

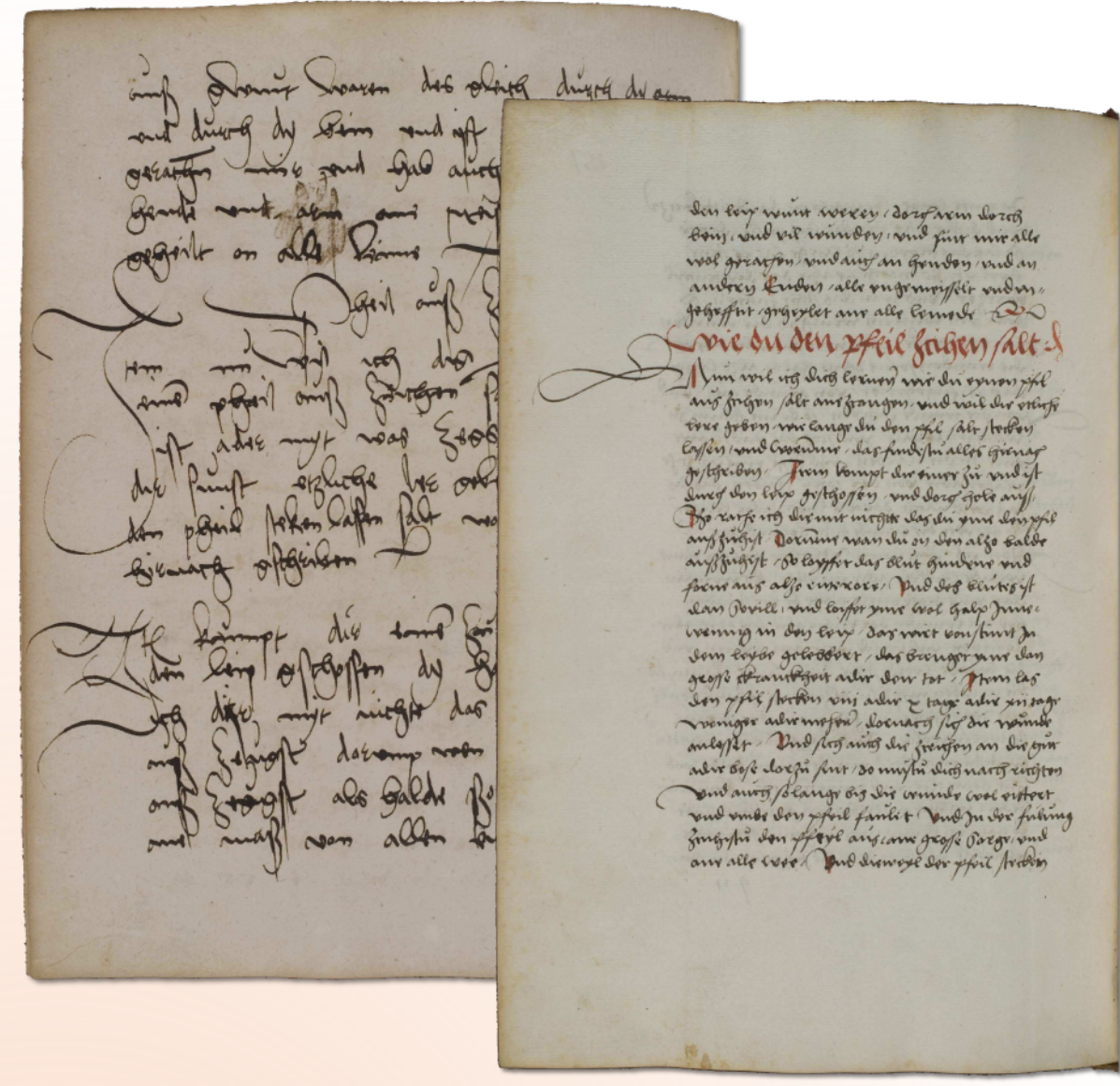
SaDA - Werkzeuge für die semi-automatische Differenzanalyse komplexer Textvarianten

SaDA ist ein vom BMBF gefördertes Gemeinschaftsprojekt von Germanisten, Romanisten und Informatikern mit dem Ziel, web-basierte Werkzeuge für den ausgefeilten Vergleich mehrerer Varianten oder Überlieferungen eines Textes zu entwickeln. Die Ergebnisse sollen philologischen Ansprüchen gerecht werden und für kritische Editionen mit Synopse und kritischem Apparat verwendbar sein. <http://informatik.uni-halle.de/SaDA>

KORPUS

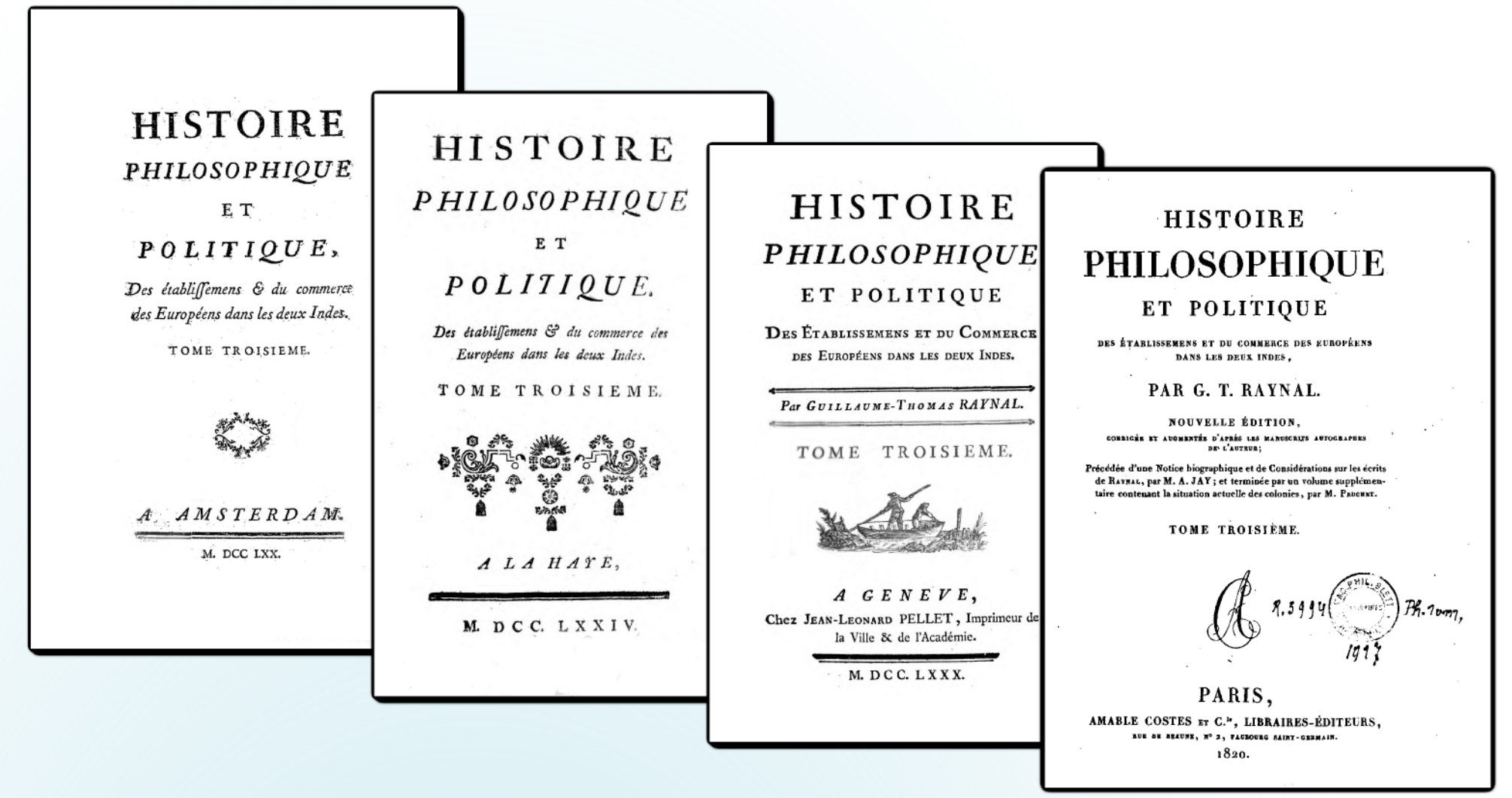
Die „Wundarzney“ des Heinrich von Pfalzpaint

Heinrich von Pfalzpaint gilt als Ahnherr der plastischen Chirurgie und ist einer der bekanntesten Wundärzte des späten Mittelalters. Von seinem berühmten medizinischen Kompendium, der „Wundarzney“ von 1460, sind elf Überlieferungen bekannt. Davon sind zehn verfügbar. Jede besteht aus bis zu 247 Kapiteln. Insgesamt haben sie einen Umfang von etwa 700.000 Wortformen.

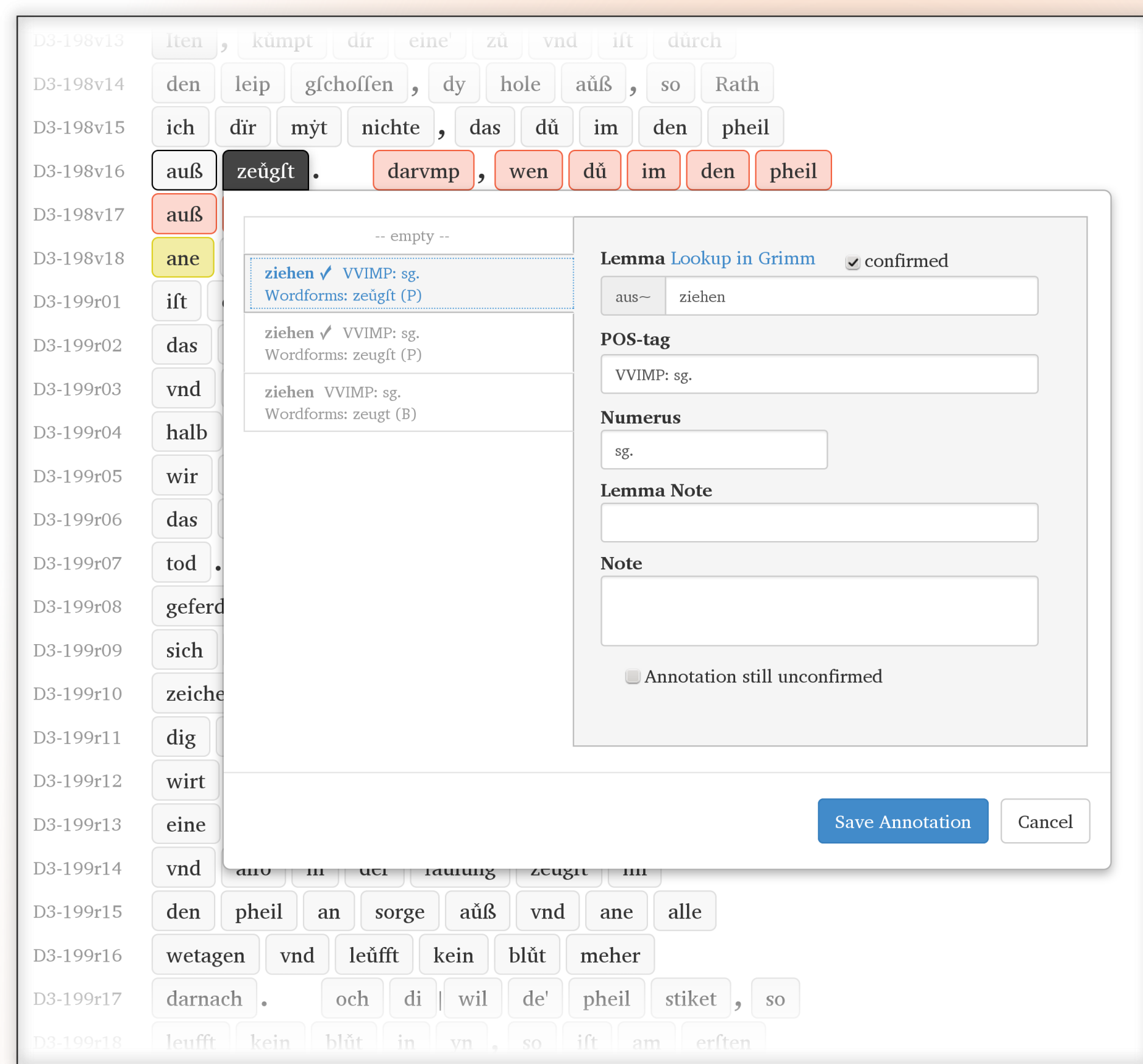


Die „Histoire philosophique et politique des établissements et du commerce des Européens dans les deux Indes“ von Guillaume T. F. Raynal

Die „Histoire“ ist eines der einflussreichsten Werke der französischen Aufklärung, welches die europäische Kolonialpolitik kritisch beschreibt. Das Werk wurde dreimal umfangreich überarbeitet. So umfasst bspw. Buch 6 der postum veröffentlichten Auflage von 1820 über 52.000 Wörter gegenüber den rund 28.500 Wörtern der verbotenen Erstausgabe von 1770.



NORMALISIERUNG



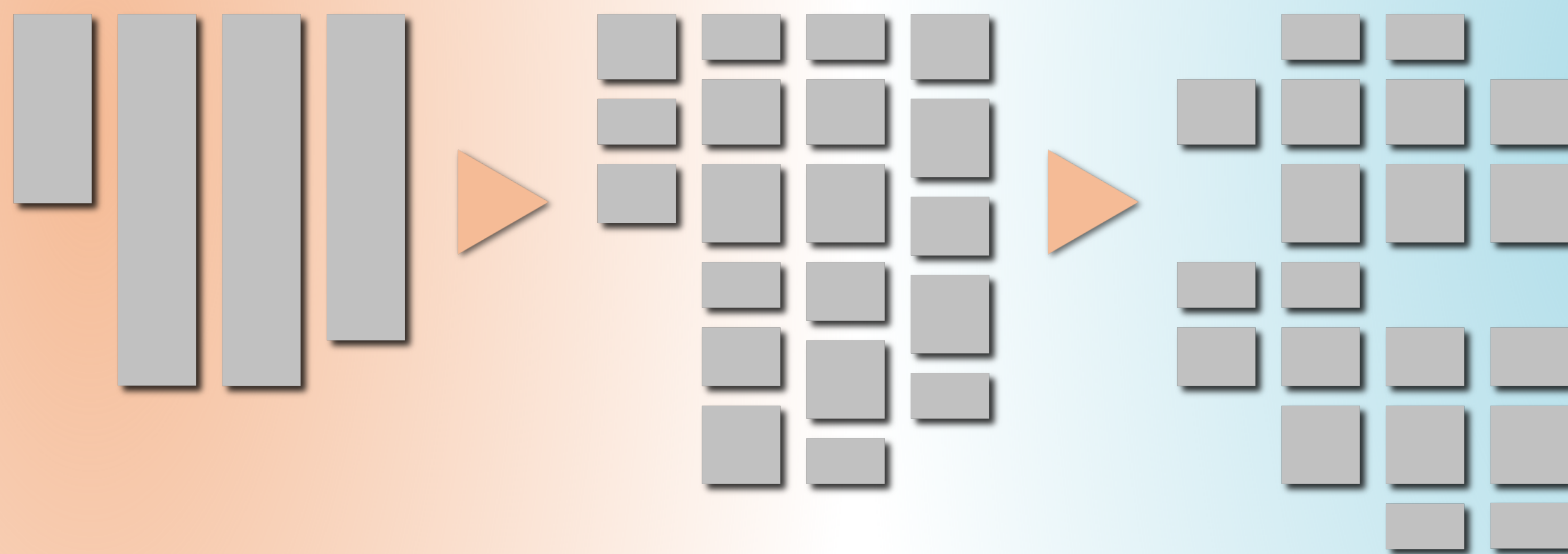
Benutzerfreundliche Lemmatisierung und morphologische Annotierung frühneuhochdeutscher Handschriften mit LAKomp

Handschriften der Sprachstufe Frühneuhochdeutsch sind eine große Herausforderung für den Textvergleich. Eine Normalisierung durch Lemmatisierung ist notwendig. Das Werkzeug LAKomp bietet zur manuellen Lemmatisierung schnelle und mitlernende Eingabemöglichkeiten, mit denen sich der Bearbeiter auf die Lemmatisierungsentscheidungen konzentrieren kann.

MAKROEBENE

1. Schritt: Segmentierung

Die Texte werden in kleinere Segmente zerlegt, welche die spätere Bearbeitung und Darstellung erleichtern. Für die „Wundarzney“ sind dies Teilsätze; für die „Histoire“ haben sich ganze Absätze als sinnvoller Kompromiss zwischen darstellbarer Größe und inhaltlich benötigtem Kontext erwiesen.



2. Schritt: Alignierung

Zur Zuordnung einander ähnlicher Segmente werden Signaturen für die Segmente bestimmt, in die bspw. signifikante Wörter oder die Position einfließen. Zum einen ist der Vergleich von Signaturen deutlich schneller als der direkte Vergleich der Segmente. Zum anderen lassen sich die Signaturwerte mittels Sortierung vorfiltern, sodass nicht mehr alle Segmente miteinander explizit verglichen werden müssen.

3. Schritt: Differenzanalyse

Auf Basis der einander zugeordneten Segmente findet eine detaillierte Aufbereitung der Unterschiede statt. Textänderungen werden farblich hervorgehoben und in einem Variantenapparat übersichtlich präsentiert. Für die „Wundarzney“ geschieht dies mit LAKomp in einer Partitursynopse; für die „Histoire“ mit Pitahaya in einer synoptischen Gegenüberstellung der Absätze. Beide Werkzeuge bieten umfangreiche Eingriffsmöglichkeiten, bspw. Filteroptionen, die Orthographie und Morphologie betreffen.

Partitursynopse in LAKomp

Zeile 18: Meister, Schüler

H Dar vmb, ¹wen ²du in ⁴5 ⁶7 ⁸alfo ⁹bald ¹⁰auß zewheft/, ¹¹12 ¹³St

Bu dorumb, wen du yn zo balde auß czogeft,

D3 darvmp, wen du im den pheil auß zeugt alsbalde,

E dar umb, wan du in so bald auß zuechtfn,

Br darvmb, wen du on also balde auß zcogeft,

B8 darvmb, wan du ym den pfil alzo balde ufz czuhift,

D2 Dorumb, wan du en den alzo balde auß zühift/,

P darumb, wan du im denn pheil alfo palde auß czetgeft/,

B9 darvñ, wan du jm den pfil vs zugift alfo bald,

St

¹ Dar vmb, | <fehlt> St dar umb, E, P darvmb, B8 darvmp, D3 darvñ, B9 dorumb, Bu Dorumb, D2
² wen | <fehlt> E, B8, D2, P, B9, St ³ <fehlt> | wan E, B8, D2, P, B9
⁴ du in | <fehlt> St du en D2 du im D3, P du jm B9 du ym B8 du yn Bu du on Br
⁵ <fehlt> | so E zo Bu ⁶ <fehlt> | den D3, B8, D2, B9 denn P
⁷ <fehlt> | pheil P pfil B8, B9 pheil D3 ⁸ alfo | <fehlt> Bu, D3, E, B9, St alzo B8, D2 alfo Br
⁹ bald | <fehlt> D3, B9, St balde Bu, Br, B8, D2 palde P
¹⁰ auß zewheft/, | <fehlt> St auß czogeft/, P auß czogeft, Bu auß zeugt D3 auß zühift/, D2 auß zuechtfn, E ufz czuhift, B8 vs zugift B9 ¹¹ <fehlt> | alfo bald, B9 ¹² <fehlt> | alsbalde, D3

Synoptische Gegenüberstellung in Pitahaya

The screenshot shows the Pitahaya interface with a toolbar at the top containing options like 'Start', 'Editionen', 'Bearbeiten', and 'Edition durchsuchen'. Below the toolbar are two filter sections: 'Vergleichsfilter' and 'Anzeigefilter'. The main area displays a synoptic comparison of text variants, with differences highlighted in various colors (green, red, blue). A sidebar on the right contains a 'Farbmodus' section with a legend for the colors used in the comparison. At the bottom, there is a 'Varianten' section with a table of variant details.



MARTIN-LUTHER-UNIVERSITÄT HALLE-WITTENBERG

Institut für Germanistik
Sylwia Kösser – sylwia.koesser@germanistik.uni-halle.de
Aletta Leipold – aletta.leipold@germanistik.uni-halle.de
Hans-Joachim Solms – hans-joachim.solms@germanistik.uni-halle.de

Institut für Informatik
André Medek – andre.medek@informatik.uni-halle.de
Paul Molitor – paul.molitor@informatik.uni-halle.de
Marcus Pöckelmann – marcus.poeckelmann@informatik.uni-halle.de
Jörg Ritter – joerg.ritter@informatik.uni-halle.de

Institut für Romanistik
Thomas Bremer – thomas.bremer@romanistik.uni-halle.de
Julia Ritter – julia.ritter@student.uni-halle.de
Susanne Schütz – susanne.schuetz@romanistik.uni-halle.de

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Projekt-Kürzel: 01UG1247