



Methodische Zugänge, epistemische Konsequenzen

2

2.1 Methodische Rahmenbedingungen

2.1.1 Themenmodelle und relationale Analysen im Kontext der bisherigen Forschung

Während im Einleitungskapitel der übersetzungstheoretische Rahmen der vorliegenden Analyse skizziert wurde, wird im zweiten Kapitel die digitale Methode vorgestellt. Hierfür werden zunächst Grundcharakteristiken der Methode erläutert sowie ausgehend davon ihre Anschlussfähigkeit an die literarische Textinterpretation i. A. und an den literarischen Übersetzungsvergleich im Besonderen diskutiert. Zusammengeführt werden diese Überlegungen schließlich im für den digital gestützten Übersetzungsvergleich entwickelten Analyseverfahren.

Ein zentrales Charakteristikum digitaler Analysetools, zu denen auch das im Rahmen dieser Analyse verwendete Themenmodell *Topic Explorer* gehört, ist das Priorisieren quantitativ belegbarer Textfakten. Da eine solche Prämisse schwerlich mit der die geisteswissenschaftliche Forschungspraxis charakterisierenden qualitativen Argumentationsführung zu vereinbaren ist, werden digitale Geisteswissenschaften häufig als methodologische Fingerübungen abgetan (Ramsay 2011: 2). Die Vielzahl der z. B. über das an der University of Alberta beheimatete Projekt *TAPoR (Text Analysis Portal for Research)* online frei zugänglichen Anwendungen (Rockwell et al.) scheint einen solchen Verdacht zunächst

Ergänzende Information Die elektronische Version dieses Kapitels enthält Zusatzmaterial, auf das über folgenden Link zugegriffen werden kann
https://doi.org/10.1007/978-3-662-69569-2_2.

zu erhärten, da die Mehrheit der Analysetools entweder nur eine eher oberflächliche Arbeit mit den Texten ermöglicht oder sich auf Überblicksfunktionen sowie auf konkret eingrenzbarere Anwendungen wie Autorattribution beschränkt. Auch die ebenfalls kostenlose Software *Voyant Tools* (Sinclair/Rockwell) bietet v. a. digitale Visualisierungswerkzeuge, die in erster Linie auf die Überblicksfunktion abzielen. Ähnliches gilt ferner für die unter dem Titel *Word Smith Tools* zusammengefasste Sammlung digitaler Analysewerkzeuge, die sowohl Prinzl (2016, 2017) im Rahmen ihrer Untersuchung englischsprachiger Thomas Mann-Retranslations als auch Bosseaux (2007) in ihrer Analyse französischsprachiger Virginia Woolf-Retranslations genutzt haben. *Word Smith Tools* sind quantitative Analyseinstrumente, mithilfe derer sich für bestimmte Begriffe und Begriffskombinationen Frequenzen, Verhältnisse und relative Häufungen ermitteln lassen (Prinzl 2017: 65). Im Vergleich hierzu gehen die im Rahmen der vorliegenden Analyse implementierten Themenmodelle insofern einen Schritt weiter, als entsprechende Frequenzen hier lediglich den Ausgangspunkt einer vollautomatisierten themenbezogenen Kategorisierung unterschiedlich skalierten Textebenen bilden. Während also mithilfe der Themenmodelle komplexe Aussagen zur thematischen Strukturierung von Texten und Textbestandteilen sowie zu abstrakteren Ähnlichkeitsverhältnissen zwischen diesen Themenstrukturen möglich sind, beschränken sich sonstige, frei zugängliche digitale Analysewerkzeuge überwiegend auf Frequenzen und wörtliche Übereinstimmungen. Dementsprechend waren die vorgestellten frei zugänglichen Anwendungen zumindest auf dem technischen Stand, den sie zum Beginn der Arbeit am vorliegenden Projekt erreicht hatten, einerseits im Vergleich zum Close Reading und andererseits im Vergleich zur seither rasant fortgeschrittenen Entwicklung im Bereich der sprachfokussierten Künstlichen Intelligenz bzw. des Natural Language Processing (NLP) definitiv unterkomplex.

So verlockend der Gedanke eines digitalen Analysewerkzeugs, das auf Knopfdruck einfache Antworten auf komplexe literaturwissenschaftliche Fragestellungen liefert, angesichts dieser technologischen Entwicklungen auch sein mag – er erweist sich nicht nur angesichts der technischen Gegebenheiten (noch) als unrealistisch, sondern steht außerdem im Widerspruch zur literaturwissenschaftlichen Argumentationspraxis als solcher. Der Hauptgrund hierfür besteht im Kriterium der Falsifizierbarkeit, das objektivierbare Forschungsthesen in der Regel auszeichnet. Während diese für gewöhnlich mit einem klaren Ja oder Nein beantwortet werden können, trifft dies auf komplexe literaturwissenschaftliche Interpretationsansätze üblicherweise nicht zu (Ramsay 2011: xi, 17). Gleiches gilt zudem für die Reliabilität solcher Ansätze (Jockers 2013: 6), die aus

rezeptionsästhetischer Perspektive bereits dadurch beeinträchtigt ist, dass unterschiedliche Rezipient*innen zu unterschiedlichen, aber gleichermaßen legitimen Interpretationsschlüssen gelangen können (Fish 1989: 141).

Dementsprechend lässt sich auch die hier untersuchte Forschungsfrage, inwiefern sich historische Kontexte auf thematische Charakteristiken der japanischen Thomas Mann-Übersetzungen auswirken, nicht sinnvollerweise bejahen oder verneinen. Stattdessen sind Auswertung und Kontextualisierung der digitalen Analyseergebnisse insofern hermeneutisch bzw. traditionell geisteswissenschaftlich, als ihre Glaubwürdigkeit nicht abschließend belegt oder widerlegt, sondern graduell erzeugt wird. Als Grundlage dieser Glaubwürdigkeit fungiert einerseits die Summe und Qualität der vorgebrachten Argumente (also in erster Linie Interpretationsbeispiele sowie Informationen aus der Sekundärliteratur) und andererseits ihre rhetorische und argumentative Aufbereitung. Dieser Modus einer graduell hergestellten Glaubwürdigkeit ist im Domänenkonflikt der Natur- und Geisteswissenschaften sukzessive in Verruf geraten, sodass man sich von den vermeintlich empirisch-objektivierenden Digital Humanities eine akademische Re-Nobilitierung der Literaturforschung erhofft hat. Vor dem Hintergrund derartiger interdisziplinärer Revierkonflikte bietet die digitale Methode einen sinnvollen Anlass, Forschungs-, Argumentations- und Interpretationspraktiken als solche hinsichtlich ihrer epistemischen Auswirkungen zu reflektieren. Ein Beispiel hierfür ist die (nicht nur) im Zusammenhang digitaler Analysezugänge relevante, aber in bisherigen literaturwissenschaftlichen Untersuchungen kaum explizit thematisierte Frage der Textsegmentierung (Bartsch et al. 2023), auf die in Abschnitt 2.3.1 ausführlicher eingegangen wird.

Vor dem inhaltlichen Einstieg ins Thema sollen jedoch, anknüpfend an den eingangs skizzierten Überblick über mögliche kostenfreie Alternativen zur Arbeit mit Themenmodellen, im Sinne eines Forschungsstandes zur digitalen Methode noch einige Projekte vorgestellt werden, von denen die vorliegende Arbeit entscheidend profitieren konnte. Wertvolle Anregungen haben in diesem Zusammenhang v. a. Lisa Rhodys Topic Modelings englischsprachiger Gedichtkorpora (2012) geliefert, durch die sie bestehende Genredefinitionen aktualisiert und die Übertragung des ursprünglich für die Sachtextanalyse entwickelten Topic Modeling auf literarische Texte problematisiert hat. Interessante Denkanstöße boten außerdem Drouins durch Netzwerkgraphen visualisierte Topic Modelings zum Kirchenmotiv bei Marcel Proust (2010; 2017) und seine Überlegungen zur digitalen Synchronisierung des ansonsten diachronen Erzählverlaufs.

Von besonderem Interesse waren außerdem digital gestützte Forschungsprojekte, die sich schwerpunktmäßig mit Beziehungen zwischen Texten befassen. Auch in Hinblick auf eine generelle Methodenreflexion zum Topic Modeling

relevant ist hierbei Thomas Weitins und Katharina Hergets Projekt zum zwischen 1871 und 1876 herausgegebenen *Deutschen Novellenschatz*, das die Beziehungen zwischen den in dieser Textsammlung enthaltenen 86 Novellen durch Topic Modeling indirekt aufarbeitet, indem spezifische, nur für bestimmte Einzeltexte oder Autor*innen charakteristische „Falkentopics“ ermittelt werden, die aus der sonstigen Themenverteilung des Korpus „herausfallen“ (Weitin/Herget 2017: 43–44; Weitin 2021: 125). Im selben Projektkontext hat ferner auch Fotis Jannidis (2017) stilometrische Distanzmaße zwischen den Texten des *Deutschen Novellenschatzes* untersucht, sich hierbei aber nicht auf Topic Modeling, sondern auf eine abgewandelte Variante der „Delta Procedure“ bzw. des Delta-Score als einer erstmalig 2002 von John Burrows vorgestellten Methode zur Berechnung stilometrischer Textähnlichkeit gestützt. Diese bezieht sich im Unterschied zur hier vorgestellten, die jeweiligen Gesamttexte berücksichtigenden Methode nur auf die am häufigsten innerhalb der Texte vorkommenden Wörter (Burrows 2002: 267, 269), bei denen es sich mehrheitlich um Funktionswörter handelt. Darüber hinaus hat sich bspw. auch Christof Schöch (2014: 147) in einem Projekt zur digital gestützten Autorattribution mit stilometrischen Ähnlichkeitsbeziehungen auseinandergesetzt und hierbei ebenso wie die vorliegende Arbeit zwischen Inhalts- und Funktionsvokabular differenziert. Mit demgegenüber vorwiegend technischer Schwerpunktsetzung haben außerdem Chang und Blei bereits 2009 relationale Topic Models vorgestellt, mithilfe derer sich zumindest in Hinblick auf nichtliterarische Korpora Prognosen z. B. für die Wahrscheinlichkeit berechnen lassen, dass innerhalb neuer Forschungsarbeiten aus einem existierenden Korpus zitiert wird (Chang/Blei 2009: 81–82). 2014 haben Maiya und Rolfe dann anhand sogenannter Topic Similarity Networks im Rückgriff auf den sogenannten Hellinger-Abstand, also einer Formel zur Berechnung der Ähnlichkeit zwischen zwei Wahrscheinlichkeitsverteilungen, die Ähnlichkeitsbeziehungen zwischen unterschiedlichen Topics rechnerisch erfasst, sich hierbei aber mit der englischsprachigen Wikipedia und der Vergabe von Forschungsfördermitteln ebenfalls auf einen dezidiert nichtliterarischen Gegenstandsbereich bezogen.

Während die genannten Arbeiten v. a. einsprachige Korpora, also nicht Übersetzungstexte thematisieren, wurden polylinguale Themenmodelle bereits 2009 von Mimno et al. vorgestellt, eine digitale Übersetzungsforschung i. e. S. ist aber erst in der jüngsten Vergangenheit erprobt worden. Anhand von Netzwerkanalysen untersucht wurden hierbei Beziehungen zwischen Übersetzenden und Publikationsorganen, d. h. die jeweilige äußere Übersetzungsgeschichte in Japan (Long 2015) bzw. Rumänien (Tanasescu 2020); ferner reflektiert Long die generellen Erkenntnispotenziale, die sich aus der quantitativen Analyse

literarischer Texte ergeben, umfassend innerhalb seiner Monografie (2021). Darüber hinaus hat Svensson thematische Ähnlichkeitsbeziehungen bzw. wörtliche Überschneidungen zwischen 51 unterschiedlichen englischsprachigen Übersetzungsvarianten des Koran mithilfe von hierfür erstellten Lexika untersucht und auf diese Weise anhand quantitativer Kriterien von der gängigen Forschungsmeinung abweichend besonders einflussreiche Übersetzungstexte identifiziert (Svensson 2019:215–216; 223–224). Ebenfalls auf wörtliche Übereinstimmungen hat sich außerdem Theimer in ihrer Arbeit zu elf unterschiedlichen Übersetzungsvarianten des Tolstoi-Romans *Anna Karenina* (Theimer 2021) konzentriert und sich hierbei auf den Jaccard-Koeffizienten zwischen Wortmengen gestützt. Die beiden zuletzt erwähnten Projekte weisen hinsichtlich ihrer Konzentration auf Ähnlichkeitsbeziehungen zwischen Übersetzungstexten zwar einige Gemeinsamkeiten zum hier vorgestellten Dissertationsprojekt auf, beschränken sich hierbei aber auf die Quantifizierung wörtlicher Überschneidungen, die nicht nur eine sehr eingeschränkte Perspektive auf Übersetzungsähnlichkeit darstellt, sondern auch keinen qualitativ interpretierenden Rückbezug auf die einzelnen Tokens, also spezifische Wörter, Wortgruppen oder Satzteile im ursprünglichen narrativen Zusammenhang ermöglicht. Diese Option bietet das Analysetool *TL-Explorer* (Zhai et al. 2020) zwar durchaus; im Zentrum steht dabei allerdings, ähnlich wie bei den eingangs vorgestellten, online frei zugänglichen Analysetools, der Überblick über alle verfügbaren Übersetzungen eines Werks. Auch die vorgestellten Beispiele einer digitalen Übersetzungsforschung konzentrieren sich folglich entweder auf die äußere oder aber auf bestimmte quantitative Aspekte der inneren Übersetzungsgeschichte, sodass die vorliegende Analyse hier ebenfalls einen wichtigen Beitrag leistet.

2.1.2 Grundbegriffe: Close, Distant, Blended, Algorithmic, Scalable

Wie konstruktiv die Herausforderung traditioneller (geisteswissenschaftlicher) Argumentationsstrukturen durch digitale Analysemethoden sein kann, hat bereits Stephen Ramsay (2011) in seinem Plädoyer für eine Synthese von Close Reading und Distant Reading im Algorithmic Criticism gezeigt: Das als methodisches Kernkonzept der Digital Humanities terminologisch durch Moretti (2013) geprägte Distant Reading wertet große Textmassen unter quantitativen Gesichtspunkten computerbasiert aus. Dagegen berücksichtigt das die traditionelle geisteswissenschaftliche Hermeneutik konstituierende Close Reading nur eine relativ geringe Textmenge, geht aber qualitativ vor. In diesem Zusammenhang beziehen sich die Attribute *close* bzw. *distant* nicht etwa (wie vom Methodendiskurs

teils suggeriert) auf prinzipielle Genauigkeit bzw. Ungenauigkeit, sondern auf die jeweilige Skalierung der Betrachtungsebene. Diese fällt im Distant Reading normalerweise sehr viel größer, in einigen Sonderfällen jedoch auch sehr viel kleiner aus als im Close Reading (Moretti 2013: 77). Die aus beiden Ansätzen resultierenden Interpretationszugänge heben dementsprechend gerade dadurch bestimmte narrative und sprachstilistische Aspekte des untersuchten Textmaterials hervor, dass sie andere vernachlässigen (ebd.: 48–49). Dies betrifft Close Reading und Distant Reading gleichermaßen: Was das eine ignoriert, gerät in den Fokus des jeweils anderen. Auf eine entsprechende Kombination unterschiedlicher Betrachtungs- und Skalierungsebenen, die im weiteren Verlauf dieses Abschnitts noch problematisiert werden soll, bezieht sich ferner auch Matthew Jockers (2013: 24) mit seinem aus den Wirtschaftswissenschaften entlehnten Konzept der Makro- und Mikroanalyse (macro/micro analysis), das qualitative Einzelinterpretation (micro) mit quantitativer Korpusanalyse (macro) verknüpft (Krautter/Willand 2020: 84).

Das vor diesem Hintergrund an Jockers Makroanalyse anzunähernde Distant Reading erlaubt also bspw. im Kontext des am MIT Digital Humanities Lab realisierten *Gender Novels Project* Rückschlüsse darauf, ob innerhalb der gesamten (kanonisierten und digitalisiert vorliegenden) literarischen Produktion einer bestimmten Epoche bestimmte Adjektive vorwiegend mit männlichen oder weiblichen Charakteren in Verbindung gebracht werden: Während die in den entsprechenden literarischen Texten auftauchenden männlichen Figuren v. a. „old“ oder „young“, „first“ oder „last“ sowie „good“ oder sogar „great“ sind, zeichnet sich die fiktive Damenwelt überwiegend durch Äußerlichkeiten wie „beautiful“, „pretty“ oder „lovely“ sowie durch zwischenmenschliche Vorzüge wie „sweet“, „Lady“ oder „dear“ aus. Im Unterschied zu diesen eher allgemeinen Aussagen bezieht sich ein weitgehend mit Jockers Mikroanalyse gleichzusetzendes Close Reading üblicherweise auf einen begrenzten Textbereich, erhebt dafür aber den Anspruch, dessen semantisches Potenzial inklusive sämtlicher Ambiguitäten so erschöpfend wie möglich zu auszuloten. Voraussetzung für dieses Ausloten ist ein Dualismus von Text und Interpretierenden, wobei die Komponente der Interpretierenden durch Konzepte wie Fishs Interpretationsgemeinschaften kontextualisierend erweitert werden kann (Fish 1980: 335). Dieser Dualismus ist im Distant Reading insofern aufgebrochen, als zwischen Text und Interpret*in die keineswegs neutrale dritte Instanz Big Data tritt, die ihrerseits die beiden Dimensionen der Digitalisierung (d. h. Erweiterung von Textmenge und Textaggregatzustand) und der Algorithmen (d. h. digitale Analyseverfahren) umfasst (Stulpe/Lemke 2016: 17, 41). Ebenso, wie zusätzlich zur Textanalyse als solcher

die Instanz der Interpretierenden insbesondere im rezeptionsästhetischen Rahmen auf ihre Kontexte hin befragt wird, muss auch die Instanz Big Data in ihren technologischen Bedingtheiten und epistemischen Konsequenzen umfassend reflektiert werden.

Bezieht man sich vor dem Hintergrund von Distant Reading und Close Reading außerdem auf die die Übersetzungsanalyse nach Kinkel konstituierenden Kategorien Inhalt, Form, Stil und Außersprachliches (Kinkel 2001: 123), lassen sich in erster Linie der sehr allgemein zu verstehende Inhalt sowie quantitativ realisierte Aspekte der Formalstilistik im Distant Reading, qualitativ realisierte Aspekte der Formalstilistik, inhaltliche Nuancen sowie Außersprachliches dagegen im historisch kontextualisierten Close Reading ermitteln. Auch hier ist jedoch nicht allein die Abdeckung aller Analysekatogorien von Bedeutung, sondern vielmehr die Frage, welche quantitativ-qualitativen Zusammenhänge zwischen ihnen z. B. im Algorithmic Criticism nachvollzogen werden können. Ein als Dialog von innerer und äußerer Übersetzungsgeschichte realisierter Algorithmic Criticism erfasst folglich auch das komplexe Zusammenspiel von rechnerisch objektivierbaren Potenzialitäten und dem, was menschliches Interpretationsvermögen auf dieser Basis hervorbringt, von sprachlichem Ausgangsmaterial und einem sich hieraus konstituierenden Werksinn, der nicht nur in Bezug auf die literarische Übersetzung mehr ist als die Summe seiner Einzelteile (Ramsay 2011: 67).

Im Zusammenhang des Algorithmic Criticism ergibt sich damit eine umgekehrte Argumentationslogik, die nicht etwa so wie traditionelle literaturwissenschaftlich-hermeneutische Interpretationsthesenbildung vom Teil oder Detail zum Ganzen führt, sondern i. d. R. eine auf das Textganze bezogene, quantitativ fundierte Deutungsansage an den Anfang der Betrachtung stellt. Damit generiert der Algorithmic Criticism gleichermaßen vorab Analyseergebnisse einer höheren, abstrakteren logischen Ordnung. Anwender*innen digital gestützter Interpretationsverfahren wie dem Algorithmic Criticism wird dementsprechend direkt zu Beginn des in Abschnitt 2.4.2 ausführlicher beschriebenen Analysevorganges eine digital generierte Schlussfolgerung vorgesetzt, die anschließend unter qualitativen Gesichtspunkten einzuordnen und zu hinterfragen ist. Dies wiederum deckt sich mit der generell in Hinblick auf die literaturwissenschaftliche Implementierung digitaler Methoden bezogenen Aussage Weitins, „dass der anfangs ersparte Aufwand am Ende doch aufgebracht werden muss, wenn man mit den Ergebnissen sinnvoll arbeiten will“ (Weitin 2021: 2): An die Stelle der kleinteiligen Erschließung einer möglichen Schlussfolgerung tritt also die ebenso kleinteilige Rekonstruktion in Hinblick auf das Zustandekommen einer ebensolchen Schlussfolgerung.

Die Frage, inwiefern sich die aus der Kombination von Distant Reading und Close Reading hervorgegangenen Ergebnisse historisch kontextualisieren lassen, bleibt in Ramsays Konzept des Algorithmic Criticism offen, sodass hier das sozialwissenschaftliche Konzept eines Blended Reading eine wertvolle Ergänzung bietet. Mit dem literaturwissenschaftlichen Algorithmic Criticism teilt es ebenso wie mit der Kombination aus Makro- und Mikroanalyse den Anspruch, Close Reading und Distant Reading „nicht als oppositionelle oder gar exklusive, sondern als synergetische Vorgehensweisen“ zu betrachten (Stulpe/Lemke 2016: 43). Konkret fordert das Blended Reading eine „konsekutive Verknüpfung mehrerer Analyseverfahren aus dem Repertoire des Text Mining (*distant reading*) mit der Notwendigkeit, immer wieder Einzeltexte aus relevanten Zeitabschnitten des Untersuchungskorpus gegenzulesen (*close reading*)“ (ebd.: 43, Hervorh. Original). Ein wesentlicher Unterschied zum Algorithmic Criticism besteht darin, dass im Blended Reading i. d. R. ganze Texte, im Algorithmic Criticism nur Textstellen oder kurze Textformen einem Close Reading unterzogen werden. Auch variieren die Kriterien, anhand derer Texte bzw. Textabschnitte für ein Close Reading ausgewählt werden, je nach Erkenntnisinteresse. Die in Bezug auf das Blended Reading vorgeschlagene Auswahl anhand relevanter Zeitabschnitte ist dabei prinzipiell auch im Rahmen eines Algorithmic Criticism denkbar; ebenso nahe liegt es im Kontext beider Ansätze, unter quantitativen Aspekten auffällige Texte bzw. Textabschnitte einer genaueren Betrachtung zu unterziehen. Auch ein relationaler Vergleich von Textkorpora und Textstellen kann in beide Konzepte integriert werden.

Gemeinsam haben Blended Reading und Algorithmic Criticism zudem, dass sie modulare Analyseprozesse sind, die eine „Rückbindung der durch die computergestützten Verfahren generierten Ergebnisse an die Texte selbst“ via Close Reading vorsehen, wobei sich die beiden Analysemodule Distant Reading und Close Reading wechselseitig bedingen (ebd.: 55): Distant Reading dient nicht nur der Vorauswahl der zu interpretierenden Textbereiche, sondern bringt als solches bereits inhaltliche Ergebnisse hervor (s. Abschnitt 5.1), die im Close Reading überprüft werden und im Idealfall weitere Analyseiterationen zur Folge haben.

Anschlussfähig in Hinblick auf Rezeptions- und Übersetzungsforschung ist Blended Reading außerdem insoweit, als der „Verweisungszusammenhang zwischen Text [...] und Kontext [...] prinzipiell eine Umkehrung der Blickrichtung erlaubt“ (ebd.: 22): Kontexte hinterlassen Spuren im Text, die sich rekonstruieren lassen. Dem entspricht die Auffassung einer deskriptiven Übersetzungsforschung, dass Übersetzungsprozesse – und damit auch ihre Kontextgeprägtheit – retrospektiv am Übersetzungstext zu erschließen sind (Toury 2014: 18). Geht demzufolge auch ein Übersetzungstext einerseits aus anderen Texten bzw. Texteinflüssen,

andererseits aus „einem zeitgenössischen Ereignishorizont und kulturellen Hintergrundgewissheiten, die auf eine konkrete historisch-soziale Realität verweisen“ (Stulpe/Lemke 2016: 21) hervor, lassen sich diese Faktoren am Text nachvollziehen. Dabei bestehen diese Texteinflüsse im Falle der japanischen Thomas Mann-Rezeption bspw. in weiteren Fach- und Übersetzungspublikationen der Übersetzenden, der „zeitgenössische Ereignishorizont“ z. B. in den Taishō-zeitlichen Bildungsreformen oder in der japanischen Kriegsniederlage und die „kulturellen Hintergrundgewissheiten“, die literatursoziologisch auch als Kontexte zweiter und dritter Ordnung bezeichnet werden, z. B. im *kyōyōshugi*-Bildungsideal.

Schwerlich mit einer Thomas Mann-Übersetzungsanalyse in Einklang zu bringen ist Blended Reading allerdings durch seine Nähe zur Diskurs- und Semantikforschung, die der Konzentration auf einen Einzelautor zuwiderläuft. Trotz der unbestreitbaren Pluralität der japanischen Thomas Mann-Übersetzung sind aus dieser Schwerpunktsetzung heraus entwickelte Vorgehensweisen des Blended Reading wie z. B. die Erstellung von Subkorpora, die einen bestimmten Diskurs repräsentieren, hier kaum operationalisierbar (ebd.: 52). Dennoch ist die Kontextorientiertheit des sozialwissenschaftlichen Blended Reading ein wichtiger gedanklicher Anknüpfungspunkt, von dem auch die literaturwissenschaftliche Rezeptions- und Übersetzungsforschung maßgeblich profitiert.

Die genannten Konzepte bleiben im Allgemeinen jedoch relativ vage in Hinblick auf die Frage, wie sich das allseits geforderte Zusammenspiel zwischen unterschiedlichen Analysemodi und Betrachtungsebenen sowie insbesondere zwischen dem qualitativ interpretierten Mikrotext und dem historisch kontextualisierten Makrotext konkret gestalten könnte. Hier erweist sich das ursprünglich 2012 vom Anglisten und Altphilologen Martin Mueller entwickelte Konzept des Scalable Reading als wertvolle Ergänzung. Es setzt sich ebenso wie das Blended Reading mit der Synthese zwischen Text- und Kontextanalyse bzw. mit „new and powerful ways of shuttling between ‘text’ and ‘context’“ auseinander (Mueller 2012), ist dabei aber im Unterschied zum sozialwissenschaftlichen Blended Reading u. a. mit dem sich aus der Fiktionalität literarischer Texte ergebenden „Problem der Übertragbarkeit von textuellem Wissen auf die Realität“ konfrontiert (Krautter/Willand 2020: 78): Während die im Rahmen sozialwissenschaftlicher Blended Readings betrachteten Texte vielfach direkte Rückschlüsse auf ihren Entstehungskontext erlauben, gestaltet sich ein entsprechender Rückbezug im Falle literarischer Texte, die ihre Kontexte ästhetisch überformen oder chiffrieren, deutlich komplizierter.

Diese Problematik tangiert im Falle der hier vorgestellten Analyse die Herstellung interpretierender Bezüge zwischen einerseits der externen Übersetzungsgeschichte der Kontexte (d. h. der institutionellen Affiliation der

Übersetzenden, der Verlagsgeschichte und im abstrakteren Sinne auch des *kyōyōshugi*-Bildungsideals) sowie andererseits der inneren Übersetzungsgeschichte der anhand der Texte nachvollzogenen Ähnlichkeits- und Einflussbeziehungen. Obwohl es unmittelbar einleuchtet, dass historische Zusammenhänge zwischen den Übersetzenden (bspw. in Form einer zeitgleichen Tätigkeit innerhalb derselben akademischen Institution oder eines erwiesenen Schüler-Mentor-Verhältnisses) auch in literarischen Texten wie den *Tonio Kröger*-Retranslations zu charakteristischen Effekten führen könnten, liegt ein solcher Zusammenhang keinesfalls in der Natur der Sache: Typologische Ähnlichkeiten zwischen unterschiedlichen literarischen Texten können auch dann zustande kommen, wenn ihre Verfasser*innen vollständig unabhängig voneinander agiert und übersetzt haben. Umgekehrt ist es ebenso möglich, dass zwei Übersetzende engen beruflichen oder privaten Kontakt miteinander gepflegt, aber dennoch ihre Texte (aus Prinzip?) gänzlich unabhängig voneinander verfasst oder sich sogar gezielt voneinander abgegrenzt haben. Beide zuletzt entworfenen Szenarien mögen etwas konstruiert wirken, veranschaulichen aber dennoch, dass das von Martin Mueller im Scalable Reading entworfene „shuttling between ‘text’ and ‘context’“ an sich zwar grundsätzlich möglich ist, hierbei aber keinesfalls ohne qualitative Interpretationsschlüsse auskommt, die ihrerseits einer eingehenden Überprüfung bedürfen. Wird also bspw. in der vorliegenden Untersuchung ein Zusammenhang zwischen der Publikationsumgebung, der institutionellen Anbindung sowie Textcharakteristiken diskutiert, nähern wir uns der Möglichkeit eines solchen Zusammenhanges lediglich interpretierend an: Sofern keine explizite Äußerung des oder der jeweiligen Übersetzenden dahingehend vorliegt, dass der jeweilige Text in Reaktion auf bestimmte, klar benannte Kontextfaktoren entstanden sei (und selbst solch eine Äußerung müsste noch hinterfragt werden!), ist und bleibt ein solcher Zusammenhang zwischen Text und Kontext eine Interpretation, die zwar auf Grundlage anderer Skalierungsschritte formuliert wird, sich aber keineswegs „automatisch“ aus diesen ergibt. Entsprechend muss auch die implizite Annahme problematisiert werden, dass sich aus stilometrischen Ähnlichkeiten automatisch auch das Vorhandensein genetischer Einflussbeziehungen schlussfolgern ließe (Weitin 2021: 30) – je mehr spezifische stilometrische Ähnlichkeitsbeziehungen zwischen zwei Texten nachgewiesen werden können, desto wahrscheinlicher wird eine über typologische Parallelen hinausgehende genetische Beziehung, aber absolute Gewissheit bietet auch hier nur ein explizites Selbstzeugnis des Verfassers oder der Verfasserin. Untersucht hat diesen Zusammenhang zwischen typologischer Ähnlichkeit und genetischem Einfluss insbesondere Matthew Jockers, der für die Berücksichtigung einer Vielzahl potenzieller, sowohl stilistischer als auch thematischer Ähnlichkeitsmarker plädiert und

in diesem Zusammenhang auch das aus den Sozialwissenschaften stammende Konzept der Informationskaskaden als Erklärung literarischer Nachahmung in Erwägung gezogen hat (Jockers 2013: 156–158).

Auch jenseits der mithin zu problematisierenden Zusammenhänge zwischen Text und Kontext bzw. zwischen typologischer Ähnlichkeit und genetischem Einfluss ergeben sich aus der im Scalable Reading implizierten „Methodenmetapher vom skalierbaren Lesen“ (Weitin 2017: 2) weitere potenziell irreführende Implikationen in Hinblick auf das Zusammenspiel von quantitativen und qualitativen Textzugängen bzw. auf die „Vorstellung eines synthetisierenden und damit stufenlosen Übergangs“ zwischen diesen unterschiedlichen Analysemodi, wobei es „Unklar bleibt, wie genau der Prozessierungs- und Medienbruch zwischen der qualitativen Betrachtung von Einzeltexten und der quantitativen Verarbeitung von Textkorpora zu überbrücken ist“ (Krautter/Willand 2020: 86–87). Das erstmalig in Abschnitt 2.4.2 dieser Arbeit ausführlicher vorgestellte Analyse-schema bietet diesbezüglich insofern wichtige Anhaltspunkte, als die Skalierung der vier Analyseschritte analog zu den der Methode Topic Modeling inhärenten Skalierungsebenen gestaltet ist und dabei zunächst eine Höherskalierung auf der Ebene der Gesamtexte (Krautter/Willand 2020: 86) sowie anschließend ein Hereinzoomen auf einzelne Topics, Topic-Paare, Absatzdokumente und schließlich auf einzelne Tokens vorsieht. Dabei gilt es jedoch zu beachten, dass dieser Analyseverlauf von der Makro- zur Mikroanalyse, der anschließend im Sinne des Algorithmic Criticism erneut auf der Makroebene ansetzt, in dieser Form aus den Skalierungsebenen eines Topic Modeling abgeleitet werden kann, aber nicht zwingend aus ihnen hervorgehen muss: Gerade die im ersten Analyseschritt betrachteten quantitativen Ähnlichkeitsbeziehungen der Gesamtexte basieren, sieht man sich die entsprechenden SQL-Kommandos genauer an, auf Durchschnittswerten, die ihrerseits auf Grundlage einer Vielzahl isolierter Ähnlichkeitsbeziehungen zwischen einzelnen Topics bzw. Themenschwerpunkten berechnet worden sind – die dafür notwendigen quantitativen Mikroanalysen erfolgen also im Hintergrund der Berechnung und somit aus der Anwendungsperspektive zunächst „unsichtbar“, konstituieren aber dennoch die Ergebnisse, auf die sich alles Weitere bezieht. Ebenso implizieren die so untersuchten Ähnlichkeitsbeziehungen der Gesamtexte bereits einzelne Ähnlichkeitswerte für konkrete Absatzdokumente, die aber erst in einem späteren, dritten Analyseschritt zum Tragen kommen. Da diese Ausführungen der eigentlichen Erläuterung des Analyse-schemas vorgeschaltet sind, müssen und können sie sich an dieser Stelle noch nicht im Detail erschließen. Wichtig ist aber, dass sich der vermeintlich „natürliche“ Analyseverlauf von der Makro- zur Mikroanalyse bzw. von den Gesamtexten zu einzelnen Tokens bzw. Wörtern, Wortgruppen oder Satzteilen

mitnichten von selbst ergibt, sondern Rechenoperationen, die nicht zwingend in der genannten Abfolge durchgeführt werden müssen, gezielt an die Logik der Anwender*innen anpasst: Orientierte sich das Analyseschema strikt an den im Topic Modeling durchgeführten Rechenoperationen, spränge es zwischen einzelnen Topics bzw. Themenkomplexen, Topicpaaren, Absatzdokumenten und den entsprechenden Durchschnittswerten hin und her. Bereits hierin verdeutlicht sich also die Relevanz der im Kontext des Scalable Reading thematisierten Skalierungsproblematik.

Selbst vermeintlich objektive Analysekatgorien und die in Entsprechung hierzu gebildeten methodischen Verfahrensmuster beruhen demzufolge bereits auf qualitativen Interpretationen des Gegenstandsbereichs (Krautter/Willand 2020: 93). Daher ist selbst den vermeintlich ausschließlich quantitativen Analyse-schritten innerhalb des hier vorgestellten Schemas insofern ein qualitativ-interpretierendes Moment eigen, als sie an die Vorstellung eines allmählichen Herunterskalierens angepasst worden sind. Die in diesem Abschnitt dargelegten Überlegungen veranschaulichen somit, dass das Zusammenspiel quantitativer und qualitativer Textzugänge mit eigenen Risiken und Herausforderungen einhergeht, die explizit thematisiert und problematisiert werden müssen. Dies resultiert jedoch nicht etwa aus der durch binäre Begriffspaare wie *distant* und *close*, „empirisch versus theoretisch, analytisch versus interpretativ, makro versus mikro sowie objektiv versus (inter-)subjektiv“ (Schruhl 2018: 250) suggerierten Unvereinbarkeit beider Analysemodi, sondern vielfach aus mangelnder Trennschärfe zwischen den nur augenscheinlich so klar abgegrenzten Konzepten: Während den vermeintlich objektiven quantitativen Zugängen bspw. in Hinblick auf die Gestaltung der Analyseschritte bereits ein qualitativ interpretierendes Moment innewohnt, kommen auch traditionelle qualitative Interpretationszugänge kaum ohne quantifizierende Überlegungen in Hinblick auf „die Bestimmung von Protagonisten oder zentralen Handlungslinien, die Bestimmung von Themen und Motiven, die Identifikation metrischer Muster“ aus und unterscheiden sich so von den „definitiven Auszählungen“ quantitativer Forschungsdesigns v. a. durch eine Praxis des schätzungsweisen, „mehr oder weniger unspezifischen Quantifizierens“ (Krautter/Willand 2020: 94–95). Hieraus folgt, dass *Distant Reading* zwar wie eingangs erwähnt mehrheitlich auf quantitativen und *Close Reading* mehrheitlich auf qualitativen Vorgehensweisen basieren, aber den jeweils komplementären Modus trotzdem nicht gänzlich ausschließen. Auch dies spricht für einen produktiven Dialog zwischen quantitativen und qualitativen Verfahren der vergleichenden Textanalyse, wie er im Folgenden dargestellt wird.

2.2 *Topic Modeling* als Modus einer digitalen Übersetzungsanalyse

2.2.1 Funktionsprinzipien und Validierung

Im Rahmen des hier vorgestellten Projektes kam eine Anwendungsmodifikation des am Institut für Informatik der Martin-Luther-Universität Halle-Wittenberg entwickelten Themenmodells Topic Explorer zum Einsatz. Dabei handelt es sich um eine Java- und SQL-basierte Software, die als in Javascript verfasste Einzelseiten-Web-Anwendung zugänglich ist. Die die Korpora konstituierenden Dokumente werden im Falle der deutschsprachigen Texte mittels der Software Treetagger (part-of-speech tagging), im Falle der japanischsprachigen Texte mit MeCab aufbereitet, während die Themenmodellierung auf der Codesammlung MALLET basiert. Die Berechnungsergebnisse werden in relationalen Datenbanken abgespeichert und können mithilfe der Datenbanksprache SQL (Structured Query Language) bzw. entsprechend konfigurierten Anfragen abgerufen werden.

Jenseits dieser technischen Spezifika sind Themenmodelle wie Topic Explorer insofern anschlussfähig an den Algorithmic Criticism, als sie nicht nur im Sinne eines Distant Reading Themenstrukturen auf der Makroebene des Gesamtkorpus erfassen, sondern auf dieser Grundlage Dokumente bzw. Textstellen kriteriengeleitet auswählen (Niekler/Wiedemann/Heyer 2014: 6). Grundsätzlich basieren Themenmodelle auf der Annahme, dass sich jeder literarische oder nichtliterarische Text aus einer begrenzten Anzahl thematischer Komplexe, sogenannter Topics, zusammensetzt (Blei 2012b: 8–9). Diese Topics sind definiert als eine Menge von Termen (alternativ zu „Term“ ist in den Sprachwissenschaften auch die Bezeichnung „Type“ gebräuchlich), d. h. lexikalischen Grundformen ohne morphosyntaktische Merkmale wie bspw. der Verbinfinitiv GEHEN. Sie sind abgegrenzt von Tokens, d. h. der in einen konkreten morphosyntaktischen Kontext eingebetteten Wortform wie „[du] gingst“. Die im Rahmen dieser Analyse durch Großschreibung der Terme typografisch visualisierte Term-Token-Differenzierung ist für die Textanalyse dahingehend essenziell, dass Terme wie z. B. GEHEN verallgemeinerbare Aussagen zum Textinhalt ermöglichen, während der Anzahl und Verteilung der Token-Iterationen („gingst“, „gehe“, „gegangen“ etc.) u. a. thematische Gewichtungen einzelner Textbereiche zu entnehmen sind.

In diesem Zusammenhang sind Topics definiert als Wortlisten bzw. als Menge von Termen, die auf Textebene durch wiederholt gemeinsam auftretende Tokens belegt sind (für ein auf *Tonio Kröger* bezogenes Beispiel s. das Topic [KÜNSTLER/GABE/WIRKUNG] in Abb. 2.1); in nicht auf literarische Texte bezogenen Topic Modelings können Topics so trivial sein wie

[VIRUS/KRANK/INFIZIEREN/etc.]. Die Wahrscheinlichkeit, mit der die drei beispielhalber genannten Terme innerhalb der untersuchten Dokumente gemeinsam auftreten, wird durch einen Wahrscheinlichkeitsalgorithmus berechnet. Die Grundlage dieser Berechnung ist in der Regel kein zusammenhängender Text, sondern ein Korpus von Textdokumenten. Dementsprechend bezieht sich die korpusbasierte Wahrscheinlichkeitsberechnung auf drei Ebenen von zunehmender semantischer Komplexität: Erstens die der Tokens (Wortformen), zweitens die der Textdokumente (z. B. Textabsätze) sowie drittens die des Gesamtkorpus (z. B. ganze Übersetzungstexte) (Blei/Ng/Jordan 2003: 996). Die Tokens der Dokumente, d. h. bspw. „[du] infiziert“, werden jeweils einzelnen Topics zugeordnet, wofür sich im Beispiel das Topic [VIRUS/KRANK/INFIZIEREN/etc.] anbietet. Jedes Dokument des Korpus ist durch diese Topic-Zuordnungen seiner Tokens als spezifisch proportionierte Zusammensetzung unterschiedlicher Topics aufzufassen (Blei 2012a: 78–79). Zugleich gibt die Software Topic Explorer für jedes Topic Top-Dokumente aus, in denen das betreffende Topic durch besonders viele Tokens vertreten ist, was einen Abgleich zwischen der abstrakteren Ebene der Topics und Terme einerseits und der durch die Dokumente repräsentierten Textebene andererseits möglich macht. Gemessen werden also Topics, Topic-Anteile innerhalb von Texten bzw. Textbereichen und die Zuordnung der die Texte konstituierenden Wortformen und Wörter zu bestimmten Topics (Weitin/Herget 2017: 35).

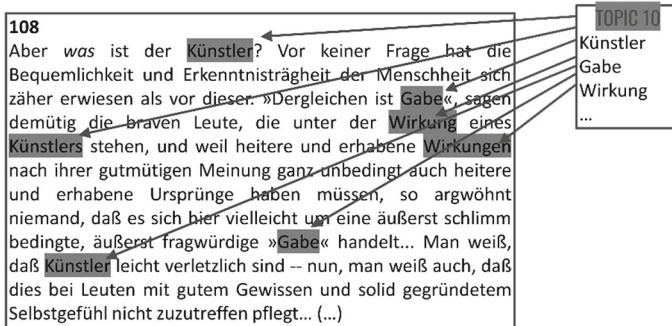


Abbildung 2.1 Zuordnung von Tokens zu Topics

Durch die jeweiligen Topic-Proportionen der einzelnen Texte entsteht im Zuge dessen ein das gesamte Korpus semantisch charakterisierendes Topic-Spektrum. Auf diesem können die Dokumente je nachdem, welche Topics in ihnen durch

besonders viele Tokens belegt sind, entsprechend verortet werden. Besteht ein beispielhaft angenommenes Korpus aus sämtlichen Zeitungsartikeln des Jahres 2020, könnte ein hypothetischer Beispielartikel (bzw. ein Dokument des Korpus) neben [VIRUS/KRANK/INFIZIEREN/etc.] auch unmittelbar oder mittelbar assoziierte Topics zu Themenbereichen wie „Schutzmaßnahmen“, „Statistik“, „Gesundheitssystem“ oder „Politik“ enthalten. Damit lässt sich der hypothetische Artikel als spezifisch gewichtete Kombination der genannten Topics beschreiben. Wird eine Vielzahl von Dokumenten bzw. Artikeln auf diese Weise analysiert, können hieraus im Falle unseres Beispiels Aussagen über die Pandemiebericht-erstattung des Jahres 2020 abgeleitet werden. Die sich im Beispiel unmittelbar erschließende Topic-Semantik sollte unterdessen nicht darüber hinwegtäuschen, dass Topics keine Themen im Sinne einer 1:1-Entsprechung sind (Blei/Ng/Jordan 2003: 996), denn der Wahrscheinlichkeitsalgorithmus von Topic Explorer erfasst die Tokens nicht nach semantischen, sondern ausschließlich nach rechnerischen Kriterien. Dementsprechend repräsentieren Topics quantitative Wahrscheinlichkeitsverteilungen, die sich sprachstrukturell ähnlich, aber nicht deckungsgleich zu qualitativen Themen i. e. S. verhalten.

Ferner wurden wahrscheinlichkeitsalgorithmische Themenmodelle wie Topic Explorer streng genommen nicht für den Zweck entwickelt, vorhandenen Text zu analysieren, sondern um Vorhersagen dazu zu treffen, wie ein sich am gegebenen Textkorpus orientierender neu generierter Text aussehen könnte. Sie sind demzufolge eigentlich prädiktiv ausgerichtet, lassen sich aber für die Analyse existierender Korpora umfunktionieren. Derartige Wahrscheinlichkeitsprognosen simulieren im Zusammenhang der Übersetzungsanalyse eine Rekonstruktion der Texterstellung: Der in das Korpus eingepflegte Text wird so in seine quantifizierbaren thematischen Bestandteile zergliedert, als sollten diese anschließend zu einem thematisch vergleichbar gewichteten neuen Text zusammengesetzt werden. Dieses Vorgehen der Themenmodelle entspricht der Auffassung der deskriptiven Übersetzungswissenschaft insofern, als in beiden Fällen der Übersetzungstext als Ausgangspunkt genutzt wird, um die seine Entstehung bedingenden Erwägungen und Entscheidungen zu rekonstruieren (Touy 2014: 18, 22).

Veranschaulicht lässt sich das allgemeine Funktionsprinzip der Themenmodelle durch Rhodys pointierten, im Folgenden weiterentwickelten Vergleich der Topics mit einem Einkauf (Rhody 2012: 21). Die Charakterisierung der einzelnen Dokumente durch die in ihnen auftretenden Topics ist hierbei gleichzusetzen mit dem Einkaufswagen anderer, den Supermarkt verlassender Kund*innen, in den man vor Betreten des Ladens hineinschaut. Ebenso, wie Topics Rückschlüsse auf die semantischen Charakteristiken des Gesamtkorpus zulassen, verrät der Blick in den fremden Einkaufswagen bzw. auf die darin enthaltenen Produkte etwas

über das so auszugsweise repräsentierte Warenangebot des Marktes. Verlässt eine Person den Laden mit Äpfeln, die nächste mit Bananen und die übernächste mit Orangen, ist es tendenziell wahrscheinlicher, dass der betreffende Supermarkt über eine gut sortierte Obstabteilung verfügt. Die einzelnen Produkte entsprechen also den Tokens, ihre Zuordnung zu einer bestimmten Abteilung oder Warengruppe den Topics und die Einkaufswagen den Dokumenten. Stünde man folglich den ganzen Tag vor dem Supermarkt und unterzöge sämtliche herauskommende Einkaufswagen einer entsprechenden Betrachtung, könnte man sich einen ähnlich umfassenden Eindruck vom Sortiment des Marktes bzw. von der Gewichtung der angebotenen Warengruppen verschaffen, wie es Topic Modeling in Hinblick auf ein Textkorpus ermöglicht. Außerdem ist der Vergleich auch in Anbetracht der prädiktiven Natur der Wahrscheinlichkeitsalgorithmen insofern tragfähig, als die Auswertung der fremden Einkaufswageninhalte den Ausgangspunkt für den eigenen Einkauf bilden könnte.

Durch diese evidenzbasierte Kategorisierung und Gewichtung unterschiedlicher Themenkomplexe kann der Eindruck entstehen, Topic Modeling ermögliche im Sinne einer *science of literature* die radikal objektive Analyse literarischer Texte (Pasanek/Sculley 2008: 358). Doch auch jenseits der Tatsache, dass die Wortwahl eines Thomas Mann anderen Regeln folgt als das Warenangebot im Supermarkt, endet hier die Vergleichbarkeit, denn weder Themenmodelle noch digitale Methoden an sich garantieren Objektivität. Bedingt ist dies allein schon dadurch, dass nicht nur die eventuelle Auswertung der Analyseergebnisse, sondern schon Entwicklung und Anwendung der digitalen Methode (wie es im vorherigen Abschnitt bereits in Hinblick auf die Skalierungsproblematik thematisiert worden ist) durch Erkenntnisinteressen beeinflusst sind (Blei 2012b: 10). Die digitalen Analyseinstrumente können folglich gezielt auf die Erfassung solcher Daten hin ausgerichtet werden, die aus menschlicher Perspektive sinnhaft erscheinen, und andere Daten herausfiltern. Auch ist die Auswahl von Texten oder Textstellen für ein Close Reading im Algorithmic Criticism bzw. Blended Reading zwar durch quantifizier- und objektivierbare Kriterien geleitet; die Entscheidung, welche quantitativen Aspekte im Einzelnen berücksichtigt werden, liegt aber nach wie vor bei den Interpretierenden (Stulpe/Lemke 2016: 47).

Demzufolge muss nicht nur die eventuelle Anpassung der digitalen Analysewerkzeuge an Erkenntnisinteressen und Erwartungshorizonte zwingend als solche benannt und reflektiert werden; auch eine adäquate Validierung der Analyseergebnisse ist ohne domänenspezifische Fachkenntnisse nahezu ausgeschlossen (Kitchin 2014: 5). Während ein nichtliterarisches Topic Modeling je nach inhaltlicher Ausrichtung u. U. auf Grundlage des Allgemeinwissens validiert werden kann, ist eine solche externe Validierung im Topic Modeling literarischer Texte

kaum sinnvoll. Dies liegt zunächst daran, dass der Algorithmic Criticism im Idealfall einen mehrfach wiederholten und aktualisierten hermeneutischen Argumentationskreislauf zwischen algorithmisch modellierten Datensätzen und der Interpretationshypothese vorsieht (Sculley/Pasanek 2008: 410–411), also methodische Fachkompetenz erfordert. Daneben bedarf es inhaltlichen Fachwissens, um die Ergebnisse literaturwissenschaftlicher Topic Modelings einzuordnen und zu bewerten.

Kein hinreichendes Validitätskriterium ist dagegen die Interpretierbarkeit der Analyseergebnisse (Sculley/Pasanek 2008: 421). Interpretierbarkeit suggerieren können bereits einzelne Tokens oder Terme (Moretti 2013: 203), zu denen bspw. auch der „unschöne, aber spannende“ Ein-Wort-Titel *Litteratur* [sic] gehört, den Thomas Mann ursprünglich für seine Erzählung *Tonio Kröger* vorgesehen hatte (Mann, in Wysling 1968: 14). Infolgedessen können *Tonio Kröger*-Topics wie [GOTT/AUGE/TANZEN] geradezu lyrisch anmuten, wobei entsprechende Assoziationen insbesondere durch die Tradition der interpretierenden Auseinandersetzung mit poetischen Wortlisten bedingt sind (Pigeot 1990). Neben der formalen Ähnlichkeit der Topics zu poetischen Listen ist es jedoch auch eine im Rezeptionsästhetischen Kontext beobachtete Bereitschaft zur Sinnstiftung, die sie u. U. literarisch relevanter erscheinen lässt, als sie eigentlich sind. Dieses Phänomen veranschaulicht Fish am Beispiel eines Literaturkurses, der quasi aus Gewohnheit den Tafelanschrieb zu interpretieren beginnt, obwohl das vermeintliche „Gedicht“ eine willkürliche, aus der vorherigen Lehrveranstaltung übrig gebliebene Stichwortsammlung ist (Fish 1980: 328). Ein solcher Interpretationseifer ist umso ausgeprägter, je mehr Interesse an der Sinnhaftigkeit der Resultate besteht. Dementsprechend kann auch eine im Kontext der digitalen Geisteswissenschaften formulierte Interpretationsthese, die sich ausschließlich auf die augenscheinliche Interpretierbarkeit weniger handverlesener Topics stützt, nicht adäquat validiert werden, da solche vermeintlich „sinnhaften“ Berechnungsergebnisse immer im Gesamtkontext der im Topic Modeling generierten Daten (zu denen vielfach „Rauschen“ gehört) betrachtet werden müssen (Weitin 2021: 119).

Validiert werden kann ein Distant Reading im Kontext des Algorithmic Criticism vielmehr in erster Linie dadurch, dass unterschiedliche Berechnungen abgeglichen und als sinnhaft eingestufte Analyseergebnisse ebenso wie „Datenrauschen“ lückenlos dokumentiert sowie konsequent hinterfragt werden (Sculley/Pasanek 2008: 422–423). Andernfalls bleibt auch hier der metaphorische „Sprung“ von quantitativen Frequenzen und themenstrukturellen patterns zur qualitativen Deutung mit dem Risiko einer subjektiven Argumentationsführung behaftet (Craig 1999: 103), da auch quantitativ generierte „Interpretationsfakten“ prinzipiell qualitativ interpretiert werden müssen (Jockers 2013: 30).

2.2.2 Topic-Klassifikation und semantische Validität

Ziel einer durch Themenmodelle gestützten Analyse literarischer Texte ist es daher keinesfalls, das, was man auf Grundlage der Sekundärliteratur ohnehin schon weiß oder vermutet, wahrscheinlichkeitsalgorithmisch absegnen zu lassen. Wenngleich Ramsays Algorithmic Criticism auch impliziert, dass die durch Distant Reading generierten Ergebnisse im direkten Anschluss durch Close Reading interpretiert werden, ist hier eigentlich noch ein validierender Zwischenschritt erforderlich, der die Ergebnisse des Distant Reading, d. h. die Topics als solche reflektiert. Obwohl Letztere auf quantifizierbaren Kriterien basieren, sollten sie nicht als objektive Argumentationsgrundlage betrachtet, sondern sowohl in Hinblick auf die sie bedingenden Berechnungsfaktoren als auch in Hinblick auf ihre daraus resultierende semantische Beschaffenheit analysiert werden.

Im Zusammenhang dieses in den Algorithmic Criticism zu integrierenden Zwischenschritts stellt Rhody (2012) ein Modell der Topic-Klassifikation sowie in Orientierung an Chang et al. (2009) zwei Testverfahren vor, durch die sich semantische Kohärenz und Interpretierbarkeit der Topics vergleichen lassen. Das erste der beiden Testverfahren, das Chang et al. als *Word Intrusion Test* bezeichnen (Chang et al. 2009: 3–4), bezieht sich zunächst auf die das untersuchte Topic konstituierenden Terme. Zu den acht Termen innerhalb eines Topics, die auf Korpusebene in der größten Anzahl von Dokumenten je mindestens einmal belegt sind, wird dabei ein willkürlich ausgewähltes Wort hinzugefügt. Sind Expert*innen in der Lage, den semantischen Fremdkörper als solchen zu erkennen, kann von einer Passung zwischen korpusexternen semantischen Konzepten und dem berechneten Topic, d. h. von Interpretierbarkeit ausgegangen werden (Rhody 2012: 28). Ist dagegen kein Kontrast zwischen den wahrscheinlichkeitsalgorithmisch berechneten Termen und der zufälligen *Intrusion* feststellbar, handelt es sich um ein semantisch unklares Topic (s. u.).

Das zweite vorgestellte Testverfahren ist der *Topic Intrusion Test* (Chang et al. 2009: 4), der zusätzlich die Dokumente berücksichtigt, innerhalb derer das untersuchte Topic durch die meisten Tokens belegt ist. Diesen Top-Dokumenten wird ein erneut willkürlich bestimmtes Fremddokument hinzugefügt. Können Experten auch hier das ursprünglich nicht dazu gehörende Dokument identifizieren (Rhody 2012: 29), sind auch die Top-Dokumente des untersuchten Topics semantisch kohärent. Dies bedeutet, dass das getestete Topic als thematisches Bindeglied zwischen solchen Dokumenten bzw. Texten fungiert, deren Zusammenhang sich auch unter qualitativem Aspekt unmittelbar erschließt. Während der *Word Intrusion Test*

dementsprechend zur Topic-Analyse i. e. S. herangezogen werden kann, ermöglicht der *Topic Intrusion Test* einen validierenden Abgleich zwischen Distant Reading und Close Reading.

Der Praktikabilität dieser Testverfahren sind allerdings insbesondere in Bezug auf das Topic Modeling literarischer Texte allein schon dadurch Grenzen gesetzt (ebd.), dass der *Topic Intrusion Test* durch umfassende Kenntnis der analysierten Dokumente an Aussagekraft verliert: Um z. B. die Erzählung *Tonio Kröger* in angemessener Weise digital untersuchen zu können, muss man sie sowie ggf. die auf ihr basierenden Übersetzungstexte so gut kennen, dass ein willkürlich hinzugefügter Textabsatz unweigerlich auffiele. Hinzu kommt außerdem, dass der *Word Intrusion Test* die Anschlussfähigkeit der Topics an bereits existierende semantische Konzepte erfasst. Diese Konzepte gehören jedoch der Alltagssprache an, von der der literarische Sprachgebrauch vielfach abweicht. Dagegen werden im Topic Modeling literarischer Texte mit höherer Wahrscheinlichkeit Topics generiert, die sich kaum mit alltagssprachlichen semantischen Konzepten vereinbaren lassen (ebd.). Hieraus folgt nicht nur, dass auf Grundlage literarischer Textkorpora berechnete Topics sowie deren Top-Dokumente am Maßstab der Alltagssprache gemessen semantisch weniger kohärent sind, sondern ebenso, dass diejenigen Topics, die sich infolgedessen einer intuitiv-alltagssprachlichen Interpretation verwehren, möglicherweise sogar diejenigen sind, die den digital analysierten literarischen Text als solchen definieren.

Darüber hinaus entwirft Rhody das Modell einer Topic-Klassifikation, das zwischen vier die jeweilige Zusammensetzung der Topics beeinflussenden Faktoren unterscheidet. Hierbei ist der erste Topic-Typ durch im Zuge der Korporaerstellung aufgetretene OCR-Fehler (s. Abschnitt 2.4.1) sowie durch fremdsprachliche Einschübe bedingt (Rhody 2012: 29–30), die, sofern sie gehäuft innerhalb derselben Dokumente auftreten, denselben Topics zugeordnet werden. Einer solchen Häufung von OCR-Fehlern wurde im Rahmen der vorliegenden Analyse durch eine manuelle Bereinigung der Textdateien vorgebeugt (s. u.).

Als zweiten Topic-Typ klassifiziert Rhody sogenannte *Large Chunk Topics*, die v. a. auf den jeweils unterschiedlichen Umfang der im Korpus enthaltenen Dokumente zurückzuführen sind (ebd.). Auch die im Rahmen dieser Analyse untersuchten Dokumente, die einzelnen Textabsätzen der Erzählung *Tonio Kröger* entsprechen, umfassen teils nur wenige Wörter, erstrecken sich mitunter aber auch über mehrere Textseiten. Dabei enthalten die umfangreicheren Absatzdokumente mehr Tokens, sodass die Wahrscheinlichkeit, dass innerhalb eines solchen Dokuments zahlreiche Tokens einem bestimmten Topic zugeordnet werden, größer ist als bei Absatzdokumenten mit geringem Umfang. Dass hierdurch auch eine

Zuordnung zu den Top-Dokumenten eines Topics bei umfangreicheren Dokumenten wahrscheinlicher ist, führt zu statistischen Verzerrungseffekten, die im Algorithmic Criticism unbedingt berücksichtigt werden müssen.

Der dritte Topic-Typ nach Rhody ist der der semantisch evidenten Topics (ebd.: 31), die sich als anschlussfähig in Hinblick auf bestehende semantische Konzepte der Alltagssprache erweisen. Hiervon unterscheidet sich die literarische Rede bspw. durch Herstellung ansonsten inakzeptabler inhaltlich-semantischer Zusammenhänge. Deshalb ist im literarischen Topic Modeling mit hoher Wahrscheinlichkeit von *Word Intrusions* und *Topic Intrusions* im oben skizzierten Sinne auszugehen (ebd.). Diese entsprechen der literaturwissenschaftlichen Diskursdefinition Bakhtins, der zufolge sich einerseits zwischen dem Wort und dem hierdurch bezeichneten Objekt und andererseits zwischen dem Wort und dem es äußernden Subjekt eine semantisch „elastische“ Umgebung anderer, „fremder“ Wörter erstreckt, die auf dasselbe Objekt referieren können (Bakhtin 1981: 293). Demzufolge besteht gerade in jenen Abweichungen von gängigen Vorstellungen dessen, was interpretierbar ist, das innovative epistemische Potenzial eines literarischen Topic Modeling.

Dies gilt umso mehr für die semantisch unklaren Topics, die Rhody als vierten und letzten Topic-Typ vorstellt (Rhody 2012: 31–33). Sie entsprechen keinem einheitlichen semantischen Konzept der Alltagssprache, sodass der oben skizzierte *Word Intrusion Test* negativ ausfällt. Kann ein semantisch unklares Topic dennoch anderweitig validiert werden (bspw. durch Konfigurationsstabilität, s. Abschnitt 2.4.3), ist es insofern anschlussfähig an die Diskursdefinition Bakhtins, als es die literarisch ausgereizte semantisch elastische Umgebung alltagssprachlicher Konzepte erfasst. Dies entspricht ferner der Auffassung Underwoods (2012), dass Topics nicht nur Themen i. e. S., sondern ebenso unterschiedliche Sprachregister repräsentieren können. Diese Sprachregister referieren indirekt (oftmals metaphorisch) auf einen Sachverhalt, ohne diesen explizit zu benennen. Rhody veranschaulicht dies am Beispiel eines Gedichts des US-Amerikaners Paul Laurence Dunbar, in dem der Term TOD kein einziges Mal belegt ist, aber trotzdem in Form eines „elegischen“, d. h. wehmütig-klagenden Sprachregisters semantisch allgegenwärtig ist (Rhody 2012: 32). Während hier von einer spezifischen Gestaltungsabsicht Dunbars ausgegangen werden darf, weist Underwood (2012) darauf hin, dass solche metaphorischen Sinnebenen nicht zwangsläufig intendiert sein müssen, sondern auch auf gesamtgesellschaftliche Diskurse zurückzuführen sein können. In diesem Falle erfasst das Themenmodell nicht das sprachlich referierte semantische Konzept an sich, sondern diskursiv bedingte Konnotationen oder spezifisch gestaltete Sprachregister,

zu denen auch der Dialog oder die freie indirekte Rede gehören können (zu Letzterem s. Abschnitt 5.4.5.5). Demnach sollten auch semantisch unklare Topics im Topic Modeling literarischer Texte keinesfalls übergangen oder durch Manipulation der Berechnungsparameter aktiv gefiltert werden. Übereinstimmend mit der oben vorgestellten Definition von Distant Reading und Close Reading fokussieren sie Textaspekte, die im Close Reading potenziell vernachlässigt werden, und vernachlässigen selbst diejenigen Textaspekte, auf die sich ein Close Reading schwerpunktmäßig konzentriert.

Demzufolge sind die beiden *Intrusion Tests* nach Chang et al. und Rhodys Modell der Topic-Klassifikation im Rahmen dieser Untersuchung insofern relevant, als sie den an die Topics herangetragenen alltagssprachlichen Maßstab des sinnvollerweise Interpretierbaren so verschieben, dass die semantische Spezifik literarischer Texte berücksichtigt wird. Dies bedeutet keineswegs, dass jedes Topic, das so gut wie keine Herstellung textorientierter semantischer Bezüge zulässt, automatisch ein besonderes literarisches Stilregister repräsentiert. Die vorgestellten Tests und Klassifikationen eröffnen lediglich heuristische Zugänge und sollten folglich mit anderen Validierungsinstrumenten wie bspw. Close Reading und Konfigurationsabgleich (s. Abschnitt 2.4.3) kombiniert werden.

2.2.3 Anschlussfähigkeit gegenüber literarischen Texten und deren Interpretation

Ebenso wie andere Themenmodelle wurde auch Topic Explorer nicht zur Analyse literarischer Texte, sondern zur Erschließung umfangreicher nichtliterarischer Textbestände entwickelt. Rhodys im vorherigen Abschnitt vorgestellte Überlegungen, sich angesichts der semantischen Spezifik literarischer Texte in besonderem Maße auf semantisch unklare Topics zu konzentrieren, illustriert hierbei den Sonderstatus literarischer Topic Modelings: Diese priorisieren gerade diejenigen Analyseergebnisse, die bei nichtliterarischen Topic Modelings wahrscheinlich verworfen würden. Literarische Topic Modelings sind dementsprechend eine Anwendungsmodifikation, deren Potenziale und Konsequenzen im folgenden Abschnitt diskutiert werden.

Möglich wird das auf literarische Korpora übertragene Topic Modeling zunächst durch seine Anschlussfähigkeit an den u. a. durch Franco Moretti begründeten quantitativen Formalismus, der literarische Stilistik als „matter of relative frequency“ definiert (Moretti 2013: 162): Insbesondere dann, wenn Wiederholung über das in alltagssprachlichen Kontexten übliche Maß hinausgeht, veranlasst sie die Rezipierenden zur Annahme von Literarizität. Solche stilistisch

relevanten Wiederholungsstrukturen werden durch Themenmodelle schwerpunktmäßig erfasst. In Kombination mit deren soeben diskutiertem subversivem Potenzial hinsichtlich etablierter semantischer Konzepte und alltagssprachlicher Interpretierbarkeit führt dies dazu, dass im Topic Modeling eine stilistische bzw. stilometrische Dimension von üblicherweise als stilistisch unmarkiert eingestuftem Vokabular erschlossen werden kann (ebd.: 206–207). Dabei bezeichnet der Begriff der Stilometrie „computergestützte Verfahren der Erhebung stilistischer Merkmale und ihrer Häufigkeiten in Texten, sowie der Nutzung dieser Merkmale und Häufigkeiten für die Klassifikation von Texten“ (Schöch 2014: 133) oder, anders ausgedrückt, die „Fortführung der Stilistik mit den Mitteln der Statistik“ (Weitin 2021: 54).

Stilometrische Analysen sind bislang überwiegend im Zusammenhang der Autorattribution eingesetzt worden, erweisen sich aber, wie die hier vorgestellte Analyse demonstriert, auch bei der Untersuchung von Ähnlichkeits- und Einflussbeziehungen zwischen unterschiedlichen Texten als produktiv: In beiden Fällen liegt der Fokus nicht auf absoluten Texteigenschaften, sondern auf „relativer Ähnlichkeit oder Differenz“, anhand derer Texte klassifiziert oder gruppiert werden (Schöch 2014: 134). Hinzu kommt, dass ebenso, wie Topic Modeling auf vollautomatischem Clustering (FAC/Fully Automated Clustering) basiert, sodass keine der Analysekategorien vorab definiert werden, auch bei stilometrischen Analysen „Zielkategorien oder Gruppen“ nicht etwa vorgegeben werden, sondern aus den festgestellten Ähnlichkeitsbeziehungen hervorgehen sollten (ebd.).

Die sowohl mit einer stilometrischen Untersuchung als auch mit Topic Modeling einhergehende relationale Perspektive hat ferner zur Folge, dass anstelle der sinnlich fassbaren, intuitive Assoziationen hervorrufenden Sprachgestaltung der Textoberfläche das im Close Reading üblicherweise hinter der Textoberfläche verborgene, auf sprachlichen Patterns und Strukturen basierende quantitative Textskelett in den Fokus der digitalen Analyse gerät (Moretti 2013: 218). Dementsprechend geht mit Distant Reading ein von herkömmlicheren Lektüren zunächst grundverschiedener Analysemodus einher: Während im Close Reading Beobachtungsfragmente argumentativ verknüpft werden, beziehen sich sowohl die wahrscheinlichkeitsalgorithmischen Berechnungen als auch die Darstellung der Analyseergebnisse im Topic Modeling auf einen im Distant Reading synchronisierten Gesamttext. Diese synchrone Erfassung sämtlicher quantitativ relevanter Texteigenschaften illustriert im Umkehrschluss die spezifischen Grenzen menschlicher Textrezeption. Für gewöhnlich lesen wir ein Buch nach dem anderen, Absatz für Absatz, Zeile für Zeile, Wort für Wort. Literatur ist dementsprechend als Zeitkunst, Lektüre als zeitlich gebundener Vorgang definiert (Fish 1980: 345), sodass sich auch Close Reading durch seine Diachronizität auszeichnet. Obwohl

die das Close Reading bedingende Interpretationsparaphrase als Synchronisierungsanstrengung aufgefasst werden kann, beinhaltet sie dennoch meist Elemente der zeitlichen Handlungsabfolge und verknüpft einzeln erschlossene Argumentationsbausteine nachträglich zu Strukturen höherer Ordnung. Dies wird auch durch die mit Close Reading typischerweise einhergehende hermeneutische Argumentationsstruktur nicht relativiert, da auch hier ein Interpretationszusammenhang zwischen anfänglich auf Textfragmente bezogenen Analysen hergestellt wird.

Im Algorithmic Criticism bzw. Distant Reading entstehen wahrscheinlichkeitsalgorithmisch generierte Datensätze, also Argumentationsstrukturen höherer Ordnung, hingegen wie oben angedeutet bereits zu Beginn der Arbeit am Text. Interpretiert wird folglich zunächst weniger, welcher argumentative Zusammenhang sich zwischen Fragmenten herstellen lässt, sondern umgekehrt, warum diese Fragmente innerhalb der Berechnungen in bestimmte Zusammenhänge integriert sind. Der Gegenstand der im Distant Reading an den Text zu stellenden Fragen ist daher nicht wie im Close Reading der Teil, sondern ein (quantifiziertes) Textganzes; die Synchronisierung des Gesamtextes ist weniger das Ergebnis als der Ausgangspunkt des Algorithmic Criticism. Die sich hieraus ergebenden literaturtheoretischen Konsequenzen thematisiert insbesondere Drouin (2011, 2017) im Zusammenhang seiner Topic Modelings zum Kirchenmotiv in Marcel Prousts Roman *À la recherche du temps perdu*. Die visualisierende Aufbereitung der Analyseergebnisse durch Netzwerkgraphen begreift er als einen „electrical circuit in which a magical being resides“, also als eine im metaphorischen Schaltkreis fixierte narrative Simultanität, die ansonsten diachron entfaltete Konzepte als Netzwerk visualisiert (Drouin 2017). Eine solche synchronisierende Visualisierung leistet auch die Benutzeroberfläche der Software Topic Explorer, in der zwar nicht der Volltext, aber das diesen thematisch konstituierende Topic-Spektrum durch farblich voneinander abgegrenzte Kästen ebenfalls visuell aufbereitet ist. Auch die thematische Schwerpunktsetzung einzelner Dokumente erschließt sich insofern unmittelbar, als die Tokens im Fließtext des Dokuments je nach Topic-Zugehörigkeit farbig unterlegt sind. Enthält ein Dokument zahlreiche Tokens, die z. B. einem rot eingefärbten Topic zugeordnet werden, erscheint der Fließtext ebenfalls überwiegend rot. Unterdessen ist die Synchronisierung der Gesamtexte im Topic Modeling keineswegs nur an diese Benutzeroberfläche gebunden, sondern ebenso an die diese konstituierende SQL-Datenbank (s. u.). Synchronisiert ist ein literarischer Text demzufolge auch dann, wenn seine quantitativ-numerische Gesamtcharakteristik in Wertetabellen eingepflegt und das Ergebnis nicht mehr im traditionellen Sinne lesbar ist.

Aus dieser synchronisierenden Erfassung der Gesamttexte ergibt sich eine ausgeprägte methodische Passung in Hinblick auf komparatistische Fragestellungen (Moretti 2013: 215): Je umfassender die verglichenen Texte analysiert werden können, desto repräsentativer ist der Vergleich an sich. Gleiches gilt insbesondere auch für übersetzungswissenschaftliche Vergleichsanalysen, in deren Rahmen eine Synchronisierung der Gesamttexte z. B. zur Erfassung kompensatorischer Übersetzungsstrategien beitragen kann. Letztere werden dann eingesetzt, wenn formale Eigenschaften einer bestimmten Textstelle (z. B. Metaphorik) nicht an der entsprechenden Textstelle der Übersetzung, aber ersatzhalber in deren syntaktischer Umgebung zielsprachlich realisiert werden (Kinkel 2001: 108). Diesem Ansatz zufolge ist ein Übersetzungstext also dann gelungen, wenn seine formalen Eigenschaften in der Summe (bspw. der metaphorischen Ausdrücke) denen des Originaltextes entsprechen. Während sich Close Reading auf den direkten Vergleich einzelner Textstellen beschränkt, bietet die Synchronisierung des Gesamttextes im Distant Reading also grundsätzlich die Möglichkeit, auch derartige Kompensationsstrategien zumindest quantitativ zu erfassen. Darüber hinaus ist eine solche vom diachronen Lektüreakt emanzipierte Analyse insofern sinnvoll, als es in Thomas Manns literarischen Werken laut Autor „keine freie Note“ (Mann, zit. nach Kurzke 1997: 85) mehr gibt: Angesichts der gestalterischen Konsequenz, mit der diese Texte mit ihren komplexen thematischen Wechselbeziehungen und Verweisstrukturen angelegt sind, lohnt es sich, den Gesamttext als „nichtaktualisierte Potentialität“ (Bode 1988: 325), d. h. als ein Rezeptionsangebot, das von menschlichen Rezipierenden kaum vollständig ausgeschöpft werden kann, quantitativ zu synchronisieren – wobei auch hier die Synchronisation nur einen spezifischen Teilbereich der Potenzialität aktualisiert.

Auf diese oder ähnliche Weise synchronisierte Repräsentationen sind eine Grundvoraussetzung sowohl der literaturwissenschaftlichen Interpretation als auch der literarischen Übersetzung. Ramsay bezeichnet sie als „hinter“ dem Ursprungstext liegende alternative Textgestalten oder Alternativtextualitäten (Ramsay 2011: 45). Dabei bezieht sich der Begriff der alternativen Textgestalt vorrangig auf die äußerlich veränderte Repräsentationsform, wogegen der Begriff der Alternativtextualität abstrahierend die hierdurch gegenüber dem Ursprungstext veränderten ästhetischen Eigenschaften bezeichnet. Da Alternativtextualitäten den Text, aus dem sie hervorgegangen sind, potenziell unterminieren oder sogar ersetzen, bedeutet eine auf ihnen basierende Interpretation eine konstruktive Destabilisierung des Ausgangstextes (Irizarry 1996: 155). Im Kontext des Algorithmic Criticism profitiert diese Destabilisierung insofern entscheidend von der Interaktivität der digitalen Methode, als sich z. B. im Topic Modeling je nach

Berechnungskonfiguration eine Vielzahl alternativer Textgestalten (zu denen auch tabellarische Auflistungen gehören) parallel erzeugen und vergleichen lässt.

Unter literaturtheoretischen Gesichtspunkten problematisch ist die im Rahmen dieser Analyse implementierte Anwendungsmodifikation des Themenmodells Topic Explorer dagegen in Hinblick auf syntagmatische Texteneigenschaften, zu denen bspw. Satzbau und narrativer Verlauf gehören. Themenmodelle repräsentieren ausschließlich Inhalte und berücksichtigen hierbei die Zugehörigkeit der Inhalte zu einzelnen Dokumenten, die Anzahl entsprechender Tokenbelege auf Dokumentenebene sowie die durch gemeinsame Tokenbelege realisierten Zusammenhänge zwischen den Inhalten. Für die Berechnung ist es jedoch vollkommen irrelevant, an welcher Stelle innerhalb der Dokumente die Inhalte belegt sind. Der Gegenstand der wahrscheinlichkeitsalgorithmischen Berechnungen ist nicht der narrative Handlungsverlauf, sondern eine Masse von ausschließlich durch die Zuteilung zu Dokumenten strukturierten Inhalten, die dementsprechend auch als *bag of words* bezeichnet werden (Blei 2012a; Mohr/Bogdanov 2013: 547).

Diese Eigenschaft der Themenmodelle ist allerdings anschlussfähig an die konstruktive Destabilisierung literarischer Texte. Das Ausblenden syntagmatischer Relationen im Topic Modeling ist demnach ebenfalls als Alternativtextualität aufzufassen, die einen spezifischen Anteil des in den Ausgangstext eingeschriebenen Interpretationspotenzials realisiert. Eine derartige Alternativtextualität ähnelt den ihrerseits an Topics erinnernden *entropic poems* Irizarrys bzw. der von ihr angeregten Interpretationstechnik einer Gedichtlektüre in umgekehrter Versfolge (Irizarry 1996; Ramsay 2011: 32–38). Während Irizarry argumentiert, dass dergestalt das semantische Potenzial derjenigen Textbestandteile freigesetzt werde, die ansonsten im narrativen Zusammenhang syntagmatisch versiegelt sind, ist einer allzu unkritischen Übernahme dieses Gedankens mit Vorbehalt zu begegnen: Isolierte Topics bzw. Terme verführen wie oben bereits problematisiert zu sich verselbständigenden Interpretationen, sodass weitere Validierungstechniken wie der Vergleich unterschiedlicher Berechnungskonfigurationen und der Rückbezug auf die Textebene im Close Reading unerlässlich sind.

2.3 Topic Modeling als Modus eines digitalen Übersetzungsvergleichs

2.3.1 Parallelisierung der Übersetzungskorpora

Als prägnantes Beispiel eines Literatur- und Kulturtransfers im globalen Maßstab profitiert auch eine Untersuchung der japanischen Thomas Mann-Übersetzung von der größtmöglichen Skalierung der Betrachtungsebene und den quantitativen Formalisierungszwängen eines Distant Reading. Auch der digitale Übersetzungsvergleich kann durch Berücksichtigung größerer Textmengen und durch die quantitative Synchronisierung der Gesamttexte tendenziell mehr Repräsentativität für sich beanspruchen.

Um den deutschsprachigen *Tonio Kröger*-Ausgangstext sowie die darauf basierenden japanischsprachigen Übersetzungstexte mit dem Themenmodell Topic Explorer vergleichend analysieren zu können, müssen digitale Korpora erstellt werden. Obwohl es dabei grundsätzlich möglich wäre, die spezifischen Ausgaben des Originaltextes zu berücksichtigen, aus denen die unterschiedlichen Übersetzungstexte im Einzelnen hervorgegangen sind, sind solche editionskritischen Erweiterungen nur eingeschränkt sinnvoll, sofern der Ausgangstext als verbindendes Element zwischen den diversen Übersetzungsvarianten aufgefasst werden soll. Aus diesem Grund bezieht sich das Topic Modeling auf eine verallgemeinernde Abstraktion unterschiedlicher Ausgaben des Ausgangstextes (Frank/Hulpe 1987: 102), sodass sämtlichen japanischsprachigen Übersetzungstexten die *Tonio Kröger*-Edition der Großen Kommentierten Frankfurter Ausgabe gegenübergestellt wird. Gerechtfertigt ist dieses Vorgehen auch dadurch, dass Unterschiede zwischen den japanischen Übersetzungstexten weniger auf unterschiedliche Editionsvarianten des deutschsprachigen Ausgangstextes, sondern in erster Linie auf durch den jeweiligen Rezeptionskontext bedingte Gestaltungsabsichten, also unterschiedliche implizite Übersetzungstheorien zurückzuführen sein dürften. In Entsprechung hierzu wird auch in Hinblick auf die Übersetzungstexte jeweils nur eine Textausgabe berücksichtigt. Obwohl man auch Neuausgaben zuvor publizierter Übersetzungstexte je nach Umfang eventueller Nachbearbeitungen als Retranslations klassifizieren könnte (Koskinen/Paloposki 2010: 294), ist das Ziel der vorliegenden Untersuchung ein möglichst umfassender, aber nach wie vor operationalisierbarer relationaler Übersetzungsvergleich. Aus diesem Grund werden graduelle Variationen individueller Übersetzungskonzepte zugunsten der historisch kontextualisierten Überblicksdarstellung vernachlässigt.

Die Vergleichbarkeit der Übersetzungskorpora wird durch Parallelisierung gewährleistet: Es werden aufeinander abgestimmte Parallelkorpora erzeugt, wie

sie ansonsten bspw. auch in Hinblick auf multilinguale Parlamentsberichte eingesetzt werden (Kenny 2011a: 61–62). Hierbei bildet jeder (Übersetzungs-)Text ein Korpus, das aus einzelnen Dokumenten besteht, die im Falle der vorliegenden Analyse den Textabsätzen des *Tonio Kröger*-Ausgangstextes entsprechen. Diese Absatzdokumente werden für jedes Korpus identisch nummeriert, sodass z. B. das Dokument 1 des Ausgangstextkorpus dem Dokument 1 der zu jedem Übersetzungstext erstellten Korpora entspricht. In diesem Zusammenhang erweist es sich als vorteilhaft, dass nicht nur der *Tonio Kröger*-Ausgangstext zusätzlich zur Einteilung in vergleichsweise umfangreiche Kapitel auch durch Absätze typografisch gegliedert ist, sondern dass auch die japanischen Übersetzungstexte diese Absatzverteilung weitgehend beibehalten.

Damit gewährleistet die Segmentierung nach Textabsätzen, die von der ansonsten in Hinblick auf Topic Modeling üblichen, auf Wort- oder Zeichenanzahlen basierenden und aufgrund ihrer Automatisierbarkeit unkomplizierteren Segmentierung abweicht (Bartsch et al. 2023: o. p.), die thematische Vergleichbarkeit der Absatzdokumente in parallelisierten Übersetzungskorpora. Die wenigen dennoch von der Absatzverteilung des Ausgangstextes abweichenden Absätze der Übersetzungen werden an den Ausgangstext angepasst, d. h. entweder entsprechend zusammengefasst oder aufgeteilt. Dieses Vorgehen entspricht der von Frank und Hulpke als erster Schritt der Übersetzungsanalyse vorgeschlagenen Untersuchung auf großflächige Veränderungen, zu denen neben einer abweichenden Absatz- oder Strophenstruktur auch umfassende Auslassungen und Erweiterungen gehören können (Frank/Hulpke 1987: 106). In den japanischen *Tonio Kröger*-Übersetzungen treten solche Abweichungen jedoch so vereinzelt auf, dass in Bezug auf keinen der analysierten Texte von großflächigen Veränderungen die Rede sein kann.

Vom Standpunkt der deskriptiven Übersetzungswissenschaft aus betrachtet hat diese Parallelisierung der Übersetzungskorpora zusätzlich den Vorteil, dass sie zugunsten eines relationalen Vergleichs von einer nur auf Ausgangstext und Übersetzung bezogenen Differenzialanalyse Abstand nimmt (Lambert/van Gorp 2014: 48). Eigenschaften unterschiedlicher Übersetzungstexte werden also nicht ausschließlich vor einem ausgangstextlichen Erwartungshorizont bewertet, sondern untereinander verglichen. Operationalisieren lässt sich dies insbesondere über quantitative Topic-Zusammenhänge, die im folgenden Abschnitt thematisiert werden.

2.3.2 Quantifizierung der Topic-Zusammenhänge als Modus des Übersetzungsvergleichs

Bedingt dadurch, dass sich der digitale Übersetzungsvergleich auf parallelisierte Korpora bezieht, bildet nicht etwa das nur einzelne Korpora bzw. Texte repräsentierende User Interface, sondern die diese Korpora konstituierende Datenbank die Grundlage der relationalen Vergleichsanalyse. Diese Datenbank kann mithilfe der Datenbanksprache SQL (Structured Query Language) abgefragt und bearbeitet werden. Die Datenbankansicht hat gegenüber dem Topic Explorer-User Interface den Vorzug, dass die tabellarischen Datensätze unterschiedlicher Korpora bzw. unterschiedlicher Übersetzungstexte flexibel navigiert werden können. In diesem Zusammenhang ermöglichen es SQL-Abfragen, korporaübergreifende Überschneidungen zwischen unterschiedlichen Topics zu quantifizieren, die sich als Zusammenhangs- oder Ähnlichkeitswerte interpretieren lassen.

Die Ermittlung dieser Zusammenhangs- oder Ähnlichkeitswerte folgt dahingehend demselben Grundprinzip wie Topic Modeling i. A., dass zwischen Textelementen (Tokens), die im Text gemeinsam belegt sind, ein thematischer Zusammenhang angenommen wird: Dies gilt sowohl für Tokens innerhalb eines Dokuments im selben Korpus, die demselben Topic zugeordnet werden, als auch für Tokens aus unterschiedlichen Korpora, die, sofern sie wiederholt im selben (parallelisierten) Absatzdokument belegt sind, ebenfalls einen thematischen Zusammenhang zwischen diesen Korpora nahelegen (s. u.). Dabei ist ein Zusammenhang zwischen Tokens jedoch nicht im Sinne unmittelbarer syntagmatischer Nachbarschaft, d. h. des direkten Aufeinanderfolgens auf Satzebene zu verstehen, sondern meint lediglich das Auftreten innerhalb derselben Dokumente bzw. Textabsätze. Folglich entsprechen die im Topic Modeling ermittelten Zusammenhänge oder Ähnlichkeitsbeziehungen nicht nur dem, was in einem Atemzug genannt wird, sondern erfassen auch weitere Teile der jeweiligen Textumgebungen und damit „latente semantische Muster“ (Weitin 2021: 130).

Ausgehend von diesen Grundprinzipien werden für die Ermittlung von quantitativen Ähnlichkeitsbeziehungen in Bezug auf zwei (oder mehr) Topics eines Korpus, die im Folgenden A und B genannt werden, diejenigen Absatzdokumente ermittelt, in denen entweder zu A oder zu B gehörende Terme durch entsprechend zugeordnete Tokens belegt sind (s. Abb. 2.2). Je mehr Dokumente zugleich Tokens von Topic A und Topic B beinhalten, desto wahrscheinlicher ist ein inhaltlich-thematischer Zusammenhang zwischen beiden Topics. Die Anzahl der entsprechenden Dokumente ist innerhalb der SQL-Berechnungen als Zusammenhangswert `count(distinct dt1, DOCUMENT_ID)` repräsentiert. Auf dieser Grundlage werden außerdem zwei weitere, als `s_min2` und `s_min3` bezeichnete

Zusammenhangswerte berechnet. Der Wert $s_{\min 2}$ referiert auf die Anzahl sämtlicher Tokens, die entweder Topic A oder Topic B zugeordnet werden und auf Dokumentenebene gemeinsam belegt sind. Während anhand von `count(distinct dt1, DOCUMENT_ID)` folglich die Verteilung der Topic-Zusammenhänge in Bezug auf das Gesamtkorpus bzw. den Gesamttext dergestalt ersichtlich wird, dass der Anteil der den betrachteten Topic-Zusammenhang realisierenden Absatzdokumente an allen Dokumenten des Korpus bestimmt wird, zeigt der Wert $s_{\min 2}$, durch wie viele Tokens der Topic-Zusammenhang in der Gesamtheit aller Absatzdokumente realisiert ist. Ist `count(distinct dt1, DOCUMENT_ID)` z. B. hoch, $s_{\min 2}$ hingegen niedrig, ist der Topic-Zusammenhang zwischen A und B demzufolge zwar in vielen Dokumenten präsent, aber quantitativ schwach ausgeprägt.

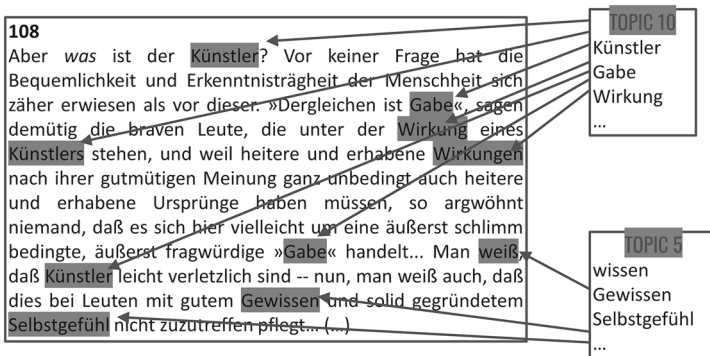


Abbildung 2.2 Topic-Zusammenhänge

Zunächst nicht berücksichtigt sind hierbei Unterschiede im Dokumentumfang, wie sie z. B. auch die *Tonio Kröger*-Textabsätze aufweisen. Diese Unterschiede beeinflussen die Berechnungen insofern, als umfangreichere Absatzdokumente mit höherer Wahrscheinlichkeit sowohl Topic A zugeordnete Tokens als auch Topic B zugeordnete Tokens enthalten, also einen Topic-Zusammenhang AB realisieren. Entsprechenden statistischen Effekten trägt der Durchschnittswert $s_{\min 3}$ Rechnung, indem er für die durch den Wert `count(distinct dt1, DOCUMENT_ID)` erfassten Dokumente die durchschnittliche Anzahl der entweder Topic A oder Topic B zugeordneten Tokens pro Absatzdokument repräsentiert. Damit zeigt auch $s_{\min 3}$, durch wie viele Tokens ein Topic-Zusammenhang AB auf Korpusebene belegt ist, berücksichtigt dabei aber im Unterschied zu $s_{\min 2}$

min2 die statistischen Auswirkungen des Dokumentumfangs. Demzufolge fällt $s_{\text{min}3}$ dann sehr hoch aus, wenn sich eine relativ große Anzahl von einem Topiczusammenhang AB realisierenden Tokens auf eine geringe Anzahl von Absatzdokumenten konzentriert. Ist $s_{\text{min}3}$ dagegen sehr niedrig, bedeutet dies, dass der Topiczusammenhang AB zwar u. U. in vielen Absatzdokumenten, aber jeweils nur durch sehr wenige Tokens nachweisbar ist. Entsprechend lässt sich ein hoher $s_{\text{min}3}$ -Wert dahingehend deuten, dass der Topiczusammenhang AB auf Dokumentenebene quantitativ ausgeprägt ist, was im Übersetzungsvergleich als intensive, aber auf einzelne Textabschnitte beschränkte Ähnlichkeitsbeziehung interpretiert werden kann. Dagegen bedeutet ein hoher Wert für $s_{\text{min}2}$, dass der Topiczusammenhang AB bzw. die dadurch repräsentierte Ähnlichkeitsbeziehung dann das gesamte Korpus bzw. den Gesamttext relativ gleichmäßig charakterisiert, wenn die Werte für $s_{\text{min}3}$ moderat ausfallen. Für den digitalen Übersetzungsvergleich sind folglich v. a. die beiden Werte $s_{\text{min}2}$ und $s_{\text{min}3}$ relevant, da diese den quantitativen Ausprägungsgrad eines Topiczusammenhangs sowohl auf der Ebene des Gesamttextes ($s_{\text{min}2}$) als auch in spezifischen Textbereichen ($s_{\text{min}3}$) widerspiegeln.

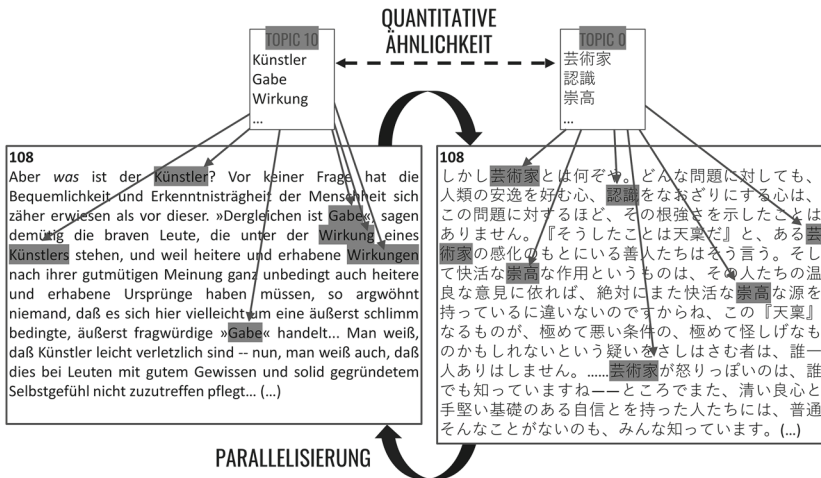


Abbildung 2.3 Korporaübergreifende Topic-Zusammenhänge

Ursprünglich beziehen sich die genannten Werte nur auf Topiczusammenhänge innerhalb eines einzigen Korpus. Die Parallelisierung der Übersetzungskorpora

ermöglicht jedoch eine für den digitalen Übersetzungsvergleich nutzbare Anwendungsmodifikation. Da die Absatzdokumente unterschiedlicher Korpora identisch nummeriert sind, kann sich die oben vorgestellte SQL-Abfrage zu `count(distinct dt1, DOCUMENT_ID), s_min2` und `s_min3` auch auf parallelisierte Dokumente beziehen. Hierbei werden z. B. ein Absatzdokument 7 des Ausgangstextes und das parallelisierte Absatzdokument 7 eines Übertextes virtuell überlagert, wodurch sich der Zusammenhang zwischen einem Topic A des Ausgangstextes und einem Topic B des Übertextes korpora- bzw. textübergreifend quantifizieren lässt (s. Abb. 2.3). Die Mechanismen, mittels derer die Werte `count(distinct dt1, DOCUMENT_ID), s_min2` und `s_min3` generiert werden, bleiben dabei unverändert: `count(distinct dt1, DOCUMENT_ID)` repräsentiert die Anzahl der parallelisierten Absatzdokumente, die den korporaübergreifenden Topiczusammenhang AB realisieren, `s_min2` gibt an, wie viele Tokens innerhalb dieser Absatzdokumente insgesamt den Topics A oder B zugeordnet sind, und `s_min3` zeigt, ob sich der Topiczusammenhang auf einzelne, möglicherweise umfangreiche Dokumente konzentriert. Entsprechende Zusammenhänge können dabei nicht nur für einzelne Topicpaare, sondern für sämtliche Topics zweier Korpora berechnet werden. Hierdurch entstehen umfangreiche tabellarische Auflistungen von Zusammenhangswerten, deren Durchschnitt Aussagen zur quantitativen Zusammenhangs- oder Ähnlichkeitsbeziehung zwischen zwei Textkorpora erlaubt. Diese quantitativen Ähnlichkeitsbeziehungen bilden den Ausgangspunkt eines vierschrigen Algorithmic Criticism, der im Abschnitt 2.4.2 ausführlich vorgestellt wird. Schwerpunktmäßig betrachtet werden in diesem Zusammenhang einerseits besonders hohe Durchschnittswerte für `s_min2` und `s_min3`, da diese einen quantitativ ausgeprägten korporaübergreifenden Topic-Zusammenhang repräsentieren, der sich als Ähnlichkeitsbeziehung interpretieren lässt. Neben ausgeprägten Ähnlichkeiten sind allerdings auch Abweichungen für den Übersetzungsvergleich von besonderem Interesse. Deshalb werden im Rahmen der Analyse auch Korporapaare untersucht, die durch besonders niedrige Werte für `s_min2` und `s_min3`, also durch eine schwach ausgeprägte quantitative Ähnlichkeitsbeziehung auffallen.

Auch Prinzl hat in ihrer digitalen Korpusanalyse mithilfe von *Word Smith Tools* Ähnlichkeitsbeziehungen zwischen unterschiedlichen Retranslations untersucht. *Word Smith Tools* erfasst jedoch ebenso wie die meisten im Kontext digitaler Übersetzungsvergleiche bisher verwendeten Analysewerkzeuge ausschließlich wörtliche Überschneidungen (Prinzl 2017: 74), Topic Explorer dagegen thematische Ähnlichkeiten. Während Prinzls Analyse also nur diejenigen Fälle berücksichtigen kann, in denen sich unterschiedliche Übersetzende exakt gleich

ausdrücken, operationalisiert Topic Explorer ein demgegenüber deutlich komplexeres Verständnis von Übersetzungsähnlichkeit.

2.3.3 Differenzierung zwischen Funktionswörtern und Inhaltswörtern

In Hinblick auf eine durch Topic Modeling quantifizierte Übersetzungsähnlichkeit ist die Unterscheidung zwischen Inhalts- und Funktionswörtern unerlässlich: Unter Inhaltswörtern versteht man die klassischerweise semantisch interpretierbaren Nomina, Verben und Adjektive; unter Funktionswörtern Konjunktionen, Artikel, Pronomen, Präpositionen und Hilfsverben ohne eigene lexikalische Semantik (DWDS 2021). Für quantitative Analysen sind diese Funktionswörter insofern besonders relevant, als sie zwar weniger als 0,1 Prozent des Gesamt vokabulars einer Sprache, aber entsprechenden Erhebungen zufolge knapp 60 Prozent des tatsächlichen Sprachgebrauchs konstituieren (Pennebaker 2011: ix) und damit bei Einnahme einer nicht anderweitig (bspw. durch Filtermechanismen) regulierten quantitativen Perspektive zwangsläufig dominieren. In der quantitativen Linguistik wird diese umgekehrte Proportionalität von Anteil am Gesamt vokabular und Verwendungshäufigkeit bzw. die entsprechende Wahrscheinlichkeitsverteilung durch das Zipf'sche Gesetz beschrieben (Weitin 2021: 57–58).

Diskutiert wurde in diesem Zusammenhang, ob in Hinblick auf die Verwendung von Funktionswörtern in literarischen Texten überhaupt eine Gestaltungsintention angenommen werden darf oder ob sie auch hier überwiegend unwillkürlich eingesetzt werden. Dies wirft letztlich die Frage auf, ob sich entsprechende Quantifizierungs- und Skalierungsansätze überhaupt auf solche sich der bewussten Kontrolle der Autor*innen entziehende Analyseeinheiten stützen sollten (Krautter/Willand 2020: 96), da Letztere möglicherweise über eine sich bewusst an Rezipient*innen richtende Aussageabsicht hinausgehen. Entsprechende Argumente sind jedoch nicht nur bei Einnahme einer rezeptionsästhetischen Perspektive müßig, bezüglich derer es keine Rolle spielt, ob ein Rezeptionsangebot intentional zustande gekommen ist oder nicht: Insbesondere dann, wenn wie im Falle der hier vorgestellten Untersuchung Einflussspuren zwischen unterschiedlichen Texten aufgedeckt werden sollen, ist es gleichgültig, ob ein eventuell stilometrisch nachvollziehbarer Einfluss den jeweiligen Übersetzenden gänzlich bewusst gewesen ist oder nicht. Ebenso, wie es Thomas Weitin in Bezug auf die stilometrische Autorattribution beschrieben hat, wird

also auch hier der literarische bzw. übersetzerische Einfluss nicht „als intentionale Kategorie operationalisiert“ (Weitin 2021: 63) – was natürlich nicht heißt, dass es nicht im bewussten Ermessen der Übersetzenden läge, inwieweit sie beim Übersetzen vorherige Übersetzungsvarianten als solche konsultieren. Das jeweilige stilistische Ausmaß, in dem eine solche Orientierung an den Texten der Vorgänger*innen stattfindet, könnte sich aber in ähnlicher Form teilweise der bewussten Kontrolle entziehen, wie man auch in alltagssprachlichen Zusammenhängen mitunter bestimmte Ausdrucksformen von Bekannten oder aus den Medien übernimmt, ohne sich einer solchen Beeinflussung gänzlich gewahr zu sein. Die Berücksichtigung von Funktionswörtern ergibt sich bei der vergleichenden Analyse von Retranslations zudem dadurch, dass diese Texte innerhalb eines über individuelle Gestaltungsabsichten der Übersetzenden hinausgehenden intertextuellen Umfelds verortet sind, das maßgeblich durch unterschiedliche historische Kontexte geprägt ist: Denkbar ist daher, dass Übersetzende bestimmte (stilistische) Elemente des *Tonio Kröger*-Ausgangstextes nicht etwa deswegen formaläquivalent-wörtlich realisiert haben, weil sie sich bewusst dafür entschieden haben, sondern weil ihnen ein entsprechendes Vorgehen vor dem Hintergrund der bisherigen, nicht nur Thomas Mann-bezogenen Übersetzungstradition selbstverständlich erschien. Auch deswegen erlaubt gerade der in sonstigen Analysen häufig vernachlässigte Umgang mit Funktionsvokabular wichtige Rückschlüsse auf derartige normativ aufgeladene „Übersetzungsselbstverständlichkeiten“.

Dafür, Funktionsvokabular keinesfalls durch entsprechende Mechanismen aus der digitalen Analyse herauszufiltern, sprechen hierbei auch erste Ergebnisse aus dem oben erwähnten Bereich der Autorattribution, welche sich auch für relationale Analysen gewinnbringend einsetzen lassen: So hat bspw. Matthew Jockers im Rahmen einer Studie zu genderspezifischem Gebrauch von Funktionsvokabular systematische Abweichungen im Pronominagebrauch und in der Interpunktion von Autorinnen und Autoren nachgewiesen (Jockers 2013: 93). Ferner hat auch Christof Schöch in einem Projekt zur digital gestützten Autorattribution versucht, zwischen Inhalts- und Funktionswörtern dahingehend zu differenzieren, dass sich Gattungszugehörigkeit in erster Linie durch Inhaltswörter, Autorschaft hingegen eher durch Funktionswörter bestimmen ließe, gelangte dabei aber zu keinen eindeutigen Ergebnissen (Schöch 2014: 147, 149). Nicht zuletzt, weil sich Thematik und Stilistik wechselseitig bedingen (Jockers 2013: 158), sollte auch Funktionsvokabular indessen keinesfalls isoliert, sondern in seinem Zusammenwirken mit Inhaltsvokabular betrachtet werden, wofür die im Folgenden vorgestellte Kombination aus quantitativen und qualitativen Textzugängen hervorragende Voraussetzungen bietet.

2.4 Rahmenbedingungen der Analyse

2.4.1 Erstellung und Aufbereitung der Übersetzungskorpora

Bevor die Übersetzungskorpora mit dem Themenmodell Topic Explorer analysiert werden konnten, bedurfte es einer Aufbereitung der Texte. Diese gestaltete sich auf dem gegenwärtigen Stand der Technik sehr arbeitsintensiv, da sowohl die deutschsprachigen Originalwerke Thomas Manns als auch die japanischsprachigen Übersetzungstexte kaum digitalisiert vorliegen. Die japanische Project Gutenberg-Variante Aozora Bunko bietet lediglich einige Erzählungen Manns in der Übersetzung Saneyoshi Hayaos als Digitalisate an, zu denen neben *Tonio Kröger* auch *Der Tod in Venedig* gehört. Da in Japan das Urheberrecht bereits 50 Jahre nach dem Tod des Autors erlischt, konnte das Digitalisat der vom bereits 1962 verstorbenen Saneyoshi Hayao verfassten *Tonio Kröger*-Erstübersetzung bedenkenlos genutzt werden. Digitalisate der 14 übrigen *Tonio Kröger*-Retranslations mussten dagegen erstellt werden.

In Deutschland ist die Rechtslage in Hinblick auf Thomas Manns Werke komplizierter. Da hier nach dem Tod des Autors nicht 50, sondern 70 Jahre vergangen sein müssen, bis das Urheberrecht erlischt, stellen online zugängliche Digitalisate noch bis 2025 eine Urheberrechtsverletzung dar. Ein Beispiel hierfür ist die US-amerikanische Plattform Gutenberg.org, gegen die S. Fischer als Thomas Manns Hausverlag vor dem Landgericht Frankfurt erfolgreich Klage eingereicht hat (S. Fischer 2023). Infolgedessen sind nicht nur die wider geltendes deutsches Urheberrecht über das Portal zugänglichen Werke Thomas Manns für sämtliche Inhaber einer deutschen IP-Adresse gesperrt, sondern das komplette Portal. Vor diesem Hintergrund konnte zunächst weder auf ein vollständiges deutschsprachiges Digitalisat der Erzählung *Tonio Kröger* zugegriffen noch eine Vollkopie bspw. durch Scannen angefertigt werden, denn solche Vollkopien sind durch die Rückausnahme (§ 53 Abs. 4b UrhG) explizit von den anderweitig legalen (§ 53 Abs. 2 Satz 1 Nr. 1 UrhG) analogen oder digitalen Teilkopien für den wissenschaftlichen Eigenbedarf abgrenzt. In diesem Zusammenhang ist es allerdings essenziell, dass die in der Analyse genutzten Übersetzungskorpora weder öffentlich zugänglich sind noch den Zugriff auf eine Vollkopie des Originaltextes ermöglichen. Stattdessen umfassen die verwendeten Dokumente nur einzelne Textabsätze, die den Umfang von Teilzitate nicht überschreiten. Dieses digitalisierte Karteikartensystem ähnelt dem nicht öffentlichen Thomas Mann-Korpus des Leibniz-Instituts für Deutsche Sprache in Mannheim sowie dem öffentlich zugänglichen Thomas Mann-Korpus der Universität Kyūshū insofern, als alle drei

Korpora nicht als lesbare Vollkopien, sondern ausschließlich als wissenschaftliche Recherchewerkzeuge genutzt werden können.

Eine weitere Absicherung bietet seit 2018 außerdem der spezifisch auf digitales Text Mining und Data Mining (TDM) bezogene und dementsprechend als „TDM-Schranke“ bezeichnete §60d des deutschen Urheberrechts (Schöch et al. 2020). Dieser gesetzlichen Regelung zufolge ist es zulässig, „das Ursprungsmaterial auch automatisiert und systematisch zu vervielfältigen, um daraus insbesondere durch Normalisierung, Strukturierung und Kategorisierung ein auszuwertendes Korpus zu erstellen“, solange dieses Korpus nur „einem bestimmten abgegrenzten Kreis von Personen für die gemeinsame wissenschaftliche Forschung sowie einzelnen Dritten zur Überprüfung der Qualität wissenschaftlicher Forschung öffentlich zugänglich“ ist. Einschränkungen gelten dabei nur dahingehend, dass keine kommerzielle Nutzung erfolgen darf und das Korpus ferner nach Abschluss der Forschungsarbeiten gelöscht werden muss, wobei es bestimmten Institutionen wie z. B. Bibliotheken zur Aufbewahrung überlassen werden kann. In Anbetracht der mithin urheberrechtlich zulässigen automatisierten Textvervielfältigung im wissenschaftlichen Verwendungszusammenhang konnten demzufolge auch die im Rahmen des vorliegenden Projekts untersuchten Texte manuell gescannt werden. Die in der Folge generierten Korpora wurden so aufbereitet, dass lediglich einzelne Textabsätze im digitalen Karteikartenformat einsehbar sind.

Die digitale Aufbereitung der Scans wurde durch OCR, d. h. *Optical Character Recognition* bzw. automatisierte Schriftzeichenerkennung, nur teilweise erleichtert. Obwohl Prinzl in Bezug auf ihre digitale Analyse englischsprachiger Thomas Mann-Übersetzungen von vergleichbaren Problemen berichtet (Prinzl 2017: 60), gestaltet sich OCR insbesondere in Hinblick auf nichtwestliche Schriftsysteme wie die japanischen Kanji, Hiragana und Katakana notorisch problematisch, sodass einzelne Zeichen falsch erfasst werden oder das vertikale Zeilenlayout im japanischen Text vertauscht wird. Um solche Fehler zu minimieren, wurden im Rahmen des vorgestellten Projekts unterschiedliche OCR-Softwares einem umfangreichen Praxistest unterzogen. Die zuverlässigsten Ergebnisse konnten mithilfe der seit 2013 vom auf OCR spezialisierten japanischen Unternehmen Mediadrive angebotenen Software *e.typist neo v. 15.0* erzielt werden, sodass diese zur Aufbereitung der japanischsprachigen Übersetzungskorpora genutzt wurde. Dennoch erzeugte auch *e.typist* bei weitem noch keine OCR-Dateien, die man bedenkenlos in ein Themenmodell einpflegen könnte. Daher mussten die OCR-Rohdateien manuell bereinigt werden. Kommentare und Fußnoten, sofern sie nicht direkt in den Fließtext der Übersetzung eingebettet waren, wurden hierbei entfernt, da sie ansonsten die korporaübergreifende Vergleichbarkeit

der Textabsätze beeinträchtigt hätten. Ferner brachte die 1928 publizierte *Tonio Kröger*-Übersetzung Mukasa Takeos durch das ausgebleichte Druckbild derart unbrauchbare OCR-Ergebnisse hervor, dass dieser Text per Hand abgetippt werden musste.

Darüber hinaus erfordert das Topic Modeling eine Aufspaltung des Gesamttextes auf einzelne Textdateien, die im Falle dieser Analyse den 269 Absätzen des deutschsprachigen Ausgangstextes entsprechen. Auch diese Dokumentdateien wurden manuell erstellt, nachdem sich ein automatisiertes Vorgehen mithilfe von text file splitter-Programmen als unpraktikabel erwiesen hatte. Da diese Programme die ursprüngliche Datei i. d. R. nach einer festgelegten Anzahl von Zeichen aufspalten, sodass Einzeldateien mit identischer Zeichenanzahl entstehen, und die Textabsätze je nach Übersetzung unterschiedlich viele Zeichen umfassen (vgl. Tab. 5.115, 5.116 EZM), eignen sich text file splitter nicht für die absatzbasierte Dokumenterstellung im Übersetzungsvergleich. Infolgedessen wurden für jeden der insgesamt 16 analysierten Texte 269 einzelne, sich an der Absatzverteilung des deutschsprachigen Ausgangstextes orientierende Textdateien erstellt.

In Anbetracht dieser sich bereits bei der Korporaerstellung ergebenden Herausforderungen ist nicht zu leugnen, dass die Nutzung vorhandener literarischer Korpora den Arbeitsaufwand beträchtlich reduziert hätte. Trotzdem lieferte die umständliche Überführung des analogen Textmaterials in eine digitale Form wichtige Erkenntnisse bspw. zum Schriftbild der untersuchten Texte. Dass hierbei das Analyseinstrument parallel zum Projektfortschritt entwickelt wurde, war ebenso Herausforderung wie auch Chance, denn nur so konnte die digitale Methode im Zuge ihrer Entwicklung an die Erfordernisse einer literaturwissenschaftlichen Übersetzungsanalyse angepasst werden. In diesem Zusammenhang hat der explorative Charakter des vorliegenden Projekts genau das erzwungen, was digitale Geisteswissenschaften leisten sollten: Die praktische Erprobung innovativer methodischer Zugänge, die Erschließung neuartiger epistemischer Potenziale und die kritisch-transparente Methodenreflexion.

2.4.2 Beschreibung der Analyseschritte

Um einen Überblick über das vierschriftige Analyseverfahren (s. Abb. 2.4) zu ermöglichen, wird im Folgenden die Abfolge der Analyseschritte in Grundzügen erläutert. Die daraus hervorgegangenen Ergebnisse werden im fünften Kapitel dieser Arbeit ausführlich vorgestellt und diskutiert. Der erste Analyseschritt erfasst die zwischen unterschiedlichen Korpora auf der Ebene der jeweiligen Gesamttexte berechneten Ähnlichkeitsrelationen. Diese werden über

eine hierfür entwickelte SQL-Abfrage (s. Anhang 9.1.1 EZM) für alle möglichen Textpaare berechnet und die Ergebnisse tabellarisch aufbereitet. Die entstehenden Ähnlichkeitshierarchien werden für unterschiedliche Sortierungen und Berechnungskonfigurationen abgeglichen; die Notwendigkeit eines solchen Abgleichs wird im folgenden Abschnitt 2.4.3 ausführlicher erörtert. Besagte Ähnlichkeitshierarchien bilden die Grundlage einer relationalen Basisklassifikation, die die japanischen *Tonio Kröger*-Übersetzungen entweder einer Kerngruppe, die durch ausgeprägte quantitative Ähnlichkeitsbeziehungen charakterisiert ist, oder einer durch besonders schwache Ähnlichkeitsbeziehungen charakterisierten Übersetzungsperipherie zuordnet.

Schritt	Skalierungsebene	Leitfrage	Ergebnis
1	Gesamttexte	Welche Texte ähneln sich quantitativ?	Vorstrukturierung durch Distant Reading
2	Topics/Topic-Paare	Welche Topics charakterisieren die quantitative Ähnlichkeit?	Nähere Bestimmung der quantitativen Ähnlichkeitsbeziehungen
3	Topic-Paare, Absatzdokumente	Auf welche Textabsätze konzentriert sich die quantitative Ähnlichkeit?	Nähere Bestimmung der quantitativen Ähnlichkeitsbeziehungen, Rückbezug auf Textebene
4	Textebene	Ist die quantitative Ähnlichkeit auf Textebene qualitativ realisiert?	Relationale Analyse, relationales Close Reading

Abbildung 2.4 Vierschrittiger Algorithmic Criticism

Diese Basisklassifikation wird in Analyseschritt 2 durch Top- und Bottom-Tops konkretisiert. Während Top-Tops ausgeprägte quantitative Ähnlichkeitsbeziehungen bedingen, spielen Bottom-Tops bei der Entstehung der besonders schwachen Ähnlichkeitsbeziehungen die unter allen Topics geringste Rolle. Hierfür wird in Hinblick auf jedes im ersten Analyseschritt ermittelte Korporapaar AB für jedes einzelne Topic aus Korpus A die durchschnittliche quantitative Ähnlichkeit zu allen Topics aus Korpus B berechnet. Daraus geht hervor, welches Top- bzw. Bottom-Topic aus Korpus A die quantitative Ähnlichkeit zu Korpus B im durchschnittlich höchsten bzw. geringsten Maße konstituiert.

In Analyseschritt 3 erfolgt auf dieser Grundlage der Rückbezug auf die Textebene. Hierfür werden für die in Schritt 2 anhand von Durchschnittswerten ermittelten Top- bzw. Bottom-Topics jeweils Bezugstopics aus Korpus B bestimmt, die die besonders stark oder schwach ausgeprägte quantitative Ähnlichkeit der Korporapaarung AB in Hinblick auf ein Top- oder Bottom-Topic aus Korpus A realisieren. Ermittelt wird also, welche konkreten Topicpaare eine quantitative Ähnlichkeit oder Unähnlichkeit (die in diesem Zusammenhang als besonders schwach ausgeprägte Ähnlichkeit definiert wird) zwischen zwei Texten schwerpunktmäßig konstituieren. Infolgedessen werden diejenigen Textabsätze bzw. Absatzdokumente bestimmt, innerhalb derer die quantitative Ähnlichkeit bzw. Unähnlichkeit dieser Topicpaare belegt ist: Für Topicpaare, die ein Top-Topic enthalten, werden parallelisierte bzw. inhaltsgleiche Absatzdokumente ermittelt, in denen die betreffende Ähnlichkeitsbeziehung mit dem höchsten quantitativen Ausprägungsgrad realisiert ist. Für die ein Bottom-Topic beinhaltenden Topicpaare werden dagegen diejenigen Absatzdokumente ermittelt, in denen die schwach ausgeprägte quantitative Ähnlichkeit durch die geringste Anzahl entsprechender Tokens realisiert ist. Ergänzend geprüft werden außerdem der jeweilige Umfang der ausgewählten Absatzdokumente sowie das Ausmaß der jeweiligen wörtlichen Überschneidung der Topicpaare auf Dokumentenebene.

Ausgehend hiervon werden im Rahmen des vierten und letzten Analyseschritts eine relationale Analyse sowie relationale Close Readings der zuvor anhand quantitativer Kriterien ausgewählten Absatzdokumente durchgeführt. Diese qualitativen Analyseverfahren beziehen sich auf eine repräsentative Auswahl von elf Absatzdokumenten in insgesamt 16 unterschiedlichen Textfassungen, also auf insgesamt 176 Absatzdokumente. Diesen stehen auf Seiten der quantitativen, auf die Gesamtexte bezogenen Analyseverfahren insgesamt 4304 Absatzdokumente (269 Absatzdokumente in je 16 Fassungen) gegenüber. Ausführlicher dargelegt werden die Kriterien für die Auswahl dieser Absatzdokumente in Abschnitt 5.3.3.

Die sich auf eine repräsentative Auswahl an Absatzdokumenten beziehende relationale Analyse ist durch die mithilfe der Analyseschritte 1 bis 3 ermittelten Ähnlichkeitsbeziehungen geleitet: Die Topics der die quantitativen Ähnlichkeits- und Unähnlichkeitsbeziehungen konstituierenden Topicpaare, d. h. die im Rahmen der Analyseschritte 2 und 3 ermittelten Top-Topics, Bottom-Topics und Bezugstopics bilden Fokustopics der relationalen Analyse, bei der die in ausgewählten Absatzdokumenten enthaltenen Tokens dieser Fokustopics mit den ihnen entsprechenden Übersetzungsvarianten der übrigen Übersetzungstexte abgeglichen werden. In der Folge werden die unterschiedlichen Übersetzungsvarianten relational, also in ihrem Verhältnis zu den übrigen Übersetzungsvarianten als Konsensübersetzungen oder als sich hiervon abgrenzende Countertranslations

klassifiziert. Für jeden Übersetzungstext werden davon ausgehend die prozentualen Anteile von Konsensübersetzungen bzw. Countertranslations innerhalb des jeweils untersuchten Absatzdokumentes ermittelt, dessen Übersetzung sich so als tendenziell konservativ oder tendenziell neuartig charakterisieren lässt. Anschließend werden die zunächst tabellarisch vorliegenden Ergebnisse dieser Analyse grafisch aufbereitet, sodass sie die jeweilige quantitative Entwicklung der relationalen Übersetzungscharakteristiken im historischen Verlauf repräsentieren. Auf Grundlage dieser Ergebnisse werden sieben Leitfragen entwickelt, anhand derer spezifische Übersetzungstokens im relationalen Close Reading interpretiert und so die zuvor ermittelten Übersetzungscharakteristiken an konkreten Textbeispielen nachvollzogen und überprüft werden. Diese Ergebnisse werden im sechsten Kapitel dieser Arbeit mit historischen Kontextinformationen zu umfassenden Übersetzungsprofilen zusammengeführt.

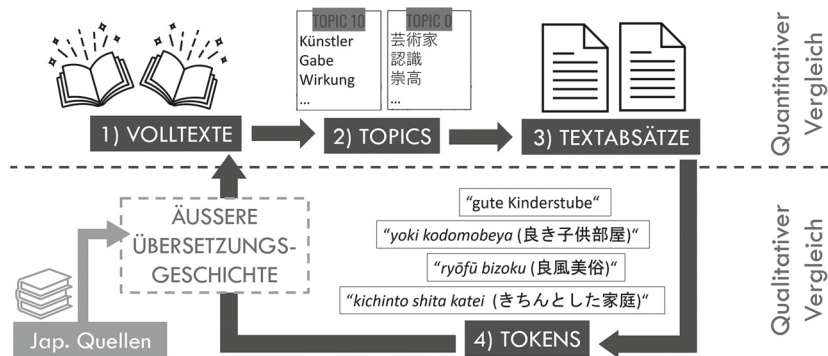


Abbildung 2.5 Analyseschema zur Aufarbeitung der äußeren und inneren Übersetzungsgeschichte

Bei der relationalen Analyse geht es also v. a. um die relationalen Ähnlichkeitscharakteristiken (Tradition vs. Abgrenzung) einzelner Absatzdokumente in unterschiedlichen Übersetzungsfassungen, beim relationalen Close Reading dagegen um unterschiedliche Übersetzungsvarianten einzelner Tokens. Die qualitative Interpretation der textuellen Mikroebene erfolgt damit auf Grundlage und im Bewusstsein der quantitativ erschlossenen Makroebene: Die digitale Methode fungiert nicht nur als Werkzeug in Hinblick auf die Auswahl der zu analysierenden Textstellen, sondern generiert einen spezifischen Erkenntnis- und Erwartungshorizont (z. B. in Hinblick auf die Ähnlichkeit zweier Texte), auf den

sich die Interpretation im relationalen Close Reading bezieht. Im Fokus der Analyse stehen folglich nicht länger nur die jeweiligen Eigenschaften der Einzeltexte, sondern die auf unterschiedlichen Skalierungsebenen betrachteten Beziehungen zwischen den Texten. Dabei umfassen diese Skalierungsebenen nicht nur die Ebenen der Gesamttexte, der Topics, der Absatzdokumente und der einzelnen Tokens, die im Sinne der historisch-deskriptiven Übersetzungsforschung (Frank 1992: 381; Frank 1988b: 195) eine innere Übersetzungsgeschichte repräsentieren, sondern darüber hinaus auch eine historische Kontextualisierung der Analyseergebnisse im Sinne der äußeren Übersetzungsgeschichte (vgl. Abb. 2.5).

2.4.3 Berechnungs- und Konfigurationsstabilität

Die Belastbarkeit der Topic Modeling-Ergebnisse ist maßgeblich bestimmt durch die Frage der Berechnungs- bzw. Konfigurationsstabilität, sodass diese vorab thematisiert werden muss. Konfigurationsstabilität meint, dass sich die Berechnungsergebnisse in Abhängigkeit von bestimmten Parametern ändern. Zu diesen Parametern gehören die berücksichtigten Wortarten, die Anzahl der berechneten Topics sowie Ober- und Untergrenzen, die festlegen, bis bzw. ab welcher Belegfrequenz ein Term in die Berechnungen einbezogen wird: Wird die Untergrenze bspw. auf 3 festgelegt, bedeutet das, dass nur Terme berücksichtigt werden, die in drei oder mehr Absatzdokumenten durch Tokens belegt sind. Werden entsprechende Konfigurationsparameter angesetzt, lassen sich gerade in Bezug auf motivisch dicht besetzte literarische Texte wie *Tonio Kröger* ohne Probleme scheinbar bedeutungsschwere Topics berechnen, die basismetaphorisch und kulturell aufgeladene Terme wie z. B. HELL, MUTTER und MEER enthalten. Dass sich diesbezügliche Interpretationen regelrecht aufdrängen, erinnert derweil mehr an Stanley Fishs im Abschnitt 2.2.1 erwähntes Interpretationsexperiment (Fish 1980: 328) und weniger an literaturwissenschaftliche Interpretationspraxis. Interpretieren lässt sich grundsätzlich alles, sofern es nur vage genug ist; derartige Interpretationen sagen allerdings mehr über die Interpretierenden als über den interpretierten Text aus. Gerade auch aus der Kenntnis des Ausgangstexts *Tonio Kröger* resultiert infolgedessen die Gefahr, auf suggestive Topics eine Sinnhaftigkeit zu projizieren, die diesen in erster Linie aufgrund ihrer Deutungsoffenheit zugeschrieben werden kann: Eine entsprechende Ansammlung von Nomina und Adjektiven mag bei entsprechender Rezeptionshaltung quasi-lyrisch wirken, entsteht in dieser Form aber möglicherweise nur durch die Festlegung der Berechnungsparameter auf spezifische Wortarten sowie auf bestimmte Ober- und Untergrenzen. Noch gezielter lassen sich durch Eingabe von stopwords bestimmte

Terme gänzlich aus den Berechnungen herausfiltern; auch die bereits erwähnten Ober- und Untergrenzen der Dokumentanzahlen können bewusst so angesetzt werden, dass die Ergebnisse bspw. kein hochfrequentes Funktionsvokabular (wie Partikeln oder Konjunktionen) enthalten. Ein solches Vorgehen legt keine verborgene Sinndimension des literarischen Textes offen, sondern ist bestenfalls naiv und schlimmstenfalls manipulativ. Um derartigen methodischen Fallstricken vorzubeugen, wurde in der vorliegenden Analyse ein relationaler Ansatz gewählt, der sich mit den quantitativen Ähnlichkeitsbeziehungen gerade nicht auf intuitiv Interpretierbares bezieht, sondern auf eine statistisch geleitete Vorstrukturierung der Übersetzungskorpora. Ferner wurden – im Unterschied zu einer ausschließlich an Inhaltsfragen interessierten quantitativen Semantik (Weitin 2021: 119) – in die Berechnungen keineswegs ausschließlich Konfigurationen einbezogen, die nur i. e. S. interpretierbare Wortarten bzw. Inhaltswörter mit eigener lexikalischer Bedeutung beinhalten. Stattdessen wurden schwerpunktmäßig sogenannte MISC-Konfigurationen betrachtet, die sämtliche bzw. diverse (miscellaneous) im Text enthaltene Wortarten berücksichtigen. Dem liegt der Anspruch zugrunde, durch Implementierung digitaler Methoden auch diejenigen stilometrischen Text-eigenschaften zu berücksichtigen, die im traditionellen Close Reading tendenziell vernachlässigt werden.

Daneben beeinflusst die Anzahl der berechneten Topics die Analyseergebnisse: Je mehr Topics berechnet werden, desto enger sind diese gefasst. Mag dies bis zu einem bestimmten Grad sinnvoll sein, führt eine semantische Überspezifizierung dazu, dass die Topics keine allgemeineren Aussagen zum untersuchten Textkorpus mehr zulassen. Gleiches gilt für eine zu niedrige Topic-Anzahl, die zu semantisch weit gefassten, quasi jedes analysierte Textdokument charakterisierenden Topics führt. Welche Topic-Anzahlen in Bezug auf die untersuchten Textkorpora sinnvoll sind, lässt sich durch Experimentieren ermitteln. Ein typischer Indikator für eine zu hohe Topic-Anzahl ist die Aufspaltung bestimmter Terme auf zahlreiche unterschiedliche Topics, die sich im Topic Explorer-User Interface durch Eingabe des betreffenden Terms in die Suchmaske ermitteln lässt. Während es aufgrund semantischer Mehrdeutigkeit keine Seltenheit ist, dass sich ein Term auf drei oder vier unterschiedliche Topics aufspaltet, legt eine Aufspaltung auf eine deutlich höhere Anzahl von Topics semantische Überspezifizierung und damit eine zu hohe Topic-Anzahl nahe. Hinsichtlich der *Tonio Kröger*-Übersetzungskorpora wurden in diesem Zusammenhang Konfigurationen mit 6, 8, 10, 12 und 20 Topics ebenso untersucht wie unterschiedlich hoch angesetzte Obergrenzen der Dokumentanzahlen. Die jeweiligen Ergebnisse wurden auf ihre Konfigurationsstabilität überprüft, also unterschiedliche Konfigurationen berechnet und

miteinander abgeglichen. In Hinblick auf die im ersten Analyseschritt untersuchten quantitativen Ähnlichkeitsbeziehungen wurden also Ähnlichkeitshierarchien für unterschiedliche Konfigurationen berechnet und verglichen sowie ihr vertikales Spektrum in der Folge gedrittelt, sodass die Korporapaarungen im oberen Drittel der Hierarchie eine hohe, die im mittleren Drittel eine moderate und die im unteren Drittel eine schwach ausgeprägte quantitative Ähnlichkeit aufweisen. Anschließend wurde einem spezifischen Korporapaar nur dann eine besonders ausgeprägte Ähnlichkeitsbeziehung attestiert, wenn dieses für alle getesteten Konfigurationen im oberen Drittel der jeweiligen Ähnlichkeitshierarchie situiert war. Diese weitgehend konfigurationsstabilen Ähnlichkeitsbeziehungen bildeten den Ausgangspunkt für die weiteren Analyseschritte.

Konfigurationsstabilität ist derweil nicht das einzige Evaluationskriterium in Hinblick auf die Berechnungsergebnisse eines Topic Modeling; vielmehr sind das konsequente Reflektieren der Konfigurationsfaktoren und -stabilität einerseits und andererseits der ergänzende Abgleich mit Textkenntnis und -Interpretation die wichtigsten Instrumente zur Einordnung der Analyseergebnisse. Dementsprechend sollten Ergebnisse, die die gängige Forschungsmeinung scheinbar bestätigen, genauso hinterfragt werden, wie hiervon abweichende Ergebnisse umgekehrt nicht pauschal verworfen werden sollten. Diesbezüglich ist es sinnvoll, dass Topic Explorer in jedem Stadium der Analyse den Rückbezug auf konkrete Textbereiche und damit eine stichprobenartige Evaluierung der Analyseergebnisse ermöglicht. Ein zusätzliches Evaluationskriterium sind außerdem statistische Effekte wie z. B. der unterschiedliche Umfang der untersuchten Absatzdokumente.

Doch selbst, wenn zweimal hintereinander exakt dieselben Parameter in Topic Explorer eingegeben werden, werden innerhalb der so neu generierten Korpora keine absolut identischen Topics berechnet. Weitin spricht diesbezüglich vom „Elefant im Raum dieser Methode: Topic Models sind relativ instabil, was fast alle Untersuchungen einfach nicht thematisieren“ (Weitin 2021: 9). Diese Berechnungsinstabilität lässt sich auf das Wesen von Topic Modeling an sich zurückführen, da keine absoluten Textcharakteristiken, sondern Wahrscheinlichkeiten berechnet werden. Diese Wahrscheinlichkeiten sind letztlich nicht mehr und nicht weniger als eine dynamisch erstellte Prognose dafür, wie ein auf Basis des vorhandenen Materials neu generierter Text aussehen könnte. Dass diese Berechnungsinstabilität der Topics im Falle literarischer Texte umso ausgeprägter erscheint, dürfte dadurch bedingt sein, dass einzelne, bspw. motivisch miteinander verwobene Themenbereiche hier im Vergleich z. B. zu wissenschaftlichen Fachaufsätzen deutlich unschärfer voneinander abgegrenzt sind. Eine weitere denkbare Ursache der Topic- und Berechnungsinstabilität besteht im

(bspw. im Vergleich zum hochgradig standardisierten Sprachgebrauch wissenschaftlicher Publikationen) breiteren, Verstöße gegen alltagssprachliche Normen umfassenden Variationsspektrum der literarischen Sprache. Dass die Berechnungsinstabilität nicht zwangsläufig auf den relativ geringen Umfang der hier analysierten Übersetzungskorpora zurückgeführt werden kann, hat ein stichprobenartiges Topic Modeling des deutlich umfangreicheren Thomas Mann-Romans *Buddenbrooks* gezeigt, da die Berechnungsinstabilität auch hier derjenigen von *Tonio Kröger* vergleichbar war.

Eine von mehreren möglichen Strategien im Umgang mit dem Problem der Berechnungs- und Konfigurationsinstabilität ist die Erstellung von Kontrollmodellen bzw. der Abgleich unterschiedlicher Berechnungsiterationen mit denselben Parametern (Weitin 2021: 132–133), der im Rahmen der hier vorgestellten Analyse (insbesondere im ersten Analyseschritt) ebenfalls durchgeführt worden ist. Auf automatisierte vielfache Berechnungsiterationen, wie sie Thomas Weitin und Katharina Herget zur Stabilisierung der Berechnungsergebnisse im literaturwissenschaftlichen Topic Modeling eingesetzt haben, musste dagegen verzichtet werden, da eine entsprechende Stabilisierung nur im bei einer an Wort- bzw. Zeichenzahlen orientierten Segmentierung funktioniert (Weitin/Herget 2017: 35). Letztere lässt sich jedoch nicht mit der für den Übersetzungsvergleich notwendigen Parallelisierung der Korpora vereinbaren. Experimentell untersucht wurde in diesem Zusammenhang jedoch, ob eine Neusegmentierung der *Tonio Kröger*-Korpora anhand der umfangreicheren Kapiteldokumente (anstelle von Absatzdokumenten) mehr Konfigurations- und Berechnungsstabilität garantiert; dies war jedoch nicht der Fall.

Dabei ist zu beachten, dass vom Problem der Berechnungsinstabilität ausschließlich die Neuerstellung von Textkorpora betroffen ist. Ist ein Textkorpus einmal erstellt, sind alle sich darauf beziehenden Berechnungen – wie z. B. die im Folgenden thematisierten Ähnlichkeitshierarchien – berechnungsstabil und replizierbar. Während die Überführung der analysierten Texte in Topic Explorer-Korpora als solche mithin eine Art nicht replizierbare Momentaufnahme darstellt, sind alle darauffolgenden Berechnungen, sofern sie sich auf exakt dasselbe Korpus beziehen, stabil. Deswegen muss sich die den Zugriff auf exakt dieselben Topic Explorer-Korpora voraussetzende Replizierbarkeit der Ergebnisse von der Korporaerstellung auf eine möglichst transparente Dokumentation und Auswertung verlagern. In Anlehnung an Ramsay und Irizarry werden die Topic Explorer-Korpora folglich als Alternativtextualitäten, als konstruktive Destabilisierungen aufgefasst, die – genau wie klassische Interpretationsparaphrasen – von anderen Interpretierenden zwar ähnlich, aber nicht identisch erzeugt werden können (Ramsay 2011: 45; Irizarry 1996: 155).

2.5 Zwischenfazit zum zweiten Kapitel

Im zweiten Kapitel wurden die methodischen Grundlagen des vorgestellten Projektes in Hinblick auf Kerncharakteristika der Themenmodelle und ihre Anschlussfähigkeit an das Konzept des Algorithmic Criticism, die literaturwissenschaftliche Interpretation und den literarischen Übersetzungsvergleich dargelegt. Gezeigt werden konnte, dass Themenmodelle im Rahmen eines Algorithmic Criticism dazu eingesetzt werden können, umfangreiche Textkorpora thematisch zu strukturieren und so eine kriteriengeleitete Auswahl von kürzeren Textabschnitten für qualitative Analysen bzw. Close Readings zu treffen. In diesem Zusammenhang können parallelisierte Übersetzungskorpora erstellt und die Ähnlichkeitsbeziehungen zwischen den Gesamttexträumen quantifiziert werden. Das zum Abschluss des Methodenkapitels vorgestellte Analyseverfahren sieht vor diesem Hintergrund einen schrittweisen Rückbezug von der Ebene der Gesamttexträume auf einzelne Übersetzungstokentoken und von quantitativen zu qualitativen Analyseverfahren vor, wobei der Übergang zwischen diesen unterschiedlichen Skalierungsebenen im Sinne des Scalable Reading problematisiert und reflektiert werden muss.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

