

Linguistic Modeling as a Technique in Forensic Authorship Attribution

The reported study was funded by RFBR, project number 19-31-27001 (19-312-90022)

expert block
=
qualitative analysis

- **Interpretative linguistics** (McMenamin G.R. Forensic Linguistics: advances in forensic stylistics);
- **Idiolect** (Coulthard M. Author identification, idiolect, and linguistic uniqueness; Vul. S.M. Forensic Authorship Examination: Methodological Basics)

computer block
=
quantitative analysis

- **sylometry;**
- **n-grams;**
- **collocation \ colligation;**
- **word \ sentence length;**
- **mathematical model of idiolect**

Параметр	Класс	
	Ω_{CO}	Ω_{CP}
	x_i	x_i
N 11	0,007022	0,141775
N 22	0,007022	0,044372
N 21	0,008427	0,045455

+ correlation matrices

Automatically made conclusion

objectivity compliance with law

q
u
a
n
t
i
t
a
t
i
v
e

q
u
a
l
i
t
a
t
i
v
e

The basis:

In the English-speaking school:

- **Labbe C.; Labbe D.** Inter-Textual Distance and Authorship Attribution. Corneille and Molière. Journal of Quantitative Linguistics;
- **De Vel O., Anderson A., Corney M. and Mohay G.** Mining Email Content for Author Identification Forensics SIGMOD Record;
- **Merriam T.** An Application of Authorship Attribution by Intertextual Distance in English;
- **Wright D.** Using word n-grams to identify authors and idiolects: A corpus approach to a forensic linguistic problem

In the Russian-speaking school:

- **Vul S.M.** Forensic Authorship Examination: Methodological Basics: Methodological Manual;
- **Komissarov A.Yu.** Forensic study of written language: textbook;
- **Goroshko E.I.** Forensic classification expertise examination: problems of establishing the gender of the author of the document.

Khomenko Anna,

Senior Lecturer, Faculty of Humanities, Research Intern,

Faculty of Informatics, **National Research University**

Higher School of Economics,

Nizhny Novgorod



the law of forensic expertise in Russia:

codes of procedure;
Federal law No. 73

the main idea –
INTEGRATION

- **Koppel, M., Schler, J.** Exploiting Stylistic Idiosyncrasies for Authorship Attribution;

- **Baranov A.N.** Introduction to Applied Linguistics: Textbook;

- **Litvinova T., Rangel F., Rosso P., Seredin P., Litvinova O.** Overview of the RusProfiling PAN at FIRE Track on Cross-genre Gender Identification in Russian