no. 4 (2000): 471–95. <http://dl.acm.org/citation.cfm?id=971883>.
- Stamatatos, Efstathios. "On the Robustness of Authorship Attribution Based on Character N-Gram Features." *Journal of Law and Policy* 11, no. 2 (2013): 420–440. <http://brooklynworks.brooklaw.edu/jlp/vol21/iss2/7>.
- Binongo, Jose Nilo G., and M.W.A. Smith. "The Application of Principal Component Analysis to Stylometry." *Literary and Linguistic Computing* 14, no. 4 (1999): 445–65. <https://doi.org/10.1093/lc/14.4.445>.
- Eder, Maciej. "Computational Stylistics and Biblical Translation: How Reliable Can a Dendrogram Be?" In *The Translator and the Computer*, edited by Tadeusz Piotrowski and Łukasz Grabowski, 155–70. Wrocław: WSF Press, 2012.
- Herrmann, J. Berenike, Christof Schöch, and Karina van Dalen-Oskam. "Revisiting Style, a Key Concept in Literary Studies." *Journal of Literary Theory* 9, no. 1 (2015): 25–52. <https://zenodo.org/record/50893>.
- Kestemont, Mike. "Function Words in Authorship Attribution. From Black Magic to Theory?" In *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)*, 59–66. Gothenburg, Sweden: Association for Computational Linguistics, 2014. <http://aclweb.org/anthology/W/W14/W14-0908.pdf>.
- Kestemont, Mike, Kim Luyckx, Walter Daelemans, and Thomas Crombez. "Cross-Genre Authorship Verification Using Unmasking." *English Studies* 93, no. 3 (2012): 340–56. <https://doi.org/10.1080/0013838X.2012.668793>.
- Kjell, Bradley. "Authorship Determination Using Letter Pair Frequency Features with Neural Network Classifiers." *Literary and Linguistic Computing* 9, no. 2 (1994): 119–24. <https://doi.org/10.1093/lc/9.2.119>.
- Hirst, Graeme, and Ol'ga Feiguina. "Bigrams of Syntactic Labels for Authorship Discrimination of Short Texts." *Literary and Linguistic Computing* 22, no. 4 (2007): 405–17. <https://doi.org/10.1093/lc/fqm023>.
- Rybicki, Jan, David Hoover, and Mike Kestemont. "Collaborative Authorship: Conrad, Ford and Rolling Delta." *Literary and Linguistic Computing* 29, no. 3 (2014): 422–31. <https://doi.org/10.1093/lc/fqu016>.

## **Applications of stylometry**

- Binongo, Jose Nilo G. "Who Wrote the 15th Book of Oz? An Application of Multivariate Analysis to Authorship Attribution." *Chance* 16, no. 2 (2003): 9–17. <http://dh.obdurodon.org/Binongo-Chance.pdf>
- Burrows, John. "The Englishing of Juvenal: Computational Stylistics and Translated Texts." *Style* 36, no. 4 (2002): 677–99.
- Craig, Hugh. "Jonsonian Chronology and the Styles of *A Tale of a Tub*." In *Re-Presenting Ben Jonson: Text, History, Performance*, edited by Martin Butler, 210–32. London: Macmillan, 1999.
- Craig, Hugh, and Arthur F. Kinney. *Shakespeare, Computers, and the Mystery of Authorship*. Cambridge: Cambridge University Press, 2009.
- Daelemans, W. "Explanation in Computational Stylometry." In *Proceedings of the 14th International Conference on Computational Linguistics and Intelligent Text Processing*, (Samos, Greece), edited by A. Gelbukh, 451–62. CICLing'13. Berlin / Heidelberg: Springer, 2013. [https://doi.org/10.1007/978-3-642-37256-8\\_37](https://doi.org/10.1007/978-3-642-37256-8_37).
- Dalen-Oskam, Karina van, and Joris van Zundert. "Delta for Middle Dutch—Author and Copyist Distinction in Walewein." *Literary and Linguistic Computing* 22, no. 3 (2007): 345–62. <https://doi.org/10.1093/llc/fqm012>.
- Forsyth, Richard S., David I. Holmes, and Emily K. Tse. "Cicero, Sigonio, and Burrows: Investigating the Authenticity of the *Consolatio*." *Consolatio* 14, no. 3 (1999): 375–400. <https://doi.org/10.1093/llc/14.3.375>.
- Halteren, Hans van, R. Harald Baayen, Fiona Tweedie, Marco Haverkort, and Anneke Neijt. "New Machine Learning Methods Demonstrate the Existence of a Human Stylome." *Journal of Quantitative Linguistics* 12, no. 1 (2005): 65–77. <http://dx.doi.org/10.1080/09296170500055350>.
- Hänlein, Heike. *Studies in Authorship Recognition: A Corpus Based Approach*. Europäische Hochschulschriften 14. Frankfurt am Main: Peter Lang, 1999.
- Jannidis, Fotis, and Gerhard Lauer. "Burrows's Delta and Its Use in German Literary History." In *Distant Readings. Topologies of German Culture in the Long Nineteenth Century*, edited by Matt Erlin and Lynne Tatlock, 29–54. Rochester: Camden House, 2014. [http://gerhardlauer.de/index.php/download\\_file/view/335/1/](http://gerhardlauer.de/index.php/download_file/view/335/1/).
- Jockers, Matthew L. *Macroanalysis - Digital Methods and Literary History*. Champaign, IL: University of Illinois Press, 2013.
- Juola, Patrick. "The Rowling Case: A Proposed Standard Protocol for Authorship Attribution." *Digital Scholarship in the Humanities* 30, suppl. 1 (2015): 100–113. <https://doi.org/10.1093/llc/fqvo40>.
- Stewart, Larry L. "Charles Brockden Brown: Quantitative Analysis and Literary Interpretation." *Literary and Linguistic Computing* 18, no. 2 (2003): 129–38. <https://doi.org/10.1093/llc/18.2.129>.