



Inhalt / Content

Close ReadingAls Close Reading wird in den Geisteswissenschaften die Technik des sehr genauen Lesens bezeichnet. Oft werden dabei Passagen im Hinblick auf mehrere Bedeutungszusammenhänge interpretiert.

Der Close-Reading-Selbsttest – die Rahmenbedingungen

Der Close-Reading-Selbsttest – die Vergleichskategorie

Quick Reading

Der Quick-Reading-Selbsttest

Eine Frage der Interpretation

Vorteile des schnellen Lesens nebenbei

Distant ReadingAls Distant Reading wird im Zusammenhang mit den digitalen Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück.

Der große Vorteil des Distant Reading – die Zeitersparnis

Der große Nachteil des Distant Readings – die fehlerhafte Erkennung

Scalable Reading – die Lösung?

Wie viel Distanz zum Text kann eine literarische Analyse verkraften?

Bibliographie

In seinem Buch "Distant Reading" (Moretti, 2013), das ich vor kurzem [hier](#) vorgestellt habe, behauptet Franco Moretti bewusst provokant, es würde nichts nützen, immer mehr zu lesen. Statt dessen müssten Literaturwissenschaftler*innen endlich die Kunst des nicht-Lesens erlernen. Tatsächlich haben wir inzwischen viele digitale Methoden zur Hand, mit denen man sich Distanz zum Text verschaffen kann, d.h. diese nicht selber lesen muss, sondern einen Computer zu Hilfe nehmen kann. Aber welche Ungenauigkeiten muss man eigentlich in Kauf nehmen, wenn man [Distant Reading](#)Als Distant Reading wird im Zusammenhang mit den digitalen Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück. betreiben möchte? Und wie lässt sich die Methode sinnvoll einsetzen? Dazu habe ich in einem Selbsttest die Methoden des Close und [Distant Reading](#)Als Distant Reading wird im Zusammenhang mit den digitalen



Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück. (hier mit Hilfe eines "out of the box"-Named Entity Recognition Tools) mit einer dritten Variante, dem lockeren Lesen nebenbei, nennen wir es Quick Reading, verglichen.

https://hnp9zs.podcaster.de/download/Podcast_Lesemethoden.mp3

Close ReadingAls Close Reading wird in den Geisteswissenschaften die Technik des sehr genauen Lesens bezeichnet. Oft werden dabei Passagen im Hinblick auf mehrere Bedeutungszusammenhänge interpretiert.

Close Reading ist eine der zentralen Methoden der Literaturwissenschaft. Der Begriff bezeichnet das Verfahren, Texte sehr genau zu lesen und Wort für Wort im Hinblick auf eine tiefer liegende Semantik zu interpretieren. Dabei kann jedes Wort in mehr als einen Bedeutungszusammenhang eingebettet sein. Sehr eng mit dem Close Reading ist die Technik der Annotation (Jacke, 2018) verbunden. Denn häufig werden von literaturwissenschaftlich Lesenden Bedeutungskategorien mit Hilfe farblicher Marker in den Text integriert oder Anmerkungen an den Rand geschrieben. **Close Reading**Als Close Reading wird in den Geisteswissenschaften die Technik des sehr genauen Lesens bezeichnet. Oft werden dabei Passagen im Hinblick auf mehrere Bedeutungszusammenhänge interpretiert. kann sowohl mit analogen Texten als auch digital durchgeführt werden. Für das digitale **Close Reading**Als Close Reading wird in den Geisteswissenschaften die Technik des sehr genauen Lesens bezeichnet. Oft werden dabei Passagen im Hinblick auf mehrere Bedeutungszusammenhänge interpretiert. bieten sich Annotationstools wie **CATMA** (Schumacher, 2019) oder **WebAnno** (Schumacher, 2018) an.

Der Close-Reading-Selbsttest – die Rahmenbedingungen

Für meinen Methodenvergleich habe ich die ersten 100 Seiten eines Romans (also effektiv die ersten 90 Seiten, da der Text wie üblich nicht auf Seite 1 beginnt) mit Hilfe des Tools



Annotationstools CATMA sorgfältig annotiert. CATMA ist eine literaturwissenschaftliche Web-Applikation, die man kostenfrei nutzen kann. Ganz ähnlich wie beim Unterstreichen eines Buches mit Markern, kann man damit Annotationskategorien anlegen und Textpassagen damit markieren. Ganz anders als beim Markern kann man allerdings am Ende quantitative Abfragen machen und die so erhaltenen Daten in Grafiken visualisieren. Ja, in CATMA kann man sogar gemeinsam annotieren, wovon ich [hier](#) auch schon einmal berichtet habe.

Der Close-Reading-Selbsttest – die Vergleichskategorie

Um eine gut vergleichbare Kategorie zu haben, die ich nicht nur beim Close Reading nutzen kann, sondern die auch von Tools zur automatischen Annotation von Texten abgedeckt wird, habe ich entschieden, Ortsnamen zu annotieren. Mehr eine linguistische als eine literaturwissenschaftliche Kategorie sind Namen von Städten, Ländern, Flüssen doch ein gutes Beispiel für einen konkreten Untersuchungsaspekt literarischer Texte, der auch tatsächlich Teil aktueller Forschung ist, wie z.B. in der Literaturgeographi (Piatti, 2008). Insgesamt habe ich übrigens 91 Ortsnennungen markiert, die hier als 100% angelegt werden sollen.



Quick Reading

Neben Close und [Distant Reading](#) Als Distant Reading wird im Zusammenhang mit den digitalen Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück. gibt es natürlich noch eine dritte Variante. Ein überfliegendes, ungenaues oder schnelles Lesen, bei dem der Text in seiner Gesamtheit, nicht aber in allen Details erfasst wird. Ich nenne es mal "Quick Reading", um es hier irgendwie begrifflich fassen zu können. Zum Teil wird dieses schnelle, überfliegende Lesen außerhalb der [Digital Humanities](#) Auch als digitale Geisteswissenschaften bezeichnet. Ein Forschungsfeld, in dem vielfältige digitale Methoden eingesetzt werden, um geisteswissenschaftliche Projekte zu bereichern. Das können z.B. Computerprogramme zur Textanalyse sein oder Software, mit der digitale Editionen zugänglich gemacht werden. Zum Feld der digitalen Geisteswissenschaften kann auch die Beschäftigung mit Phänomenen der Digitalisierung und die digitale Wissenschaftskommunikation gezählt werden. allerdings auch als [Distant](#)



ReadingAls Distant Reading wird im Zusammenhang mit den digitalen Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück. bezeichnet.

Der Quick-Reading-Selbsttest

Der Schritt, der hier als zweites aufgeführt wird, war eigentlich mein erster. In einem spontanen Einfall, begann ich, einfach mal drauf los zu lesen und alle Ortsnennungen in eine Tabelle einzutragen. Dies geschah auf dem Sofa, in der U-Bahn oder wo auch immer ich gerade ein paar Minuten Zeit fand. Ergebnis war, dass ich von den 91 später beim **Close Reading**Als Close Reading wird in den Geisteswissenschaften die Technik des sehr genauen Lesens bezeichnet. Oft werden dabei Passagen im Hinblick auf mehrere Bedeutungszusammenhänge interpretiert. gefundenen Ortsnennungen bereits 67 erfasst hatte, also etwa 74%. Diese zugegebener Maßen nicht besonders gute Quote kann natürlich einerseits durch mangelnde Sorgfalt begründet werden. Allein dadurch lässt sie sich aber nicht erklären.

Eine Frage der Interpretation

Ein zweites Problem dessen, was ich hier „Quick Reading“ nenne, ist, dass beim spontanen Draufloslesen und Annotieren Kategorien oft noch nicht genau genug gefasst sind. Selbst bei einer recht präzise scheinenden Kategorie wie der des Ortes gibt es immer wieder Unschärfen. Ein Beispiel wie „er wohnt in der Rothenbaumchaussee 71 in der Nähe der Hamburger Universität“ kann z.B. als eine Ortsnennung oder als zwei betrachtet werden. Interessiert dich z.B. die Verknüpfung von Orten mit kultureller Bedeutung, so machst du hier vielleicht zwei Notizen: „wohnt “+ „Rothenbaumchaussee“ und „Hamburg“ + „Universität“. Eine Interpretation als eine Ortsreferenz, in der wohnt/Rothenbaumchaussee/Hamburg/Universität insgesamt verknüpft wird, ist aber auch plausibel. Es kommt ganz auf die Perspektive an.

Vorteile des schnellen Lesens nebenbei

Ein großer Vorteil des „Quick Readings“ gegenüber dem **Distant Reading**Als Distant Reading wird im Zusammenhang mit den digitalen Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen



gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück. ist, dass hier in der Regel nichts falsch markiert (in den digitalen geisteswissenschaften spricht man von False Positives), sondern nur Vorkommnisse einer Kategorie übersehen werden (False Negatives). So werden zwar nicht alle Erwähnungen erfasst, aber immerhin wird nichts markiert, das nicht wirklich ein Ortsname ist. Auch kommt man mit dieser Technik schnell voran und kann statt kleiner Textmengen, mittelgroße Korpora bewältigen.

Distant Reading Als Distant Reading wird im Zusammenhang mit den digitalen Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück.

Distant Reading ist ein Begriff, der von Franco Moretti geprägt wurde, der in seinem gleichnamigen Buch feststellt, dass Literaturwissenschaftler*innen bisher wahrscheinlich nicht mehr als 2% aller literarischen Texte erforscht haben (Moretti, 2013). Um den Rest, die 98%, das, was Moretti "the great unread" nennt, betrachten zu können, müssten wir lernen nicht zu lesen (Moretti, 2013). Denn große Textmengen kann man nur dann in den Blick bekommen, wenn man eine nicht unerhebliche Distanz zu den Texten aufbaut.

Hier teste ich nur eine von zahlreichen Distant-Reading-Methoden und zwar die **Named Entity Recognition** Auch als NER bezeichnet, ist eine Methode bei der eindeutig benannte Größen (Entitäten) automatisch vom Computer erkannt und im Text markiert werden.

Typische Beispiele für NER-Kategorien sind Personen, Orte und Organisationen.. In meinem

Schumacher, Mareike auf <https://lebelieberliterarisch.de>

Lizenz: cc-by 4.0 (Creative Commons Namensnennung)



letzten [Artikel](#) habe ich den Stanford Named Entity Recognizer (Finkel, Genager and Manning, 2005) bereits kurz vorgestellt. StanfordNER erkennt mit einer Genauigkeit von etwa 60-70% Entitäten in Texten, also klar benennbare Größen wie Orte, Organisationen oder Personen. Gemäß meinem hier angelegten Verständnis von Ortsnennungen als Vorkommnisse konkreter Ortsnamen hat das Tool 53 Entitäten richtig erkannt, also etwa 58%. Das ist etwas weniger als die 60-70% Erkennungsgenauigkeit, die das Tool durchschnittlich erreicht.

Eine mögliche Erklärung ist die relativ enge Auslegung der Ortskategorie. Denn die Formulierung "im Rinnstein", die z.B. vom Tool erkannt wird, ist kein Ortsname, fällt also nicht in meine Kategorie, obwohl es hier auch um einen Ort geht, an dem etwas sich befinden kann. Eine andere mögliche Erklärung ist die Domäne für die das Tool optimiert wurde. Der StanfordNER wurde für Sachtexte entwickelt und erreicht darum als "out of the box"-Lösung für die Literaturanalyse vergleichsweise schlechte Ergebnisse (Jannidis *et al.*, 2015). Auf jeden Fall bringt mich das schnelle Lesen nebenbei insgesamt zu weit mehr nach meinem angelegten Verständnis richtig markierten Orten.

Der große Vorteil des Distant Reading – die Zeitersparnis

Nun muss man natürlich lobend und staunend erwähnen, dass StanfordNER für das Lesen des gesamten Textes (etwa 700 Seiten) schlappe 4 Minuten brauchte, während ich damit sicher zwei Wochen verbrachte und dabei ja noch nicht einmal 1/7 des Textes mit Annotationen versehen habe. Wenn ich täglich auch nicht viel – vielleicht 15 Minuten – las, so brauchte ich insgesamt wohl immer noch ca. 52 Mal länger als die Software. Bei großen Korpora sollte also der Zeitfaktor in eine Entscheidung zur Methodik definitiv mit einbezogen werden.

Der große Nachteil des Distant Readings – die fehlerhafte Erkennung

Der StanfordNER hat in meiner Stichprobe 33 Mal Entitäten falsch als Orte markiert (False Positives), was 34% aller als Orte annotierten Passagen entspricht. Diese Fehlerquote ist in meinen Augen recht hoch. Sie ist allerdings auch relativ leicht zu beheben, indem man die Liste der gefundenen Entitäten einfach nachbearbeitet. So wird man einige Zeit, die das Tool einem verschafft hat zwar wieder verbrauchen. Je nach Größe des Korpus nimmt diese Nachbearbeitung der Datenbasis evtl. einen vertretbaren Zeitaufwand in Anspruch. Eine zweite Möglichkeit, die Tools wie StanfordNER bieten, ist selbst an der Optimierung der



Software für die eigene Domäne zu arbeiten. Bei einem Machine-Learning-Tool wie StanfordNER kann man das über einen eigenen Trainingprozess umsetzen.

Scalable Reading – die Lösung?

Neben Close und Distant Reading gibt es noch eine weitere Methode, das Scalable Reading (Mueller, 2013). Scalable Reading ist eine Fusion aus Close und Distant Reading, der Begriff inspiriert von Google earth's Möglichkeiten des Rein- und Rauszoomens und des damit verbundenen Perspektivwechsels. Müller, der den Begriff geprägt hat, beschreibt, wie beim Rauszoomen vom Close zum Distant bzw. Nicht-Lesen mehr und mehr Kontext sichtbar wird, der in Analysen mit einbezogen wird (Mueller, 2020). Gerade durch das Zoomen innerhalb eines Analyseprojekts kann sehr viel Dynamik entstehen, die häufigen Perspektivwechsel können einen am Ende sehr nah an ein Korpus heran führen.

Verbindet man nun die Idee des Skalaren mit der Idee einer Methodenkombination (im Sinne eines Mixed Methods Ansatz), so wird ein weiteres Projektsetting denkbar. Um ein mittelgroßes bis großes Korpus bewältigen und trotzdem sehr tief ins Thema einsteigen zu können, können Teilkorpora gebildet werden. Ein kleines Kernkorpus kann mit [Close Reading](#) als Close Reading wird in den Geisteswissenschaften die Technik des sehr genauen Lesens bezeichnet. Oft werden dabei Passagen im Hinblick auf mehrere Bedeutungszusammenhänge interpretiert. bewältigt werden. Ein Mittelgroßes Erweiterungskorpus kann mit Hilfe von Quick Reading einen Hinweis darauf liefern, ob das betrachtete Phänomen auch auf weitere Texte zutrifft. Die so erstellten Daten können evtl. sogar zum Training des [Distant Reading](#) als Distant Reading wird im Zusammenhang mit den digitalen Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück. Tools verwendet werden, welches dann auf ein für das Projekt eher peripheres Rahmenkorpus angewendet werden kann. Auf diese Weise setzt du nicht nur einen klaren Schwerpunkt für dein Projekt, den du sehr genau betrachtest, sondern bekommst auch ein Gefühl dafür, ob deine Analyse übertragbar ist.

Wie viel Distanz zum Text kann eine



literarische Analyse verkraften?

Du siehst also, dass die Frage nach der Distanz, die du bei der Analyse zu deinen Texten einnehmen solltest oder vertretbarer Weise einnehmen kannst, am Ende eine Abwägungsfrage ist. Bis zu 10 Texte kann man in einem größeren Projekt wie z.B. einer Dissertation per **Close Reading** Als Close Reading wird in den Geisteswissenschaften die Technik des sehr genauen Lesens bezeichnet. Oft werden dabei Passagen im Hinblick auf mehrere Bedeutungszusammenhänge interpretiert. bewältigen. 10 - 50 oder, wenn man sehr schnell liest, vielleicht auch bis zu 100 Texte sind bei einem Projektzeitraum von etwa drei Jahren mit Quick Reading nebenbei zu schaffen. Spätestens ab einem Korpus-Umfang von 100 Texten müssen dann Distant-Reading-Methoden her. Vielleicht wäre **Scalable Reading** Eine Lesemethode, bei der Close Reading und Distant Reading dynamisch verbunden werden. Wie beim Herauszoomen aus einem Bildausschnitt wird durch zunehmende Distanz der Kontext mehr mit einbezogen, wie beim Hereinzoomen in ein Bild werden mit zunehmender Nähe mehr Details sichtbar. Der Begriff geht auf Martin Mueller zurück. für dich aber auch ein Königsweg, auf dem du munter in deine Texte hinein- und wieder hinauszoomen und so Textkenntnis und Kontextwissen verbinden kannst.

Diesen Artikel zitieren / how to cite: Mareike K Schumacher, "Distant Reading: Wie viel Distanz zum Text ist gesund?," in *Lebe lieber literarisch*, September 24, 2014, <https://lebelieberliterarisch.de/distant-reading/>, [zuletzt geprüft / Access: August 11, 2020].

Abonniere

Lebe lieber literarisch

Blog Blog ist kurz für Web-Log und steht für ein online Publikationsformat. Man kann sowohl der als auch das Blog sagen. Es gibt Blogs aller Sparten, von Linklisten über Tagebuchartige Formate bis hin zu wissenschaftlichen Blogs. Die Veröffentlichung kann schnell und unkompliziert erfolgen oder redaktionellen Standards entsprechen. In den Geisteswissenschaften etablieren sich Blogs zunehmend als Alternative zur langwierigeren wissenschaftlichen Publikation. **Lebe lieber literarisch** ist ein populärwissenschaftlicher Literaturblog. und **Podcast** Der Podcast ist ein auditives Veröffentlichungsformat. Podcasts entwickelten sich mit der Einführung von iPods und der Möglichkeit, die



Plattform iTunes auch zur Veröffentlichung eigener audio-Formate zu nutzen. Podcasts erfreuen sich seit Beginn des 21. Jahrhunderts zunehmender Beliebtheit und werden zum Teil auch zur Wissenschaftskommunikation eingesetzt. Im Gegensatz zum Vlog (Video-Blog) haben Podcasts keine visuelle Komponente, können aber trotzdem zum Teil auch über die Video-Plattform YouTube gehört werden. Die Sound-Dateien in den Lebe-lieber-literarisch-Blogartikeln stammen aus dem gleichnamigen Podcast, der auch unabhängig von diesem Blog auf iTunes, Spotify oder anderen Podcast-Plattformen gehört werden kann.

Wöchentlich News zu den Themen Literatur, Kultur, [Digital Humanities](#) Auch als digitale Geisteswissenschaften bezeichnet. Ein Forschungsfeld, in dem vielfältige digitale Methoden eingesetzt werden, um geisteswissenschaftliche Projekte zu bereichern. Das können z.B. Computerprogramme zur Textanalyse sein oder Software, mit der digitale Editionen zugänglich gemacht werden. Zum Feld der digitalen Geisteswissenschaften kann auch die Beschäftigung mit Phänomenen der Digitalisierung und die digitale Wissenschaftskommunikation gezählt werden. und Bloggen für Geisteswissenschaftler*innen

Invalid email address

Abonnieren

Take Action Now!

Probiere es einfach mal aus. Du kannst dich jederzeit vom Newsletter abmelden.

Bibliographie

Finkel, J. R., Genager, T. and Manning, C. (2005) 'Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. ', in. *43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005)*, Michigan.

Jacke, J. (2018) *Manuelle Annotation, forTEXT*. Available at:

<https://fortext.net/routinen/methoden/manuelle-annotation> (Accessed: 22 June 2020).

Jannidis, F. *et al.* (2015) 'Automatische Erkennung von Figuren in deutschsprachigen Romanen ', in. *DHD 2015*, Graz.

Moretti, F. (2013) *Distant Reading* Als Distant Reading wird im Zusammenhang mit den

Schumacher, Mareike auf <https://lebelieberliterarisch.de>

Lizenz: cc-by 4.0 (Creative Commons Namensnennung)



digitalen Geisteswissenschaften eine Technik des nicht-Lesens verstanden. Bücher werden von einem oder mehreren Computerprogrammen gelesen, die nach bestimmten Kriterien informative Daten aus den Texten bereitstellen. Der Begriff des Distant Readings geht auf Franco Moretti zurück.. London: Verso.

Mueller, M. (2013) *Morgenstern's Spectacles or the Importance of Not-Reading.*, [Scalable Reading](#) Eine Lesemethode, bei der Close Reading und Distant Reading dynamisch verbunden werden. Wie beim Herauszoomen aus einem Bildausschnitt wird durch zunehmende Distanz der Kontext mehr mit einbezogen, wie beim Hereinzoomen in ein Bild werden mit zunehmender Nähe mehr Details sichtbar. Der Begriff geht auf Martin Mueller zurück.. Available at:

<https://scalablereading.northwestern.edu/2013/01/21/morgensternsspectacles-or-the-importance-of-not-reading/>. (Accessed: 23 June 2020).

Mueller, M. (2020) [Scalable Reading](#) Eine Lesemethode, bei der Close Reading und Distant Reading dynamisch verbunden werden. Wie beim Herauszoomen aus einem Bildausschnitt wird durch zunehmende Distanz der Kontext mehr mit einbezogen, wie beim Hereinzoomen in ein Bild werden mit zunehmender Nähe mehr Details sichtbar. Der Begriff geht auf Martin Mueller zurück., [Scalable Reading](#) Eine Lesemethode, bei der Close Reading und Distant Reading dynamisch verbunden werden. Wie beim Herauszoomen aus einem Bildausschnitt wird durch zunehmende Distanz der Kontext mehr mit einbezogen, wie beim Hereinzoomen in ein Bild werden mit zunehmender Nähe mehr Details sichtbar. Der Begriff geht auf Martin Mueller zurück.. Available at:

<https://sites.northwestern.edu/scablereading/2020/04/26/scalable-reading/>.

Piatti, B. (2008) *Die Geographie der Literatur*. Göttingen: Wallstein.

Schumacher, M. (2018) *WebAnno, forTEXT*. Available at:

<https://fortext.net/tools/tools/webanno> (Accessed: 22 June 2020).

Schumacher, M. (2019) *CATMA, forTEXT*. Available at: <https://fortext.net/tools/tools/catma> (Accessed: 22 June 2020).





Distant Reading: Wie viel Distanz zum Text ist gesund?

