

Spiele im Spiel – Datenbankbasiertes Arbeiten zur interaktionale Sprache im Dramenwerk von Andreas Gryphius

Eggert, Lisa

lisa.eggert@uni-due.de

Universität Duisburg-Essen, Deutschland

Müller, Melissa

melissa.mueller@uni-hamburg.de

Universität Hamburg

Die Bühnensprache des Barocktheaters – insbesondere der Trauerspiele – ist normiert. Der Alexandriner als Sprechvers mit seiner festen Anzahl an Hebungen und dem Paarreim scheint trotz der dialogischen Struktur der Texte kaum Spielräume für eine Mündlichkeit im heutigen Sinne zu lassen. Die gebundene und dadurch disziplinierte Sprache der Dramen steht in Spannung zu gängigen Vorstellungen von Mündlichkeit, gilt diese doch als spontan, wenig an Normen orientiert und in der Tendenz individuell statt standardisiert.

Welche Spielräume lässt die Versifizierung dennoch zu? Welche Formen konzeptioneller Mündlichkeit lassen sich in barocken Dramentexten finden? Lassen sich Korrelationen zwischen linguistischen Phänomenen interaktionaler Sprache und den verschiedenen Möglichkeiten der Vergestaltung feststellen? Diese und weitere Fragen werden im Rahmen des DFG-Projektes *Interaktionale Sprache bei Andreas Gryphius – datenbankbasiertes Arbeiten zum Dramenwerk aus linguistisch-literaturwissenschaftlicher Perspektive* bearbeitet.

Beschäftigt sich das Projekt insgesamt inhaltlich mit Spielräumen, die im hochnormierten Trauerspiel gefunden werden können, so nimmt der Vortrag einen strukturellen Aspekt der Projektarbeit in den Blick: die Annotation von Versmaßen und ihren Abweichungen. Dabei sollen jene Spielräume in den Blick genommen werden, die an der Schnittstelle von fachlichen Anforderungen und technischen Sachzwängen entstehen und ausgehandelt werden müssen.

Projektbeschreibung und Forschungsgegenstand

Das Ziel des genannten Projektes ist es, mit Hilfe einer annotierten Datenbank (auf der Basis der

Datenbankarchitektur ANNIS3; <http://annis-tools.org/> – Krause/Zeldes (2016)), die das vollständige Dramenwerk von Andreas Gryphius enthält, die oben ausgeführten und weitere Fragen korpusbasiert aus literatur- und sprachwissenschaftlich übergreifender Sicht zu klären.

Hierbei wird interaktionale Sprache verstanden als Form sprachlichen Handelns, das unabhängig von seiner medialen Realisation (mündlich oder schriftlich) durch die gemeinschaftliche Erzeugung von Bedeutung von zwei oder mehr Sprecher*innen gekennzeichnet und sequenziell organisiert ist.

Das Projektkorpus besteht aus sämtlichen Dramen von Andreas Gryphius (1616-1664). Die Verschriftlichung folgt der historisch-kritischen Dramen-Ausgabe Eberhard Mannacks (1991). Die Annotation findet mithilfe des Annotationstools INCEpTION (<https://inception-project.github.io/>) statt.

Durch die Annotation von sowohl linguistischen als auch literaturwissenschaftlichen Phänomenen kann ein umfassenderer Blick auf interaktionale Sprache geworfen werden, da hier die sprachlichen Merkmale konzeptioneller Mündlichkeit mit literarischen Darstellungselementen konsequent eingeführt werden. Mit der Anwendung einer korpusbasierten Methodik werden aus literaturwissenschaftlicher Sicht neuartige empirische Analysemöglichkeiten erprobt, indem beispielsweise systematisch nach Regiebemerkungen, Sprecherwechseln, Versmaßen oder Reimphänomenen gesucht werden und Verteilungen erhoben werden können.

Neben den jeweiligen teilfachspezifischen Fragestellungen eröffnet das Projekt die Möglichkeit zu übergreifenden Untersuchungen, indem die zuvor jeweils disziplinspezifisch vorgenommenen Annotationen in der Abfrage miteinander kombiniert werden. Hier ergeben sich folgende mögliche Fragestellungen: (i) Welche aus der synchron orientierten Forschung zur konzeptionell mündlichen Sprachverwendung bekannten Phänomene (vgl. Schwitalla 2006, Fiehler 2016, Hennig 2006) lassen sich in den Dramentexten überhaupt lokalisieren? (ii) Welche tauchen dagegen nicht auf? Welche Gründe lassen sich für das Auftreten bzw. Ausbleiben der Phänomene finden? (iii) Wie korrelieren v.a. im hoch normierten Bereich des Trauerspiels konzeptionelle Mündlichkeit und artifizielle Versifikation?

Datenerhebung und -aufbereitung

In Vorbereitung auf die Analyse waren drei Arbeitsschritte notwendig:

1. Datenerhebung
2. Datenaufbereitung
3. Datenexport zur weiteren Analyse

Datenerhebung

Im Rahmen der Datenerhebung wurden die Texte als Abschrift digitalisiert. Dazu wurden Vorgaben erarbeitet, wie mit Zeilen- und Versschreibweise verfahren und formale Besonderheiten verschriftlicht wurden. Nach der Abschrift wurden die Daten in ein Standarddateiformat gebracht werden, sodass die Daten annotiert werden können.

Datenaufbereitung

Gängige Tokenisierungsdienste, wie sie beispielsweise WebLicht bereitstellt, sind dabei primär auf wohlgeformte Satzstrukturen oder gesprochen-sprachliche Daten ausgelegt. Die dem Projekt zu Grunde liegenden Daten stellen allerdings eine Mischung von Vers- und Prosatexten mit Besonderheiten der Literatur des 16. Jh. dar, wodurch eine individuelle Tokenisierung notwendig wurde. Der hier entwickelte Tokenizer teilt die Sätze zeilenbasiert ein und sieht die Interpunktion somit nicht als maßgeblich für das Satzende an. Außerdem werden Satz- und Sonderzeichen separiert und der Datelexport ist in einem validen tcf-Format möglich.

Auf Basis der Rohdaten im TCF-Dateiformat wurden mittels des DTA::CAB Web Service weitere Annotationen hinzugefügt. CAB („Cascaded Analysis Broker“) wurde vom Deutschen Textarchiv speziell für die linguistische Analyse historischer Texte entwickelt, um historische Schreibvarianten auf äquivalente „kanonische“ moderne Wortformen abzubilden. Somit werden die Daten in einem Schritt um Lemmatisierung, PoS-Tagging (STTS) und eine Normalisierung (orthography correction) ergänzt.

Die so angereicherten Daten wurden zunächst mit webAnno im späteren Verlauf mit dem auf webAnno-basierenden Tool INCEpTION korrigiert und auf verschiedenen literatur- und sprachwissenschaftlichen Ebenen annotiert.

Datenexport

Bislang gibt es zwischen dem ANNIS- und TSV3-Dateiformat, welches dem CoNNL-Format ähnlich ist, noch keine Konvertierungsmöglichkeiten. Um die annotierten Daten zur weiteren Analyse in ANNIS übertragen zu können, wird auf die Pepper-Bibliothek zurückgegriffen, welche in Zusammenarbeit mit der HU Berlin um einen Import von TSV3 erweitert wurde.

Spielräume

Das Projekt sieht sich in mehrfacher Hinsicht mit Fragen nach Spielräumen konfrontiert.

1. Wie einleitend erwähnt, ist es gerade die Suche nach Abweichungen und Spielräumen, die das Projekt inhaltlich bestimmt. Hierzu wird nach der linguistischen wie literaturwissenschaftlichen

Strukturanalyse für prominente Stellen auch die jeweilige Redesituation betrachtet und so die formale und inhaltliche Untersuchung enggeführt. Ferner lassen sich aber auch über das Gesamtkorpus hinweg statistische Aussagen bspw. über die Häufigkeit von Abweichungen im Versmaß oder das Auftauchen bestimmter Formen von Gesprächspartikeln treffen. Erste Ergebnisse zum Versmaß werden im Rahmen des Vortrags vorgestellt.

2. Darüber hinaus betrifft die Frage nach Spielräumen und Interpretation auch das Design von Tagsets. Im Gegensatz zur erprobten linguistischen Methodik sind bislang kaum literaturwissenschaftliche Verfahren zur empirischen Erhebung und Interpretation interaktionaler Sprache eingeführt. Aus diesem Grund wurde in der Projektarbeit ein Set von Annotationskategorien sowie Kriterien der literaturwissenschaftlichen Auswertung solcher Phänomene neu entwickelt und kritisch reflektiert. Wesentlich sind dabei die etablierten Kategorien der Dramenanalyse, Rhetorik und Stilistik (Heudecker/Wesche 2009) sowie der Metrik. Zurückgegriffen werden kann in der methodischen Exploration auf Erkenntnisse und gängige Verfahren der Computerphilologie (Jannidis 2010a, b; Jannidis/Smith 2013), der Figuren- und Dialoganalyse im Drama (Pfister 2001: 196-264) sowie der Abweichungspoetik und Spielraumanalyse zur Barockzeit (Wesche 2004).

Dabei mussten die Phänomene so gewählt werden, dass sie einerseits interaktionale Momente der Damentexte erfassen und andererseits für die manuelle Annotation handhabbar sind. Während sich Interaktionalität aus literaturwissenschaftlicher Sicht mit verschiedensten Kategorien auf unterschiedlichen Ebenen eines Textes beschreiben lässt, werden die Möglichkeiten der Analyse durch Annotations- und Analysetools technisch limitiert. So lassen sich beispielsweise bestimmte Phänomene der Personenkonstellation oder des Drameninhalts nur eingeschränkt einzelnen Token zuordnen und dadurch in einem Tool wie INCEpTION schwer annotieren. Auch erlaubt ANNIS zwar komplexe Suchanfragen, kann aber bestimmte Aspekte der Textgestaltung wie Einrückungen oder ähnliches nicht darstellen. Der Vortrag gibt hier einen kurzen Überblick über die Tagsets der beiden Projektteile.

1. Die Wahl der Tools unterlag hierbei allerdings insofern Sachzwängen, als diese für die linguistische Auswertung notwendig sind. So mussten innerhalb der Möglichkeiten, die diese Tools bieten, Spielräume und Lösungen gefunden werden, die für beide Projektteile zufriedenstellend sind.
2. Schließlich mussten um eine handhabbare und möglichst konsistente Annotation zu gewährleisten Guidelines entwickelt werden, die den Annotator*innen eine eindeutige Zuordnung einzelner Tags erlauben. Hier entstand häufig ein Trilemma, das auch Evelyn Gius, Nils Reiter und Marcus Willand

in ihrem Shared-Task 2017 zur Erprobung von Guidelines (<https://sharedtasksinthehd.github.io/>) beschrieben haben: Die Kriterien „Konzeptionelle Angemessenheit“ (Wird aus der Guideline ersichtlich um welches Phänomen es geht? Orientiert sich die Guideline an gängigen Definitionen? Sind die in der Guideline beschriebenen Phänomene der Komplexität des literaturwissenschaftlichen Konzepts angemessen?), „Anwendbarkeit“ (Wie einfach kann die Guideline angewendet werden? Wie hoch ist das Inter-Annotator-Agreement?) und „Nützlichkeit“ (Wie können die nach der Guideline annotierten Daten weiterverwendet werden? Welche Fragen kann man an die Annotation stellen – und wie?) lassen sich nicht gemeinsam in gleich hohem Maß realisieren.

Am Beispiel der Annotation von Versmaßen und sowie deren Abweichungen zeigt der Vortrag, wie im Rahmen der Projektarbeit mit den letzten beiden Punkten verfahren wurde. Hierzu werden im Anschluss an eine knappe Projektvorstellung ausgewählte Phänomene, die annotiert werden, kurz erläutert. Danach wird der Workflow der Guideline-Entwicklung für die metrische Gestaltung der Dramentexte dargestellt. Da sich das Projekt zum Zeitpunkt des Vortrages in der letzten Phase befindet, kann hier neben Ergebnissen auch eine kritische Rückschau gezeigt werden, die mit Blick auf andere Projektentwicklungen fruchtbar sein kann.

Bibliographie

- Fiehler, Reinhard** (2016): Gesprochene Sprache. In: Wöllstein, A. und P. Eisenberg (Hrsg.): Duden - die Grammatik. Berlin: Dudenverlag 1181-1260.
- Gryphius, Andreas** (1991): Dramen. Hrsg. v. Eberhard Mannack. Frankfurt/M.: Bibliothek deutscher Klassiker.
- Hennig, Mathilde** (2006): Grammatik der gesprochenen Sprache in Theorie und Praxis. Kassel: Kassel Univ. Press.
- Heudecker, Sylvia / Wesche, Jörg** (2009): Rhetorik und Stilistik der deutschsprachigen Länder in der Zeit des Barock. In: Fix, U., A. Gardt und J. Knappe (Hrsg.): Rhetorik und Stilistik. Bd. 1. Berlin: de Gruyter 97-112.
- Jannidis, Fotis** (2010a): Digital Editions on the Net. Perspectives for Scholarly Editing in a Digital World. In: Schäfer, J. und P. Gendolla (Hrsg.): Beyond the Screen. Bielefeld: transcript 543-560.
- Jannidis, Fotis** (2010b): Methoden der computergestützten Textanalyse. In: Nünning, V. (Hrsg.): Methoden der literatur- und kulturwissenschaftlichen Textanalyse. Stuttgart: Metzler 109-132.
- Jannidis, Fotis / Smith, Kathleen** (2013): The Specification of User Requirements in the Design of Virtual Research Environments for the Arts and Humanities: Programming for the Future? In: Lossau, N. und H. Neuroth (Hrsg.): Evolution der Informationsinfrastruktur. Glückstadt: Verlag Werner Hülsbusch 71-84.
- Krause, Thomas / Zeldes, Amir** (2016): *ANNIS3: A new architecture for generic corpus query and visualization*. in: Digital Scholarship in the Humanities 2016 (31)). <http://dsh.oxfordjournals.org/content/31/1/118> [zuletzt abgerufen am 14.06.2018]
- Pfister, Manfred** (2001): Das Drama. Theorie und Analyse. München: Fink.
- Schwitalla, Johannes** (2012): Gesprochenes Deutsch. Berlin: E. Schmidt.
- Wesche, Jörg** (2004): Literarische Diversität: Abweichungen, Lizenzen und Spielräume in der deutschen Poesie und Poetik der Barockzeit. Tübingen: Niemeyer.