

# Textvergleich und Intertextualität

Dr. Jörg Ritter

Institut für Informatik, Martin-Luther-Universität Halle-Wittenberg

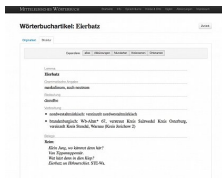
23. Februar 2023

# Themenfelder unserer Arbeitsgruppe...

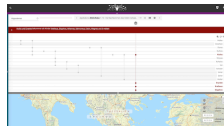
## Interaktive Werkzeuge



## Wörterbuch-Editionen



## Editionen / Portale



# Textvergleich - Herausforderungen

zwei, mehrere oder viele Textzeugen/Überlieferungen

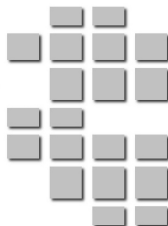
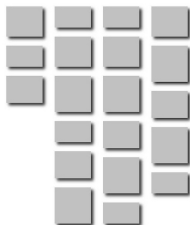
kurze, lange oder sehr lange Texte

unterschiedlichste Granularität der Segmente

eine oder mehrere Sprachen oder fremdsprachige Einschübe

ältere Sprachstufen mit nicht genormter Grafie

# Textvergleich - Ansatz



# LERA: Französischer Roman

Raynal: „Histoire Philosophique [...]“, 4 Auflagen: 1770, 1774, 1780, 1820

## HISTOIRE PHILOSOPHIQUE ET

*POLITIQUE,*  
*Des établissements & du commerce*  
*des Européens dans les deux Indes.*

TOME TROISIEME.



A. AMSTERDAM.

M. DCC. LXX.

## HISTOIRE PHILOSOPHIQUE ET *POLITIQUE.*

*Des établissements & du commerce des*  
*Européens dans les deux Indes.*

TOME TROISIEME.



A LA HAYE,

M. DCC. LXXIV.

## HISTOIRE PHILOSOPHIQUE ET POLITIQUE

DES ÉTABLISSEMENTS ET DU COMMERCE  
DES EUROPÉENS DANS LES DEUX INDES.

Par GUILLAUME-THOMAS RAYNAL.

TOME TROISIEME.



A GENEVE,

Chez JEAN-LEONARD PELLET, Imprimeur de  
la Ville & de l'Académie.

M. DCC. LXXX.

## HISTOIRE PHILOSOPHIQUE ET POLITIQUE

DES ÉTABLISSEMENTS ET DU COMMERCE DES EUROPÉENS  
DANS LES DEUX INDES.

PAR G. T. RAYNAL.

DEUXIÈME ÉDITION,  
CORRIGÉE ET AUGMENTÉE D'UN LIVRE DE SUPPLÉMENT  
DE COMMERCE.

Principes d'une Méthode Géographique et de Constatations sur les droits  
de Peuples, par M. A. JAY; et traduites par un volume supplémentaire  
contenant la doctrine générale des colonies, par M. TAYLOR.

TOME TROISIEME.



PARIS,

AMABLE COSTES & C<sup>o</sup>, LIBRAIRES-ÉDITEURS,  
RUE DE HARVEY, N<sup>o</sup> 17. MDCCLXXXIIII.  
1820.

# LERA: analoge Vorgabe

[22] *Weltbild im 1490*

[23] -4 - (+1492) *Weltbild*

95.1 [... dans des mers éloignées.]

[24] Un homme obscur, plus avancé que son siècle dans la connaissance de l'astronomie & de la navigation, sembla veiller à l'agrandissement de l'Empire: Christophe Colomb feroit comme par instinct, qu'il devoit y avoir un autre continent, & que c'étoit à lui de le découvrir. Les Antipodes, que la raison même traitoit de chimère, & la superstition d'erreur & d'impénétrable étoit aux yeux de cet homme de génie, une vérité incontestable. Plein de cette idée, la plus saine qui soit entrée dans l'esprit humain, il proposa à Oïtesa, sa patrie, de mettre sous ses loix un autre hémisphere méridional par cette petite république, & par l'Angleterre même, qu'il devoit trouver ouverte à toutes ces entreprises maritimes; il porta ses vues & ses projets à Isabelle.

siècle

s. 14

[24]

[23] III. Colomb eut le projet de découvrir l'Amérique.

Ce fut dans ces circonstances glorieuses, qu'un homme obscur, plus avancé que son siècle dans la connaissance de l'astronomie & de la navigation, proposa à l'Espagne heureuse au-dedans de s'agrandir au-dehors, Christophe Colomb feroit comme par instinct, qu'il devoit y avoir un autre continent, & que c'étoit à lui de le découvrir. Les Antipodes, que la raison même traitoit de chimère, & la superstition d'erreur & d'impénétrable étoit aux yeux de cet homme de génie, une vérité incontestable. Plein de cette idée, l'une des plus grandes qui soient entrées dans l'esprit humain, il proposa à Gènes sa patrie, de mettre sous ses loix un autre hémisphere méridional par cette petite république, & par le Portugal, où il vivoit, & par l'Angleterre même, qu'il devoit trouver disposée à toutes les entrepri-

siècle

[23]

s. 15

[24]

Ce fut dans ces circonstances glorieuses, qu'un homme obscur, plus avancé que son siècle dans la connaissance de l'astronomie & de la navigation, proposa à l'Espagne heureuse au-dedans de s'agrandir au-dehors. Christophe Colomb feroit comme par instinct qu'il devoit y avoir un autre continent, & que c'étoit à lui de le découvrir. Les Antipodes, que la raison même traitoit de chimère, & la superstition d'erreur & d'impénétrable étoit aux yeux de cet homme de génie, une vérité incontestable. Plein de cette idée, l'une des plus grandes qui soient entrées dans

siècle

## DES DEUX INDES. 337

l'esprit humain, il proposa à Gènes sa patrie, de mettre sous ses loix un autre hémisphere méridional par cette petite république, & par le Portugal où il vivoit, & par l'Angleterre même, qu'il devoit trouver disposée à toutes les entreprises maritimes; il porta ses vues & ses projets à Isabelle.

## PHILOS. ET POLITIQUE. II

Fes maritimes; il porta ses vues & ses projets à Isabelle.

H70: Un homme obscur,

74: Ce fut <sup>dans</sup> ces circonstances glorieuses qu'un homme obscur,

80: Ce fut dans ces circonstances glorieuses, qu'un homme obscur,

H70: navigation, sembla veiller à l'agrandissement de l'Espagne.

74: navigation, proposa à l'Espagne heureuse au-dedans de s'agrandir au-dehors.

80: navigation, proposa à l'Espagne heureuse au-dedans de s'agrandir au-dehors.

H70: idée, la plus saine qui soit entrée dans l'esprit

74,80: idée, l'une des plus grandes qui soient entrées dans l'esprit

H70: trouver ouverte à toutes ces

74,80: trouver disposée à toutes les

H70: maritimes;

74,80: maritimes;

H70: étoit  
74,80: étoient

H70,74: chimère,  
80: chimère,

H70,74: instinct,  
80: instinct

H70: hémisphere méridional

74: hémisphere méridional

80: hémisphere méridional (unmöglich)

# LERA: Werkzeug

Demoversion: <https://lera.uzi.uni-halle.de>



**Features:** automatische Alignierung, Segmentvergleich, Nachbearbeitung, Analyse, Suche, Apparate ...

**Sprachen:** deutsch, englisch, französisch, spanisch, italienisch, norwegisch, arabisch, hebräisch, slowakisch, vietnamesisch

**Ausgewählte Projekte:** Keter Shem Tov, Hannah Ahrendt digital, Edinburgh Conan Doyle Project, Kalīla and Dimna, HyperAzpilcueta, Editionsprojekt Karl Gutzkow

# Optimale Alignierungen?

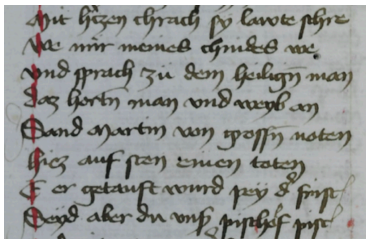
$$\min \sum_{p=1}^n \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} (c_1 \cdot d_{i,j} \cdot z_{i,j,p} + c_2 \cdot (1 - z_{i,j,p}))$$

unter den Nebenbedingungen ...

mit den magischen Konstanten  $c_1$  und  $c_2$



# Frühneuhochdeutsch



F024-003vb,09	*(M*)it h'czen chrach \$y\ ' lawte \$chre(:)
F024-003vb,10	(")*(W*)e mi\ r meines chindes(,) we(!)(")(,)
F024-003vb,11	*(v*)nd \$prach zü dem heilign\ - man(.)
F024-003vb,12	*(d*)az hortn\ - *1b man vnd weyb an*2b(:)
F024-003vb,13	(")*(S*)and Martin von gro\$\$n\ - noten
F024-003vb,14	*(h*)iez auf \$ten*1a ei\ :nen toten(,)
F024-003vb,15	*(E*)er getauft wurd pey\ ' d' fri\$t(.)
F024-003vb,16	*(S*)ey\ 'd aber du vn\$' pi\$cho%lf pi\$t

# Frühneuhochdeutsch

## Varianz der Wortformen am Beispiel von **Pfeil**

pfeil      pfeyl      pfeyle      pffeil      pheil      pfeileÿfn      fpil  
pfilyfen      pffhiel      pfeyleÿfn̄      pfeileiffen      pheilschafft  
pfeyleÿßen      pffeile      pfeill      pfeil      pfil      fpÿl      fpfil  
pfiljfen      pfiljffen      ppheiles      pfeilschaffth      pfeilschafft  
Pfeil      pfeileifsenn      pffheill      phiell      phiell      phyell  
    pfeyll      pfeyles      pfeyllschafft      bremfzenpfeyleÿfn̄  
pfeyleÿfen      pfeyleÿfzenn      bremfzpfeyl      pfeyleifzen      pfÿl  
    pfeÿl      pfeileÿffen      pfÿl      pfyl      pheyles      Pheil  
phileyfen      bremßenpheil      pfeille      Pheüll      pfeillschafft  
Brembspfeil      pfeÿle      Pfeÿle      pfeÿls      pfeyleÿßen  
    bremßenpfeÿl      bremßenpfeÿleÿßen      pffeileÿfzen



# Rezepte in Frühneuhochdeutsch

**Hypothese:** Handlungsanweisungen in Rezepten kann man an einem Imperativ gefolgt von einem Substantiv erkennen...

**Fragestellung:** Wie weit kommt man, ohne manuell zu annotieren?

The screenshot shows a software interface for text analysis. The main window is titled 'Texte und Rezepte' and 'Text hochladen'. It contains a list of recipes, each with a unique ID and a line of text. Several words in the text are highlighted in green, indicating they have been identified as keywords. A search filter panel is visible on the right side, showing a search bar and a table of results.

Text to be analyzed:

- E:R:310 Ala guttur salis vni<sup>us</sup> sals
- E:R:311 Diebe salz koch mit **schilff** und die selb ge<sup>u</sup> (**u**)
- E:R:312 **Wen** drey **weyß** ge<sup>u</sup> alt Schneyr (**u**) die drey salte
- E:R:313 rugel (**u**) Ala ph<sup>u</sup>nd wache (**u**) So ist kurz / (**u**) So **ist** hockereckel
- E:R:314 / (**u**) die **weyß** zu<sup>u</sup>stehen Vnd die hensch
- E:R:315 gelochte **weyß** / (**u**) Ala R<sup>u</sup>, hockelich magel
- E:R:316 vnd ein geloch<sup>u</sup> hockelich we<sup>u</sup>ndweyß<sup>u</sup> / (**u**) So ist
- E:R:317 per<sup>u</sup>wandels mit dem Rote Steng<sup>u</sup> / (**u**) So ist **weyß** (**u**)
- E:R:318 Alsdah ph<sup>u</sup>nd koch<sup>u</sup>weyß<sup>u</sup>rotweyß<sup>u</sup> / (**u**) So **ist** per<sup>u</sup>wandels
- E:R:319 mit dem roten **weyß** / (**u**) So **ist** **weyß** mit
- E:R:320 das falbe **weyß** / (**u**) Sich der koch<sup>u</sup>weyß<sup>u</sup> ganz gleich (**u**)
- E:R:321 die kreuter in einemucker Stoffs / (**u**) ges<sup>u</sup> in
- E:R:322 ein wenig kreut<sup>u</sup> drey<sup>u</sup> / (**u**) So koch<sup>u</sup> es sich ganz Stoffs (**u**)
- E:R:323 **hoch** die **weyß** ver / (**u**) **die** syl<sup>u</sup> Stoffs klan / (**u**) drey
- E:R:324 weyß ein ph<sup>u</sup>nd koch<sup>u</sup> drey<sup>u</sup> am Stoffs (**u**) So ist<sup>u</sup> in St
- E:R:325 drey<sup>u</sup> luter Stoffs (**u**) Die Schneyr Stes hockelich / (**u**) in
- E:R:326 drey<sup>u</sup> in drey<sup>u</sup> drey<sup>u</sup> / (**u**) oder Stes die kreuter drey<sup>u</sup> (**u**)
- E:R:327 Jk die selb / (**u**) So hockelich<sup>u</sup> v<sup>u</sup> die ander koch<sup>u</sup> / (**u**) St
- E:R:328 die selb mit den kreuter in einem koch<sup>u</sup> / (**u**) als lang
- E:R:329 drey<sup>u</sup> ercken koch<sup>u</sup> / (**u**) die die kreuter wol wach
- E:R:330 gelochte Stes / (**u**) drey<sup>u</sup> **weyß** u<sup>u</sup> drey<sup>u</sup> an **weyß** / (**u**) **weyß**
- E:R:331 die **weyß** kreuter / (**u**) **weyß** an **weyß** oder vcher aus Stes / (**u**)
- E:R:332 ganz ercken wach<sup>u</sup> weyß<sup>u</sup> drey<sup>u</sup> mit die hensch
- E:R:333 gelochte ph<sup>u</sup>weyß<sup>u</sup> / (**u**) koch<sup>u</sup> gelochte / (**u**) die weyß drey<sup>u</sup>
- E:R:334 / (**u**) gepreyt ein<sup>u</sup>weyß<sup>u</sup> / (**u**) koch<sup>u</sup>weyß<sup>u</sup> / (**u**) weyß<sup>u</sup> drey<sup>u</sup>
- E:R:335 gelochte / (**u**) jgliche wol geweg<sup>u</sup> an ket / (**u**) in
- E:R:336 ein wenig drey<sup>u</sup> Stes / (**u**) man drey<sup>u</sup> syl<sup>u</sup> von drey
- E:R:337 koch<sup>u</sup> / (**u**) So ist syl<sup>u</sup> mit giend<sup>u</sup>weyß<sup>u</sup> / (**u**) die syl<sup>u</sup> wol gies
- E:R:338 wach / (**u**) **weyß** syl<sup>u</sup> / (**u**) syl<sup>u</sup> wach gies / (**u**) drey<sup>u</sup> **weyß**
- E:R:339 syl<sup>u</sup> mit drey<sup>u</sup> ein **weyß** drey<sup>u</sup>weyß<sup>u</sup>weyß<sup>u</sup> / (**u**)

Search filter panel:

Rezept	Datum	Typ	Benutzer	Status
Salko	E-47:10			
Ala	10			
Schaden	E-57:16		bestenrig	

# Intertextualität: Paraphrasensuche

## Ideen

via (multilingualen) Worteinbettungen und (Relaxed) Word Mover Distance

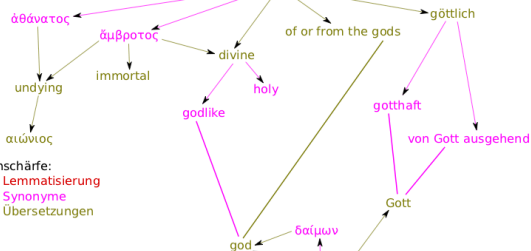
via Boolescher Suche mit „ähnlichen“ Wörtern bzw. wortweisen Übersetzungen

Unvereinbar aber ist das Schändliche mit allem Göttlichen, aber das Schöne ist vereinbar.

Plato (*symp. Pl.* 206 d 1-2)

**ἀνάρμωστος δ' ἐστὶ τὸ αἰσχρὸν παντὶ τῷ θεῷ, τὸ δὲ καλὸν ἀρμόττων.**

ἀνάρμωστος εἶναι αἰσχρὸς πᾶς θεῖος καλὸς ἀρμόζω/ἀρμόττω



Unschärfe:

- Lemmatisierung
- Synonyme
- Übersetzungen

αἰσχρὸς ἐναντίος φύσις θεός  
τὸ μὲν γὰρ αἰσχρὸν ἐναντίον καὶ τῇ φύσει καὶ τῷ θεῷ.

Plotin (*Plot. enn.* III 5, 1, 19-20)

Denn das Schändliche ist sowohl der Natur als auch dem Gott entgegengesetzt. \_ \_ \_

Vielen Dank für Ihre Aufmerksamkeit!