

Universität Duisburg-Essen/ Campus Essen

Institut für Germanistik

Fakultät für Geisteswissenschaften

Seminar: Empirische Schreibforschung

Dozentin: Dr. Ulrike Behrens

Semester: Wintersemester 17/18

Evaluation eines Kriterienrasters zum Wortschatz

Ann-Christin Chmill

ann-christin.chmill@stud.uni-duisburg-essen.de

Matrikelnummer: 000000

Studiengang: LA MA Gr

Anzahl der Studiensemester: 4

Abgabedatum: 29.3.2019

Anzahl der Zeichen: 37.825

Inhalt

1) Einleitung	2
2) Theoretischer Rahmen.....	3
2.1) Zur Bewertung von Schülerleistungen.....	3
2.2) Zum Wortschatz.....	5
2.3) Zur Interraterreliabilität.....	6
3) Das Kriterienraster	8
3.1) Aufbau.....	8
3.2) Anwendung	9
4) Darlegung der Studie.....	10
4.2) Aufbau der Studie	10
4.3) Methodische Durchführung	10
5) Darstellung der Ergebnisse.....	12
6) Interpretation der Ergebnisse.....	12
7) Evaluation des Kriterienrasters.....	15
8) Fazit und Perspektive	16
9) Literatur	18
10) Anhang	19
10.1) Das Kriterienraster	19
10.2) Das Textkorpus	19
10.3) Ergebnisse der Type-Token-Ratio im Vorfeld der Kodierung.....	23
10.4) Die Kodierergebnisse	24

1) Einleitung

Vermutlich hat sich jede Schülerin und jeder Schüler schon einmal ungerecht bewertet gefühlt. Das führt nicht nur zu viel Frust, sondern kann auch die Motivation hemmen sich weiterhin anzustrengen. Bewertungen sind für SchülerInnen nicht selten intransparent und wirken daher manchmal nicht objektiv und fair. Aus diesen Erfahrungen heraus stellt sich die Frage, wie Schülerleistungen, gerechter, angemessener und präziser bewertet werden können. Um eine objektivere Bewertung für den Wortschatz als Teilbereich von Schreibkompetenz zu ermöglichen, wird in dieser Arbeit ein Kriterienraster entwickelt, das den Wortschatz in verschiedenen Teilaspekten misst und somit präzise erfassen soll. Die Evaluation dieses Kriterienrasters bildet das zentrale Thema dieser Arbeit. Am Ende der Arbeit soll die daraus resultierende Forschungsfrage beantwortet werden, inwiefern das Kriterienraster geeignet ist, um den Wortschatz von SchülerInnen zu bewerten.

Im Theorieteil dieser Arbeit geht es zunächst um den Begriff *Bewerten* und die allgemeine Problematik bei der Bewertung von Schülerleistungen (Kapitel 2.1). Um in Kapitel 3 das Kriterienraster zur Bewertung des Wortschatzes zu entwickeln, wird es in Kapitel 2.2 zunächst um den Wortschatz allgemein gehen. An dieser Stelle wird deutlich, wie sich Wortschatz überhaupt definieren lässt und wie man ihn erheben kann. Kapitel 2.3 widmet sich dem Begriff *Interraterreliabilität*. In Kapitel 3 werden dann der Aufbau und die Anwendung des Kriterienrasters thematisiert. In Kapitel 4 wird die eigentliche Studie dargelegt, denn es geht um die einzelnen Schritte, die für die Evaluation des Kriterienrasters nötig sind. Das entwickelte Kriterienraster wird zunächst praktisch angewendet, indem drei Lehramtsanwärterinnen mit diesem insgesamt 22 Schülertexte hinsichtlich des Wortschatzes bewerten. Die Bewertungen der Lehramtsanwärterinnen werden anschließend verglichen und deren Übereinstimmung berechnet. Die Ergebnisse dieser Bewertungen werden in Kapitel 5 zunächst dargestellt und in Kapitel 6 ausführlich interpretiert und hinterfragt. In Kapitel 7 findet die abschließende Evaluation des Kriterienrasters statt, indem auf Stärken und Schwächen des Kriterienrasters eingegangen wird. Welche Schlussfolgerungen sich aus der Interpretation und Evaluation ergeben, wird in Kapitel 8 thematisiert.

2) Theoretischer Rahmen

Bevor die Kriterien zur Bewertung des Wortschatzes von SchülerInnen festgelegt werden können, müssen zwei weitere Schritte berücksichtigt werden. Zum einen muss in Kapitel 2.1 zunächst der Begriff *Bewertung* definiert werden und welche Rolle dabei dem Kriterienraster zukommt. Um ein möglichst präzises Kriterienraster zu entwickeln, geht es in Kapitel 2.2 darum, wie der Wortschatz im Deutschen aufgebaut ist und was dieser beinhaltet. Nur so ist es möglich, genaue Kriterien zu formulieren, die den Maßstab für die spätere Bewertung bilden werden.

2.1) Zur Bewertung von Schülerleistungen

Der Begriff *Bewerten* wird in der Fachliteratur oft mit den Begriffen *Beurteilen* und *Benoten* in Zusammenhang gebracht und nicht selten synonym verwendet. Allgemein ist mit *Bewerten* ein Prozess gemeint, „in dem der Wert einer Sache eingeschätzt oder festgelegt wird“ (Behrens 2013, 20). In Bezug auf Leistungen von SchülerInnen lässt sich *Bewerten* als „ein kognitiver Akt bzw. ein mentaler Prozess des Einschätzens“ (Becker-Mrotzek/Böttcher 2009, 50) definieren. Dieses Einschätzen erfolgt anhand von Wertmaßstäben, die dem Bewertenden jedoch nicht bewusst sein müssen (vgl. ebd.). Die Maßstäbe, an denen sich eine Bewertung orientiert, können unterschiedliche Formen annehmen. So kann der Bewertende zwischen einer sachlichen bzw. kriterialen Bezugsnorm, einer individuellen Bezugsnorm und einer sozialen Bezugsnorm als Grundlage für seine Bewertung wählen. Während bei der Bewertung anhand einer individuellen Bezugsnorm die individuellen Lernfortschritte von SchülerInnen entscheiden, orientiert sich eine Bewertung anhand einer sozialen Bezugsnorm daran, was eine Schülerin oder ein Schüler im Vergleich zur Klasse oder einer anderen Vergleichsgruppe leistet. Die Bezugsnorm, die in dieser Arbeit relevant ist, ist dagegen die sachliche bzw. kriteriale Bezugsnorm, bei der sich die Bewertung an zuvor festgelegten Kriterien orientiert. Diese Kriterien können z.B. in einem Kriterienraster festgehalten werden. (Vgl. Behrens 2013, 22f) Ein solches Kriterienraster wird in dieser Arbeit evaluiert. In diesem Fall geht es um den Wortschatz von SchülerInnen der dritten Klasse, für dessen Bewertung im folgenden Kapitel ein Kriterienraster entwickelt wird. Hierbei müssen die Kriterien „möglichst unmissverständlich formuliert sein“ (ebd., 23), damit die Leistungen von SchülerInnen eindeutig bewertet werden können. Verschiedene Studien, wie z.B. die von Birkel und Birkel (2002), haben gezeigt, dass Lehrkräfte Schwierigkeiten damit haben, Schülertexte gleich zu bewerten, weshalb man zu der An-

nahme kommt, dass [d]ie Bewertung [...] also offensichtlich nicht nur von der Leistung des Schülers ab[hängt], sondern sehr wesentlich auch davon, von wem die Bewertung vorgenommen wird.“ (Behrens 2013, 25) In Anbetracht der stark voneinander abweichenden Bewertungen durch Lehrkräfte gehen Birkel & Birkel davon aus, dass „wenn eine Leistung schwer zu identifizieren ist, [...] der Lehrer [...] außerhalb der eigentlich zu beurteilenden Leistung nach Hinweisreizen sucht, die ihm eine Einordnung der Leistung [...] ermöglicht“ (2002, 223). Das können zum Beispiel die Handschrift, das Geschlecht oder die Rechtschreibung sein. In Fällen wie diesen kommt es zum sogenannten „Halo-Effekt“ (ebd., 224). Unterschiedliche Bewertungen ein und derselben Schülerleistungen lassen sich also psychologisch erklären. Bei der Bewertung von Schülerleistungen tritt neben dem Halo-Effekt auch oftmals ein sogenannter „Rosenthal- oder Pygmalion-Effekt“ (Behrens 2013, 25) ein, wonach die Bewertenden durch Leistungserwartungen beeinflusst werden. Ein Kind, über das man die Information hat, dass es generell eher gute Leistungen erbringt, wird demnach auch besser bewertet und umgekehrt. Ungerechte Bewertungen, die z.B. durch die eben beschriebenen Effekte entstehen können, führen zu Unmut seitens der SchülerInnen und sind alles andere als objektiv. Daher stellt sich die Frage, wie Schülertexte gerechter bewertet werden können. Eine Möglichkeit ist die „analytische bzw. kriteriale Bewertung“ (ebd., 26). Auch Birkel und Birkel kommen zu dem Schluss: „Kriterienkataloge könnten [...] Abhilfe schaffen!“ (2002, 223) Hiermit sind Bewertungsverfahren gemeint, „die verschiedene Teilaspekte einer komplexeren Leistung separat bewerten“ (Behrens 2013, 26), wobei die Gesamtbewertung der Leistung am Schluss aus den Einzelbewertungen zusammengesetzt werden kann. Ein solches Bewertungsverfahren wird in dieser Arbeit entwickelt. Die komplexe Leistung stellt hierbei der Wortschatz der SchülerInnen dar. In einem Kriterienraster werden die Teilaspekte der Leistung aufgesplittet, so dass zunächst nur diese bewertet werden. Der Vorteil eines solchen Kriterienrasters ist vor allem die „Transparenz der Bewertungsgrundlagen“ (ebd., 27). So ist es für SchülerInnen und Eltern einfacher, die Bewertungen nachzuvollziehen. Zudem können sich Lehrkräfte durch die Anwendung solcher Kriterienraster stets selbst kontrollieren und die Gefahr, Schülertexte zu subjektiv zu bewerten, sinkt.

2.2) Zum Wortschatz

Um den Begriff *Wortschatz* zu definieren, muss man zunächst den allgemeinen Wortschatz einer Sprache von dem individuellen Wortschatz eines Menschen abgrenzen. Grundsätzlich hat jede Sprache ihren eigenen Wortschatz, wobei „[d]as Lexikon einer Sprache [...] ein offenes und instabiles System [ist], in das ständig neue Wörter aufgenommen werden. Gleichzeitig werden Wörter ungebräuchlich.“ (Römer/Matzke 2010, 1) Um den Umfang des deutschen Wortschatzes zu bestimmen, muss man sich die Frage stellen, welche Wortformen und Wortbildungen gezählt werden sollen. „Ohne die Fachwörter und morphologischen Wortformen nimmt man 300000 – 500000 deutsche Wörter an“ (ebd.) Der aktive Wortschatz eines Erwachsenen umfasst dagegen 20000 bis 50000 Wörter. Vom aktiven Wortschatz lässt sich der passive bzw. rezeptive Wortschatz abgrenzen, der sogar mehr als das Doppelte vom aktiven Wortschatz umfassen kann. (Vgl. Rothweiler/Meibauer 1999, S.9) Das sogenannte „mentale Lexikon“ (ebd., 10) ist der Speicher, in dem all diese lexikalischen Einheiten gespeichert werden. Zusätzlich gibt es nach Rothweiler und Meibauer (ebd., 11) verschiedene Informationen, die mit einem Lexem gespeichert werden: Dazu zählen „semantische Repräsentationen“, also was das jeweilige Lexem bedeutet, die „lexikalische Kategorie“, also die Wortart, „[s]yntaktische Eigenschaften“, wie z.B. das Genus, „[m]orphologische Eigenschaften“, wie z.B., dass es sich um ein Kompositum handelt oder wie die Pluralform gebildet wird, und die „phonetisch-phonologische Form“, die z.B. die Silbenzahl beinhaltet.

Um eine Aussage über den Wortschatz eines Menschen machen zu können, kann man sowohl mündliche als auch schriftliche Äußerungen untersuchen. Dabei wird oftmals die sogenannte Type-Token-Ratio¹ errechnet, welche „vor allem an zusammenhängend gesprochenen oder geschriebenen Sprachproben vorzunehmen [ist].“ (Gloy 1973, 12) In diesem Kontext muss jedoch erwähnt werden, dass hierbei niemals die gesamte Kapazität eines Wortschatzes ermittelt wird, „sondern [lediglich] ein selektiver Ausschnitt des lexikalischen Wissens“ (Kauschke 1999, 136). Weiterhin darf nicht vergessen werden, dass bei schriftlichen Leistungen von SchülerInnen oft nur ein kleiner Wortschatz verwendet wird, was teilweise mit der Unkenntnis von Wörtern oder Ausdrücken zu erklären, jedoch auch auf Vermeidungsstrategien zurückzuführen ist (vgl. Kurtz 2015, 120). Viele Kinder greifen auf einen „sicheren Kernbestand gewohnter Ausdrucksweisen“ (ebd.) zurück und umgehen damit Fehlerrisiken.

¹ Im Folgenden abgekürzt als TTR

Möchte man die TTR einer Sprachprobe wie z.B. eines Textes ermitteln, werden zwei Daten erhoben: „[D]ie Gesamtzahl [...] aller definierten Einheiten, aus denen das untersuchte Sprachsample besteht, und die Anzahl [...] der verschiedenen Einheiten, die in dem Text auftreten“ (Gloy 1973, 14), welche auch Types genannt wird. Die Gesamtzahl aller definierten Einheiten, auch Token genannt, kann sich hierbei einerseits auf die Gesamtanzahl aller Wörter beziehen, womit dann eine generelle Aussage über die lexikalische Vielfalt des untersuchten Textes getroffen werden kann. Bezieht sich diese Gesamtzahl andererseits auf einzelne Wortarten, so kann die TTR auch für einzelne Wortarten ermittelt werden, weshalb man in diesen Fällen auch von einer qualitativen Analyse sprechen kann (vgl. Gloy 1973, S.16). „[D]as geschieht durch Division der Summe aller Types einer Wortart durch die Summe aller Tokens derselben Wortart.“ (ebd., S.17) Um den Wortschatz von SchülerInnen anhand eines Textes zu bewerten, wird die TTR in diesem Fall für einige Wortarten einzeln durchgeführt. Bevor eine TTR jedoch ermittelt werden kann, „ist es notwendig, daß (sic!) man sich eindeutig für formale oder semantische Kriterien entscheidet“ (ebd., 44) Diese Entscheidungen hängen davon ab, unter welcher Fragestellung ein Textkorpus untersucht werden soll (vgl. ebd.). Die formalen und semantischen Kriterien werden in Kapitel 3.2 in Form von Anwendungshinweisen erläutert.

2.3) Zur Interraterreliabilität

„Reliabilität kennzeichnet die Zuverlässigkeit oder Genauigkeit einer Messung.“ (Wirtz/Caspar 2002, 15) Die Interraterreliabilität oder auch Beurteilerreliabilität ist eine Art der Reliabilität und bezeichnet die „Korrelation der Urteile zweier beliebiger Rater“ (ebd., 18). Eine Interraterreliabilität ist nur dann erreichbar, „wenn sich alle Rater an demselben Merkmal bei der Einschätzung orientieren und dieses präzise erfassen“ (ebd., 18). Um diese Art der Reliabilität wird es in dieser Studie gehen. Als reliabel bezeichnet wird eine Beurteilung dann, „wenn andere Beurteiler mit gleichem Wissensstand zu einem ähnlichen Urteil kommen“ (ebd., 15). In diesem Kontext ist es wichtig zu betonen, dass die Reliabilität einer Messung nicht automatisch bedeutet, dass diese auch valide ist. „Schätzen zwei Urteiler eine Eigenschaft jeweils derselben Personen grundsätzlich verschieden ein, so muss mindestens einer der Rater ein nicht valides Urteil abgeben.“ (ebd., 16) Wenn jedoch mehrere Rater sehr ähnlich oder gleich urteilen, haben sie vermutlich dasselbe Merkmal erfasst (vgl. ebd., 16).

Zur Reliabilität allgemein lässt sich weiterhin sagen, dass diese kein Merkmal von Ratingskalen ist, sondern immer von dem bestimmten Gebrauch einer Ratingskala abhängt. „Spricht man von der Zuverlässigkeit einer Skala, so meint man damit natürlich die Zuverlässigkeit des Urteilsprozesses, der zur Festlegung bestimmter Skalenwerte führt“ (Wirtz/Caspar 2002, 24). Reliabilität hängt somit immer auch von den Rahmenbedingungen des Testverfahrens ab. Unterscheidet sich eine zweite Ratergruppe z.B. durch unterschiedlichen Wissensstand von der ersten Ratergruppe, so kann man die vorher gewonnenen Ergebnisse nicht einfach auf die zweite Ratergruppe übertragen, denn „[u]m Ergebnisse auf andere Beurteiler verallgemeinern zu können, muss [...] sichergestellt sein, dass sich diese neuen Beurteiler hinsichtlich der wesentlichen Merkmale [...] nicht von der bereits untersuchten Beurteilerstichprobe unterscheiden“ (ebd.). Natürlich darf sich auch die Stichprobe des gerateten Objekts nicht von der ursprünglichen unterscheiden (vgl. ebd.).

Für die Berechnung der Interraterreliabilität gibt es verschiedene Messverfahren, die unterschiedliche Vor- und Nachteile aufweisen. Den meisten Messverfahren liegen allgemeingültige Kriterien für die Einschätzung der Testergebnisse bei. Diese sollten jedoch nur als grobe Richtlinie verstanden werden. Es gilt: Für jede Untersuchung sollte ein individueller Wert festgelegt werden, ab dem die berechnete Reliabilität als zufriedenstellend gewertet werden kann. (Vgl. ebd., 25) Ergeben sich bei der Auswertung der Kodiererergebnisse Werte, die nicht als zufriedenstellend betrachtet werden können, so deutet dies auf eine mangelnde Reliabilität hin. Es gibt verschiedene Gründe, die zu einem solchen Ergebnis führen können: Zum einen entstehen Inkongruenzen dadurch, „dass die Rater nicht einig darüber sind, welches Merkmal beurteilt werden soll“ (ebd., 31), da viele zu beurteilende Merkmale als Konstrukte begriffen werden müssen. Diese sind dann durch Untermerkmale definiert und müssen von den KodiererInnen gleich verstanden werden. Zum anderen entstehen unterschiedliche Urteile von KodiererInnen, „weil sie die Ausprägung auf der zu beurteilenden Merkmalsdimension unterschiedlich einschätzen“ (ebd.). Obwohl beiden KodiererInnen hier dieselben Informationen zur Verfügung stehen, entscheiden sie sich unterschiedlich, da sie unterschiedlichen Kriterien für ihre Entscheidungsfindung nachgehen (vgl. ebd.). Unterschiedliche Ergebnisse entstehen auch dadurch, dass den KodiererInnen keine einheitlichen oder standardisierten Informationen vorliegen. Dies ist jedoch eine der entscheidenden Voraussetzungen für Reliabilität. Es müssen möglichst exakt gleiche Informationen an die KodiererInnen weitergegeben werden, da diese ihre Entscheidung ansonsten auf der Basis unterschiedlicher Informationen treffen. Auf die Beurteilung des

Wortschatzes bezogen kann es schon zu unterschiedlichen Ergebnissen kommen, wenn einer Person der Stadtteil, in dem die Schule der Kinder liegt, genannt wird und der anderen nicht. Die eine Person könnte daraus schließen, dass es sich um einen sozialen Brennpunkt handelt und macht möglicherweise Entscheidungen von dieser Information abhängig.

3) Das Kriterienraster

Das Kriterienraster, das in dieser Arbeit evaluiert wird, stellt eine Kombination aus der TTR und einem herkömmlichen Kriterienraster dar.² Wie dieses aufgebaut ist, wird in Kapitel 3.1 erläutert. Welche Anwendungshinweise hierbei beachtet werden müssen, wird in Kapitel 3.2 definiert. Hierbei geht es vorrangig um die Problematik, welche Wörter als Types bzw. Tokens gezählt werden dürfen und welche nicht.

3.1) Aufbau

Das Kriterienraster besteht aus insgesamt acht Kategorien, die auf einer dreistufigen Skala eingeschätzt werden müssen. Die neun Kategorien gliedern sich in einen allgemeinen Teil und in einen wortartenspezifischen Teil. Im allgemeinen Teil muss zunächst der Gesamteindruck des Wortschatzes eingeschätzt werden. Anschließend folgt die Einschätzung der Treffsicherheit der verwendeten Wörter. Dieser Teil ist neben der Bewertung der wortartenspezifischen Vielfalt im zweiten Teil des Kriterienrasters von großer Bedeutung, da eine TTR für die einzelnen Wortarten an sich noch keine Aussage darüber zulässt, inwiefern der angewendete Wortschatz überhaupt angemessen und im richtigen Kontext angewendet wird. Im zweiten Teil des Kriterienrasters werden dann die Wortarten Nomen, Adjektive, Verben, Adverbien, Konjunktionen und Präpositionen hinsichtlich ihrer TTR eingeschätzt. Dies geschieht auf einer dreistufigen Skala, wobei die erste Stufe gewählt werden muss, wenn die TTR für diese Wortart unter einem bestimmten Wert liegt. Die zweite Stufe muss ausgewählt werden, wenn die TTR zwischen zwei definierten Werten liegt und die dritte Stufe, wenn die TTR über einem definierten Wert liegt. Die einzelnen Werte für die Skala unterscheiden sich für jede Wortart, da die Wortarten unterschiedlich umfangreich erwartet werden. Dies hängt einerseits vom Alter der Kinder und andererseits von der Textsorte ab, die untersucht wird. In einer spannenden Erzählung würde man von einem Kind beispielsweise mehr Adjektive erwarten als in einem sachlichen Bericht, weshalb hierbei entspre-

² Vgl. Anhang 10.1

chend nicht die gleiche TTR erwartet werden kann.

Auf Beschreibungen der Merkmale wie beispielsweise „Das Kind verwendet viele Nomen.“ wird im Bereich der Wortarten verzichtet, um den KodiererInnen möglichst wenig Interpretationsspielraum zu lassen. Begriffe wie *viele*, *oft*, *häufig* etc. können unterschiedlich verstanden werden, was zu einer geringeren Beurteilerübereinstimmung führen könnte. Im allgemeinen Teil kann auf ähnliche Formulierungen jedoch nicht verzichtet werden, da man ein Merkmal wie z.B. Treffsicherheit nicht messen kann.

3.2) Anwendung

Da zum Zeitpunkt der Texterhebung die Überschrift für die Schülertexte an der Tafel vorgegeben war, muss der Umgang mit den darin vorkommenden Wörtern festgelegt werden. Einige Kinder haben die Überschrift *So feiern wir Weihnachten* übernommen, bei anderen kommt diese gar nicht oder nur abgewandelt (z.B. *Wir feiern Weihnachten so*) vor. Da alle Kinder die Möglichkeit hatten, diese Wörter der Tafel zu entnehmen, werden diese aufgegriffenen Wörter im Text gezählt. Die Überschriften der Texte sollten jedoch bei der Zählung der Types und Tokens nicht berücksichtigt werden, da diese nicht alle Kinder verwendet haben und das Ergebnis sonst leicht verfälscht werden könnte. Eine Ausnahme hierbei sind Wörter, die Kinder der Überschrift hinzugefügt haben wie z.B. bei *So feierten wir früher Weihnachten*, bei dem das Wort *früher* eigenständig ergänzt wurde. Dieses Wort wird normal mitgezählt.

Da unklar sein könnte, wie mit Tokens wie z.B. *essen*, *gegessen*, *aßen* umgegangen werden soll, wird vorgegeben, dass die Wörter, die einem Lemma angehören, als ein Type gezählt werden müssen. Wörter wie z.B. *singen* und *Gesang*, die einer Wortfamilie angehören, müssen jedoch als einzelne Types gezählt werden.

Auch der Umgang mit Hilfsverben wie z.B. *haben* oder *sein* muss einheitlich gestaltet werden. Für die Zählung der Types und Tokens gilt, dass diese Verben als Vollverb wie z.B. in *Dann haben wir Hunger*. gezählt werden. Werden diese Verben jedoch als finite Formen mit einer infiniten Form eines Vollverbs verwendet, verlieren sie ihre lexikalische Bedeutung und werden daher nicht als Types oder Tokens gezählt. Das wäre z.B. bei dem Satz *Dann haben wir meine Großeltern besucht*. der Fall. Das Verb *besucht* wird gezählt, das Hilfsverb *haben* jedoch nicht. Da Modalverben die Aussage eines Satzes verändern, wird ihnen eine lexikalische Bedeutung zugerechnet und sie werden wie andere Verben normal gezählt. Trennbare Verben wie z.B. *einkaufen* werden als ein Type bzw. Token

gezählt, auch wenn der Verbzusatz nachgestellt wird, wie z.B. bei *Mama kauft immer Käse ein*.

Da alle Vornamen von Familienangehörigen durch *[Name]* ersetzt wurden, können diese nicht unterschieden werden. Aus diesem Grund werden sie weder als Types noch als Tokens gezählt.

4) Darlegung der Studie

Im Folgenden wird erläutert, welche Schritte nötig sind, um schlussendlich die Evaluation des Kriterienrasters durchzuführen. Hierbei geht es zunächst um den allgemeinen Aufbau der Studie und im weiteren Verlauf um die methodische Durchführung, wobei auf die Datenaufbereitung, die Kodierschulung, die Datenerhebung und die Auswertung eingegangen wird.

4.2) Aufbau der Studie

Um zu testen, inwiefern das entwickelte Kriterienraster für die Bewertung des Wortschatzes von SchülerInnen geeignet ist, wurde dieses drei Lehramtsanwärterinnen vorgelegt, die anhand dieses Kriterienrasters insgesamt 22 Texte von SchülerInnen des dritten Schuljahres bewertet haben. Alle Kodiererinnen haben bereits einen Masterabschluss und befinden sich im Referendariat an verschiedenen Grundschulen. Da sie beruflich immer wieder mit der Bewertung von Schülertexten konfrontiert werden und sich mit den verschiedenen Fachtermini bezüglich des Wortschatzes und der Wortarten auskennen, eignen sie sich für die Rolle der Kodiererinnen.

4.3) Methodische Durchführung

Datenaufbereitung

Um sicherzustellen, dass die Kodiererinnen nichts anderes als den Wortschatz der SchülerInnen bewerten, wurde das Textkorpus im Vorfeld transkribiert und orthografisch korrigiert.³ So besteht nicht die Gefahr, dass die Kodiererinnen durch Orthografiefehler oder Handschrift, im Sinne des Halo-Effektes, beeinflusst werden. Die Gefahr der Beeinflussung durch die Textlänge besteht jedoch weiterhin, da die vorliegenden Schülertexte unterschiedlich lang sind. Um den Rosenthaleffekt zu vermeiden, wurden den Kodiererinnen keinerlei Informationen über die SchreiberInnen der Texte mitgeteilt. Einige Kinder haben

³ Vgl. Anhang 10.3

ihre eigenen Namen in den Texten erwähnt. Diese wurden durch *[Name]* ersetzt, damit das Geschlecht des Kindes die Kodiererinnen nicht beeinflusst. Da nicht auszuschließen ist, dass die Kodiererinnen hinter einigen Namen von den Kindern oder auch deren Familienmitgliedern einen Migrationshintergrund vermuten könnten, wurden auch alle weiteren Namen durch *[Name]* ersetzt.

Kodierschulung

Um zu gewährleisten, dass die Kodiererinnen die zu beurteilenden Merkmale exakt erfassen, wurden diese in einer vorangegangenen Kodierschulung erläutert. Hierbei wurden in der Gruppe mehrere Wortartenbestimmungen durchgeführt und diskutiert. Die drei Kodiererinnen wurden zudem auf die Anwendungshinweise für das Kriterienraster hingewiesen. Beispiele, die nicht Teil der Schülertexte waren, wurden hierzu erklärt. Fragen der Kodiererinnen bezüglich der Kinder, der Schule, der Lage der Schule, der Verteilung der Geschlechter etc. wurden nicht beantwortet.

Datenerhebung

Um die Daten für die Berechnung der Interraterreliabilität zu erheben, wurde den Kodiererinnen das Textkorpus von 22 Schülertexten in Papierform und in digitaler Form zur Verfügung gestellt. Weiterhin wurde den Kodiererinnen ein Stift, ein Lineal und verschiedenfarbige Textmarker bereitgestellt, um sich gegebenenfalls Wörter oder Sätze in den Texten zu markieren. Zudem lagen den Kodiererinnen jeweils die in Kapitel 3.2) definierten Anwendungshinweise in Papierform vor, sodass sie diese noch einmal lesen konnten, sollten sich Unsicherheiten im Umgang mit bestimmten Wörtern ergeben. Die Ergebnisse der Beurteilung wurden von den Kodiererinnen in Form einer Tabelle festgehalten.

Auswertung

Um die Interraterreliabilität zu berechnen, gibt es verschiedene Messverfahren. Unter Berücksichtigung der erhobenen Daten, des Skalenniveaus und der Anzahl von drei Kodiererinnen wurde die Berechnung von Krippendorff's Alpha ausgewählt. Die Berechnung des Reliabilitätskoeffizienten nach Krippendorff hat den Vorteil, dass diese bei einer beliebigen Anzahl von Kodierenden, einem beliebigen Skalenniveau und unabhängig von der Stichprobengröße angewandt werden kann. Zudem wird Krippendorff's Alpha nicht von fehlenden Werten oder der Anzahl der Ausprägungen beeinflusst. (Vgl. Hayes/Krippendorff 2007, 82) Bei der Berechnung von Krippendorff's Alpha erhält man einen Wert zwischen 0 und 1, wobei 0 für keine oder eine zufällige Übereinstimmung bei den Kodiererergebnissen steht und 1 eine absolute Übereinstimmung. Geht man nach Krip-

pendorff, so darf keine Variable mit einem Wert unter .67 akzeptiert werden. Variablen mit einem Wert von $\alpha \geq .8$ können dagegen als gut bezeichnet werden. (Vgl. Krippendorff 1980, 146).

5) Darstellung der Ergebnisse

Merkmal	K (α)
Teil I	
1. Gesamteindruck	.47
2. Treffsicherheit	.66
Teil II	
3. Nomen	1
4. Adjektive	.86
5. Verben	.94
6. Adverbien	.98
7. Konjunktionen	.89
8. Präpositionen	.97

Im Folgenden werden die Werte für Krippendorff's Alpha erläutert, die sich aus den einzelnen Kodierergebnissen⁴ der drei Kodiererinnen ergeben haben.

Im ersten Teil des Kriterienrasters wurden keine zufriedenstellenden Werte erreicht. Insgesamt liegt für das Merkmal Gesamteindruck mit $K(\alpha)=.47$ die niedrigste Interraterreliabilität vor. In nur neun von 22 Fällen urteilten die Kodiererinnen gleich. Eine mit $K(\alpha)=.66$ etwas

höhere Interraterreliabilität wird für das Merkmal Treffsicherheit erreicht.

Für den zweiten Teil des Kriterienrasters liegen die errechneten Werte alle in einem guten Bereich. Die größte Interraterreliabilität wird für das Merkmal Nomen erreicht, bei der alle Kodierergebnisse identisch sind und somit der Wert $K(\alpha)=1$ erreicht wird. Besonders hoch ist auch die Interraterreliabilität für die Merkmale Verben, Adverbien und Präpositionen, bei denen für Krippendorff's Alpha Werte im Bereich zwischen .94 und .98 erlangt wurden, welche sehr nah an einer absoluten Übereinstimmung liegen. Ebenfalls im guten Bereich liegen die Werte für die Variablen Adjektive ($K(\alpha)=.86$) und Konjunktionen ($K(\alpha)=.89$). Die aus den Kodierergebnissen errechneten Werte für Krippendorff's Alpha zeigen also zunächst einmal, dass es im zweiten Teil des Kriterienrasters eine gute bis absolute Übereinstimmung zwischen den Kodiererinnen gibt und sich dieser Teil stark vom ersten Teil unterscheidet, bei dem keine zufriedenstellenden Ergebnisse erzielt wurden. Im Folgenden werden die Ergebnisse einzeln interpretiert und hinterfragt.

6) Interpretation der Ergebnisse

Die geringe Interraterreliabilität für das Merkmal Gesamteindruck lässt sich vermutlich auf die Formulierung der drei Ausprägungen zurückführen. Ausdrücke wie *wenig* oder *sehr* wurden von den drei Kodiererinnen ggf. abweichend interpretiert. Auch das Adjektiv *um-*

⁴ Vgl. Anhang 10.4)

fangreich kann unterschiedlich verstanden werden. Dass der Wert für $K(\alpha)$ für das Merkmal Treffsicherheit höher ausgefallen ist, könnte auf die besseren Formulierungen der drei Ausprägungen zurückzuführen sein. Hier lag ein weniger großer Interpretationsspielraum vor, da die Ausprägungen 1 und 3 mit *nicht angemessen* und *immer angemessen* leichter zu erfassen sind. Die Ausprägung 2 mit der Formulierung *teilweise angemessen* war in diesem Fall das „Sammelbecken“ für alle Texte, die ein oder zwei dem Kontext nicht angemessene Wörter oder Formulierungen enthielten. Schwierig wäre es geworden, wenn in einigen Texten viele unangemessene Wörter oder Formulierungen verwendet worden wären, da nicht deutlich ist, ab wann ein Text bezüglich der Treffsicherheit der Ausprägung 1 zugeordnet werden muss.

Für den zweiten Teil lässt sich anders auf die Nichtübereinstimmungen zwischen den Kodiererinnen blicken. Denn da die Kodiererinnen bei der Einschätzung auf der dreistufigen Skala keinen Interpretationsspielraum hatten und sich im Hinblick auf ihre ermittelte TTR für die jeweilige Wortart entscheiden mussten, sind mangelnde Übereinstimmungen am ehesten darauf zurückzuführen, dass den Kodiererinnen Fehler bei der Zählung der Tokens und Types unterlaufen sind und sie das Merkmal nicht richtig erfasst haben. Fehler können zum einen durch das einfache Verzählen oder eine falsche Wortartenbestimmung entstehen.

Eine unterschiedliche Einschätzung auf der dreistufigen Skala findet sich bei Text 16 für die Variable Adjektive. Während die erste Kodiererin (K1) und die dritte Kodiererin (K3) den Text auf der Skala bei 1 eingeordnet haben, hat die zweite Kodiererin (K2) den Text bei 3 eingeordnet. Ein Blick in die Auflistung der gezählten Tokens und Types zeigt, dass K2 das Adjektiv *kurz* als Token und Type gezählt und somit eine $TTR=1$ errechnet hat, die der Ausprägung 3 entspricht. K1 und K3 dagegen haben das einzige Adjektiv im Text entweder bewusst oder unbewusst nicht gezählt und haben daher eine $TTR=0$ errechnet und diese entsprechend der Ausprägung 1 zugeordnet.

Uneinig waren sich die Kodiererinnen auch bei Text 19. Hier haben K2 und K3 den Text hinsichtlich der TTR für Adjektive der Ausprägung 1 zugeordnet, während K1 sich für die Ausprägung 3 entschieden hat. Dies ist entstanden, da K2 und K3 das Wort *gemeinsam* nicht als Adjektiv identifiziert haben, während K1 dieses in ihre Zählung aufgenommen hat. Demnach haben K2 und K3 eine $TTR=0$ errechnet, die der Ausprägung 1 entspricht, während K1 eine $TTR=1$ errechnet hat und diese der Ausprägung 3 zugeordnet hat.

Das Merkmal Verben wurde von den Kodiererinnen nur bei Text 20 unterschiedlich einge-

schätzt. Während K1 und K2 den Text der Ausprägung 2 zugeordnet haben, hat sich K3 für Ausprägung 3 entschieden. Uneinig waren sich die Kodiererinnen hier bezüglich der Zählung des Verbs *sind*, also dem Lemma sein. Das Verb wird in dem Satz *Wenn wir dann im Zimmer sind, versteckt Mama die Geschenke unter dem Tannenbaum.* als Vollverb eingesetzt. Da es in den meisten Texten als Hilfsverb fungiert und daher nicht gezählt werden darf, scheinen sich K1 und K2 getäuscht zu haben, da diese das Verb als Token und Type gezählt haben und somit eine TTR=0,8 erreichen, die noch knapp der Ausprägung 2 entspricht, dass für TTR zwischen 0,7 und 0,8 vorgesehen ist. K3 dagegen, die das Verb korrekt gezählt hat, hat eine TTR=0,81 errechnet und damit die Ausprägung 3 ausgewählt.

Das Merkmal Adverbien wurde ebenfalls nur einmal unterschiedlich eingeschätzt. Grund dafür ist die Wortkombination *und so weiter* aus Text 7, das von K2 als eine Konjunktion gezählt und von K1 und K3 in die drei Bestandteile *und* (Konjunktion), *so* (Adverb) und *weiter* (Adverb) zerlegt und entsprechend gezählt wurde. Insgesamt herrschte auch Uneinigkeit darüber, ob es sich bei *und so weiter* nicht sogar um einen Phraseologismus handle. K2 erreicht durch ihre Zählung eine TTR, die der Ausprägung 2 entspricht, während die TTR von K1 und K3 der Ausprägung 3 entspricht. Diese unterschiedlichen Ansichten bezüglich der richtigen Zählung führen dazu, dass sich auch die Einschätzungen für das Merkmal Konjunktionen für den Text 7 unterscheiden. K1 und K3 haben zwei Adverbien mehr gezählt als K2 und aufgrund der dadurch öfter vorkommenden Konjunktion *und* eine geringere TTR, die der Ausprägung 1 entspricht. K2, die *und so weiter* als eigenständige Konjunktion gezählt hat, hat demnach eine höhere TTR errechnet, die der Ausprägung 3 entspricht.

Bei dem Merkmal Präpositionen gab es lediglich eine Nichtübereinstimmung. Diese ist auf das Wort *seit* in Text 19 zurückzuführen, das von K2 als Konjunktion und von K1 und K3 als Präposition gezählt wurde. Tatsächlich kann das Wort *seit* sowohl Konjunktion als auch Präposition sein, weshalb in diesem Fall der Kontext entscheidend ist. Der zugehörige Satz lautet: *Seit zwei Jahren bin ich auch im Chor.* Das Wort *seit* „dient [hier] zur Angabe des Zeitpunkts, zu dem, oder der Zeitspanne, bei deren Beginn ein noch anhaltender Zustand, Vorgang begonnen hat“ (Duden online), weshalb es sich um eine Präposition handelt.

Dass die Kodiererinnen die meisten Texte in Bezug auf die einzelnen Merkmale identisch eingeschätzt haben, bedeutet nicht, dass sie dieselben Werte für die TTR ermittelt haben. Da es auf der dreistufigen Skala einen gewissen Spielraum für die Einschätzung der TTR gibt, ist nicht auszuschließen, dass Fehler bei der Wortartenbestimmung oder Zählung sei-

tens der Kodiererinnen unentdeckt bleiben. Hätte die Einschätzung anhand einer zehnstufigen Skala durchgeführt werden müssen, so wären diese schon eher aufgefallen und die Werte für $K(\alpha)$ wären insgesamt wohlmöglich geringer ausgefallen.

7) Evaluation des Kriterienrasters

Die Durchführung, Auswertung und Interpretation der Kodierergebnisse machen deutlich, dass das vorliegende Kriterienraster sowohl Stärken als auch Schwächen aufweist, die im Folgenden nacheinander erläutert werden:

Da im zweiten Teil des Kriterienrasters auf Formulierungen verzichtet wird, die viel Interpretationsspielraum zulassen, ist die Einschätzung auf der dreistufigen Skala nach der Ermittlung der TTR eindeutig. Sofern die KodiererInnen das Merkmal richtig erfassen, lässt sich eine hohe Interraterreliabilität erwarten. Dies ist auch im Fall der drei Kodiererinnen gelungen, die im zweiten Teil des Kriterienrasters sehr gute Werte erreicht haben. Die Nutzung der TTR für das Kriterienraster hat den Vorteil, dass sowohl die Vielfalt als auch die Verwendungshäufigkeit bestimmter Wortarten erfasst werden können. Diese beiden Aspekte machen die Stärken des Kriterienrasters deutlich.

Die Schwäche des Kriterienrasters im ersten Teil liegt in der Formulierung der Ausprägungen und in dem zugehörigen Teil der Kodierschulung. Die Ausprägungen wurden unterschiedlich interpretiert und die Schülertexte entsprechend verschieden bewertet.

Zu den Schwächen des Kriterienrasters im zweiten Teil zählt die Abhängigkeit der TTR von der Textlänge. Da mit zunehmender Textlänge auch die Wahrscheinlichkeit zunimmt, dass sich Wörter im Text wiederholen, sinken die TTR-Werte bei längeren Texten, sodass diese von den KodiererInnen schlechter bewertet werden. Dies entspricht aber nicht unbedingt der eigentlichen Qualität hinsichtlich des Wortschatzes und muss somit überdacht werden. Vergleichbar wären die TTR-Werte nur, wenn es sich um ungefähr gleich lange Schülertexte handelt. Da dies in einem Klassensatz an Texten aufgrund der heterogenen Schülerschaft jedoch selten der Fall sein wird, ist die TTR als Maß in Kombination mit einem Kriterienraster eher ungeeignet.

Ein weiterer Nachteil des Kriterienrasters ist die Anwendbarkeit auf unterschiedliche Textsorten. Je nach Textsorte lassen sich manche Wortarten vermehrt erwarten, weshalb die festgelegten TTR-Werte für die drei Ausprägungen jeweils angepasst werden müssten.

Die wohl größte Schwäche des Kriterienrasters ist die Anwendung. Zum einen benötigt man sehr viel Zeit, um zunächst alle Tokens und Types zu zählen und die entsprechende

TTR zu ermitteln. Zum anderen ist die erfolgreiche Anwendung des Kriterienrasters stark abhängig von den Vorkenntnissen der KodiererInnen. Einige Wörter sind nicht eindeutig einer Wortart zuzuordnen, was eine intensive Auseinandersetzung mit den Schülertexten erfordert und daher sehr zeitintensiv ist.

8) Fazit und Perspektive

Die detaillierte Beschreibung der Stärken und Schwächen des Kriterienrasters zeigt eindeutig, dass die Schwächen hierbei überwiegen. Daher ist das Kriterienraster aus den genannten Gründen nicht zufriedenstellend.

Für den ersten Teil, in dem eher niedrigere Werte erreicht wurden, müsste die Kodierschulung mehr darauf ausgerichtet sein, Formulierungen wie *umfangreich* mit Inhalt zu füllen, damit die KodiererInnen eine einheitliche Vorstellung davon bekommen, wann ein Wortschatz als umfangreich betrachtet werden kann und wann eben nicht. Zudem sollte bezüglich des Merkmals Treffsicherheit im Vorhinein festgelegt werden, wann genau eine Formulierung als unangemessen gilt und wann nicht. Grammatikalische Fehler dürfen nicht mit kontextunangemessenen Formulierungen oder Wörtern verwechselt werden.

Um das Problem mit der Textlänge und den davon abhängenden TTR-Werten zu lösen, könnte man andere Maße nutzen, die die Textlänge als Variable eliminieren. Hierfür macht Mathiebe (2018, 72-73) einige Vorschläge. Für die meisten Maße benötigt man jedoch Schülertexte mit mindestens hundert Wörtern. Bezogen auf das vorliegende Textkorpus hätte man ein solches Maß lediglich bei vier der Schülertexte anwenden können, da alle anderen zu kurz sind. Bei einem Textkorpus, dessen Schülertexte länger sind, könnte man jedoch ein solches Maß anwenden.

Ein weiterer Nachteil ist die Anwendbarkeit des Kriterienrasters auf andere Textsorten. Eine Lösung wäre, ein Repertoire über verschiedene Kriterienraster anzulegen, die für eine bestimmte Textsorte geeignet sind. Diskussionswürdig wäre in diesem Fall noch, wie die TTR-Werte für die einzelnen Ausprägungen definiert sein müssen.

Was sich nicht beheben lässt, ist der Aufwand, den das Kriterienraster mit sich bringt. Durch die zeitintensive Wortartenbestimmung, die Zählung der Tokens und Types und die Errechnung der verschiedenen TTR für sämtliche Wortarten ist das Kriterienraster für den Schulalltag ungeeignet. Den Anspruch, als Kriterienraster für die Beurteilung von Schülertexten im Schulalltag zu fungieren, stellt das Kriterienraster jedoch nicht. Es geht vorrangig darum zu evaluieren, inwiefern der Wortschatz durch ein Kriterienraster beurteilt wer-

den kann. Die Antwort darauf ist komplex. Jedes Kriterienraster bringt Vor- und Nachteile mit sich. Insgesamt scheint jedoch ein Kriterienraster besser geeignet, dass die zu erfassenden Merkmale eindeutig definiert, jedoch auf aufwendige Berechnungen von beispielsweise der TTR verzichtet.

9) Literatur

Sekundärliteratur

- BEHRENS, U. (2013): Bewerten. In: Björn Rothstein, Claudia Müller (Hrsg.): Kernbegriffe der Sprachdidaktik Deutsch. Ein Handbuch. Baltmannsweiler, 20-28.
- BECKER-MROTZEK, M. / BÖTTCHER, I. (2009): Texte bearbeiten, bewerten und benoten. *Lehrer-Bücherei Grundschule*. Berlin.
- BIRKEL, C. / BIRKEL, P. (2002): Wie einig sind sich Lehrer bei der Aufsatzbeurteilung? Eine Replikationsstudie zur Untersuchung von Rudolf Weiss. In: *Psychologie in Erziehung und Unterricht* 49, 219-224.
- GLOY, K. (1973): Die Type-Token-Ratio als Instrument der quantitativen Linguistik. Ein Forschungsbericht. Konstanz.
- HAYES, A. F. / KRIPPENDORFF, K. (2007): Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77-89.
- KAUSCHKE, C. (1999): Früher Wortschatzerwerb im Deutschen: Eine Empirische Studie zum Entwicklungsverlauf und zur Komposition des kindlichen Lexikons. In: Jörg Meibauer, Monika Rothweiler (Hgg.): *Das Lexikon im Spracherwerb*. Tübingen, 128-156.
- KRIPPENDORFF, K. (1980): *Content analysis. An introduction to its methodology*. Beverly Hills: Sage.
- KURTZ, G. (2015): „Warum ist es für uns mühsam? Das ist eigentlich für ihr.“ – Kinder erarbeiten sich Dimensionen lexikalischer Einheiten im Dialog. In: Katharina Kuhs, Stephan Merten (Hgg.): *Arbeiten am Wortschatz. Sprache und Sprachgebrauch untersuchen. Koblenz-Landauer Studien zu Geistes-, Kultur- und Bildungswissenschaften* Bd. 14, 117-136.
- MATHIEBE, M. (2018): *Wortschatz und Schreibkompetenz. Bildungssprachliche Mittel in Schülertexten der Sekundarstufe I*, Münster.
- Römer, C. / Matzke, B. (2010): *Der deutsche Wortschatz. Struktur, Regeln und Merkmale*. Tübingen.
- ROTHWEILER, M. / MEIBAUER, J. (1999): *Das Lexikon im Spracherwerb – Ein Überblick*. In: Jörg Meibauer, Monika Rothweiler (Hgg.): *Das Lexikon im Spracherwerb*. Tübingen, 9-31.
- WIRTZ, M. / CASPAR, F. (2002): Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen.

Internetquellen:

„seit“ auf Duden online https://www.duden.de/rechtschreibung/seit_von_da_an_spaeter_als

10) Anhang

10.1) Das Kriterienraster

Kriterienraster zur Beurteilung des Wortschatzes		
Teil I: Allgemeine Kriterien		
1. Gesamteindruck des Wortschatzes	Zu bewerten ist der Gesamteindruck des Wortschatzes bezogen auf dessen Umfang.	(1) Der verwendete Wortschatz ist nicht umfangreich. (2) Der verwendete Wortschatz ist wenig umfangreich. (3) Der verwendete Wortschatz ist sehr umfangreich.
2. Treffsicherheit	Zu bewerten ist die kontextabhängige Angemessenheit der gewählten Wörter.	(1) Die verwendeten Wörter sind oft unangemessen. (2) Die meisten Wörter werden treffsicher verwendet. (3) Alle Wörter werden treffsicher verwendet.
Teil II: Wortartenspezifische Vielfalt		
3. Nomen	Die Anzahl der verschiedenen Wörter einer Wortart muss durch die Anzahl aller Wörter einer Wortart dividiert werden, um die TTR zu erhalten.	(1) Die TTR liegt unter einem Wert von 0,7 (2) Die TTR liegt in einem Bereich zwischen 0,7 und 0,8 (3) Die TTR liegt über einem Wert von 0,8
4. Adjektive		(1) Die TTR liegt unter einem Wert von 0,5 (2) Die TTR liegt in einem Bereich zwischen 0,5 und 0,8 (3) Die TTR liegt über einem Wert von 0,8
5. Verben		(1) Die TTR liegt unter einem Wert von 0,7 (2) Die TTR liegt in einem Bereich zwischen 0,7 und 0,8 (3) Die TTR liegt über einem Wert von 0,8
6. Adverbien		(1) Die TTR liegt unter einem Wert von 0,7 (2) Die TTR liegt in einem Bereich zwischen 0,7 und 0,8 (3) Die TTR liegt über einem Wert von 0,8
7. Konjunktionen		(1) Die TTR liegt unter einem Wert von 0,3 (2) Die TTR liegt in einem Bereich zwischen 0,3 und 0,5 (3) Die TTR liegt über einem Wert von 0,5
8. Präpositionen		(1) Die TTR liegt unter einem Wert von 0,7 (2) Die TTR liegt in einem Bereich zwischen 0,7 und 0,8 (3) Die TTR liegt über einem Wert von 0,8

10.2) Das Textkorpus

Schülertext 1

So feiern wir Weihnachten // Wir stellen unsere Tanne auf, die / wir schon sieben Jahre haben. Dann / gehen wir zu meinem Onkel und Tante / und feiern bei ihnen. Manchmal bei / uns. Dann essen wir da und machen die / Geschenke auf und Wichtelgeschenke / auf. Und wir gehen nach Hause.

Schülertext 2

So feiern wir Weihnachten // Bei uns darf man erst ab dem / ersten Advent Weihnachtssachen / kaufen. Am vierten Advent stellten / wir den Tannenbaum auf und ich durfte ihn schmücken. Am zweiten / Advent ging ich in die Kirche / und auch am dritten, zweiten und ersten. / Auch an Heiligabend ging ich / in die Kirche. Am Heiligabend / essen wir Raclette.

Schülertext 3

Wir feiern Weihnachten mit Gulasch, / einen kleinen weißen Tannenbaum. / Es kommen auch meine Oma, Opa, Uroma, / Uropa, erste Tante, zweite Tante, dritte Tante zu Weihnachten / zu uns. Die Geschenke packen wir / am Ende aus.

Schülertext 4

So feiern wir Weihnachten // Jeder macht ein Geschenk für / jeden. Und dann ruft die Mutter / uns und wir verteilen die / Geschenke. Später macht unser Papa / den Kamin an und wir singen / Weihnachtslieder. Dann dürfen wir / noch ein bisschen mit / unseren neuen Spielen spielen.

Schülertext 5

Ich feier so Weihnachten. Ich und meine / Familie gehen in ein Zimmer und / warten bis das Glöckchen klingelt. Dann / gehen wir in das Wohnzimmer und / packen die Geschenke aus. So feier / ich Weihnachten. Am nächsten Tag / frühstücken wir immer mit anderen / Familien in einem Restaurant. Da gibt es / sogar einen Schokobrunnen.

Schülertext 6

So feiern wir Weihnachten // Ich und meine Mutter gehen mit meiner / Oma Essen. Und wir haben eine / Weihnachtsfeier gemacht mit den Freunden und / Freundinnen von meiner Mutter. Da haben / wir gegessen und getrunken. Und an / Heiligabend schmücken wir unsere Wohnung / und stellen einen Weihnachtsbaum auf.

Schülertext 7

Morgens bereiten meine Eltern / alles vor: Das Essen, den Schmuck und / so weiter. Abends ziehen wir uns dann / an und gehen in die Kirche. Nach der / Kirche gehen wir dann nach Hause. / Anschließend kocht mein Papa / Vorspeise, Nachspeise und Nachtisch. / Nach dem Essen muss ich in mein Zimmer gehen und warten. Solange / schließen meine Eltern alle Türen. Ungefähr zehn Minuten später holen / mich meine Eltern wieder. Und dann / packen wir die Geschenke aus.

Schülertext 8

So feiern wir Weihnachten // Ich und meine Schwester und meine andere Schwester sind in / mein Zimmer gegangen. Dann haben wir solange / gespielt, bis unsere Mama uns gerufen hat und dann durften wir / kommen. Und als erstes mussten wir ein Lied singen. / Meine große Schwester auf dem Saxophon und ich und meine / kleine Schwester mussten singen und wenn das Lied zu Ende / war, durften wir die Geschenke auspacken. / Einen Tag später das Ganze was wir bekommen haben spielen / da kommen alle Verwandten und beschenken uns und dann gibt es / immer Essen.

Schülertext 9

So feiern wir Weihnachten // Meine Familie und ich gehen in die Weihnachtsmesse. / Dann fahren wir nach Hause. / Meine Mama klingelt mit einer Glocke. / Dann dürfen wir ins Wohnzimmer / kommen. Nun ist bei uns dann Bescherung. / Jetzt packen wir die Geschenke aus. / Danach essen wir leckeres Essen. / Am ersten Weihnachtstag gehen wir nach / Bottrop. Dann feiern wir mit einem / Teil meiner Familie. Am zweiten Weihnachtstag / feiere ich mit dem anderen Teil meiner / Familie.

Schülertext 10

Wir feiern Weihnachten zu Hause. Jedes / Jahr haben wir einen schön geschmückten / Weihnachtsbaum bei uns stehen. Wenn / alle da sind, gibt es eine Suppe zu essen. / Dann sagt Mama immer: „Jetzt werden / Geschenke ausgepackt.“ An den zwei / anderen Weihnachtstagen machen wir nichts / Besonderes mehr.

Schülertext 11

So feiern wir Weihnachten // Wenn wir bei meiner Tante, Cousine und / Onkel waren, haben wir eine Glocke genommen. / Dann mussten meine Cousine und ich / rausgehen bis mein Papa, Mama, Tante und Onkel / riefen: „Kommt herein!“ Dann haben wir / ausgepackt. Am ersten Weihnachtsfeiertag essen / wir die Essensreste auf: z.B. Ente, Nudelsalat und / Spaghetti. Und sechs Tage nach Weihnachten / habe ich Geburtstag.

Schülertext 12

So feierten wir früher Weihnachten // Wir sind aus dem Zimmer geflitzt / und wollten ins Wohnzimmer. Aber die Tür war abgeschlossen / bis 18.00 Uhr mussten wir warten. Danach haben wir / eine Gans gegessen. Dann mussten wir ins Bett.

Schülertext 13

So feiern wir Weihnachten // Jedes Jahr kommt unsere ganze Familie. Wir essen / immer etwas und dann müssen Linus und ich ins / Kinderzimmer. Wenn unser kleines Glöckchen klingelt, / stürmen wir aus dem Zimmer direkt auf den schön / geschmückten Tannenbaum zu und dürfen die Geschenke / auspacken. Ich packe immer alles aus und danach / spiele ich mit den Sachen. Am ersten Weihnachtstag / gehen wir mit Oma und meinem Stiefopa essen. Am / zweiten Weihnachtstag fahre ich immer / mit meinem Bruder und meinem Vater in den Skiurlaub. / Das macht viel Spaß.

Schülertext 14

So feiern wir Weihnachten // Ich, [Name] und [Name] sitzen in / [Name] Zimmer, spielen ein Spiel / und warten bis es klingelt. / Dann klingelt es und Mama sagt: „Kommt herein.“ / Nun singen wir und dann dürfen / wir die Geschenke auspacken. / Danach essen wir und probieren / die Geschenke aus. / Am ersten Weihnachtstag gehen / wir mit Papas Familie essen. / Und am zweiten Weihnachtstag gehen wir mit Mamas Familie essen.

Schülertext 15

Wir kriegen ein Weihnachtsgeschenk / und meine Mama versteckt das / Geschenk. Dann suchen wir das Geschenk. / Wenn wir es gefunden haben, dürfen / wir es auspacken. Wir dürfen mit / dem Spielzeug spielen. Danach holt / meine Mama mich zum Essen. Zum / Nachtschisch gibt es ein Lebkuchenhaus.

Schülertext 16

So feiern wir Weihnachten // Wir verstecken unsere Geschenke immer / in einer Schublade und sagen es nur / unserem Vater. Kurz bevor wir zur / Weihnachtsmesse gehen, sagt mein Vater: / „Ich muss noch duschen, geht schon / einmal vor.“ Dann gehen die anderen / zur Kirche. Mein Vater duscht sich aber / nicht, sondern er legt die Geschenke / in das Wohnzimmer vor das Klavier. / Anschließend macht er die Küchentür / auf und legt Sterne auf den Boden. / Diese führen zu den Geschenken und / dann zu der Haustür. Nun geht er / auch zu der Kirche. Wenn wir wieder / zu Hause sind, singen wir Lieder. / Dann geht Papa in die Küche und sagt: / „Das Christkind war hier!“ Am Ende / der Lieder packen wir die Geschenke aus. / Am zweiten Weihnachtstag kommen mein Opa / und meine Tante.

Schülertext 17

Bei uns geht Weihnachten so: Wir / holen den Weihnachtsbaum. Danach / schmücken wir ihn. Wenn der / vierundzwanzigste Dezember gekommen / ist, dann legt Mama die Geschenke / darunter. Am Abend essen wir Hähnchen, / Rotkohl und Kartoffeln. Nach dem Abendessen / dürfen wir dann die Geschenke / auspacken. Am ersten / Weihnachtstag

kriegen wir die Geschenke / von Oma und Opa. Danach essen / wir noch die Reste von gestern auf. / Am zweiten Weihnachtstag kocht / Mama Partysuppe. Nach der Partysuppe / räumen wir ein bisschen auf und / ruhen uns aus.

Schülertext 18

Bei uns ist die Wohnzimmertür am / Morgen des vierundzwanzigsten Dezember immer / abgeschlossen. Um 16.00 Uhr gehen / wir dann in die Kirche. Wenn wir / wieder nach Hause kommen, spielt / Mama auf dem Klavier, [Name] / und [Name] auf der Blockflöte, ich / auf der Querflöte und Oma und / Papa singen Weihnachtslieder. Nun / bimmelt immer eine Glocke und / wir dürfen rein ins Wohnzimmer. / Dort leuchtet der Tannenbaum und / die Geschenke liegen auf den Sofas. / Am ersten Weihnachtstag spielen wir / mit den Geschenken. Das ist immer / richtig schön und am zweiten / Weihnachtstag fahren wir immer zu Oma. / Wie wird Weihnachten wohl dieses / Jahr werden? Genauso?

Schülertext 19

So feiern wir Weihnachten // Jedes Jahr fängt Weihnachten / bei uns in der Kirche an. / Seit zwei Jahren bin ich auch / im Chor. Danach gehen wir zu / Oma [Name]. Mamas zwei / Schwestern kommen mit ihrer / Familie auch. Wenn alle da sind / müssen wir warten bis Oma / aus dem Wohnzimmer „Kommt / rein!“ ruft. Auf den Geschenken / stehen Namen damit wir keine Geschenke / vertauschen. Nachdem alles / ausgepackt ist, essen alle gemeinsam. / Am zweiten Weihnachtstag gehen wir / mit Papas Bruder und seiner / Familie zu Oma [Name]. Dort / stehen die Geschenke schon auf / dem Wohnzimmertisch. / Wenn alle ausgepackt sind, essen / wir erst etwas. Meistens bleiben wir / noch ein wenig dort. Am letzten / Weihnachtstag feiern ich und meine / Familie immer zu Hause. Was wir / den Rest des Tages machen ist / immer anders.

Schülertext 20

So feiern wir Weihnachten // Wir feiern zu Hause. / Zuerst gehen wir in die Kirche. / Eine Weile später müssen mein Bruder / und ich in unser Zimmer. Wenn wir / dann im Zimmer sind, versteckt Mama / die Geschenke unter dem Tannenbaum. / Und wenn Mama die Glocke klingelt, / dürfen wir aus dem Zimmer kommen / und die Geschenke auspacken. Wenn wir / dann die Geschenke ausgepackt haben, / essen wir zu Abend. Nach dem / Essen dürfen wir mit den Geschenken / spielen. Am 25.12. fahren wir zu meiner / Oma nach Leverkusen. Und am 26.12. / fahren wir nach Essen zu [Name] und [Name]. Da gibt es auch noch kleine / Geschenke.

Schülertext 21

So feiern wir Weihnachten // Jedes Jahr steht in unserem / Wohnzimmer ein Weihnachtsbaum. / Außer letztes Jahr, da haben / wir einfach das kleine Sofa genommen / und die Geschenke darauf gelegt. / Vorher kam Oma. Meine Schwester / und ich haben ihr etwas vorgetanzt. / Ihr hat es sehr gut gefallen. Dann / sind wir ins Wohnzimmer gegangen / und alles vor voller Geschenke. Meine kleine / Schwester [Name] hat gesagt: „Sieh mal, / [Name]! Der Weihnachtsmann war da!“ / Nun haben wir die ganzen Geschenke / ausgepackt. Jetzt mussten wir ins Bett, / weil es schon sehr spät war.

Schülertext 22

So feiern wir Weihnachten // Wir feiern Weihnachten, dass wir Kinder / in unsere Zimmer müssen und dann tun / unsere Mama, Papa und Oma die Geschenke unter / den Tannenbaum. Und wenn sie fertig sind / rufen die uns und dann packen / wir unsere Geschenke

aus. Und wenn / wir das gemacht haben, haben wir / noch ein paar Süßigkeiten gekriegt. Danach / hat unser Vater noch ein paar Fotos / von uns mit unseren Geschenken gemacht. Aber nun / mussten wir auch schon ins Bett. / Und am zweiten Tag sind wir nach / draußen gegangen und haben uns / mit Schnee einen Schneemann gebaut und / haben ihn bekleidet. Und danach / haben wir uns mit Schneekugeln / abgeworfen. Und am dritten / Tag sind wir auf die Eisfläche / gegangen. Und einmal bin ich hingefallen. / Ein paar Minuten später / sind wir nach Hause gegangen und / haben auf dem Rückweg eine Schneeballschlacht / gemacht. Nun sind wir nach Hause gegangen / und haben einen heißen Kakao gemacht.

10.3) Ergebnisse der Type-Token-Ratio im Vorfeld der Kodierung

Text	Anzahl Wörter	Nomen			Adjektive			Verben			Adverbien			Konjunktionen			Präpositionen		
		Token	Types	TTR	Token	Types	TTR	Token	Types	TTR	Token	Types	TTR	Token	Types	TTR	Token	Types	TTR
1	43	7	7	1	0	0	0	7	6	0,86	5	4	0,8	5	1	0,2	4	3	0,75
2	52	10	6	0,6	0	0	0	8	6	0,75	3	2	0,67	3	1	0,33	9	5	0,56
3	34	13	10	0,77	2	2	1	3	3	1	1	1	1	0	0	0	4	3	0,75
4	39	7	6	0,86	1	1	1	7	7	1	4	3	0,75	3	1	0,33	2	2	1
5	47	10	9	0,9	0	0	0	8	7	0,875	5	5	1	3	1	0,33	6	4	0,67
6	41	9	8	0,89	0	0	0	7	6	0,86	1	1	1	6	1	0,17	3	2	0,67
7	73	17	13	0,76	0	0	0	11	9	0,82	11	9	0,82	5	1	0,2	5	2	0,4
8	86	14	10	0,71	2	2	1	15	11	0,73	6	4	0,67	11	4	0,36	4	4	1
9	68	15	12	0,8	1	1	1	11	9	0,82	7	4	0,57	1	1	1	9	5	0,56
10	44	9	9	1	2	2	1	8	8	1	4	4	1	1	1	1	3	2	0,67
11	57	17	14	0,82	0	0	0	9	9	1	2	1	0,5	6	2	0,33	4	4	1
12	33	6	6	1	0	0	0	7	6	0,86	3	3	1	2	2	1	4	3	0,75
13	82	16	15	0,94	3	3	1	13	11	0,85	6	4	0,67	7	2	0,29	7	4	0,57
14	61	11	8	0,73	0	0	0	15	11	0,79	4	3	0,75	6	1	0,17	6	4	0,67
15	43	9	7	0,78	0	0	0	10	9	0,9	2	2	1	2	2	1	1	1	1
16	126	25	19	0,76	1	1	1	21	12	0,57	8	6	0,75	8	4	0,5	11	7	0,64
17	78	19	14	0,74	0	0	0	13	13	1	6	4	0,67	4	2	0,5	6	3	0,5
18	100	22	19	0,86	1	1	1	13	12	0,92	7	5	0,71	7	2	0,29	12	7	0,58
19	122	26	18	0,69	1	1	1	21	15	0,71	13	10	0,77	6	4	0,67	10	7	0,7
20	97	21	14	0,67	1	1	1	16	13	0,81	7	6	0,86	8	2	0,25	15	8	0,53
21	83	12	9	0,75	6	5	0,83	14	12	0,86	8	6	0,75	5	3	0,6	3	3	1
22	146	26	22	0,85	1	1	1	20	13	0,65	13	9	0,69	16	4	0,25	13	7	0,54

10.4) Die Kodierergebnisse

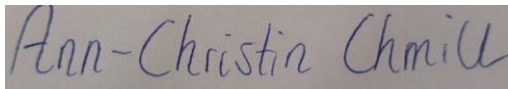
Text	Gesamteindruck			Treffericherheit		
	K1	K2	K3	K1	K2	K3
1	2	2	2	2	3	2
2	2	3	2	2	3	2
3	2	2	2	3	3	3
4	2	3	2	3	3	3
5	2	3	2	3	3	2
6	2	3	2	3	3	3
7	3	3	2	2	2	2
8	3	2	2	2	3	3
9	3	2	3	3	3	3
10	2	2	2	3	3	3
11	2	2	2	3	3	3
12	2	1	2	3	3	3
13	3	3	3	3	3	3
14	1	2	1	3	3	3
15	1	2	1	2	2	2
16	3	3	2	3	3	3
17	3	3	3	3	3	3
18	3	3	3	3	3	3
19	3	3	3	3	3	3
20	2	3	2	3	3	3
21	2	3	2	3	3	3
22	3	3	3	2	2	2

Text	Nomen			Adjektive			Verben			Adverbien			Konjunktionen			Präpositionen		
	K1	K2	K3	K1	K2	K3	K1	K2	K3	K1	K2	K3	K1	K2	K3	K1	K2	K3
1	3	3	3	1	1	1	3	3	3	2	2	2	1	1	1	2	2	2
2	1	1	1	1	1	1	2	2	2	1	1	1	2	2	2	1	1	1
3	2	2	2	3	3	3	3	3	3	3	3	3	1	1	1	2	2	2
4	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	3	3	3
5	3	3	3	1	1	1	3	3	3	3	3	3	2	2	2	1	1	1
6	3	3	3	1	1	1	3	3	3	3	3	3	1	1	1	1	1	1
7	2	2	2	1	1	1	3	3	3	3	2	3	1	3	1	1	1	1
8	2	2	2	3	3	3	2	2	2	1	1	1	2	2	2	3	3	3
9	2	2	2	3	3	3	3	3	3	1	1	1	3	3	3	1	1	1
10	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	1	1	1
11	3	3	3	1	1	1	3	3	3	1	1	1	2	2	2	3	3	3
12	3	3	3	1	1	1	3	3	3	3	3	3	3	3	3	2	2	2
13	3	3	3	3	3	3	3	3	3	1	1	1	1	1	1	1	1	1
14	2	2	2	1	1	1	2	2	2	2	2	2	1	1	1	1	1	1
15	2	2	2	1	1	1	3	3	3	3	3	3	3	3	3	3	3	3
16	2	2	2	1	3	1	1	1	1	2	2	2	2	2	2	1	1	1
17	2	2	2	1	1	1	3	3	3	1	1	1	2	2	2	1	1	1
18	3	3	3	3	3	3	3	3	3	2	2	2	1	1	1	1	1	1
19	1	1	1	3	1	1	2	2	2	2	2	2	3	3	3	2	1	2
20	1	1	1	3	3	3	2	2	3	3	3	3	1	1	1	1	1	1
21	2	2	2	2	3	3	3	3	3	2	2	2	2	3	3	3	3	3
22	3	3	3	3	3	3	1	1	1	1	1	1	1	1	1	1	1	1

Eidesstaatliche Versicherung

Ich versichere, dass ich diese Hausarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Stellen und Formulierungen, die dem Wortlaut oder Sinn nach anderen Werken entnommen sind, habe ich in jedem einzelnen Fall unter genauer Angabe der Quelle kenntlich gemacht. Ich bin darüber informiert, dass ein Zuwiderhandeln die Exmatrikulation zur Folge haben und mit einem Bußgeld bis zu 50.000 Euro belegt werden kann.

Essen, den 29.3.2019

A handwritten signature in blue ink on a grey rectangular background. The signature reads "Ann-Christin Chmill".

(Ann-Christin Chmill)