

Animate

Ein Tool zum Datenexport aus den historischen Referenzkorpora

Abstract *Animate* (<https://github.com/matthias-stemmler/animate>, Stand: 26.11.2025) was developed to facilitate the export of results from the reference corpora for the historical German language levels (e. g. ReA (750–1050), ReM (1050–1350), ReN (1200–1650) and ReF (1350–1650)), whose data is ANNIS-based. The tool provides a user interface that allows users to export search results from all these corpora directly, along with the desired additional information. This article demonstrates the functions of *Animate* and its potential for investigating language change. As a case study, we present a corpus study on the diachrony of the interjection *ei/ey*. In addition, we demonstrate how corpora can be evaluated using *Animate*. In conclusion, the program offers a solution to a variety of practical problems in synchronous and diachronic corpus linguistic studies, and can also contribute to the evaluation of existing corpus annotations.

Keywords reference corpora, data export, corpus linguistics, language change, interjections

1. Ausgangslage: Zum Wert der Referenzkorpora in der Germanistischen Linguistik

Für diachrone Untersuchungen von Sprachwandelphänomenen im Deutschen sind die Referenzkorpora für die historischen Sprachstufen zu einer grundlegenden Ressource geworden, um sprachhistorische wie diachrone Korpusstudien durchzuführen. Mit dem Referenzkorpus Altdeutsch (ReA; 750–1050), dem Referenzkorpus Mittelhochdeutsch (ReM; 1050–1350), dem Referenzkorpus Frühneuhochdeutsch (ReF; 1350–1650) und dem Referenzkorpus Mittelniederdeutsch/Niederrheinisch (ReN; 1200–1650) decken die Korpora sowohl eine große Zeitspanne als auch verschiedene Sprachräume ab.¹ Zeitlich ergänzt werden die Referenzkorpora inzwischen noch durch das syntaktisch annotierte GiesKaNe-Korpus (1650–1950),² das bislang 24 Texte enthält (GiesKaNe 0.3) und noch auf 72 Texte erweitert werden soll. Weitere Ressourcen für diachrone korpusbasierte Untersuchungen bilden registerspezifische Korpora zu bspw. Hexenverhörprotokollen, Predigten, wissenschaftlichen Texten etc. (vgl. die Übersicht unter <https://corpus-tools.org/annis/corpora.html>, Stand: 26.11.2025). Allen genannten Datenressourcen ist gemein, dass sie im ANNIS-Format vorliegen. ANNIS („ANNotation of Information Structure“) ist ein open source sowie plattformübergreifendes Such- und Visualisierungstool (Krause/Zeldes 2016). Es wurde für komplexe, mehrschichtige linguistische Korpora mit verschiedenen Arten von Annotationen entwickelt und ermöglicht sowohl die Suche nach Daten durch ein web-

1 Zugang zu den Referenzkorpora über www.deutschdiachrondigital.de (Stand: 31.5.2025).

2 www.uni-giessen.de/de/fbz/fb05/germanistik/forschung/sprache/gieskane (Stand: 31.5.2025).

basiertes Such-Interface und die formale Abfragesprache ANNIS Query Language (AQL) als auch verschiedene Darstellungsmöglichkeiten wie beispielsweise in Tabellenansicht und als Graphen. Über ANNIS können damit in den Korpora komplexe Suchabfragen durchgeführt werden, die als empirische Datengrundlage für vielfältige linguistische Fragestellungen genutzt werden können.

Die Datenlage für die Untersuchung von Sprachwandelphänomenen im Deutschen hat sich mit der Erschließung der Referenzkorpora bedeutend verbessert. Der Bezug auf die Referenzkorpora hat Vorteile, die bereits von Wegera (1990) im Vorfeld der Entstehung des Referenzkorpus Mittelhochdeutsch (ReM) zusammengefasst wurden. Für dieses Korpus, das ca. 2 Mio. Wortformen umfasst, wurden Originalhandschriften zugrunde gelegt, da gerade für das Mittelhochdeutsche die häufig literaturwissenschaftlich orientierten Editionen für linguistische Untersuchungen nicht als verlässlich angesehen werden können. Zur Transparenz sind im Referenzkorpus die Texte sowohl als diplomatischer Lesetext, als modernisierter Lesetext als auch als normalisierter Lesetext hinterlegt. Zudem ist das Korpus weitgehend ausgewogen in Bezug auf die Faktoren Zeit, Raum und Textsorte zusammengestellt. Untersucht man ein bestimmtes Sprachwandelphänomen, lässt sich etwa nach bestimmten Zeitperioden und Schreiborten suchen bzw. filtern. Ebenso können die Ergebnisse zu den verschiedenen Textbereichen (Wissenschaft, Alltag, Religion, Poesie etc.), Textsorten (Arzneibuch, Minnesang, höfische Epik, Gebet etc.) und Textart (Vers, Prosa, Mischformen und Urkunde) zugeordnet werden. Zudem sind sie nach verschiedenen morphologischen und syntaktischen Parametern annotiert (vgl. für einen detaillierten Überblick Petran et al. 2016; Klein/Dipper 2016). Mit der Initiative „Deutsch Diachron Digital“ (www.deutschdiachrondigital.de/, Stand: 26.11.2025) besteht ein Verbund der Referenzkorpora für die verschiedenen Sprachstufen, der die technologische Kooperation und den Austausch in Bezug auf Nutzungsszenarien für Forschung und Lehre fördert.

Die Referenzkorpora bieten zweifellos große Vorteile bei diachronen Untersuchungen des Deutschen. Durch die Nutzung der Korpora für verschiedene sprachliche Phänomene ergibt sich zudem der Vorteil der Vergleichbarkeit der Ergebnisse. Gerade vor dem Hintergrund, dass die textsortenspezifischen Charakteristika einen großen Einfluss auf die Verwendung grammatischer Elemente haben (vgl. etwa zum Einfluss der Unterscheidung Prosa vs. Vers exemplarisch Prell 2005; zur Unterscheidung zwischen Nähe- vs. Distanzsprache Ágel/Hennig (Hg.) 2006; Elspaß 2005 und zum Einfluss regionaler Variation Elspaß 2010), ist die registerbezogene und dialektale Ausgewogenheit eine relevante Eigenschaft der Referenzkorpora. Einzeluntersuchungen, die sich auf nur eine Textsorte oder nur einen Text stützen, sind in ihrer Aussagekraft eingeschränkt, da erst zu untersuchen ist, ob die Ergebnisse auch für andere Texte gelten. Legen zwei Studien jeweils die gleiche Datengrundlage zugrunde, lassen sich die jeweiligen Ergebnisse zudem besser aufeinander beziehen.

Da die historischen Referenzkorpora aufgrund der jeweiligen Spezifika der zu annotierenden Sprachstufen in eigenständigen Projekten verwirklicht wurden, besteht in Bezug auf die Nutzung der Korpora allerdings das Problem, dass für jedes Korpus eigene Annotationsrichtlinien bestehen und die Termini nicht vereinheitlicht sind. Gerade mit Blick auf die Untersuchung diachroner Fragestellungen erschwert das die Nutzung der Korpora. Ein weiteres bekanntes arbeitspraktisches Problem stellt zudem der Export der Suchergebnisse dar. Zur Aufbereitung der Daten für eine anschließende qualitative Analyse ist es in der Regel wünschenswert, die Treffer sowohl in ihrem Kontext als auch mit den jeweiligen Annotationen (bspw. Lemma, Wortart, grammatische Flexion) zu extrahieren. Zusätzlich sollten auch die entsprechenden Metadaten zur Textquelle wie Dokumentname, Zeitraum, Sprachregion etc. erfasst werden. In ANNIS stehen dafür verschiedene Export-Funktionen

zur Verfügung, die dieses Ziel aber nicht ohne Weiteres ermöglichen. Der CSVExporter exportiert beispielsweise nur die gefundenen Treffer mit den jeweiligen Annotationen und ignoriert den Kontext. Eine einfache Suche nach dem Lexem „Himmel“ (AQL-Abfrage: `norm="himel"`) im ReM (Version 2.1) führt zu einer CSV-Liste, in der der Treffer (1_span) mit den jeweiligen Annotationen aber ohne Kontext ausgegeben wird, vgl. Abbildung 1:

1_id	1_span	1_anno_grammar::inflection	1_anno_grammar::inflectionClass	1_anno_grammar::lemma	1_anno_grammar::pos
salt:/11-12_1-obd-PV-X/M040#t262_m1	himil	Akk.Sg	st_Masc	himel	NA
salt:/11-12_1-obd-PV-X/M095#t158_m2	himil	Dat.Sg	st_Masc	himel	NA
salt:/11-12_1-obd-PV-X/M137#t215_m1	himel	Nom.Sg	st_Masc	himel	NA

Abb. 1: Drei Suchergebnisse der mit dem CSVExporter exportierten Daten importiert in Excel (die Abb. zeigt aus Platzgründen nur eine Auswahl der Spalten an)

Der TextColumnExporter erlaubt den Export mit Kontext vor und nach dem Treffer. Diese Exportmöglichkeit gibt die Treffer aber ohne weitere Metadaten und Annotationen aus. Es können zwar noch „Annotation Keys“ eingegeben werden. Ohne weitere Kenntnisse über das jeweilige Korpus ist aber nicht offenkundig, welche Art von Annotation Keys eingegeben werden können. Eingaben wie „lemma“ oder „pos“ werden ignoriert. Bei „Parameters“ können Metadaten wie „title“ mit ausgegeben werden. Auch dafür ist es aber notwendig, die jeweiligen Parameter bereits zu kennen. Wird ein ungültiger Parameter (bspw. `meta-keys=doc`) eingegeben, führt der Export ohne weitere Fehlermeldung zu einer leeren Liste. Dabei ist zu berücksichtigen, dass die jeweiligen Annotationskategorien in den verschiedenen Referenzkorpora unterschiedlich sind und insofern für jede Sprachstufe separat die jeweiligen Handbücher und Annotations-Guidelines konsultiert werden müssen.

match_number	speaker	left_context	match_column	right_context
1 sText1		enperinnit, unta alumbē šin ungewiteri michila .(.) er geladata den	himil	hin uf unta die erdan geunteršceiden(.) [...] zerucke .(.) oba
2 sText1		meninšche ., vn\ er an de\me uierzgošte\me tage . zi	himil	vóir(:) dannan gelo\vbi ich ín chunftich z ir# teilin ubir
3 sText1		*(S*)c\~a *(M*)aRia <:>.(.) *(D*)o gehit ime šo werde . der	himel	zu\o der erde „, da der ešil unde daz rint

Abb. 2: Drei Suchergebnisse der mit dem TextColumnExporter exportierten Daten importiert in Excel (alle Spalten)

In Abbildung 2 zeigt sich als weiteres Problem, dass der Text verschiedene Sonderzeichen enthält (vgl. *an de\me uierzgošte\me tage* statt *an dēme uierzgoftēme tage*). Das liegt daran, dass sowohl der CSVExporter als auch der TextColumnExporter nicht auf die Ebene der handschriftlichen Tokenisierung (`tok_dipl`) oder der den modernen Erwartungen entsprechenden Tokenisierung (`tok_anno`) zugreifen, sondern auf eine dritte Default-Ebene. Dieses Problem lässt sich mit dem GridExporter lösen, bei dem die entsprechende Ebene (`tok_dipl` bzw. `tok_anno`) explizit im Feld „Annotation Keys“ angegeben werden kann. In den Daten ist dann aber nicht mehr der eigentliche Treffer ersichtlich. Das bedeutet, dass sich für den jeweiligen Treffer nicht gezielt weitere Annotationen ausgeben lassen, vgl. Abbildung 3, die die Ausgabe mit dem GridExporter zeigt.

1 tok_dipl	enperinnit unta alumbē šin ungewiteri michila . er geladata den himil hin uf unta die erdan geunteršceiden (...) zerucke . oba
2 tok_dipl	meninšche . vn\ er an dēme uierzgoftēme tage . zihimil vóir dannan gelóbi ich ín chunftich zir teilin ubir
3 tok_dipl	Sca MaRia . Do gehit ime fo werde . der himel zú der erde . da der efil unde daz rint

Abb. 3: Drei Suchergebnisse der mit dem GridExporter exportierten Daten importiert in Excel (alle Spalten)

Die dargestellten Probleme wurden auch bereits in Hartmann (2022) zusammengestellt und adressiert. In einem umfangreichen Tutorial wird dort ein mehrschrittiges Verfahren unter der Verwendung mehrerer Exportfunktionen und einem Excel-Makro vorgeschlagen. Der Datenexport ist so aber nur über Umwege möglich. Hinzu kommt, dass für den

Export der Metadaten die entsprechenden Bezeichnungen in den Korpora unterschiedlich sind und für die jeweiligen Abfragen erst eruiert werden müssen.

2. Animate

Das von Matthias Stemmler in Kooperation mit dem Lehrstuhl für Deutsche Sprachwissenschaft der Universität Augsburg entwickelte Tool *Animate* (URL: <https://github.com/matthias-stemmler/animate>, Stand: 26.11.2025) begegnet diesen Problemen, indem es eine komfortable Benutzer-Oberfläche bereitstellt, über die Daten zu Suchergebnissen in allen ANNIS-Korpora direkt und mit den jeweils gewünschten Zusatzinformationen exportiert werden können. Auf der Basis der Suchabfrage erstellt das Desktop-Programm wahlweise eine CSV- oder Excel-Datei mit einer Zeile pro Beleg, die die jeweiligen Treffer (bzw. Match-Knoten) in ihrem Kontext im KWIC-Format (KeyWord In Context) sowie zusätzliche Annotationen und Metadaten auf Teilkorpus- und Dokumentebene anzeigt. Die Funktionalität ähnelt einer Kombination aus dem TextColumnExporter und dem CSVExporter in ANNIS, bietet jedoch eine benutzerfreundlichere Oberfläche (vgl. den Screenshot in Abb. 4).

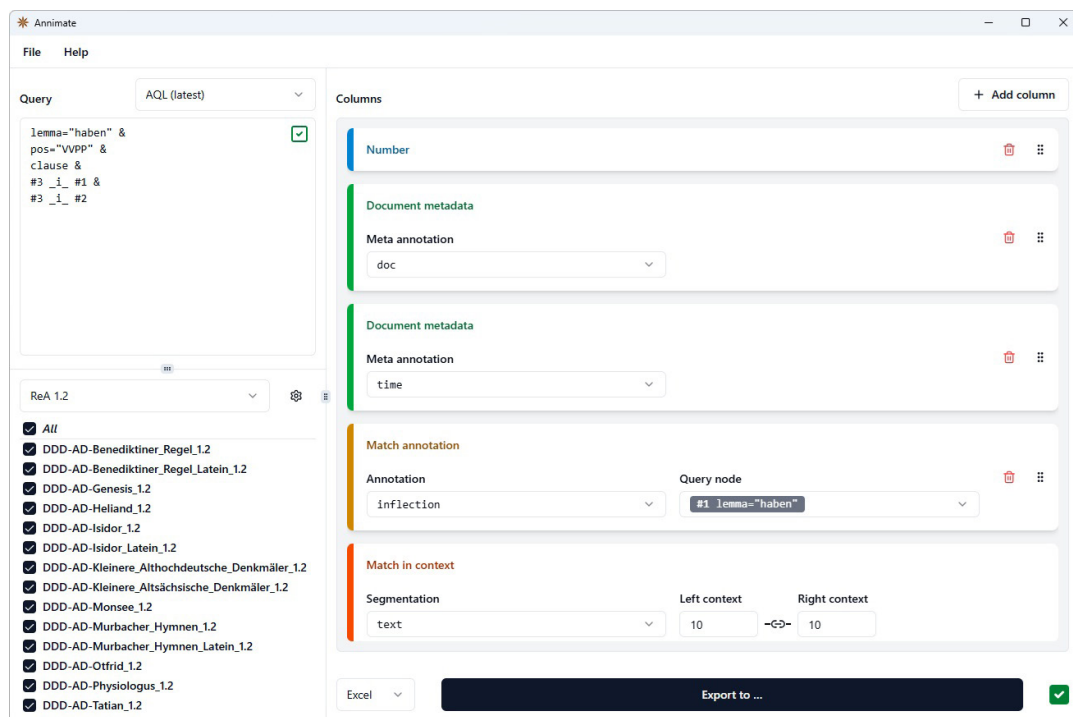


Abb. 4: Screenshot: Benutzeroberfläche mit einer Beispielabfrage in *Animate* für das Referenzkorpus Altdeutsch

Im Rahmen dieses Artikels ist es nicht möglich, die gesamte Funktionalität des frei verfügbaren Programms zu erläutern. Wir verweisen diesbezüglich auf das ausführliche und detaillierte Benutzerhandbuch (<https://matthias-stemmler.github.io/animate/user-guide/>, Stand: 26.11.2025). Im Folgenden kann nur die allgemeine Funktionsweise erklärt und eine Übersicht der wichtigsten Features gegeben werden.

Um das Programm nutzen zu können, muss *Animate* zunächst über die GitHub-Seite heruntergeladen und installiert werden. Hierfür stehen Optionen für Linux, Windows und Mac zur Verfügung. Vor der ersten Verwendung müssen zudem die gewünschten Daten aus den ANNIS-basierten Korpora in das Programm importiert werden. Um etwa Abfragen

im Referenzkorpus Altdeutsch durchführen zu können, muss das Korpus in seiner aktuellen Version zunächst heruntergeladen werden, bspw. über die Webseite www.laudatio-repository.org/ (Stand: 26.11.2025), auf der verschiedene annotierte Korpora zur Verfügung gestellt sind. Daten im Format graphANNIS oder relANNIS können direkt in *Animate* importiert werden. Dieser Schritt muss nur einmalig durchgeführt werden. Bei erneuten Korpusabfragen kann das Korpus direkt in *Animate* aufgerufen werden. Mit *Animate* können Korpusabfragen daher auch ohne Internetzugang durchgeführt werden. Es sollte aber gelegentlich überprüft werden, ob das Korpus nach einem eventuellen Release in einer aktuelleren Version zur Verfügung steht. Um mit verschiedenen Korpora, wie beispielsweise dem ReM und dem ReA, in derselben *Animate*-Installation arbeiten zu können, können deren Teilkorpora entsprechend gruppiert werden. Damit kann über die gleiche Benutzeroberfläche auf die verschiedenen Korpora zugegriffen werden. Die Suche kann dann mit der gleichen Abfragesprache wie auf der Benutzeroberfläche der Referenzkorpora durchgeführt werden. Arbeitstechnisch ist es dabei von Vorteil, die Suchabfrage zunächst auf der Webseite des entsprechenden Korpus durchzuführen und *Animate* mit der entsprechenden Suchabfrage für den Export zu verwenden. Das stellt durch den Vergleich der Zahl der gefundenen Treffer auch sicher, dass die jeweils aktuelle Version des jeweiligen Korpus importiert ist.

Für den Export mit *Animate* wird das gewünschte Korpus ausgewählt, wobei es auch möglich ist, Teilkorpora (im Fall des ReA beispielsweise nur das Teilkorpus Tatian) zu durchsuchen. Analog zur webbasierten Benutzeroberfläche der Referenzkorpora wird die Suchanfrage in einem Feld links oben eingegeben, deren Gültigkeit durch das Programm geprüft und angezeigt wird. Bei einer gültigen Abfrage wird ein grünes Häkchen angezeigt, bei einer ungültigen Abfrage ein rotes Kreuz und eine Beschreibung des Fehlers. Auf der rechten Seite der Benutzeroberfläche können dann die jeweiligen Informationen zusammengestellt werden, die zu den Suchergebnissen ausgegeben werden sollen. Die Informationen sind in die folgenden Kategorien unterteilt, vgl. Tabelle 1.

Number	Zuweisung einer Referenznummer für jeden Treffer
Match in context	KWIC/Anpassung der Kontextlänge (= Anzahl der Tokens, die links bzw. rechts des Treffers ausgegeben werden sollen). Bei komplexeren Suchabfragen muss hier noch angegeben werden, für welchen Treffer der Kontext ausgegeben werden soll.
Match annotation	Information aus den Annotationen. Beispielsweise können morphologisch-syntaktische Merkmale oder das jeweilige Lemma ausgegeben werden.
Document metadata	Metadaten zum Dokument (bspw. Zeitraum, Region, etc.)
Corpus metadata	Metadaten zum Teilkorpus

Tab. 1: Übersicht über die jeweiligen Kategorien, zu denen über *Animate* Informationen zu den jeweiligen Treffern ausgewählt werden können

Ein Vorteil des Programms ist es, dass insbesondere für die Kategorien „Match annotation“ und „Document metadata“ aus einem Vorauswahl-Menü der in ANNIS suchbaren (Meta-) Annotationskategorien ausgewählt werden kann. Bei den morphologischen Annotationen muss demnach nicht zuerst auf der Basis der Annotationen des jeweiligen Korpus und des

Benutzerhandbuchs die jeweils annotierte Kategorie (bspw. lemma, inflection, language etc.) eruiert werden, sondern diese kann direkt aus dem Auswahl-Menü angeklickt werden. Ab der neueren Version 1.5.0 von *Animate* wird neben dem Exportbutton eine Information darüber angezeigt, ob die Angaben für die Parameterauswahl vollständig sind bzw. welche Angaben fehlen.

Mit Ausnahme der Kategorie „Match in context“ erstellt *Animate* für jede der ausgewählten Kategorien eine Spalte in der Ausgabedatei. Für „Match in context“ werden für jeden Treffer zusätzlich zwei Spalten (eine für den linken, eine für den rechten Kontext) ausgegeben. Nach Angabe der benötigten Informationen können die Ergebnisse der Suchabfrage mit den erwünschten Informationen wahlweise im CSV- oder Excel-Format exportiert werden. Neben den Suchergebnissen (vgl. Abb. 5) werden auf einem zweiten Excel-Sheet die Informationen zur Abfrage dokumentiert.

Number	Document	Document time	ei/ey inflection	Context 1 (text)	Match 1 (text)	Context 2 (text)	Match 2 (text)
1	AGO_AlemGlaubensBeichte	12-13	IND_PRES_SG_1	minen werchen: den lon den furht ich ser, wan ich dicke	gesundot		han
2	BGG3_BenediktbeurerGlaubensBeichte	12-13	SUBJ_PRES_SG_3	an disiu wort dencket, wie vortlich dei sin, der sich	versumt		habe
3	BGG3_BenediktbeurerGlaubensBeichte	12-13	SUBJ_PRES_SG_2	der sich versumt habe nemir durch sine tracht, das er sin miut	glenet		habe
4	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_2	, er chome ze rāve. EXORTATIO AD CONFESSIÖNEM. Nu	habet	ir luch	gewestenet
5	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_2	erchennen alles iwers unvretts, uch sol vil harte riwen swaz ir wider gotis hilden	habet		getan
6	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_3	sele wisen in die gnade unde in die vroude, die si selbe	besezen		habent
7	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_2	Nach so getane lichte unde nach dem gehelze, den ir unsern herren got	gehessen		habet
8	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_3	. Vil gueten liute, so getanu bihte hilfet einigoten die ir bihte tougliche	habont		getan
9	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_3	bihte tougliche habent getan unde die ouch tougliche suntint, die auer offentlich	habent		gesumtit
10	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_3	manslachte, uberhuor, sippehuor, swelhe die sint, die houphafte sunte	habent		getan
11	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_3	manslachte, uberhuor, sippehuor, swelhe die sint, die houphafte sunte	habent		getan
12	BGG3_BenediktbeurerGlaubensBeichte	12-13	IND_PRES_PL_3	manslachte, uberhuor, sippehuor, swelhe die sint, die houphafte sunte	habent		getan
13	BR1_BaslerRezept1	6	SUBJ_PRES_SG_3	ni neouuht ni uoite, nripz de geshe, de imo daz tranc gebe enti simplum	pliuartan		haben
14	BGG3_BenediktbeurerGlaubensBeichte	12	SUBJ_PRES_SG_3	aller mēniglich uere sich selban gode reda geban sō, suie s er	glēbet		haben
15	DD_DeDefinitione	11-12	IND_PRES_SG_2	animal rationale et mortale. Taz ist imo gemache, mit thu	habot	tu in	genoman
16	GG83_SangallerBeichte_1	11-12	IND_PRES_SG_1	unde fone sancte Petre unde fone allen gotes heiligen, so filo hich keuualdes	haben		anfangen
17	GG83_SangallerGlaubensBeichte	11-12	IND_PAST_PL_3	begrabin wart unde er uon der helle löste alle, die sinin willin	hāton		gitan
18	GG83_SangallerGlaubensBeichte	11-12	IND_PRES_SG_1	nach minin werchin. Den lon vurht ich sere, wand ich	gesundot		han
19	MM_MementoMori	11-12	IND_PRES_PL_3	heit, si ne dāltu sie nie sō minnesen, si	habent	si re-uh	verlaeren
20	MM_MementoMori	11-12	IND_PRES_PL_2	sie des gedāhton, war sie ze iungest vān solton i nō	habint	sku iz	beuonden
21	MM_MementoMori	11-12	IND_PRES_PL_2	ze wesinne, taz ir wārint als ein man: taz	hānt	ir uber	gangen
22	MM_MementoMori	11-12	SUBJ_PAST_PL_2	als ein man: taz hānt ir uher gangen,	habetint	ir anders niuwt	getan
23	MM_MementoMori	11-12	IND_PRES_SG_3	die hella: dā muoz er iemer inne wesen: got selben	habot	er hin	gegeben
24	MM_MementoMori	11-12	IND_PRES_SG_3	tes ne mag imo der scāz ze guote werden:	habot	er sinin richtuom sō	geleit
25	MM_MementoMori	11-12	IND_PRES_SG_3	er hie verleitet, taz wirt imo ubilo geteilt.	habot	er iet hina	gegeben
26	MM_MementoMori	11-12	IND_PRES_SG_3	er wol mac: hie noh chumit der fac:	habot	er is tenne niuwt	geschin
27	MM_MementoMori	11-12	IND_PRES_SG_2	vil ubelir munde, wie betrugt tō uns sus i dō	habot	uns	geben
28	MM_MementoMori	11-12	IND_PRES_SG_3	unde lb, alō lango sō wir hie leibin, got	habot	uns selbwa	geben
29	M_Muspilli	9-12	IND_PAST_SG_3	Dar scal er vora demo rihē ar rahu stantan, pi daz er in uuertio eo	klauerkot		hapeta
30	M_Muspilli	9-12	IND_PRES_SG_3	enti mit fastun dō virma kipuazt. Demne der paldet der	kipuazt		hapet

Abb. 5: Screenshot: Beispiel für die Ausgabe der Treffer und Zusatzinformationen im Excel-Format

Als weitere arbeitspraktische Funktion ermöglicht *Animate* zudem die Speicherung der Suchabfragen in Projekten. Auf diese Weise kann die Suchabfrage mit der entsprechenden Konfiguration zu einem späteren Zeitpunkt wiederholt bzw. modifiziert durchgeführt werden. Auch macht es diese Funktion möglich, die Suchabfragen innerhalb einer Arbeitsgruppe zu teilen und zu reproduzieren.

Insgesamt ist *Animate* mit seinen verschiedenen Funktionen – Überprüfung der Suchabfrage mit Fehleranzeige, das personalisierte Zusammenstellen und Speichern von Teilkorpora, die automatische Dokumentation der Suchabfrage, die Ausgabe der Belege mit variabel anpassbarer Kontext-Länge, ein Vorauswahl-Menü der in ANNIS suchbaren (Meta-) Annotationskategorien und die Speicherung der Such-Konfiguration – ein Programm, das den Export von Ergebnissen aus ANNIS-basierten Daten entscheidend vereinfacht. Im folgenden Abschnitt soll das an einer kleinen diachronen Studie zur Verwendung der Interjektion *ei/ey* im Deutschen demonstriert werden.

3. Fallstudie

Das Potenzial des Programms demonstrieren wir im Folgenden anhand einer kleinen Korpusstudie zur Diachronie der Interjektion *ei/ey*. Als Merkmal der gesprochenen Sprache und als Subjektivitätsmarker, die evaluative oder emotive Äußerungen einer Sprechinstanz indizieren, sind Interjektionen allgemein ein interessanter Untersuchungsgegenstand zur Erforschung historischer Nähesprachlichkeit (vgl. im Detail Zeman im Ersch.). Wir greifen hier exemplarisch die Interjektion *ei/ey* heraus, da diesbezüglich in der Forschungsliteratur bereits einige Beobachtungen zu ihrer diachronen Entwicklung vorliegen (vgl. Kilian 2012; Müller 2022; Imo/Müller 2023), die an den historischen Korpora überprüft werden können.

ei/ey ist laut dem Grimmschen Wörterbuch in althochdeutscher Zeit nicht belegt, kommt aber ab dem Mhd. häufig vor³ und dient dort dem Mittelhochdeutschen Handwörterbuch von Matthias Lexer zufolge als Ausruf der Verwunderung, Freude und Klage.⁴ Das Grimmsche Wörterbuch attestiert zudem die Bedeutungen des Staunens und des Liebkosens. In Bezug auf ihren Gebrauch ist bekannt, dass die Interjektion häufig mit dem Vokativ und dem Imperativ (*ei mutter seht doch her!*) sowie in Fragen (*ei was?*), Ausrufen mit *w*-Phrasen (*ei wie gern!*), Wunschsätzen und Verwünschungen verwendet wird.⁵ Spätestens im 18. Jahrhundert zeichnet sich dann ein Sprachwandel ab, insofern als die Interjektion Kilian (2012) zufolge in diesem Zeitraum als gesprächsschritteinleitende Partikel (*Ei, was redst du doch*) verwendet wird, die häufig „ablehnende bzw. widersprechende Funktion des Gesprächsschritts ankündigt“ (Kilian 2012, S. 117).⁶

Mit dem Fokus auf dem Tool *Animate* kann es in der Kürze des vorliegenden Artikels nicht das Ziel sein, den Sprachwandel der Interjektion in allen Details nachzuzeichnen (vgl. dazu insbesondere die umfangreiche Studie von Imo/Müller 2023). Im Folgenden geht es dagegen um die Frage, ob die Interjektion auch in den historischen Sprachstufen als „gesprächsschritteinleitende“ bzw. rede-einleitende Partikel verwendet wird.

Zu diesem Zweck vergleichen wir die Vorkommen von *ei/ey* im mittelhochdeutschen Referenzkorpus (ReM, 1050–1350) und dem GiesKaNe-Korpus (1650–1950). Die Belege im Mittelhochdeutschen Referenzkorpus elizitieren wir mit *Animate* über die Suchabfrage `pos="ITJ"`. Für unsere Fragestellung hätte es ausgereicht, nach dem entsprechenden Lemma zu suchen. Uns ist es aber wichtig, *ei/ey* im Vergleich zu anderen Interjektionen einschätzen zu können. Daher präferieren wir die Suche nach allen Belegen, die als Interjektionen annotiert sind.

Dazu wird die Suchabfrage `pos="ITJ"` auf der Benutzeroberfläche in *Animate* eingegeben und das ReM-Korpus in der Version 2.1 ausgewählt, das wir bereits vorher importiert haben. Wir klicken auf „All“, um alle Teilkorpora des ReM bei der Suche zu berücksichtigen. Die Suchabfrage wird dann automatisch überprüft. Wird sie als gültig angezeigt, dann lassen sich die entsprechenden Parameter auswählen, die am Ende in der ausgegebenen Export-Liste erscheinen sollen. In diesem Fall wählen wir die Parameter mit folgenden Werten, vgl. Tabelle 2:

-
- 3 „ei“, Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm, digitalisierte Fassung im Wörterbuchnetz des Trier Center for Digital Humanities, Version 01/25. www.woerterbuchnetz.de/DWB?lemid=E00867, (Stand: 31.5.2025).
 - 4 „ei, interj.“, Mittelhochdeutsches Handwörterbuch von Matthias Lexer, digitalisierte Fassung im Wörterbuchnetz des Trier Center for Digital Humanities, Version 01/25. www.woerterbuchnetz.de/Lexer?lemid=E00323, (Stand: 31.5.2025).
 - 5 „ei“, Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm, digitalisierte Fassung im Wörterbuchnetz des Trier Center for Digital Humanities, Version 01/25. www.woerterbuchnetz.de/DWB?lemid=E00867, (Stand: 31.5.2025).
 - 6 Die häufig als Merkmal der Jugendsprache angeführte Partikel *ey* teilt die Funktionen des attention gettings, der Redeeinleitung und der emotionalen Färbung der Anrede (vgl. die Funktionsbeschreibung in Bahlo et al. 2019, S. 62). Sie geht dem Duden zufolge jedoch auf eine englische Entlehnung zurück.

Number	
Document metadata	Meta annotation = doc
Document metadata	Meta annotation = time
Document metadata	Meta annotation = language-type
Match in context	Segmentation = tok_anno; Left/Right context = 20
Match annotation	Annotation = norm; Query node = #1 pos="ITJ"

Tab. 2: Übersicht über die ausgewählten Werte in *Animate*

Die Ergebnisse der Suchabfrage lassen wir uns als Excel-Liste ausgeben. Von den gefundenen Interjektionen beschränken wir unsere Analyse dann auf diejenigen Interjektionen, die in der „match“-Spalte als „ei“, „ay“ oder „ey“ aufgeführt sind, sowie diejenigen, die als norm=„ei“ annotiert sind. Nach einer zusätzlichen manuellen Überprüfung ergeben sich insgesamt 200 Treffer.⁷

Für die qualitative Klassifikation ergänzen wir die Excel-Liste noch um folgende Spalten, in denen wir manuell die Werte für folgende Parameter annotieren.

Verwendung	+ Vokativ, + Imperativ, + w-Phrase, sonstige
Rede-Einleitung	RE (rede-initial) – NRE (nicht rede-initial)

Tab. 3: Übersicht über die Annotations-Parameter, die manuell klassifiziert werden

Die quantitative Auswertung der Belege offenbart in Übereinstimmung mit der Beschreibung im Grimmschen Wörterbuch folgende Verwendungsweisen.

Die häufigste Verwendung (53,5%) sind Exklamationen mit anschließendem Vokativ, dem teilweise noch ein Imperativ folgt. Der Kontext ist durch sprecher-orientierte Verweise und Bezüge auf die erste und zweite Person charakterisiert, vgl. (1).

- (1) **ey selig wip des volgit mime rade** [M331 – Hessische Reimpredigten, 14. Jh.]
 ‚Ei, selige Frau, folgt meinem Rat‘

Eine weitere häufige Verwendung (26%) sind Exklamationen im Kontext von w-Phrasen.

- (2) **Ei wie gerne weres tu gestorven. [...] ei wie gerne heddes tu din leven. vmbe dines svones leven gegeven.** [M335 – Rheinisches Marienlob, 13. Jh.]
 ‚Ei wie gerne wärest du gestorben. [...] Ei wie gerne hättest du dein Leben um deines Sohnes Leben gegeben.‘

Unter die Rubrik „sonstige“ fallen Verwendungen in hypothetischen Sätzen (*ei wüßte ich...*), in Wunschsätzen (*ei, möge es ihm wohl ergehen*), in Konditionalsätzen und weiteren Satzarten.

Mit wenigen Ausnahmen steht *ey* satz-initial, aber nicht unbedingt rede-initial. Das zeigt sich in folgendem Beleg aus einer längeren direkten Rede. *Ey* markiert nicht den Beginn

7 Neben *ei/ey* findet sich im ReM etwa auch die Interjektion *eia/eya*, die aber funktional nicht ganz synonym zu sein scheint und die wir daher nicht berücksichtigen. Auch das fragliche Verhältnis zwischen *ei/ey* und *hei/hey* (Imo/Müller 2023) blenden wir im begrenzten Umfang dieses Beitrags aus.

der direkten Rede. Vielmehr ist die Verwendung deutlich durch die Exklamation in Kombination mit dem Vokativ *herre* motiviert.

- (3) *Nv han ich disem lieben man Ni kein vngemach getan war vm zvrnet er wid' mich ey herre weder bin ich Im zv iunc od' zv alt od' bin ich als vngestalt daz er min nicht zv wiebe wil* [M311 – Heinrich von Freiberg: Tristan, 14. Jh.]
 ‚Nun habe ich diesem lieben Mann niemals ein Ungemach angetan. Warum zürnt er gegen mich? Ei Herr, weder bin ich ihm zu jung oder zu alt oder bin ich so ungestalt, dass er mich nicht zur Frau nehmen will.‘

Die Setzung von *ey* kann allerdings auch mit einem neuen Gedanken verknüpft sein. Es zeichnet sich damit bereits das Potenzial der funktionalen Verschiebung zu einem „rede-schritteinleitenden“ Element ab. Gerade in narrativen Texten wird *ey* als rede-einleitendes Element nach Verba dicendi verwendet (4) oder als Einleitung eines neuen Erzählabschnitts durch den Erzähler (5).

- (4) *Zv dem ritter sprach er do Ey dv vngetruwer man Waz has tu leides mir getan* [M326 – Passional, 14. Jh.]
 ‚Zu dem Ritter sprach er da: „Ei, du untreuer Mann. Was hast du mir Leid angetan.“‘
- (5) *ey waz tut nv her tristan da kerte er ab' drate*
 [M311 – Heinrich von Freiberg: Tristan, 14. Jh.]
 ‚Ei, was tut nun Herr Tristan? Da kehrte er schnell um...‘

Die Verwendungen in der Funktion als Rede-Einleitung machen insgesamt aber nur einen geringen Anteil von 16% aus (32/200).

Ei/ey im ReM	200	davon rede-einleitend: 32/16%
Vokativ (teils mit Imp.)	107	53,5%
Imperativ (ohne Vok.)	12	6%
w-Phrasen	54	27%
Andere	27	13,5%

Tab. 4: Übersicht über die Funktionen der Interjektion *ei/ey* im ReM

Im Vergleich zum ReM ergibt sich im GiesKaNe-Korpus⁸ eine andere funktionale Verteilung für die Interjektion *ey*. Um zunächst wieder die entsprechenden Belege zu elizitieren, formulieren wir die entsprechende Suchabfrage *wa="i t j"*.⁹ Die Suche liefert insgesamt 193 Treffer. Aufgrund von fünf Belegen mit der doppelten Interjektion *ey ey*, die wir jeweils als einen Treffer werten, ergibt sich eine Gesamtanzahl von 188 Belegen. Quantitativ ergibt sich, dass *ey* bzw. *ey ey* mit 21 von 188 Belegen neben *ach* (63/188) und *o/oh* (54/188) zu den drei frequentesten Interjektionen im Korpus zählt. Dieses Ergebnis korreliert mit der Verteilung der Interjektionen in den Dramen von Andreas Gryphius, wie sie von Müller (2022) untersucht wurden. Auch hier sind *ach*, *ey*, und *o* mit Abstand die frequentesten Interjektionen. Das deutet darauf hin, dass die literarischen Texte der Zeit durchaus den Gebrauch der Interjektionen in den Gebrauchstexten widerspiegeln. Zu berücksichtigen ist, dass die Treffer-Anzahl von 21 *ei/ey*-Belegen sowohl im Vergleich zum ReM (n=200) als auch zur

8 Wir danken Mathilde Hennig und Volker Emmrich für die Bereitstellung der Korpusdaten im ANNIS-Format.

9 Zur Vergleichbarkeit der Daten mit dem ReM präferieren wir auch hier eine Suchabfrage nach allen Vorkommen der Interjektionen.

Studie von Müller (2022) (n=107) gering ist, was einerseits an der Auswahl der Textsorten und berücksichtigten Regionen liegen kann, andererseits an der Größe des noch weiter im Aufbau befindlichen Korpus. Das GiesKaNe-Korpus umfasst zum derzeitigen Stand (Juli 2025) ca. 300.000 Tokens, das ReM dagegen 2 Mio. Wortformen.

Die qualitative Klassifikation der Belege zeigt, dass *ey* bzw. *ey ey* zu 90% in der direkten Rede verwendet wird (19 von 21 Belegen). Die meisten Belege (17/21) kommen im Text Nehrlich¹⁰ vor, der sich durch einen hohen Anteil an direkter Rede (23%, vgl. Zeman im Ersch.) auszeichnet. *Ey* hat in diesem Text eine appellative, aber auch eine diskursstrukturierende Funktion, indem es als Rede-Einleitung verwendet wird, vgl. (6). Diese Funktion ist besonders in dialogischen Passagen mit Sprecherwechseln zu erkennen, in denen *ey* auch als Signal für ‚turn-taking‘ verwendet wird, vgl. (7) und (8), in denen *ey* jeweils einen Sprecherwechsel anzeigt.

- (6) [...] und fing an mit beweglichen wortten *Ey ey*, du lieber Hanß Ludwig was machstu mit dem buch [Nehrlich, 18. Jh.]
- (7) und eröffnete mir: Meine zwey Geissen seyen auf und da von. „*Ey* so fahre denn alles hin“! sagt’ ich, „wenn’s so seyn muß“ [Nehrlich, 18. Jh.]
- (8) ich sagte Herr Suprint ich weiß nicht, was Cloßen sind, sagen sie mir doch teutsch, *ey* ob ihr s wist oder nicht das hilft euch nichts. [Nehrlich, 18. Jh.]

Tabelle 5 zeigt die unterschiedliche funktionale Verteilung.

<i>Ei/ey</i> in GiesKaNe	21	davon rede-einleitend: 19/90%
Vokativ	9	43%
Imperativ	1	5%
w-Fragen	4	19%
Andere	7	33%

Tab. 5: Übersicht über die Funktionen der Interjektion *ei/ey* im GiesKaNe-Korpus

Die geringe Zahl von 21 Belegen im Gesamtkorpus ist zwar nur eingeschränkt aussagekräftig. Auch ist die Funktion der Rede-Einleitung hauptsächlich in einem Text der GiesKaNe belegt. Allgemeine Aussagen zum Sprachwandel der Form können daher kaum auf der vorliegenden Datenbasis getroffen werden. Vor dem Hintergrund bereits vorliegender Studien (vgl. Kilian 2012; Müller 2022; Imo/Müller 2023) zeigt die Verteilung aber durchaus ein deutliches Muster. Die häufige Verwendung im Nehrlich-Text deutet darauf hin, dass die Funktion der Partikel als Rede-Einleitung etabliert ist: In 19 von 21 Belegen wird *ey* bzw. *ey ey* in dieser Funktion verwendet.¹¹

Weitere empirische Untersuchungen könnten die Hypothese eines Funktionswandels von der emotiven Funktion zu einem rede-einleitenden Element weiter überprüfen. Das Muster der Verwendung im Mittelhochdeutschen deutet zudem darauf hin, wie der Funktions-

10 Nehrlich, Hans Ludwig. 1723. Erlebnisse eines frommen Handwerkers im späten 17. Jahrhundert. Herausgegeben von Rainer Lächele. Tübingen: Niemeyer 1997.

11 Dieser Befund passt auch zur Studie von Müller (2022, S. 300), die *ey* eine „sehr eindeutige Komödienspezifika“ attestiert. Aufgrund ihres Bezugs zur dialogischen Kommunikation wäre es naheliegend, dass *ey* in den Dramen von Gryphius als Indikator gesprochener Sprache verwendet wird (vgl. auch Zeman im Ersch.).

wandel motiviert sein könnte: Die enge Verbindung mit dem Vokativ bedingt dort, dass *ei/ey* gelegentlich turn-initial und rede-einleitend gebraucht wird. Auch im GiesKaNe-Korpus wird *ey* häufig in der Umgebung des Vokativs verwendet. Der Vokativ steht aber nicht mehr direkt hinter der Partikel, was darauf hindeutet, dass *ey* nicht mehr vornehmlich als Anrede-Formel mit dem Vokativ verbunden ist, sondern die Rede-Sequenz eröffnet, in der dann auch ein Vokativ erscheinen kann.

(9) *eÿ eÿ sehe er doch Herr gefatter, was der Man in das buch geschriben.* [Nehrlich, 18. Jh.]

Auch ist der Anteil von rede-einleitenden Verwendungen mit 90% deutlich höher als im Referenzkorpus mit 16%. Wenngleich berücksichtigt werden muss, dass sich die Textzusammenstellungen im GiesKaNe-Korpus und im ReM sowohl qualitativ als auch quantitativ nicht problemlos vergleichen lassen, deutet sich hier eine klare funktionale Verschiebung an, die mit Hilfe von *Animate* leicht sichtbar gemacht werden kann.

Als interessante Nebenergebnisse für das ReM lassen sich zudem folgende Befunde festhalten, die wir uns mit *Animate* über den Parameter „Document metadata“ für die Werte „time“ und „language-type“ ausgeben haben lassen: Im Vergleich der Jahrhunderte ergibt sich, dass der Anteil der Verwendungen von *ei/ey* als Rede-Einleitung zunimmt (von 4% im 13. Jh. zu 35% im 14. Jh. – für das 11. Jh. wird kein entsprechender Treffer gefunden, für das 12. Jh. lediglich einer). Zudem zeichnet sich ab, dass die Interjektion zu 89% im mitteldeutschen Sprachraum verwendet wird. Dagegen sind nur 22 der 200 Belege im oberdeutschen Sprachraum belegt.

Diese Korpusbeobachtungen können natürlich nicht ohne eine sorgfältige Diskussion der Repräsentativität und Vergleichbarkeit der Daten interpretiert werden. Deutlich ist in jedem Fall geworden, dass sich aus den mit *Animate* erzielten Suchergebnissen nicht nur auf komfortable Weise Korpus-Untersuchungen durchführen lassen, die zu aufschlussreichen Ergebnissen bei der Untersuchung von Sprachwandel führen können, sondern sich aus den gewonnenen Ergebnissen wiederum weiterführende Anschlussfragen und neue Hypothesen ableiten lassen.

4. Evaluierung von Korpora

Animate ist ein Tool für den Export von Belegen. Exportiert werden können demnach auch nur Informationen, die in den entsprechenden Korpora vorhanden sind. Das bedeutet, dass auch nicht-korrekte Annotationen, wie sie in allen größeren Korpora zu unterschiedlichen Graden vorhanden sind, zu den jeweiligen Treffern ausgegeben werden. *Animate* kann aber auch dazu eingesetzt werden, Korpora zu evaluieren und damit die Qualität der Daten zu verbessern. Wir demonstrieren dieses Potenzial anhand von Daten des GiesKaNe-Korpus.

In der Fallstudie zur Interjektion *ey* in Abschnitt 3 wurde überprüft, ob *ey* rede-initial verwendet wird. Für das mittelhochdeutsche Referenzkorpus lässt sich diese Information nicht aus den Korpusdaten gewinnen. Die Angabe ‚rede-initial – nicht-rede-initial‘ wurde daher in der mit *Animate* ausgegebenen Excel-Liste manuell annotiert. Im GiesKaNe-Korpus werden Zitate dagegen auf der textgrammatischen Makro-Ebene auf der Annotations-ebene „Sonstige Strukturen“ (*strukt*) annotiert. Anhand der Annotation lassen sich rede-einleitende Interjektionen damit durch eine jeweilige Suchabfrage elizitieren. Voraussetzung dafür ist jedoch die korrekte Annotation. Um die Annotationen für *strukt* systematisch zu überprüfen, lassen sich mit *Animate* die gesamten Tokens eines Textes und die jeweils zu überprüfende Annotation im Excel-Format ausgeben. In unserem Fall führen

wir die Suche `tok @* doc="nehrlich"` aus und wählen in *Animate* die folgenden Parameter, vgl. Tabelle 6.

Number	
Match annotation	Annotation = tok; Query node = #1 tok
Match annotation	Annotation = strukt; Query node = #1 tok

Tab. 6: Übersicht über die ausgewählten Werte in *Animate*

Animate gibt dann eine Datei aus, die eine Spalte mit den gesamten Tokens des Textes Nehrlich sowie eine Spalte mit den Annotationen enthält, die für die Annotationsebene `strukt` hinterlegt sind (z.B. Zitate, Parenthesen, Vokative). Diese Spalten können dann systematisch im jeweiligen Textzusammenhang überprüft werden.

5. Fazit

Animate vereinfacht den Datenexport aus den historischen Referenzkorpora signifikant und ist damit ein intuitives und zeitsparendes Tool zur Datenaufbereitung, das nicht nur bei jeder Art von sprachhistorischen sowie diachronen Untersuchungen hilfreich ist, die auf Daten von historischen Korpora im ANNIS-Format basieren, sondern auch einen vielfältigen Einsatz in der Lehre ermöglicht.

Lassen sich diese Studien auch ohne *Animate* durchführen? Zweifellos. Der bislang schwierige Export führt allerdings dazu, dass manchmal nicht auf die Referenzkorpora zurückgegriffen wird, sondern beispielsweise eigene Textkorpora erstellt und den Untersuchungen zugrunde gelegt werden. Das wird für einige Fragestellungen der historischen Sprachwissenschaft auch weiterhin unausweichlich sein. Für andere Fragestellungen stellt gerade aber auch die Vergleichbarkeit der Ergebnisse auf der Datenbasis der Referenzkorpora einen Vorteil dar, der in der historischen Sprachwissenschaft entsprechend genutzt werden können sollte. Hier stellt die Verwendung von *Animate* nicht nur für die individuelle Arbeitspraxis, sondern auch für die Vergleichbarkeit der Studien einen großen Mehrwert dar.

Animate macht zudem den Export so schnell und einfach, dass sich das Programm auch in der Lehre einsetzen lässt. Durch die Praktikabilität kann das Programm auch für Korpusabfragen in studentischen Arbeiten (nicht nur) zur historischen Sprachwissenschaft verwendet werden. Wir hoffen daher, dass das Tool nicht nur ein arbeitspraktisches Problem löst, sondern auch die Attraktivität diachroner Studien in den historischen Referenzkorpora im Vergleich zu weiteren ANNIS-basierten Korpora erhöhen und damit positive Auswirkungen auf die historische Sprachwissenschaftsforschung insgesamt haben wird. Ein weiteres großes Potenzial sehen wir zudem im Einsatz von *Animate* bei der Evaluierung der Korpora. *Animate* kann auch dazu beitragen, die Qualität der Korpora zu erhöhen. Gerade weil die Referenzkorpora inzwischen einen Status als Standard-Ressourcen innerhalb der Sprachwissenschaft des Deutschen innehaben, ist es unabdingbar, die Qualität der Annotationen zu überprüfen und zu verbessern. *Animate* bietet auch hierfür ein großes Potenzial.

Ressourcen

Animate zum Download auf GitHub: <https://github.com/matthias-stemmler/animate> (Stand: 1.6.2025)

Repository zum Download von Korpora: www.laudatio-repository.org (Stand: 1.6.2025).

Stemmler, Matthias (o. D.): *Animate User Guide*. <https://matthias-stemmler.github.io/animate/user-guide/> (Stand: 1.6.2025).

Literatur

Ägel, Vilmos/Hennig, Mathilde (Hg.) (2006): *Grammatik aus Nähe und Distanz. Theorie und Praxis am Beispiel von Nähetexten 1650–2000*. Tübingen: Niemeyer.

Bahlo, Nils U./Becker, Tabea/Kalkavan-Aydın, Zeynep/Lotze, Netaya/Marx, Konstanze/Schwarz, Christian/Şimşek, Yazgül (2019): *Jugendsprache. Eine Einführung*. Stuttgart: Metzler.

Elspaß, Stephan (2005): *Sprachgeschichte von unten. Untersuchungen zum geschriebenen Alltagsdeutsch im 19. Jahrhundert*. (= Reihe Germanistische Linguistik 263). Tübingen: Niemeyer.

Elspaß, Stephan (2010): Zum Verhältnis von ‚Nähegrammatik‘ und Regionalsprachlichkeit in historischen Texten. In: Ägel, Vilmos/Hennig, Mathilde (Hg.): *Nähe und Distanz im Kontext variations-linguistischer Forschung*. (= Linguistik – Impulse & Tendenzen 35). Berlin/New York: De Gruyter, S. 63–84.

Imo, Wolfgang/Müller, Melissa (2023): „Von „Ey Pickelhåring/das ist wider Ehr und Redligkeit“ zu „ey Timo; lass ma RISCHtisch laut (.) öh schrElen“ – ey und ei gestern und heute. In: Imo, Wolfgang/Wesche, Jörg (Hg.): *Interaktionale Sprache im Dramenwerk von Andreas Gryphius. Literatur- und sprachwissenschaftliche Studien*. (= Sprache – Literatur und Geschichte 53). Heidelberg: Winter, S. 99–186.

Kilian, Jörg (2012): »Man spricht hier in Meißen oft: Je nu!« Historische Gesprächswörter vom 17.–21. Jahrhundert. In: Leupold, Gabriele/Passet, Eveline (Hg.): *Im Bergwerk der Sprache: Eine Geschichte des Deutschen in Episoden*. Göttingen: Wallstein, S. 102–123.

Klein, Thomas/Dipper, Stefanie (2016): *Handbuch zum Referenzkorpus Mittelhochdeutsch*. In: *Bochumer Linguistische Arbeitsberichte* 19. <https://linguistics.rub.de/forschung/arbeitsberichte/19.pdf> (Stand: 31.5.2025).

Krause, Thomas/Zeldes, Amir (2016): ANNIS3: A new architecture for generic corpus query and visualization. In: *Digital Scholarship in the Humanities* 2016, 31, S. 118–139. <http://dsh.oxfordjournals.org/content/31/1/118> (Stand: 31.5.2025).

Müller, Melissa (2022): *Partikeln im Frühneuhochdeutschen. Eine korpuslinguistisch-interaktionale Untersuchung des Dramenwerks von A. Gryphius*. Diss. Hamburg: Universität Hamburg. <https://ediss.sub.uni-hamburg.de/bitstream/ediss/10841/1/2024-04-02%20Diss%20M%C3%BCller%20Melissa.pdf> (Stand: 31.5.2025).

Petran, Florian/Bollmann, Marcel/Dipper, Stefanie/Klein, Thomas (2016): ReM: A reference corpus of Middle High German – corpus compilation, annotation, and access. In: *Journal for Language Technology and Computational Linguistics* 31, 2, S. 1–15. <https://jllc.org/article/view/208/206> (Stand: 25.3.2025).

Prell, Heinz-Peter (2005): Zur Neubearbeitung der mittelhochdeutschen Syntax von Ingeborg Schröbler. Mit einem Vergleich von Vers- und Prosatexten am Beispiel der Nominalphrase. In: Simmler, Franz (Hg.): *Syntax. Althochdeutsch – Mittelhochdeutsch. Eine Gegenüberstellung von Metrik und Prosa. Akten zum Internationalen Kongress an der Freien Universität Berlin 26. bis 29. Mai 2004*. (= Berliner sprachwissenschaftliche Studien 7). Berlin: Weidler, S. 209–222.

Wegera, Klaus-Peter (1990): Mittelhochdeutsche Grammatik und Sprachgeschichte. In: Besch, Werner (Hg.): Deutsche Sprachgeschichte. Grundlagen, Methoden, Perspektiven. Festschrift für Johannes Erben zum 65. Geburtstag. Frankfurt a. M.: Lang, S. 103–113.

Zeman, Sonja (im Ersch.): Explorationen in GiesKaNe: Grammatik des Neuhochdeutschen zwischen Nähe, Distanz und Narration. In: Ágel, Vilmos/Linnenkohl, Marcel/Schäfer, Karolin/Sievers, Laura (Hg.): Grammatik des Neuhochdeutschen zwischen Gegenwart und Geschichte. Berlin/New York: De Gruyter.

Kontaktinformation

Matthias Stemmler

E-Mail: matthias.stemmler@gmail.com

Sonja Zeman

Lehrstuhl für Deutsche Sprachwissenschaft

Universität Augsburg

Universitätsstraße 10

86159 Augsburg

E-Mail: sonja.zeman@uni-a.de

Bibliografische Angaben

Dieser Text ist Teil der Publikation: Brunner, Annelen/Hansen, Sandra/Lang, Christian/Tu, Ngoc Duyen Tanja/Wolfer, Sascha (Hg.) (2026): Deutsch im Wandel – Tools, Methoden, Ressourcen. Beiträge zur Methodenmesse der IDS-Jahrestagung 1. (= *IDSopen* 16). Mannheim: IDS-Verlag. <https://doi.org/10.21248/idsopen.16.2026.63>.