

Allgemeiner Beitrag

Ewout van der Knaap*

Zur Bedeutung der Lexik für das literarische Lesen: Didaktische Erkenntnisse einer Textdeckungsuntersuchung

The Significance of Vocabulary for Reading Literature. Didactic Evidence of an Analysis of Text Coverage

<https://doi.org/10.1515/infodaf-2018-0069>

Zusammenfassung: Bildungserfolg hängt stark mit der Lesekompetenz zusammen. Für die Sprachdidaktik sind Lesefertigkeit und Wortschatzerwerb wesentliche Fokuspunkte. Lexikalische Häufigkeit spielt dabei eine erhebliche Rolle. Dieser Beitrag verknüpft den Fokus auf den Wortschatz mit dem literarischen Lesen in Deutsch als Fremdsprache. Es wurden 26 literarische Titel aus dem literaturdidaktischen Katalog für Deutsch auf www.lezenvoordelijst.nl ausgewählt; die Hälfte davon war auf B2-Niveau eingestuft. Textdeckungsuntersuchungen mit den Häufigkeitslisten von Nation/Tschirner (2006) und Tschirner (2008b) ergaben Einblick in die didaktische Nutzbarkeit dieser Texte. Es wurde untersucht, welchen Textdeckungsgrad die ausgewählten literarischen Werke haben, wie sich diese Textdeckung zu dem eingeschätzten GER-Niveau verhält und welche (literatur)didaktischen Maßnahmen sich aus der Ermittlung der lexikalischen Brauchbarkeit der ausgewählten literarischen Werke ableiten lassen. Im Durchschnitt lag der Textdeckungsgrad im B2-Bereich auf 87 %.

Stichwörter: Vokabular, Textdeckung, Lesen, Literatur

Abstract: Educational success is strongly related to competence in reading. Didactical reading and vocabulary acquisition are essential focus points. Lexical frequency plays a prominent role in this spectrum. This contribution combines the focus on vocabulary with literary reading in German as a foreign language. 26 literary titles were selected from the didactical catalogue for German on www.lezenvoordelijst.nl

*Kontaktperson: Ewout van der Knaap, E-Mail: E.W.vanderKnaap@uu.nl

lezenvoordelijst.nl, half of which were classified at B2 level. Examinations of text coverage, using the frequency lists of Nation/Tschirner (2006) and Tschirner (2008b), gave insight into the didactical use of these texts. It was examined which rate of text coverage the chosen literary works have, how this text coverage acts to the estimated CEFR-level, and which didactical measures can be derived from the determination of the lexical usability of the selected literary works. The average of the text coverage at B2 level was 87 %.

Keywords: vocabulary, text coverage, lexical frequency, reading, literature

Erwin Tschirner zum 60. Geburtstag

1 Einleitung

Die Lesefertigkeit spielt in der Sprachdidaktik eine zentrale Rolle, weil Bildungserfolg stark mit der Lesekompetenz zusammenhängt. Dreh- und Angelpunkt im Fremdsprachenunterricht ist das Lesen: „Seit der kommunikativen Wende im Fremdsprachenunterricht wird dem Lesen als Sinnentnahme mehr Aufmerksamkeit gewidmet.“ (Lutjeharms 2010: 976). Zum einen kann die Lesekompetenz durch die Intensität der Lesezeit, zum anderen durch die Qualität der Lesezeit gefördert werden. Wesentlicher Fokuspunkt für Sprachforscher und Lehrenden ist der Wortschatzerwerb, der stark mit dem Lese-Akt in Verbindung steht, hinzu kommt aber auch die Rolle des Kontextes. Es lohnt sich aber auch, den Fokus der Lernenden an den Anfang texttypologischer Überlegungen zu stellen. Über die Bereitschaft der Lerner im Rahmen des Sprachunterrichts zu lesen entscheiden mehrere Faktoren (Day/Bamford 1998: 24, 28):

- die Anziehungskraft, der Schwierigkeitsgrad und die Verfügbarkeit des Materials,
- die Lesefertigkeit in der Fremdsprache,
- die Haltung dem Lesen im Fremdsprachenunterricht gegenüber,
- soziokulturelles Umfeld (Eltern, Verwandtschaft, Freunde),
- die Haltung dem Lesen in der Muttersprache gegenüber,
- frühere Erfahrungen mit dem Lesen in einer Fremdsprache,
- die Haltung der Kultur und Bevölkerung gegenüber,
- die Beziehung zum Dozenten und zu Mitschüler/innen,
- die im Unterricht verwendeten didaktischen Methoden und die Unterstützung der Lehrkraft.

In der Fremdsprachendidaktik sind all dies relevante Motivationsfaktoren, die auf den Leseunterricht Einfluss nehmen. Bei der Auswahl von Sachtexten und literarischen Texten spielt hinsichtlich des Schwierigkeitsgrads die lexikalische Häufigkeit zwar eine wesentliche Rolle, aber die wenigsten Lehrenden wissen, welche Wortschatzfrequenz die Texte, die sie im Unterricht einsetzen wollen, besitzen. Wortschatz kommt sowohl beim intensiven als beim extensiven Lesen erhebliche Bedeutung zu, auch im DaF-Unterricht. Im Allgemeinen gilt der Grundsatz, dass es wichtig sei, im Unterricht extensives Lesen zu ermöglichen, weil dadurch positive Effekte auf die Sprachentwicklung entstehen (vgl. die Meta-Studie von Nakanishi 2015 und das EU-Projekt Lehrkompetenzentwicklung für extensiven Leseunterricht: www.leelu.eu). Dadurch aber, dass im Schulalltag für die Textauswahl als Hauptquelle die Einstufung, die Verlage festgesetzt haben, schlichtweg übernommen wird, werden beim extensiven und literarischen Lesen Texte auf Basis von ungenauen Kriterien gewählt. Wieviel Wortschatz für gelungene Lektüre passiv notwendig ist, ist Forschungsobjekt mehrerer Studien; die Wortschatzfrequenz von literarischen Texten findet allerdings nur wenig Aufmerksamkeit (Nation/Wang Ming-tzu 1999, Nation 2006), wobei bislang nichts über deutsche Texte vorzufinden ist.

Um authentische Texte selbständig lesen zu können, sollten Lesende idealerweise 98 % der Lexik in Texten verstehen (Hu/Nation 2000, Laufer et al. 2010, Schmitt et al. 2011), für intensives Lesen sind weniger als 95 % notwendig, für extensives Lesen hingegen zwischen 95 und 98 % (Nation 2001: 150). Dass für erfolgreiche Lektüre nur ein neues Wort auf 20 Wörter (95 %) oder ein neues Wort auf 50 Wörter (98 %) vorkommen soll, bedeutet, dass Wortschatz eine Schlüsselstelle im Verstehensprozess spielt. Im Englischen sind passive Kenntnisse von 5000 Wortfamilien eine Bedingung für autonome Lektüre von Texten jeglicher Art ohne lexikalische Probleme (Laufer et al. 2010). Im Deutschen liegt das professionelle Wortschatzminimum bei 5000 Wörtern (Tschirner 2008a).

Aus didaktischer Sicht ist es notwendig, dass Wortschatzlernprogramme ambitionierte Ziele verfolgen, damit Lernende sich einen relevanten und großen Wortschatz aneignen können. Der nachfolgenden Übersicht (Tabelle 1) ist zu entnehmen, welches Wortschatzminimum auf den GER-Niveaus erreicht werden sollte (Tschirner 2009):

Tabelle 1: Wortschatzminimum per Niveau des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GER).

A1	1000
A2	2000
B1	3000
B2	5000
C1	7500
C2	10000

Im niederländischen Abitur sollen im Deutschen als Fremdsprache Lesekompetenzen auf dem B2-Niveau erreicht werden. Extensives Lesen gehört zwar zu den erprobten Mitteln, den Wortschatz zu erweitern, offen ist dabei jedoch, mit welchen Texten diese Lernziele erreicht werden können. Im Fremdsprachenunterricht im Allgemeinen und an Sekundarschulen im Besonderen ist in den Niederlanden weder freies (extensives) noch literarisches Lesen selbstverständlich. Ein Stufenmodell, das für die Fremdsprachendidaktik erweitert wurde (van der Knaap 2014), bietet Möglichkeiten, die Leseentwicklung an Sekundarschulen besser zu begleiten. Um Lehrerinnen und Lehrern sowie Schülerinnen und Schülern bei der Titelauswahl nach dem Niveaumodell behilflich zu sein, wurde ein Bücherkatalog entwickelt, bei dessen Zusammenstellung jedoch das Problem auftauchte, dass der Wortschatz einzelner Texte nur sehr pauschal in Kategorien des Gemeinsamen Europäischen Referenzrahmens für Sprachen beschrieben werden konnte. Dadurch kam die Frage nach der Textdeckung auf, d. h. dem Anteil der Wörter eines Textes, der von einer Häufigkeitwörterliste abgedeckt wird. Daraus entstanden drei Forschungsfragen:

- 1 Welchen Textdeckungsgrad haben die ausgewählten literarischen Werke?
- 2 Wie verhält sich die Textdeckung der ausgewählten literarischen Werke zu dem eingeschätzten GER-Niveau?
- 3 Welche (literatur)didaktischen Maßnahmen lassen sich aus der Ermittlung der lexikalischen Brauchbarkeit der ausgewählten literarischen Werke ableiten?

2 Zur Methode

Die Erkenntnis vom Nutzen „die häufigsten Wörter einer Sprache zu lernen“ (Tschirner 2008a: 195), nahm Tschirner zum Ausgangspunkt, ein Korpus der deutschen Gegenwartssprache und ein Häufigkeitwörterbuch der deutschen

Sprache (Jones/Tschirner 2006) zu erstellen. Um ermitteln zu können, welche Schwierigkeiten das Vokabular in den Texten bietet, die im Katalog deutscher Literatur www.lezenvoordelijst.nl aufgenommen sind, wurden Frequenzwortlisten für die häufigsten 1–1000, 1001–3000 und 3001–5000 Wörter benutzt. Es wurden 26 Titel aus dem Katalog von 75 Titeln ausgewählt, wobei die literarische Kompetenzstufe, das GER-Niveau und der Grad der Kanonisierung die primären Auswahlkriterien waren. Die Skalierung von sechs unterschiedlichen Stufen der literarischen Kompetenz beruht auf wahrnehmbaren Unterschieden von Profilen in der niederländischen Schulpraxis (van der Knaap 2014); reflektierendes Lesen (Niveau 3) und interpretierendes Lesen (Niveau 4) werden im Abitur erreicht. Sekundär wurde die Auswahl von der Zustimmung der Verlage zur Weiterverwendung dieser Texte bestimmt. Die GER-Niveaus, siehe Tabelle 2, waren zwar ungleich verteilt: B1 (7), B2 (15, d.h. 57,7%), C1 (4), dies entspricht aber den Niveaus der Abschlussprüfungen. Weil niederländische Schülerinnen und Schüler an Realschulen und Gymnasien mehrheitlich das literarische Kompetenzniveau 3 bewältigen können, wurden besonders Titel dieses Niveaus für die Analyse ausgewählt. Titel der Niveaus im unmittelbaren Bereich darüber (4) und darunter (2) sind im Vergleich zum Einstiegsniveau (1) und angestrebten Niveau (5) in der Auswahl verhältnismäßig stark vertreten:

Tabelle 2: Streuung der ausgewählten Titel über die literarischen Kompetenzniveaus und die GER-Niveaus.

literarisches Kompetenzniveau	GER				Total
	A2	B1	B2	C1	
1	0	1	0	0	2
2	0	4	1	0	5
3	0	1	9	1	11
4	0	0	4	1	5
5	0	0	1	2	3
Total	1	6	15	4	26

Auf Niveau 1 wurden zwei Titel, auf Niveau 2 und 4 jeweils fünf Titel ausgewählt; auf Niveau 3 sind 11 Titel eingestuft. Weil nur wenige Schülerinnen und Schüler das Niveau 6 erreichen, wurde dieses Niveau nicht miteinbezogen.

Die Auswahl wurde geringfügig geprägt von der Verfügbarkeit digitaler Textfassungen bei den Verlagen, die zum Teil unentgeltlich, zum Teil kostenpflichtig kooperierten. Besonders die finanzielle Beschränkung führte dazu, dass nicht alle Titel des literaturdidaktischen Katalogs einer Analyse unterzogen werden konn-

ten; bei Projektbeginn waren aber 25 Titel verarbeitet. Die Textdateien wurden mit Hilfe des Programms „antwortprofiler“ mit den Wortlisten abgeglichen.¹ Die Auswahl literarischer deutschsprachiger Prosa kann insofern als repräsentativ gelten, dass der Katalog bereits aus sorgfältig ausgewählten Titeln besteht, die eine Mischung darstellt von bewährten, an Schulen gelesenen Titeln und von neueren Werken. Auch die 26 Titel (siehe Abbildung 1) sind repräsentativ für das, was an niederländischen Schulen gelesen wird. Mit einem Klassiker von 1826 (Eichendorffs *Aus dem Leben eines Taugenichts*) und drei Klassikern aus den Jahren 1910–1950 (Kafkas *Die Verwandlung*, Remarques *Im Westen nichts Neues* und Zweigs *Schachnovelle*) besteht das Korpus aus zwei Titeln nach 1970, fünf nach 1990 und 15 Titeln nach 2000.

3 Analyse

Die sehr heterogene Textauswahl von 26 Titeln hat eine Durchschnittslänge von 41 408 Tokens (laufende Wörter) und 6265 Types (individuelle Wortformen); insgesamt haben die 26 Titel 1 076 619 Tokens. Die Durchschnittswerte der Deckungsgrade zu den unterschiedlichen Wortlisten sind für den Überblick informativ, auch ergänzt um die Standardabweichung: 74,6 % (1000; SD = 2,8), 84,2 % (3000; SD = 2,3) und 87,4 % (5000; SD = 2,1). Graphisch lässt sich zeigen, wie die Textdeckung bei 26 literarischen Titeln verteilt ist; hier unter Benutzung der Frequenzliste bis 1000 Wörter (Abbildung 1):

¹ Die Anregung und die Bereitstellung der Herder/BYU-Häufigkeitslisten verdanke ich Erwin Tschirner, dem Autor der Listen. Die häufigsten 4000 Wörter sind veröffentlicht (Jones/Tschirner 2006) und nach Themen sortiert (Tschirner 2008b); die bis 5000 erweiterte Liste wurde veröffentlicht in Tschirner 2008a. Jupp Möhring (Herder-Institut, Universität Leipzig) danke ich für technische Ratschläge und für die Errechnung der STTR. Allen Autorinnen und Autoren sowie folgenden Verlagen bin ich dankbar für ihre Bereitschaft, digitale Versionen der Texte zur Verfügung zu stellen: Arena Verlag, Beltz, Carlsen, CBT-Verlag (Random House), Eichborn, Fischer Verlag, Hanser Verlag, Kiepenheuer & Witsch, Klett-Langenscheidt, Piper Verlag, Ravensburger, Thienemann-Esslinger. Besonderer Dank gebührt Judith Schalansky, die sofort und persönlich kooperierte. Dem Forschungsinstitut Institute for Cultural Inquiry (ICON) der Universität Utrecht verdanke ich die Ressourcen für die Förderung der Verlagskosten sowie für die Hilfskraft Theresa Redl, der ich für ihre kompetente Arbeit im Studienjahr 2015/16 herzlich danke.

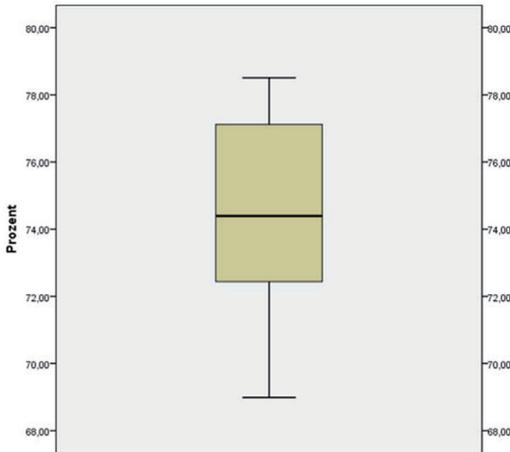


Abbildung 1: Die Textdeckung von 26 literarischen Titeln bei der Frequenzliste bis 1000 Wörter.

Wie sehr sich die Textdeckung zwischen etwa 73 und 77 % konzentriert, und dass es Ausreißer nach oben und unten gibt, ist ersichtlich.

Abbildung 2 zeigt die Resultate der Textdeckungsuntersuchung, die im Folgenden noch zu detaillieren ist:

Titel	Autor	litKomp	GeR	Tokens	Types	STTR 2000	STTR 500	< 1000 - Tok	< 3000 - Tok	< 5000 - Tok	Außerhalb - Tok
Bitterschokolade	Pressler Mirjam	1	A2	29190	4274	36,59	52,87	74,12	84,26	87,43	12,57
Yildiz heißt Stern	Heyne Isolde	1	B1	25100	3341	34,6	53,19	78,45	87,25	89,85	10,15
Love Alice	Savina Nataly Elisabeth	2	B1	31430	5640	41,38	56,79	72,34	83,22	86,47	13,53
Täglich die Angst	Theisen Manfred	2	B1	34807	5076	37,3	54,45	74,67	84,04	87,09	12,91
Lilly unter den Linden	Voorhoeve Anne C.	2	B2	66272	9359	41,75	58,27	74,73	84,83	87,94	12,06
Der Klassenfeind und ich (ohne Glossar)	Bollwahn Barbara	2	B1	54555	7337	38,01	54,76	77,38	86,38	89,15	10,85
Novemberschnee	Banscherus Jürgen	2	B1	22471	3747	38,41	55,74	77,12	86,99	90	10
Mein erstes T-shirt	Hein Jakob	3	B1	29830	6692	43,65	58,8	73,23	82,69	86,06	13,94
Russendisko	Kaminer Wladimir	3	B2	32617	6883	44,38	59,92	71,87	82,75	86,08	13,92
Der Fall Collini	Schirach Ferdinand von	3	B2	30214	5345	40,57	56,38	71,08	83,01	86,57	13,43
Meine freie deutsche Jugend	Rusch Claudia	3	B2	28088	6103	43,45	58,98	73,57	83,39	86,54	13,46
Herrn Kukas Empfehlungen	Knapp Radek	3	B2	53233	7323	39,48	56,07	76,14	85,4	88,56	11,45
Wir Kinder vom Bahnhof Zoo	F. Christiane	3	B2	101372	9497	35,3	52,08	77,54	85,72	88,46	11,54
Tschick	Herndorf Wolfgang	3	B2	58193	7257	35,86	52,15	76,41	85,77	88,52	11,48
Rotkäppchen muss weinen	Hanika Beate Teresa	3	B2	48126	5890	36,62	52,75	77,05	86,18	89,21	10,79
Herr Lehmann	Regener Sven	3	B2	72710	7372	33,22	49,83	78,05	86,27	88,65	11,35
Wie Licht schmeckt	Ani Friedrich	3	B2	44493	5453	35,75	51,87	78,17	88,05	91,13	8,87
Am kürzeren Ende der Sonnenallee	Brussig Thomas	3	C1	32050	5991	39,36	54,26	71	79,41	82,39	17,62
Der Hals der Giraffe	Schalansky Judith	4	B2	60905	9821	46,89	61,4	68,99	79,42	82,94	17,05
Schachnovelle	Zweig Stefan	4	B2	18260	4617	44,43	60,17	70,22	81,13	85,24	14,75
Am Beispiel meines Bruders	Timm Uwe	4	C1	33994	7265	43,08	57,21	72,44	82,05	85,44	14,56
Im Westen nichts Neues	Remarque Erich Maria	4	B2	57185	8509	41,4	56,88	73,64	83,26	86,73	13,26
Agnes	Stamm Peter	4	B2	25719	4164	37,22	53,18	78,51	87,76	90,39	9,62
Der Zimmerspringbrunnen	Sparschuh Jens	5	B2	35524	7205	43,15	59,4	73,15	82,75	85,96	14,04
Aus dem Leben eines Taugenichts	Eichendorff Josef von	5	C1	31132	5000	39,92	56,73	72,8	82,85	86,38	13,62
Die Verwandlung	Kafka Franz	5	C1	19149	3735	38,97	55,82	76,81	85,59	89,46	10,53

Abbildung 2: Übersicht der Textdeckung von 26 literarischen Titeln aus dem deutschen Katalog von www.lezenvoordelijst.nl, geordnet nach dem literarischen Kompetenz-Niveau („litKomp“).

Die Daten enthalten außerdem das GER-Niveau („GER“), die Gesamtzahlen der Tokens und Types in den Texten, das errechnete standardisierte Type-Token-Verhältnis („STTR“) für die Blockgrößen 2000 bzw. 500 und letzten Endes die prozentuellen Deckungsgrade der Titel für die Frequenzlisten 1–1000, 1001–3000 und 3001–5000 Wörter mit der Prozentzahl der Tokens die außerhalb der Liste bis 5000 Wörter gehen.

Die nähere Betrachtung der Ergebnisse führt zur Erkenntnis, dass die Textdeckung bei sieben Titeln am höchsten ist, wenn die Wortliste bis 5000 Wörter benutzt wird (Tabelle 3). In der Tabelle sind die Titel so geordnet, dass die Textdeckung in der letzten Spalte steigt:

Tabelle 3: Übersicht von Titeln mit höchster Textdeckung bei der 3000/5000 Wortliste. Titel, die auch im Licht der 3000er Liste zu den deckungsreichsten Titeln gehören sind mit einem Asterisk gekennzeichnet.

Autor, Buchtitel	Literarisches Kompetenzniveau	GER	Textdeckung bei 3000 Wortliste	Textdeckung bei 5000 Wortliste
Bollwahn, Barbara: <i>Der Klassenfeind und ich</i> (ohne Glossar)	2	B1	86,4 %*	89,2 %
Hanika, Beate Teresa: <i>Rotkäppchen muss weinen</i>	3	B2	86,2 %	89,2 %
Kafka, Franz: <i>Die Verwandlung</i>	5	C1	85,6 %	89,5 %
Heyne, Isolde: <i>Yildiz heißt Stern</i>	1	B1	87,3 %*	89,9 %
Banscherus, Jürgen: <i>Novemberschnee</i>	2	B1	87,0 %*	90,0 %
Stamm, Peter: Agnes	4	B2	87,8 %*	90,4 %
Ani, Friedrich: <i>Wie Licht schmeckt</i>	3	B2	88,1 %*	91,1 %

Dass mehrere Titel, die auf B1-Niveau eingestuft sind, zu den deckungsreichsten Titeln gehören, wenn sie mit der Frequenzwörterliste von 5000 abgeglichen werden, dürfte nicht verwundern, denn diese sollten entsprechend der Einstufung eigentlich zu der 3000er Liste passen. Umgekehrt gehören einige Titel, die auf B2-Niveau eingestuft sind, auch nach Abgleichung mit der 3000er Liste zu den ranghöchsten. Fünf der genannten Titel gehören auch im Licht der 3000er Liste zu den deckungsreichsten Titeln; sie sind dort mit einem Asterisk gekennzeichnet.

Die Wortschatzvariabilität wurde ermittelt, indem mit WordSmith Tools 6.0 ein standardisiertes Type-Token-Verhältnis (STTR) berechnet wurde. Es wurden dafür „Blockgrößen“ von 2000 und 500 Tokens festgelegt. Je niedriger der STTR-Wert, um so höher ist die lexikalische Vielfalt des Textes. Bei 2000 Tokens liegt die STTR im Durchschnitt bei 39,8. Die Vielfalt ist am höchsten bei *Herr Lehmann* (33,2 bei 2000 Tokens) und am geringsten bei *Der Hals der Giraffe* (46,9 bei 2000 Tokens). Es gibt dabei offensichtlich einen Bezug zwischen der STTR und dem Textdeckungsgrad, was am Beispiel von *Herr Lehmann* und *Der Hals der Giraffe* gezeigt werden kann. Benutzen wir die 5000er Wortliste, dann ist der Textdeckungsgrad von *Herr Lehmann* 88,7% und von *Der Hals der Giraffe* 82,9%. Brussigs Roman *Am anderen Ende der Sonnenallee* hat kurz vor *Der Hals der Giraffe* mit 82,4% den geringsten Textdeckungsgrad, dabei aber eine höhere Wortschatzbreite (39,4 bei 2.000 Tokens) als *Der Hals der Giraffe*.

Bei *Bitterschokolade* wurde die Originalausgabe benutzt, die Ausgabe in der Klett-Reihe „Leichte Lektüren für Jugendliche“ wird vom Verlag auf A2–B1 eingestuft. Es fällt bei diesem Titel auf, dass bei der Wortliste von 3000 (B1-Niveau) nur 84,3% der Token erfasst werden und dass lexikalische Hilfestellungen erforderlich sind. Hingegen weisen andere Titel, die auf B1-Niveau eingestuft sind, bei derselben Wortliste einen höheren Textdeckungsgrad auf: 87,3% (*Yildiz heißt Stern*) und 87,0% (*Novemberschnee*). Andererseits ist bei einigen B1-Titeln der Textdeckungsgrad niedriger: 83,2% (*Love Alice*).

Um den didaktischen Einsatz der Titel in lexikalischer Hinsicht besser einschätzen zu können, wurde in Betracht gezogen, wieviel zusätzliche Vokabel notwendig wären, um unbekannte Tokens auf 5% reduzieren und dadurch einen Deckungsgrad von 95% zu erreichen (Tabelle 4). Bei den Titeln, die den höchsten Textdeckungsgrad (vgl. Tabelle 3) haben, sind freilich am wenigsten Vokabeln notwendig, um diese 95%-Grenze erreichen zu können:

Tabelle 4: Zahl der notwendigen zusätzlichen Vokabeln, um 95% Textdeckung bei den angegebenen literarischen Titeln zu erlangen.

Buchtitel	Zahl der Vokabeln
<i>Yildiz heißt Stern</i>	24
<i>Novemberschnee</i>	157
<i>Agnes</i>	184
<i>Die Verwandlung</i>	268
<i>Wie Licht schmeckt</i>	273
<i>Herr Lehmann</i>	300
...	

Tabelle 4: (fortgesetzt)

Buchtitel	Zahl der Vokabeln
<i>Am anderen Ende der Sonnenallee</i>	1146
<i>Russendisko</i>	1231
<i>Im Westen nichts Neues</i>	1338
<i>Der Zimmerspringbrunnen</i>	1617
<i>Am Beispiel meines Bruders</i>	1799
<i>Lilly unter den Linden</i>	2372
<i>Der Hals der Giraffe</i>	2607

Die Tabelle zeigt die Titel, die die wenigsten und die Titel, die die meisten Vokabeln brauchen, um die 95 %-Textdeckungsgrenze zu erreichen. Dabei fällt auf, dass *Der Hals der Giraffe* (82,9 %) bei der 5000er Wortliste zwar einen nahezu identischen Deckungsgrad hat wie *Am anderen Ende der Sonnenallee* (82,4 %), Brussigs Buch aber weniger (1146) Wörter braucht, um unbekannte Tokens auf 5 % der gesamten Lexemmenge reduzieren zu können. Bei den restlichen Titeln, wofür zwischen 334 und 1005 Types zusätzlich notwendig wären, um die 95 %-Schwelle zu erreichen, ergibt sich einen Durchschnitt von 675 Vokabeln, die pro Titel zusätzlich notwendig sind, um das 95 %-Optimum an Textdeckung zu erreichen.

Auf Basis der vorgenommenen lexikalischen Analyse wurde außerdem die Plausibilität der GER-Zuordnung überprüft. Es stellt sich heraus, dass die GER-Zuordnung von vier Titeln (*Bitterschokolade*, *Love Alice*, *Der Hals der Giraffe*, *Der Zimmerspringbrunnen*) im Katalog auf der Webseite www.lezenvoordelijst.nl eventuell nach Berücksichtigung anderer Textelemente anzupassen wäre, weil sie auf dem angedeuteten Niveau unterhalb der Grenze von 85 % der Textdeckung liegen (siehe Abbildung 1). Das Kriterium dieser Zäsur erfolgt auf Basis des Durchschnitts der anderen Titel. Durch den im Vergleich niedrigeren Deckungswert gibt es Anlass einer Korrektur bei den erwähnten Titeln.

Als Kontrastfolie dürfte die Analyse der Titel mit Hilfe des Lesbarkeitsindex „LIX“ dienen. Dadurch können u. a. die Satzlänge, die Prozentzahl langer Wörter und der Lesbarkeitsindex erfasst werden. Der LIX „ergibt sich aus der Summe der durchschnittlichen Satzlänge eines Textes und des prozentualen Anteils langer Wörter (mehr als sechs Buchstaben)“ (vgl. <https://www.psychometrica.de/lix.html>). Je niedriger die LIX-Zahl, um so lesbarer ist der Text. Mit Tabelle 5 soll dies dargestellt werden:

Tabelle 5: Resultate des LIX bei 26 literarischen Titeln, mit Indikation der literarischen Kompetenzstufe („LK“), des GER, der durchschnittlichen Satzlänge, der Prozentzahl langer Wörter und des LIX-Wertes.

Buchtitel	LK	GER	Satzlänge	% lange Wörter	LIX
<i>Yildiz heißt Stern</i>	1	B1	8	16,5	24,5
<i>Täglich die Angst</i>	2	B1	8	17,3	25,3
<i>Bitterschokolade</i>	1	B1	8	18	26
<i>Wie Licht schmeckt</i>	3	B2	9	17,6	26,6
<i>Novemberschnee</i>	2	B1	8	19,8	27,8
<i>Agnes</i>	4	B2	10	18,8	28,8
<i>Wir Kinder vom Bahnhof Zoo</i>	3	B2	10	19,4	29,4
<i>Love Alice</i>	2	B1	10	19,5	29,5
<i>Herr Lehmann</i>	3	B2	12	18	30
<i>Herrn Kukas Empfehlungen</i>	3	B2	10	20,8	30,8
<i>Der Klassenfeind und ich (ohne Glossar)</i>	2	B1	11	21,4	32,4
<i>Im Westen nichts Neues</i>	4	B2	12	20,9	32,9
<i>Rotkäppchen muss weinen</i>	3	B2	16	17,4	33,4
<i>Tschick</i>	3	B2	13	20,5	33,5
<i>Der Hals der Giraffe</i>	4	B2	6	27,6	33,6
<i>Meine freie deutsche Jugend</i>	3	B2	10	24	34
<i>Der Fall Collini</i>	3	B2	11	24,4	35,4
<i>Lilly unter den Linden</i>	2	B2	14	21,4	35,4
<i>Am kürzeren Ende der Sonnenallee</i>	3	C1	13	23,2	36,2
<i>Der Zimmerspring- brunnen</i>	5	B2	12	25,5	37,5
<i>Am Beispiel meines Bruders</i>	4	C1	15	25,2	40,2
<i>Russendisko</i>	3	B2	14	26,4	40,4
<i>Mein erstes T-shirt</i>	3	B1	15	25,7	40,7
<i>Aus dem Leben eines Taugenichts</i>	5	C1	22	19,6	41,6

Tabelle 5: (fortgesetzt)

Buchtitel	LK	GER Satzlänge	% lange Wörter		LIX
<i>Die Verwandlung</i>	5	C1	26	23,3	49,3
<i>Schachnovelle</i>	4	B2	22	29,1	51,1
<i>Der Klassenfeind und ich</i> (nur Glossar)	2	B1	16	38,4	54,4

Der Lesbarkeitsindex soll eine rudimentäre Einschätzung des Textschwierigkeitsgrads bieten. Sofort ergeben sich aus den LIX-Werten Probleme. Gemäß der LIX-Klassifizierung (<https://www.psychometrica.de/lix.html>) deuten LIX-Werte unter 40 auf Kinder- und Jugendliteratur hin, und das kann bei modernen Klassikern wie *Der Hals der Giraffe*, *Agnes*, *Herrn Kukas Empfehlungen*, *Der Fall Collini*, *Der Zimmerspringbrunnen*, *Im Westen nichts Neues*, *Herr Lehmann*, *Am kürzeren Ende der Sonnenallee* und vielleicht auch bei *Tschick* nicht stimmen. LIX-Werte von 40 bis 50 deuten auf Belletristik hin, was nur für 5 Titel der Analyse gelten würde. Das didaktisch sehr hilfreiche Glossar im Anhang von *Der Klassenfeind und ich* entfällt auf die Kategorie *Sachliteratur* (50–60). Überraschen mag die Tatsache, dass *Der Hals der Giraffe* zwar die geringste Satzlänge hat, aber, wie sich oben erwiesen hatte, die geringste Wortschatzbreite und den geringsten Wortdeckungsgrad aufweist.

Bei der LIX-Analyse ist auffällig, dass mehrere der Titel, die einen hohen Textdeckungsgrad aufweisen, die niedrigste Prozentzahl langer Wörter enthalten. Dies hängt mit der Tatsache zusammen, dass die häufigsten Wörter kurz sind, was auch dadurch bedingt wird, dass unter den häufigsten Wörtern die meisten Funktionswörter vorkommen. Texte, die hochfrequenten Wortschatz aufweisen sind in Tabelle 6 mit einem Sternchen markiert. Hervorzuheben ist, dass mehrere Titel im unteren Bereich der literarischen Kompetenz angesiedelt sind, wodurch eine geringe Wortlänge erwartet werden könnte. Atypisch aber ist Peter Stamms Roman *Agnes*, dessen literarischer Anspruch (Niveau 4) mit einer geringen Wortlänge und einem hohen Textdeckungsgrad einhergeht.

Tabelle 6: Literarische Texte mit einer durchschnittlich geringen Wortlänge

Werk	LK	GER	% lange Wörter
<i>Yildiz heißt Stern*</i>	1	B1	16,5
<i>Täglich die Angst</i>	2	B1	17,3
<i>Rotkäppchen muss weinen*</i>	3	B2	17,4

Tabelle 6: (fortgesetzt)

Werk	LK	GER	% lange Wörter
<i>Wie Licht schmeckt*</i>	3	B2	17,6
<i>Bitterschokolade*</i>	1	B1	18
<i>Herr Lehmann</i>	3	B2	18
<i>Agnes*</i>	4	B2	18,8

Aus einer Korrelationsanalyse, die mit dem Programm SPSS durchgeführt wurde, ergab sich ein negativer Bezug zwischen dem literarischen Kompetenzniveau und dem Wortschatzdeckungsgrad der ausgewählten 26 literarischen Texte bei der 5000er Liste: $r = -.268$, $n = 26$. Dieser Zusammenhang ist allerdings nicht signifikant ($p > 0,1$). Das bedeutet, dass das literarische Kompetenzniveau unabhängig von der gemessenen lexikalischen Dichte eines literarischen Textes ist; der Befund bestätigt die Erkenntnis, dass die Lexik nicht den einzigen Faktor des literarischen Schwierigkeitsgrads darstellt.

Abschließend soll an drei ausgewählten Fällen gezeigt werden, zu welcher Problematik eine allgemeine Textdeckungsuntersuchung führen kann.

Die Frage, ob eine Vereinfachung des Wortschatzes für Lesende förderlich ist, beschäftigt die Didaktik. Es kommt beim Textverstehen natürlich nicht nur auf den Wortschatz an. *Im Westen nichts Neues* steht von den ausgewählten 26 Titeln beim Deckungsgrad an 15. Stelle. Die Zugänglichkeit trägt sicherlich dazu bei, dass dieser Klassiker sich nach wie vor an Schulen behaupten kann. Der Roman ist syntaktisch und bis auf einige Militärausdrücke auch semantisch nicht allzu kompliziert. Durch den Stil müssen Leser aber für das Textverstehen häufig inferieren, was unausgesprochen geblieben ist, was zu Problemen führen kann. Eine Edition für Leseschwache, erschienen im Spaß am Lesen Verlag, wurde zum Vergleich herangezogen. Es konnte allerdings nur ein Textauszug herangezogen werden, weil der Verlag umfassendere Textausschnitte nicht zur Verfügung gestellt hatte. Die Edition kürzt den Text erheblich, ändert die Syntax und vereinfacht den Wortschatz. Aus der Gegenüberstellung (Tabelle 7) wird ersichtlich, welche Konsequenzen das für den Textdeckungsgrad hat:

Tabelle 7: Textdeckungsgegenüberstellung zweier Ausgaben von *Im Westen nichts Neues*

	<1000 Tokens	<3000 Tokens	<5000 Tokens
<i>Im Westen nichts Neues</i> (Original, Ganztext)	73,7 %	83,3 %	86,7 %
<i>Im Westen nichts Neues</i> (vereinfacht, Auszug)	76,1 %	84,4 %	87,4 %

Die Erhöhung der Textdeckung scheint auf Basis der Analyse eines Auszugs der vereinfachten Edition minimal zu sein. Die TTR-Werte können nicht verglichen werden, weil die Korpora von unterschiedlicher Größe sind. Auch von *Tschick* bietet der Spaß am Lesen Verlag eine Ausgabe in einfacher Sprache: auf 64 statt auf 254 Seiten wird die Geschichte durch eine genaue Gliederung und gut strukturierte Chronologie nacherzählt. Ein Vorwort leitet in das Geschehen ein und Kapitelüberschriften erleichtern die Lektüre. Die einfachere Version besteht aus kürzeren, einfacher formulierten Sätzen. Für die Äußerungen in der Umgangssprache bietet der Anhang des Buches eine Wörterliste, wo die Wörter wieder in einer einfachen Sprache erläutert werden. Nur ein gründlicher Vergleich der Textfassungen könnte die Befunde präzisieren.

Den zweiten ausgewählten Fall stellt Peter Stamms Roman *Agnes* dar. In ihm kommen zahlreiche englische Wörter vor. Sie könnten, je nach sprachlichem Hintergrund der lesenden Person, den Schwierigkeitsgrad erhöhen oder verringern. Es sind relativ gängige Begriffe: Neben Ortsnamen (New York, Chicago) tauchen Wörter auf, die, wie zwischen Klammern angedeutet ist, unterschiedlichen Frequenzlisten zugehören (vgl. <http://www.victoria.ac.nz/lals/about/staff/paul-nation#vocab-lists>): beispielsweise *can* (1000), *day* (1000), *other* (1000), *survive* (2000), *mourn* (4000), und *anthology* (8000). In *Agnes* wird aber auch aus dem gehobenen englischsprachigen Kanon zitiert: u.a. aus Shakespeares 18. Sonnet (“Shall I compare thee to a summer’s day”, Kapitel 9), aus einem Gedicht von Dylan Thomas (“Refusal to Mourn the Death, by Fire, of a Child in London”, Kapitel 30), und aus einem Gedicht von Robert Burns (“Auld Lang Syne”, Kapitel 31). Würde man diesen literarischen Wortschatz entfernen, steigert sich die Wortschatzdeckung allerdings nur gering, wie aus Tabelle 8 hervorgeht:

Tabelle 8: Wortschatzdeckung in *Agnes* mit und ohne eingerückten englischen Text. geordnet nach den Gesamtzahlen der Tokens und Types in den Texten, und die prozentuellen Deckungsgrade für die Frequenzlisten 1–1000, 1001–3000 und 3001–5000 Wörter.

Werk	Tokens	Types	< 1000	< 3000	< 5000
Agnes	25719	4164	78,5 %	87,8 %	90,4 %
Agnes (ohne englische Texte)	25635	4114	78,7 %	88,0 %	90,7 %

Fremdsprachige Textelemente bedürfen aus lexikalischer Hinsicht also kaum besondere Berücksichtigung bei der Textauswahl; bei interpretativen Handlungen sind sie aber miteinzubeziehen.

Im dritten Beispiel geht es um das Verhältnis zwischen Haupttext und Nebentext. Nicht untypisch für literarische Texte, die im schulischen Kontext gelesen werden, ist, dass Glossare eingesetzt werden. *Der Klassenfeind und ich* enthält ein Glossar, das kaum ein Zehntel des gesamten Textes umfasst. Durch den lexikalischen Schwierigkeitsgrad des Glossars senkt sich der Textdeckungsgrad des Gesamttextes; die Erläuterungen hemmen also den Lesefluss. Aus einer Gegenüberstellung beliebiger Stellen lässt sich der unterschiedliche Stil und Schwierigkeitsgrad mühelos erkennen:

Haupttext (Bollwahn 2009: 35): „Mutti und Papi fragen mich in letzter Zeit immer öfter, ob ich denn schon weiß, was ich nach dem Abitur machen will. Gute Frage. Papi würde mich am liebsten überreden, Lehrerin zu werden. ‚Die werden dringend gesucht‘, liegt er mir ständig in den Ohren. ‚Da kriegst du sofort einen Studienplatz.‘ Aber ich hab schon tausendmal gesagt, dass ich nicht Lehrerin werden will.“

Begriffserklärungen (Bollwahn 2009: 246): „ADN: Der Allgemeine Deutsche Nachrichtendienst war die einzige zugelassene Nachrichten- und Bildagentur und nahm eine Monopolstellung bei der Belieferung fast aller Zeitungen sowie von Rundfunk und Fernsehen ein. Der ADN unterstand offiziell dem Ministerrat und wurde vom zentralen Parteiapparat der Sozialistischen Einheitspartei Deutschlands (SED) beaufsichtigt.“

Ohne Glossar steigert sich die lexikalische Lesbarkeit. Bollwahns Jugendroman gehört, wie wir am Anfang bereits feststellten, zu den deckungsreichsten Romanen der analysierten Liste.

Tabelle 9: Deckungsgrad von *Der Klassenfeind und ich* mit und ohne Glossar, geordnet nach den Gesamtzahlen der Tokens und Types in den Textfassungen, das errechnete standardisierte Type-Token-Verhältnis („STTR“) für die Blockgröße 2000 und die prozentuellen Deckungsgrade für die Frequenzlisten 1–1000, 1001–3000 und 3001–5000 Wörter.

	Tokens	Types	STTR (2000)	< 1000	< 3000	< 5000
<i>Der Klassenfeind und ich</i> (mit Glossar)	59 952	8505	38,9	76,0 %	85,2 %	88,1 %
<i>Der Klassenfeind und ich</i> , Bollwahn (nur Glossar)	5 397	2112	47,4	61,5 %	73,0 %	77,5 %
<i>Der Klassenfeind und ich</i> , Bollwahn (ohne Glossar)	54 555	7337	38,0	77,4 %	86,4 %	89,2 %

4 Literaturdidaktische Schlussfolgerungen

In der Fremdsprachendidaktik wird nicht zuletzt bei der Arbeit mit literarischen Texten die Bedeutung der Auswahlkriterien betont: Es sei zunächst zu prüfen, ob die Texte dem Sprachstand und den Lernzielen der Lernergruppen angemessen sind (vgl. Rösler 2012: 228). Die Frage nach den (literatur)didaktischen Maßnahmen, die sich aus der Ermittlung der lexikalischen Brauchbarkeit der ausgewählten literarischen Werke ableiten lassen, soll zunächst mit dem Hinweis auf die unterschiedlichen Werte der einzelnen 26 Titel beantwortet werden. Es liegt auf der Hand, für die jeweils extensiven oder intensiven Lektürepräsen solche Titel auszuwählen oder bereitzustellen, die nicht nur den Anforderungen der Thematik und der literarischen Kompetenz genügen, sondern die auch lexikalisch am besten geeignet sind.

Literaturdidaktisches Handeln sollte auch lexikalisches Bewusstsein miteinschließen, und kann je nach Lernziel in sprachliches Handeln übergehen. Es ist sinnvoll, Schülerinnen und Schüler auch Aufgaben zur Lexik ausführen zu lassen, damit sie aktiv Sprache lernen. Wortschatzaktivitäten wie beispielsweise die Erstellung von Vokabellisten oder die Arbeit in Kleingruppen können zum Lernen aktivieren. Die Textdeckungsuntersuchung ermöglicht es, Vokabellisten zu erstellen, und dadurch die Schülerinnen und Schüler bei der Lektüre zu unterstützen.

Will man die Lektüre literarischer Texte vorbereiten, dann könnten im Vorfeld lexikalische Einheiten eingeführt und geübt werden. Auf der Basis einer

Textdeckungsuntersuchung ließen sich, als Advanced organizer vor dem Lesen, Wortschatzübungen erstellen. Es stellt sich die Frage, ob man die Leserinnen und Leserauf die häufigen oder gerade die weniger häufigen Wörter in einem Text vorbereiten sollte; dies ist eine didaktische Entscheidung, die, wenn begründet, in beiden Fällen den Lernprozess unterstützen kann. Unten ist das Beispiel einer Wortschatzübung wiedergegeben, die auf die häufigsten Wörter in *Agnes* von Peter Stamm vorbereitet. Die Häufigkeit des Wortes im Text wird durch Klammern angegeben, danach folgen die Übersetzung ins Niederländische und die deutsche Umschreibung. Im Lückentext sind diese Wörter auszufüllen:

betrunken <8x> (drunken) 2x: von Alkohol berauscht und beeinträchtigt sein
das Arbeitszimmer / die Arbeitszimmer <7x> (studeerkamer): der Raum, vor allem in einer Wohnung, in dem man geistige Arbeit verrichtet
das Stockwerk / die Stockwerke <6x> (verdieping): die Etage eines Gebäudes
bleich (bleek) <5x>: sehr blass oder gar weiß
die Klimaanlage / die Klimaanlagen (airco) <4x>: ein Gerät, das Räume kühlt

Hans hatte an diesem Abend wieder einmal einen über den Durst getrunken. So kam es, dass er _____ in seinem _____ saß. „Da sitzt du nun, du _____ alter Mann“, sagte er zu sich selbst und seufzte. Er blickte aus dem Fenster. Seine Wohnung lag im neunten _____, so konnte er über alle Häuser in der Umgebung hinwegsehen. Er liebte diesen Ausblick. Hans ließ seine Gedanken schweifen und musste an seine erste Freundin denken. Was für eine schöne Zeit war das doch gewesen. Doch plötzlich wurde Hans ein wenig unwohl. Er wurde _____ im Gesicht und kalt war ihm auch. Er lief zu dem Schalter an der Wand, der die _____ steuerte und schaltete sie aus. Hoffentlich sollte ihm so schnell wärmer werden. Nun war es aber wirklich an der Zeit ins Bett zu gehen. Hans versprach sich, in Zukunft weniger zu trinken, denn dann wurde er immer so furchtbar melancholisch.

5 Schlussbetrachtung

Aus den Analysen der 26 ausgewählten literarischen Titel geht hervor, dass für eine sinnvolle Leseerfahrung optimale lexikalische Deckungsgrade nicht erreicht werden. Dennoch bilden Nations Kriterien (2001: 150) für intensives Lesen (< 95%) und extensives Lesen (95–98%) eine gute Orientierung für den didaktischen Einsatz der analysierten Titel. Besonders im DaF-Unterricht, aber auch im muttersprachlichen Unterricht, ist die Niveaubestimmung der Titel aus bildungssprachlichem und didaktischem Gesichtspunkt relevant, damit der Leseprozess und der Lernerfolg optimal gestaltet werden können.

Die Analysen geben Anlass, die GER-Zuordnungen einiger Titel zu überprüfen und sie im literaturdidaktischen Katalog auf www.lezenvoordelijst.nl anzu-

passen, weil sie auf dem angedeuteten Niveau unterhalb der Grenze von 85 % Textdeckung liegen. Das bedeutet, dass Textdeckungsuntersuchungen zur Präzisierung des literaturdidaktischen Instruments beitragen können. Die Ergebnisse ermöglichen genauere Angaben über Wortschatzschwierigkeiten in diesen Texten und auch profunde Aufgaben im lexikalischen Bereich. Durch die Erforschung der Textdeckung wurde herausgearbeitet, welcher Lexik didaktische Aufmerksamkeit geschenkt werden sollte, damit der fehlende Prozentanteil bei der Textdeckung keine Lesehürde mehr bedeutet. Freilich wird Textverstehen nicht nur durch lexikalische Kenntnisse, sondern auch durch andere sprachliche Fertigkeiten, Metakognition und Weltwissen angeregt.

Ersichtlich wurde am Beispiel des Jugendromans *Der Klassenfeind und ich*, dass Titel, denen ein Glossar angehängt ist, einen höheren Schwierigkeitsgrad haben, wenn man die Erläuterungen mitberücksichtigt. Das Beispiel einer vereinfachten Edition von *Im Westen nichts Neues* zeigte, dass zumindest lexikalisch nicht unbedingt die Zugänglichkeit erniedrigt wird. Außerdem konnte am Beispiel von Peter Stamm's Roman *Agnes* gezeigt werden, dass fremdsprachlicher Wortschatz, der im Roman aufgenommen wird, nicht zwangsläufig eine Hürde bildet.

An dieser Stelle ist zu erörtern, dass die bei der Textdeckung angestrebten Prozentzahlen von 95 % und 98 % aus unterschiedlichen didaktischen Gründen zwar besonders lohnenswert sind, sie aber im Hinblick auf Sprachverwandtschaft noch zu anderen Überlegungen führen. Wenn beispielsweise in *Wie Licht schmeckt*, einem Roman mit hohem Deckungsgrad, Wörter wie *bleich*, *bluten*, *Besteck*, *Bedienung*, *Surflehrer* nicht auf den Häufigkeitslisten bis 5000 vorkommen, so dürfte klar sein, dass diese für niederländische Leserinnen und Leser durch die Sprachnähe unproblematisch sind und dass Wörter wie *Bikini*, *Gangster* und *Cornetto* international bekannt sind. Ähnlich spielen Vorteile der Sprachnähe in *Der Hals der Giraffe* eine Rolle (*Applaus*, *Bauernkind*, *Testgebiet*, *Unterrock*, *Volltreffer*, *Wildschwein*) und auch Begriffe aus dem Englischen (*Trenchcoat*) und aus der Gesellschaft (*Homosexualität*) sind für Leser eines solchen Romans nicht problematisch. Das bedeutet, dass die genannten Deckungsgrade noch ein wenig, schätzungsweise zwischen ein und fünf Prozent, höher sind. Ebenfalls bewirken auch Komposita Vorteile. Am Beispiel von *Wie Licht schmeckt* lässt sich zeigen, dass *Balkontür* nicht nur durch Sprachverwandtschaft verständlich ist, sondern auch, weil *Tür* auf der Liste bis 1000 und *Balkon* auf der Liste bis 5000 vorkommt.

Was diese Studie nicht überprüft hat, ist, ob alle Wörter der 5000er Liste vorkommen und auch wiederholt vorkommen, wenn alle 26 oder gar weniger Titel gelesen sind. Dies wäre noch zu leisten. Es scheint aber nicht üblich, dass Schülerinnen und Schüler so viele literarische Texte in einer Fremdsprache lesen. Der Wortschatz wird auf vielen Wegen aufgebaut, nicht nur durch literarisches Lesen.

Gezeigt werden konnte, dass der Lesbarkeitsindex „LIX“ nützliche Einsichten in Wort- und Satzlänge erlaubt, und ein brauchbares heuristisches Instrument ist.

Mit vorliegender Studie wurde erstmals die Wortschatzdeckung in deutschen literarischen Texten untersucht. Dies ist ausbaufähig. Zum einen wären mehrere der ausgewählten Texte noch näher zu untersuchen, zum anderen müsste neben diesem Grundstock die Sammlung noch erweitert werden. Insbesondere könnten leichte Lektüren-Editionen überprüft, aber auch zusammengestellt werden. Im Allgemeinen ist aus der lexikalischen Beschäftigung auch für Literaturdidaktiker die Erkenntnis zu gewinnen, dass Wortschatzfrequenz bei der Lektüre eine Rolle spielt und die Kooperation mit Sprachwissenschaftlern sich lohnen kann. Tut man dies nicht, dann tritt das ein, was im modernen Klassiker *Tschick* passiert: „Man sollte meinen, wenn man immer abwechselnd rechts und links fährt, könnte man nicht im Kreis fahren, aber wir schafften es.“ (Herrndorf 2010: 107).

Literatur

- Bollwahn, Barbara (2009): *Der Klassenfeind und ich*. Hamburg: Carlsen.
- Herrndorf, Wolfgang (2010): *Tschick*. Berlin: Rowohlt.
- Jones, Randall J.; Tschirner, Erwin (2006): *Frequency Dictionary of German: Core Vocabulary for Learners*. London: Routledge.
- Laufer, Batia; Ravenhorst-Kalovski, Geke C. (2010): „Lexical Threshold Revisited, Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension“. In: *Reading in a Foreign Language*, 22, 15–30.
- Lutjeharms, Madeline (2010): „Vermittlung der Lesefertigkeit“. In: Krumm, Hans-Jürgen; Fandrych, Christian; Hufeisen, Britta; Riemer, Claudia (Hrsg.): *Deutsch als Fremd- und Zweitsprache: Ein internationales Handbuch*. Berlin: De Gruyter Mouton, 976–986.
- Nakanishi, Takayuki (2015): „A Meta-Analysis of Extensive Reading Research“. In: *TESOL Quarterly*, 49, (1), 6–37.
- Nation, Paul; Wang-Ming-tzu (1999): „Graded Readers and Vocabulary“. In: *Reading in a Foreign Language*, 12 (2), 355–380.
- Nation, Paul (2001): *Learning Vocabulary in Another Language*. Cambridge: CUP.
- Nation, I. S.P. (2006): „How Large a Vocabulary Is Needed For Reading and Listening?“ In: *The Canadian Modern Language Review / La Revue canadienne des langues vivantes*, 63 (1), 59–82.
- Rösler, Dietmar (2012): *Deutsch als Fremdsprache: Eine Einführung*. Stuttgart: Metzler (Sammlung Metzler 280).
- Schmitt, Norbert; Jiang, Xiangying; Grabe, William (2011): „The Percentage of Words Known in a Text and Reading Comprehension“. In: *The Modern Language Journal*, 95, 26–43.
- Tschirner, Erwin (2008a): „Das professionelle Wortschatzminimum im Deutschen als Fremdsprache“. In: *Deutsch als Fremdsprache* 45, 195–208.
- Tschirner, Erwin (2008b): *Grund- und Aufbauwortschatz Deutsch als Fremdsprache nach Themen*. Berlin: Cornelsen.

- van der Knaap, Ewout (2014): „Möglichkeiten eines literaturdidaktischen Kompetenzmodells für den fremdsprachlichen Unterricht“. In: *Deutsch als Fremdsprache*, 51, 215–225.
- Yamashita, Junko (2015): „In Search of the Nature of Extensive Reading in L2: Cognitive, Affective, and Pedagogical Perspectives“. In: *Reading in a Foreign Language*, 27 (1), 168–181.

Biographische Angaben

Ewout van der Knaap

hat die Associate Professur für neuere deutsche Literatur an der Universität Utrecht. Er führt den literaturdidaktischen Katalog Deutsch auf www.lezenvoordelijst.nl und kooperiert in dem Erasmus+Projekt „Lehrkompetenzentwicklung für Extensiven Leseunterricht“ (www.leelu.eu). Veröffentlichungen über deutschsprachige Literatur, über erinnerungskulturelle Prozesse und über Literaturdidaktik.