



# ePoetics

## Korpuserschließung und Visualisierung deutschsprachiger Poetiken (1770-1960) für den „Algorithmic Criticism“

### Wissenschaftliche Ansätze

- Literaturwissenschaftliche Grundlage: Hermeneutische Studie ‚A History of Poetics‘ (1).
- Algorithmic Criticism (nach Stephen Ramsay) (2): Kombination aus hermeneutischen und informationstechnologischen Textanalyse- und Visualisierungsverfahren.
- Kombination aus Distant Reading (nach Franco Moretti) (3) und Close Reading: Distanzierter Blick ermöglicht zunächst besseren Überblick, um dann gezielt Details näher untersuchen und vergleichen zu können.

### Daten

- Zentrales Textkorpus der Geisteswissenschaften: Poetiken und Ästhetiken von 1770 bis 1960 – enthalten als ‚Lehren von der Dichtkunst‘ ein systematisches Basiswissen für die Literaturwissenschaft aus der Anfangszeit der wissenschaftlichen Poetik bis zu ihrem Ende bzw. Übergang in die Literaturtheorie.
- Untersuchtes Korpus: 20 repräsentative Poetiken aus der 1000 Werke umfassenden Bibliographie der Ausgangsstudie werden als Image-Digitalisate erfasst, durch das Double Keying-Verfahren transkribiert und schließlich in durch umfangreiche (TEI-konforme) Metadaten angereicherte digitale Volltexte transformiert.

### Ziele

- Erkenntnisse über quantitative und qualitative Komplementaritäten sowie historische Zusammenhänge und Dynamiken des Korpus, die über die grundlegenden hermeneutischen Hypothesen hinausgehen und durch die Kombination von Distant und Close Reading, traditionell hermeneutischen und algorithmischen Methoden sichtbar und erforschbar gemacht werden.
- Interaktive visuelle Analysewerkzeuge werden konzipiert und entwickelt, die die schnelle Verifikation bzw. Falsifikation von hermeneutischen Hypothesen erlauben.
- Über die Visualisierung wird ein schneller Zugriff auf die ursprünglichen Textquellen ermöglicht, um Fehler, die durch die algorithmische Aufbereitung induziert werden können, schnell zu erkennen und zu beseitigen.
- Wissenschaftliche Fragestellungen sollen durch einen iterativen Prozess konkretisiert und (computer-) linguistische Analysemethoden verbessert werden.
- Integration des aufgearbeiteten Textkorpus und der entwickelten Analyse- und Annotationswerkzeuge in virtuelle Forschungs-Infrastrukturen wie TextGrid, DARIAH und CLARIN.
- Eröffnung neuer Forschungsfragen durch die Ergebnisse der Analysen und das Wissen über Methoden, Metadaten und Ontologien, von denen andere Projekte auf dem Gebiet der Digital Humanities profitieren können.
- Erweiterungsmöglichkeit auf andere Textkorpora.

**Digitale Erschließung**

- Korpusaufbau
- Metadaten-Konzept
- Einbindung in digitale Infrastrukturen (TextGrid, DARIAH)

**Algorithmic Criticism**

- Begriffsexplikation für Metadaten
- Hermeneutische Hypothesen
- Verbindung mit Analysewerkzeugen

**Korpusanalyse**

- Entwicklung Analysewerkzeuge
- Entwicklung trainierbarer Annotationswerkzeuge
- Werkzeugeinbindung in CLARIN

**Visualisierung**

- Entwicklung visueller Analysewerkzeuge
- Interaktive Explorationsmethoden
- Visual Analytics-Techniken

### Methoden

- Nutzung von Synergieeffekten der qualitativen und quantitativen Analyseverfahren zur Weiterentwicklung der Ansätze auf unterschiedlichsten Gebieten:
- Überprüfung und Weiterentwicklung hermeneutischer Hypothesen
  - sprachtechnologische Korpusanalyse (Analyse der Sprache an sich)
  - statistische Verfahren (Berechnung quantitativer Aussagen zur Unterstützung hermeneutischer Theorien)
  - Textmining (Named Entity Recognition, Erkennung von benutzerdefinierten Konzepten)
  - Informationsvisualisierung (schnelle Verifikation bzw. Falsifikation hermeneutischer Hypothesen)
  - Text- und Corpusviewing (schneller Zugriff auf Texte und Annotationen)
  - interaktive Analysewerkzeuge (interaktive Feedback- und Trainingsmechanismen)
  - semi-supervised Machine Learning

### Referenzen und Vorarbeiten

- (1) Richter, Sandra: A History of Poetics: German Scholarly Aesthetics and Poetics in International Context, 1770-1960. With Bibliographies by Anja Zenk, Jasmin Azazmah, Eva Jost and Sandra Richter. Berlin, New York: de Gruyter 2010.
- (2) Ramsay, Stephen: Algorithmic Criticism. In: A Companion to Digital Literary Studies. Hg. v. Ray Siemens u. Susan Schreibman. Malden, MA: Blackwell Publishing 2007, S. 477-491.
- (3) Moretti, Franco: Kurven, Karten, Stammbäume. Abstrakte Modelle für die Literaturgeschichte. Übers. v. Florian Kessler. Frankfurt am Main: Suhrkamp 2009
- (4) Rapp, Andrea: Das Projekt "TextGrid. Modulare Plattform für verteilte und kooperative wissenschaftliche Textdatenverarbeitung – ein Community-Grid für die Geisteswissenschaften". Chancen und Perspektiven für eine neue Wissenschaftskultur in den Geisteswissenschaften. In: Jahrbuch der historischen Forschung in der Bundesrepublik Deutschland: Berichtsjahr 2006, hrsg. von der Arbeitsgemeinschaft historischer Forschungseinrichtungen in der Bundesrepublik Deutschland. München: Oldenbourg 2007, S. 61-68.
- (5) www.sprache-und-genome.de, www.metalemmaliste.de
- (6) Koch, S. / Bosch, H. / Giereth, M. / Ertl, T.: Iterative integration of visual insights during patent search and analysis, Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on, 2009, 203-210.

GEFÖRDERT VOM



Bundesministerium für Bildung und Forschung