

“Wie Wölkchen im Morgenlicht”

Zur automatisierten Metaphern-Erkennung und der Datenbank literarischer Raummetaphern laRa

Schumacher, Mareike

schumacher@linglit.tu-darmstadt.de
Technische Universität Darmstadt, Germany

Bisher waren Bemühungen, theoretische Ansätze zur Metaphernforschung in praktischen informationstechnologischen Anwendungen, wie z.B. künstlichen Intelligenzen, zu implementieren wenig erfolgreich (vgl. Thaller 2021: 90). Auch der vorliegende Beitrag wird diese Lücke nicht schließen, dokumentiert aber den Versuch und zeigt auf, warum dieser scheitern musste. Mit dem Aufbau der relationalen Graphdatenbank literarischer Raummetaphern *laRa* wurde ein alternatives Recherche-Tool entworfen, das einerseits Einblicke in die Entwicklung, Etablierung und Konventionalisierung literarischer Raummetaphern gibt und andererseits literaturwissenschaftliche Fallstudien, bei denen Raummetaphern eine Rolle spielen, sinnvoll ergänzt. Der vorliegende Beitrag dokumentiert Training und Tests eines auf maschinellem Lernen basierenden Raum-Classifiers (Schumacher 2021a), der auch die Kategorie “Raummetapher” umfasst, und führt beispielhaft eine Metadaten-Netzwerkanalyse anhand der Raummetapher “Weg” durch. Auf diese Weise werden zwei digitale methodische Zugänge zum Phänomen literarischer Raummetaphern kontrastiert und gezeigt, wie Mixed-Methods-Ansätze und Cross-Validierung einer Methode durch eine andere gewinnbringend eingesetzt werden können.

Über (Raum-)Metaphern

Zwei grundlegende Probleme bei der Betrachtung von Metaphern im Vergleich zu nicht-metaphorischen Ausdrücken sind Uneigentlichkeit und Variabilität. Metaphern werden grundsätzlich aus drei Größen konstruiert: dem sprachlichen Ausdruck, dem, was der sprachliche Ausdruck im Wortsinne bezeichnet und etwas Ähnlichem (vgl. Wenz 1997: 32). Dabei unterscheidet sich die Grammatik metaphorischer Ausdrücke meist in nichts von wörtlich gemeinten Phrasen (vgl. Thaller 2021: 91). Metaphern sind dynamische Konstrukte, die eine Entwicklung durchlaufen von Einführung, über Etablierung zur Konventionalisierung und schließlich bis hin zum Übergang in den eigentlichen Sprachgebrauch (vgl. Thaller 2021), d.h. Ausdrücke können die einmal aufgebaute Metaphorik auch wieder verlieren. Metaphern sind aber nicht nur ein interessantes sprachliches Phänomen, sondern können auch als prägende Ausdrücke menschlichen Handelns fungieren (vgl. Blumenberg 1971: 213 und Wenz 1997: 33). Sie sind eine wesentliche Basis menschlichen Denkens. Der methodische Fokus der Digital-Humanities-Forschung zu Metaphern in deutschsprachigen Texten liegt derzeit auf manueller Annotation (vgl. Herrmann 2018), halbautomatischem Annotieren (vgl. Majoros 2013) und Automatisierungsvorbereitenden Ansätzen (z.B. Do, Gerloff und Nunez 2016). Mit MIPVU (Metaphor Identification Procedure VU University Amsterdam)

wurde eine sprachunabhängige Methodik zur manuellen Annotation von Metaphern entwickelt, die Annotator*innen dabei unterstützt, intersubjektiv Metaphern zu erkennen und zu annotieren (vgl. Steen et al 2010, Majoros 2013: 68, Herrmann 2018: 185). Eine vollautomatische Erkennung literarischer Metaphern, die sowohl konventionalisierte Phrasen als auch kreative Neuschöpfungen umfasst, wurde bisher nicht entwickelt und vereinzelt sogar als unmöglich bezeichnet (vgl. Gehring und Gurevych 2014: 103). Dennoch ist die Problematik ausgesprochen wichtig, da sich Metaphern bei der computationellen Erfassung anderer Phänomene als erheblicher Störfaktor erweisen können.

Raummetaphern als Störfaktor automatisierter Erkennung und Kategorisierung literarischen Raumes

Im Rahmen meiner Dissertation *Orte und Räume im Roman* (Schumacher im Erscheinen) habe ich eine Methode zur automatisierten Erkennung und Klassifizierung von Raumreferenzen in Erzähltexten entwickelt. Die Automatisierung fußt auf einem theoriebasierten Modell, das sieben Kategorien umfasst: Orte, Relationen, relationale Verben, Raumthemen, Raumbeschreibungen, Raumhinweise und Raummetaphern. Raummetaphern beinhalten zwar Raumausdrücke, bezeichnen aber nicht-räumliche Phänomene. So steht das hier gewählte Beispiel der Raummetapher “Weg” (wie in “Weg zum Glück” oder “Lebensweg”) für eine Reihe von Entscheidungen, nicht für eine geographische begeh- oder befahrbare Strecke. Im Gegensatz zu Raumsymbolen und -Motiven, bei denen räumliche Aspekte zwar zusätzlich zu ihrer Raumreferenz mit nicht-räumlicher Bedeutung aufgeladen werden, weisen Raummetaphern also keinerlei semantische Verbindung mit räumlichen (wie z.B. geografischen) Größen auf. Steht in einem Erzähltext z.B. eine Figur auf einem Berg und dieser Berg steht symbolisch für einen Erkenntnisgewinn, so ist die Geografie des Berges innerhalb der erzählten Welt nach wie vor gegeben. Wird von einer Figur in einem Roman gesagt, sie habe “einen Berg von Arbeit vor sich”, so handelt es sich bei “Berg” um eine Raummetapher, die eigentlich für eine große Menge steht. Da Raummetaphern in ihrer grammatischen Form und im sprachlichen Kontext der Referenzierung von Raum aber sehr ähneln, wurde zunächst der Versuch unternommen, sie mithilfe eines kontextsensitiven Machine-Learning-Trainings in die automatische Erkennung zu integrieren. Genutzt wurde der methodische Rahmen der Named Entity Recognition (vgl. Schumacher 2018). Entwickelt wurde ein Classifier (Schumacher 2021a), der Ausdrücke erkennt, die mehr oder weniger explizit Raum referenzieren und sie den oben genannten Kategorien zuweist. Über alle Kategorien hinweg wurde eine durchschnittliche Gesamterkennungsgenauigkeit von 75,65 % F1-Score¹ (vgl. Schumacher 2021a) erreicht. Mit 7,74 % hat die Kategorie der Raummetaphern mit Abstand die schlechteste Quote.

Trainingsprozess und Testergebnisse

Named Entity Recognition (NER) ist ursprünglich eine computeringuistische Methode zur automatischen Erkennung und Klassifizierung klar benannter Einheiten (vgl. Schumacher 2018: §1). Die am häufigsten in Named-Entity-Recognition-Tools imple-

mentierten Kategorien sind Personen, Orte und Organisationen. Für die literaturwissenschaftliche Nutzung von NER bedarf es allerdings einer Domänenadaption, bei der sowohl die implementierten Kategorien als auch die Trainingsdaten angepasst werden müssen. Die Methode wurde bereits erfolgreich für die Erkennung literarischer Figuren adaptiert (vgl. Jannidis et al. 2015) und auch eine Unterklassifizierung nach Genderzuweisungen ist möglich und für literaturwissenschaftliche Forschung gewinnbringend (vgl. Schumacher und Flüh 2020). Die Kategorie des Ortes ist für literarische Texte nahezu ebenso relevant wie die der Person bzw. Figur. Statt eines komplexen Raumkonzeptes, das in mehrere Unterkategorien aufgeteilt wird, werden bei der linguistischen Nutzung von NER-Tools lediglich Ortsnamen erkannt. Um ein NER-Tool so zu adaptieren, dass es Raumreferenzen erkennen und in eine von sieben Kategorien literarischen Raumes einsortieren kann, wurde ein Machine-Learning-Training durchgeführt. Im Folgenden werden diejenigen Ausschnitte des Trainings vorgestellt, die zeigen, inwiefern die automatisierte Erkennung von Raummetaphern dabei gescheitert ist.

Das NER Training – Testumgebung

Die Wahl des Tools fiel auf den in den Digital Humanities gut etablierten Stanford Named Entity Recognizer (Finkel, Grenager und Manning 2005), in dem kontextsensitive Conditional-Random-Fields-Algorithmen (zu CRF-Algorithmen vgl. Sutton und McCallum 2010) implementiert sind (Manning et al. 2014). Das Trainingskorpus besteht aus Ausschnitten aus 80 Romanen aus vier Jahrhunderten (18–21). Aus jedem Jahrhundert wurden 20 Erzähltexte integriert, sodass das Trainingskorpus einen gleichmäßigen Aufbau aufweist². Aus jedem Text wurden 4.000 Tokens extrahiert und ins Trainingskorpus überführt. Dabei handelt es sich um Anfangspassagen, da Romananfänge eine expositorische Funktion haben (vgl. z.B. Miller 1965: 9, Retsch 2000: 138, Richardson 2008: 4 und Herrmann 2018: 171) und eine Häufung von Raumreferenzen dadurch besonders wahrscheinlich ist³. Das Trainingskorpus und insgesamt acht Testtexte wurden in einem iterativen Prozess anhand detaillierter Guidelines (vgl. Schumacher forthcoming) manuell annotiert⁴. Aus den Testtexten wurden jeweils 10.000 Tokens betrachtet. Anschließend an Jannidis et al. (2015) wurde zunächst mit einem kleinen Trainingskorpus von 40.000 Tokens gearbeitet, das dann wiederum in einem iterativen Prozess auf 320.000 Tokens ausgeweitet wurde⁵. Insgesamt wurden 3 unterschiedliche Varianten des Trainings durchgeführt: mit jahrhundertspezifischen Daten (vgl. Abb. 1 links), mit Daten aus zwei Jahrhunderten (dem, aus dem der Testtext stammt und dem folgenden - vgl. Abb. 1 Mitte) und mit kumulierten Trainingsdaten aus allen vier Jahrhunderten (vgl. Abb. 1 rechts). Aus jedem Jahrhundert wurden insgesamt 80.000 Tokens ins Trainingskorpus überführt. Die Daten wurden diachron hinzugefügt.

Die Testergebnisse können wie in Abb. 1 visualisiert werden:

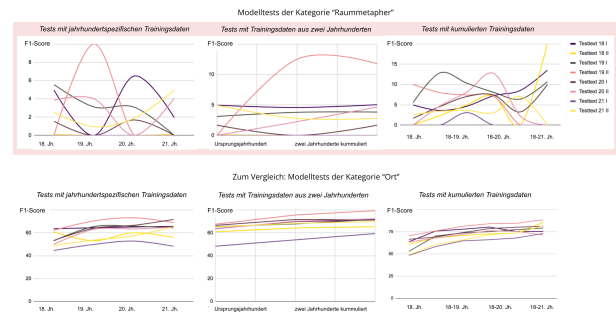


Abb. 1: Vergleich einiger Testergebnisse für das Training von Raummetaphern- und Ortskennung

Der Vergleich mit der insgesamt am besten erkannten Kategorie "Ort" zeigt, dass sich für die Testergebnisse der Kategorie "Metapher" keine Regelmäßigkeiten ergeben. Für einzelne Testtexte zeigen jahrhundertspezifische Trainingsdaten oder Trainingsdaten aus zwei Jahrhunderten den größten Trainingserfolg. Für andere ist ein Trainingskorpus aus einem ganz anderen zeitlichen Kontext passender. Eine schrittweise Ausweitung des alle vier Jahrhunderte umfassenden Trainingsmaterials zeigt keinerlei regelmäßigen Zuwachs der Erkennungsgenauigkeit. Einen viel typischeren Trainingsverlauf zeigt die Ortskategorie. Hier ist hauptsächlich die Größe des Trainingskorpus ausschlaggebend. Je mehr Trainingsdaten eingesetzt werden, desto höher die Erkennungsquote. Auch der zeitliche Kontext ist hier nur bei wenigen Testtexten bedeutend. Insgesamt zeigt die Kummulierung der Trainingsdaten eine zwar langsame aber kontinuierliche Steigerung. Das Training der Metaphererkennung gleicht hingegen einem Glücksspiel: Mal führt das Hinzufügen neuer Trainingsdaten zu einer Verbesserung des Classifiers, mal wird dadurch alles buchstäblich zurück auf Null gesetzt.

laRa - Datenbank literarischer Raummetaphern

Als alternatives Recherche-Tool und um besser zu verstehen, warum die Automatisierung von Metaphererkennung so problematisch ist, wurde die relationale Graphdatenbank literarischer Raummetaphern *laRa* (Schumacher 2021b) aufgebaut. Dazu wurden die im Trainingsprozess des maschinellen Lernens generierten Daten genutzt. Das heißt, die manuell annotierten Raummetaphern im Trainingskorpus wurden in eine Neo4J-Graphdatenbank (Graph Data Modeling Concepts and Techniques for Neo4J 2021) übertragen und dabei manuell mit Metadaten angereichert. Zu jeder Metapher wurde im Sinne der Kernmetapher von Lakoff und Johnson (1998: 9) ein Kernwort bestimmt, das dem am stärksten raumreferentiell besetzten Ausdruck entspricht. Außerdem wurde festgehalten, ob es sich um eine Ein-Wort-Metapher (Metapher, die aus einem einzigen Wort besteht), eine raummetaphorische Phrase (kurze metaphorische Phrase von ca. 3-5 Wörtern) oder ein Raumbild (komplexe metaphorische Konstruktion, die sich mindestens über einen halben Satz erstreckt) handelt. Zu jeder Metapher wurde ebenfalls manuell und in Interpretation des Kontextes eine Deutungsmöglichkeit hinzugefügt. Auf diese Weise wurde aus jeder Metapher ein Knotenpunkt mit drei Property's. Jeder dieser Knotenpunkte wurde über Relationen mit den Quellen verbunden, in denen die jeweilige Raummetapher vorkommt. Die Quellen bilden also ebenfalls Knotenpunkte, die mit den Proper-

tys “Titel”, “Autor*in” und “Erscheinungsjahr” versehen wurden (vgl. Abb. 2).

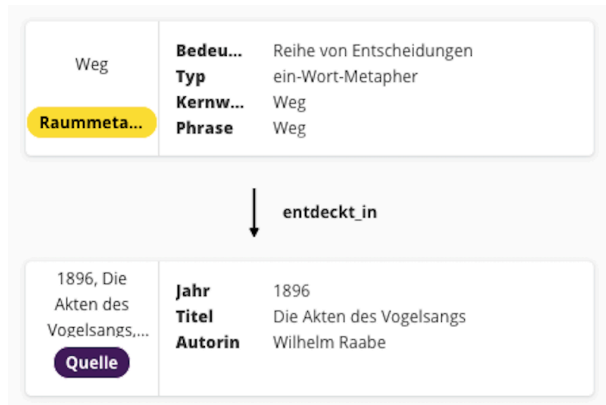


Abb. 2: Beispiel der Verbindung einer Raummetapher mit vier Propertyts mit einer Quelle mit drei Propertyts

Durch die Anreicherung mit Propertyts kann die Datenbank auf vielfältige Weise durchsucht werden. Wenn sowohl das Wortmaterial als auch die Bedeutung einer Metapher extrem ähnlich waren, wie z.B. der Fall bei "in schlechte Hände geraten" und "in schlechten Händen sein", wurden die Phrasen mit einem zweiten Typ von Relation untereinander verbunden (Relationstyp *Variation*). Insgesamt sind in *laRa* rund 800 Raummetaphern mit über 1.000 Relationen verzeichnet, die aus 80 Romanen stammen, die die Zeitspanne vom 18.–21. Jahrhundert regelmäßig abbilden, d.h. aus jedem Jahrhundert sind 20 Romane in den Aufbau der Datenbank einbezogen worden. Der Vorteil dieses systematischen Aufbaus ist, dass sich eine zwar kleine aber gleichmäßige Datenbasis ergibt, die den diachronen Vergleich der Nutzung von Raummetaphern unterstützt. Die in *laRa* verzeichneten Raummetaphern können als Netzwerk wie folgt visualisiert werden:

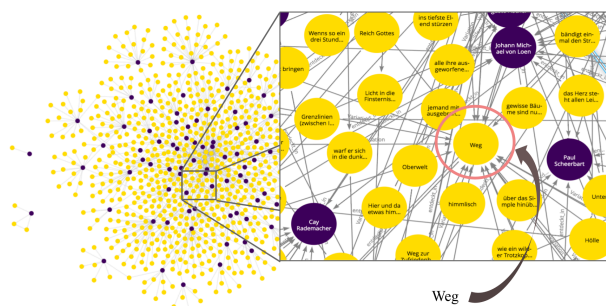


Abb. 3: Netzwerk der literarischen Raummetaphern mit Fokus auf "Weg" (helle Knoten stehen für Raummetaphern, dunkle für Quellen)

Das Netzwerk der Raummetaphern in Abb. 3 zeigt, dass es nur wenige zentrale Metaphern gibt, d.h. Metaphern, die sowohl viele Varianten aufweisen als auch in vielen Texten vorkommen. Viele Texte bilden mit ihren Raummetaphern eigene Cluster, die vielfach nur über Variationen mit anderen Metaphern, also nur indirekt mit anderen Texten verknüpft sind. Manche Text-Metaphern-Cluster sind gar nicht mit dem Hauptnetzwerk verbunden. Noch deutlicher wird das Gefüge, wenn eine Raummetapher einzeln betrachtet wird. Die Wahl fiel auf die Ein-Wort-Metapher "Weg", die keine der zentralsten Metaphern ist, sondern einen

mittleren Vernetzungsgrad aufweist. Abb. 4 zeigt diese Raummetapher mit ihren Varianten und den Erzähltexten, in denen sie vorkommen:



Abb. 4: Raummetapher "Weg" und Varianten in diachronen Vergleich

Die *laRa*-Abfrage der Metaphern mit dem Kernwort "Weg" zeigt sowohl den Variantenreichtum als auch das Vorkommen von Weg-Metaphern in literarischen Texten im diachronen Zeitverlauf. Der Variantenreichtum ist mit neun Varianten im Teilkorpus des 18. Jahrhunderts am höchsten. Im 19. Jahrhundert sind "Weg"-Raummetaphern in vergleichsweise vielen Texten zu finden. Abgesehen von der Ein-Wort-Metapher "Weg" kommt kaum eine Variante jahrhundert- oder auch nur textübergreifend vor. Nur die Metapher des "Lebensweges" kann über die Varianten "Lebenspfad" und "Pfad des Lebens" sowohl dem 18. als auch dem 19. und 20. Jahrhundert zugeordnet werden. Diese Charakteristik von Raummetaphern bleibt auch über Abfragen anderer zentraler Knotenpunkte stabil. Immer sind Häufigkeit und Variantenreichtum in den frühen Jahrhunderten am höchsten. Selten wirken Metaphern jahrhundertübergreifend. Meistens findet sich eine Variante einer Metapher nur in einem einzelnen Text. Die einzigen über alle vier Jahrhunderte wirkenden Raummetaphern sind generische Ein-Wort-Metaphern wie "Welt", "Weg" oder "Himmel".

Zusammenführung von Machine-Learning-Tests und Datenbank-Abfragen

Die Analysen, die mit Hilfe der Datenbank literarischer Raummetaphern *laRa* durchgeführt wurden, bieten eine grundlegende Erklärung für das Scheitern des Machine-Learning-Prozesses: Literarische Raummetaphern sind sehr spezifisch für begrenzte sprach- und literaturgeschichtliche Phasen. Oft führen Autor*innen eigene Varianten ein oder entwickeln komplexe metaphorische Raumbilder. Es reicht darum nicht aus, das Trainingsmaterial in kleinerem Umfang auszuweiten. Ist zufällig innerhalb eines neuen Abschnittes der Trainingsdaten dieselbe oder eine ähnliche Raummetapher vorhanden wie im Testtext, so kann die Erkennung zwar sprunghaft ansteigen, schon das Hinzufügen eines weiteren Datensatzes kann die Erkennung aber wieder absinken, wenn hier eine ähnliche Phrase in einem nicht-metaphorischen Kontext steht.

Mit der Datenbank *laRa* ist ein erster Schritt zu einer systematischen Erfassung literarischer Raummetaphern getan. Durch den gleichmäßigen Aufbau der Datenbasis gibt *laRa* schon jetzt gute Hinweise auf Konstruktion und Funktionsweise literarischer Raummetaphern. Sie kann dazu genutzt werden, andere methodische Zugänge einer Cross-Validierung (oder Cross-Falsifizierung)

rung) zu unterziehen. *laRa* kann außerdem die literaturwissenschaftliche Analyse einzelner Begriffe oder Raumsymbole wie z.B. dem des Weges gut ergänzen.

Fußnoten

1. Beim F1-Score handelt es sich um ein in der Computerlinguistik gängiges Maß, um die Erkennungsgenauigkeit automatischer Klassifikation zu erfassen. Mathematisch kombiniert der F1-Score die beiden Werte Precision (wie viele der klassifizierten Ausdrücke wurden richtig annotiert?) und Recall (wie viele der im Text befindlichen relevanten Ausdrücke wurden annotiert?).
2. Eine tabellarische Auflistung der zum Training genutzten Romane und der Quellen ihrer txt-Versionen findet sich in der Datei *Texte Trainingsdaten Raum-Classifizier.xlsx* im GitHub-Repository zum Raum-Classifizier (vgl. Schumacher 2021a)
3. Außerdem zeigt Herrmann (2018), dass in Erzählanfängen zuverlässig Metaphern vorkommen. In dem von ihr untersuchten Erzählankorpus werden in keinem Text weniger als 5,79% der Wörter metaphorisch gebraucht, im Durchschnitt sind es 14,1% (vgl. Herrmann 2018: 188).
4. Da in der Trainingsphase die Methode der manuellen Annotation genutzt wurde, konnten historische Varianzen der Schreibweise mit berücksichtigt werden, wenn diese in den generell meist normalisierten Texten aus dem TextGrid-Repository noch vorhanden waren.
5. Soweit rechtlich möglich wurden die annotierten Trainings- und Testdaten in einem GitHub-Repository zugänglich gemacht (vgl. Schumacher 2021a).

Bibliographie

- Do Dinh, Erik-Lân / Malte Gerloff / Alexandra Núñez** (2017): „Metaphern digital – Auf dem Weg von der Annotation zur automatischen Detektion“ in: *Modellierung - Vernetzung – Visualisierung: Die Digital Humanities als fächerübergreifendes Forschungsparadigma*. 3. Tagung des Verbands „Digital Humanities im deutschsprachigen Raum“ (DHD 2016), Leipzig, Mai 2. <https://doi.org/10.5281/zenodo.4645046>.
- Blumenberg, Hans** (1971): "Beobachtungen an Metaphern" in: *Archiv Für Begriffsgeschichte* 15: 161-214 <http://www.jstor.org/stable/24358391> [letzter Zugriff 8. Juli 2021].
- Dunn, Jonathan** (2013): „Evaluating the Premises and Results of Four Metaphor Identification Systems“ in: *Lecture Notes in Computer Science* 471–86.
- Finkel, Jenny Rose / Trond Grenager / Christopher Manning** (2005): "Incorporating Non-Local Information into Information Extraction Systems by Gibbs Sampling" in: *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics* 363–370.
- Gehring, Petra, und Iryna Gurevych** (2014): „Suchen als Methode? Zu einigen Problemen digitaler Metaphernerkennung“ in: *Journal Phänomenologie - Schwerpunktthema: Metaphern als strenge Wissenschaft* 41: 99-109.
- Herrmann, J Berenike** (2018): „Eine quantitative Metaphernanalyse deutschsprachiger Erzählanfänge zwischen 1880 und 1926“ in: Köppe, Tilmann / Singer, Rüdiger (eds.): *Show, don't tell: Konzepte und Strategien anschaulichen Erzählens* 167-212.
- Jannidis, Fotis / Krug, Markus / Reger, Isabella / Toepfer, Martin / Weimer, Lukas / Puppe, Frank** (2015): "Automatische Erkennung von Figuren in Deutschsprachigen Romanen" in: *Von Daten Zu Erkenntnissen* 1–6 <http://gams.uni-graz.at/o:dh-d2015.abstracts-vortraege> [letzter Zugriff 8. Juli 2021].
- Lakoff, George / Johnson, Mark** (1998): *Leben in Metaphern Konstruktion und Gebrauch von Sprachbildern*. Heidelberg: Carl-Auer-Verlag.
- Majoros, Krisztián** (2013): „Metapher und Kookkurrenz. Eine alternative ‚Trichter‘-Methode zur korpus- basierten Untersuchung metaphorischer Ausdrücke in öffentlich zugänglichen elektronischen Zeitungskorpora am Beispiel der Wissenschaftsberichterstattung im Bereich der Zellbiologie“ in *Sprachtheorie und germanistische Linguistik* 23: 65–110.
- Manning, Christopher D. / Surdeanu, Mihai / Bauer, John / Finkel, Jenny / Bethard, Steven J. / McClosky, David** (2014): "The Stanford CoreNLP Natural Language Processing Toolkit" in: *Association for Computational Linguistics (ACL) System Demonstrations* 55–60. <http://www.aclweb.org/anthology/P/P14/P14-5010> [letzter Zugriff 8. Juli 2021].
- Mason, Zachary J.** (2004): „CorMet: A Computational, Corpus-Based Conventional Metaphor Extraction System“ in: *Computational Linguistics* 30: 23–44. <https://doi.org/10.1162/089120104773633376>.
- Miller, Norbert** (1965): *Romananfänge*. Berlin: Verlag Literarisches Colloquium.
- Retsch, Annette** (2000): *Paratext und Textanfang*. Würzburg: Königshausen & Neumann.
- Richardson, Brian** (2008) *Narrative Beginnings: Theories and Practices*. Lincoln: University of Nebraska Press.
- Schumacher, Mareike** (2018): "Named Entity Recognition (NER) | ForTEXT" in: Gius, Evelyn / Meister, Jan Christoph / Schumacher, Mareike / Gerstorfer, Dominik / Meister, Malte / Blaess, Sandra / Flüh, Marie / Horstmann, Jan / Jacke, Janina (eds.): *ForTEXT - Literatur digital erforschen*. <https://fortext.net/routinen/methoden/named-entity-recognition-ner> [letzter Zugriff 8. Juli 2021]
- Schumacher, Mareike** (2021a): *Raum-Classifizier (kompatibel Mit StanfordNER)*. Zenodo: doi:10.5281/zenodo.4992662.
- Schumacher, Mareike** (2021b): *laRa - Die Datenbank für literarische Raummetaphern*. Zenodo. doi:10.5281/zenodo.4987844.
- Schumacher, Mareike / Flüh, Marie** (2020): "m*w Figurengender zwischen Stereotypisierung und Literarischen und theoretischen Spielräumen. Genderstereotype und -bewertungen in der Literatur des 19. Jahrhunderts". Zenodo. doi:10.5281/zenodo.4621892.
- Schumacher, Mareike** (im Erscheinen): *Orte und Räume im Roman. Ein Beitrag zur digitalen Literaturwissenschaft*.
- Steen, Gerard / Dorst, Lettie / Herrmann, Berenike / Kaal, Anna / Krennmayr, Tina / Pasma, Trijntje** (2010): *A Method for Linguistic Metaphor Identification: From MIP to MIPVU*. Amsterdam: John Benjamins Publishing. <https://doi.org/10.1075/celcr.14>.
- Sutton, Charles / McCallum, Andrew** (2010): *An Introduction to Conditional Random Fields*. <http://arxiv.org/abs/1011.4088> [letzter Zugriff 8. Juli 2021].
- Thaller, Manfred** (2021): "Über Metaphern und die Voraussetzungen für ihre Verwendung in der Informationstechnologie". In: Flüh, Marie, Horstmann, Jan, Jacke, Janina und Schumacher, Mareike (Hrsg.): *Toward Undogmatic Reading. Narratology, Digital Humanities and Beyond*. Hamburg: Hamburg University Press.
- Wenz, Karin** (1997): *Raum, Raumsprache und Sprachraum. Zur Textsemiotik der Raumbeschreibung*. Tübingen: Narr.