

Allgemeiner Beitrag

Jupp Möhring* und Olaf Bärenfänger*

Hochschulzugangsprüfungen und die Studienrealität: Eine empirische Untersuchung zu Lese- und Wortschatzanforderungen in der Studieneingangsphase

University Entrance Tests and the College World: A corpus-based study on lexical requirements in textbooks compared to linguistic requirements in university entrance tests

<https://doi.org/10.1515/infodaf-2018-0073>

Zusammenfassung: In diesem Beitrag werden die Leseanforderungen in der Studieneingangsphase am Beispiel der Textsorte *Lehrbuch* im Vergleich zu ausgewählten Hochschulzulassungsprüfungen (DSH, telc C1 Deutsch Hochschule und TestDaF) untersucht. Nach einem Forschungsüberblick wird eine korpusbasierte Studie zur lexikalischen Komplexität der untersuchten Texte vorgestellt. Abschließend werden Schlussfolgerungen für die Entwicklung studienvorbereitender und studienbegleitender Curricula diskutiert.

Schlüsselwörter: Sprachtest, Wortschatz, Lesekompetenz, Korpuslinguistik

Abstract: In this article, we will present a corpus-based analysis of the reading proficiency levels required in three exams as operationalized by various measures of lexical complexity (DSH, telc C1 Deutsch Hochschule, and TestDaF). These results will be compared to the difficulty levels of actual reading samples from textbooks used by undergraduate students in three subjects (chemistry, econo-

*Kontaktpersonen: Jupp Möhring, E-Mail: jupp.moehring@uni-leipzig.de
Olaf Bärenfänger, E-Mail: baerenfaenger@uni-leipzig.de

mics, and medicine). Some implications for language curricula will also be discussed.

Keywords: Language Assessments, Vocabulary, Reading Competence, Corpus Linguistics

1 Einleitung

Prüfungen wie die *Deutsche Sprachprüfung für den Hochschulzugang* (DSH), der *Test Deutsch als Fremdsprache* (TestDaF) oder *telc C1 Deutsch Hochschule*, die bei nicht-muttersprachlichen Studienbewerbern die sprachlichen Voraussetzungen für die Aufnahme eines Studiums erheben, spielen für Bildungssysteme eine essenzielle Rolle. Sie erlauben es den Hochschulen zum einen, solche Studienbewerber zu identifizieren, die den sprachlichen Anforderungen eines Hochschulstudiums gewachsen sind und aus diesem Grund ihr Studium voraussichtlich erfolgreich zu Ende bringen werden. Zum anderen verbindet sich mit den genannten Prüfungen ein umfassendes Netz an studienvorbereitenden Kursen, die nicht-muttersprachliche Bewerber sprachlich für ein Studium an einer deutschen Hochschule qualifizieren. Selbstverständlich haben Prüfungen, die studienbezogene Sprachkompetenzen zum Gegenstand haben, massiven Einfluss auf Art und Umfang dieser Kurse ebenso wie auf die in ihnen verwendeten Lehr- und Lernmaterialien. Neben dieser ausgeprägten didaktischen Wirksamkeit weisen Hochschulzugangsprüfungen auch eine ökonomische Dimension auf. Erfüllen Hochschulzugangsprüfungen ihre Funktion, trägt dies zumindest potenziell zur Effektivität der Hochschulbildung bei. Im umgekehrten Fall gelangen sprachlich nicht ausreichend qualifizierte Studienbewerber an die Hochschulen, die ihr Studium ggf. nicht abschließen. Die öffentlich finanzierten Hochschulen vergeuden dann ohnehin schon knappe Ressourcen, ebenso die Studierenden selbst im Falle ausbleibenden Studienerfolgs.

Welche Studienbewerber ihre sprachlichen Kompetenzen vor der Aufnahme eines Studiums überprüfen lassen müssen, ist durch die „Rahmenordnung über Deutsche Sprachprüfungen für das Studium an deutschen Hochschulen (RO-DT)“ klar geregelt. § 1 legt fest: „Von Studienbewerberinnen und Studienbewerbern, die ihre Studienqualifikation nicht an einer deutschsprachigen Einrichtung erworben haben, werden deutsche Sprachkenntnisse verlangt, die zum Studium an einer Hochschule befähigen (sprachliche Studierfähigkeit). Der Nachweis der Sprachkenntnisse ist nach Landesrecht eine Voraussetzung für die Zulassung oder Einschreibung zum Studium“ (HRK/KMK 2015: 2). Die Rahmenordnung legt zugleich fest, welche Instrumente zur Überprüfung studienbezogener Deutsch-

kenntnisse zugelassen sind. Es handelt sich hierbei um die DSH, den TestDaF, den Prüfungsteil Deutsch in den Feststellungsprüfungen an Studienkollegs sowie das Deutsche Sprachdiplom Stufe 2 der Kultusministerkonferenz. Eine Reihe weiterer Prüfungen befreien vom Ablegen einer der vier genannten Tests: das Große Deutsche Sprachdiplom (Goethe-Institut), das Goethe-Zertifikat C2, die Zentrale Oberstufenprüfung (Goethe-Institut), das Kleine Deutsche Sprachdiplom (Goethe-Institut) sowie telc C1 Deutsch Hochschule. Ebenfalls entbinden laut Beschluss der Kultusministerkonferenz vom 02.06.1995 einige ausländische Prüfungen nicht-muttersprachliche Studienbewerber vom Ablegen der Hochschulzugangsprüfungen. Schließlich liegt es im Ermessen der einzelnen Hochschulen, bei Studienbewerbern auf das Ablegen der Hochschulzugangsprüfungen zu verzichten, etwa wenn diese bereits über ein abgeschlossenes Germanistikstudium im Ausland verfügen oder sich nur befristet an der Hochschule aufhalten, ohne dabei einen formellen Abschluss anzustreben (z. B. Erasmus-Studierende).

Der großen Bedeutung und Wirksamkeit der Hochschulzugangsprüfungen geschuldet, räumt die Rahmenordnung der Qualitätssicherung dieser Prüfungen einen großen Stellenwert ein. § 9 spezifiziert etwa: „Sie [die HRK und die KMK; die Autoren] geben zur Sicherung der Qualität und Vergleichbarkeit Empfehlungen, insbesondere zur Prüfungserstellung, Prüfungsdurchführung und Prüfungskorrektur“ (HRK/KMK 2015: 6). Die genannten Empfehlungen sind dabei für die Testanbieter und die von ihnen lizenzierten Prüfungszentren bindend.

Das konsequente Bemühen um die Qualität der Hochschulzugangsprüfungen ist in diesem Zusammenhang keineswegs Selbstzweck. Im Jahr 2016 waren 251 542 ausländische Studierende an deutschen Hochschulen immatrikuliert (vgl. Chehab et al. 2017). Aus der Gruppe der ausländischen Bachelor-Studierenden brachen im Bezugsjahr 2012 41 % ihr Studium ab (Heublein et al. 2014: 10); bei den deutschen Bachelorstudierenden war dies nur bei 28 % der Fall (ebenda: 3). Die Gründe für einen Studienabbruch sind komplex (vgl. Kliegl/Müller 2012), jedoch gaben bei der 20. Sozialbefragung des Deutschen Studentenwerks mit 985 Teilnehmenden (Apolinarski/Poskowsky 2013) immerhin 32 % der ausländischen Studierenden an, große oder sehr große Schwierigkeiten bei der Verständigung auf Deutsch zu haben. Auch der Kontakt zu deutschen Hochschullehrern stellt 28 % der Befragten vor Probleme, ebenso wie der Kontakt zu deutschen Studierenden (41 %) und zur deutschen Bevölkerung im Allgemeinen (37 %). Aus der hohen Zahl der Studienabbrüche resultieren gravierende Folgen für eine große Zahl individueller Bildungsbiographien. Gleichzeitig ist der volkswirtschaftliche Schaden beträchtlich, wenn in etwa 100 000 Studierende die Hochschulen ohne Abschluss verlassen. Auch dem Arbeitsmarkt stehen somit nicht die Kräfte zur Verfügung, die er dringend benötigt.

Sicher wäre es überspitzt, allein die Hochschulzugangsprüfungen für hohe Studienabbruchquoten verantwortlich zu machen. Gleichwohl stellt sich die empirische Frage, ob die bestehenden Prüfungen denn in ausgewogener Weise tatsächlich diejenigen sprachlichen Anforderungen abbilden, denen sich Studierende in den ersten Semestern typischerweise ausgesetzt sehen. Dies stellt zweifellos eine Hauptanforderung an eine Hochschulzugangsprüfung dar (vgl. Marks 2015: 21). Bereits vor geraumer Zeit hatte Mehlhorn (2005: 7) moniert, die sprachlichen Anforderungen eines Hochschulstudiums gingen meist deutlich über das hinaus, was in den sprachlichen Zulassungsprüfungen abgetestet wird. Zu bedenken ist in diesem Zusammenhang zudem, dass die sprachlichen Anforderungen je nach Studienfach variieren (vgl. Zimmermann 2010: 2).

Vor diesem Hintergrund unternimmt der vorliegende Beitrag den Versuch, am Beispiel von studienbezogener Lesekompetenz und anhand der Hochschulzugangsprüfungen DSH, telc C1 Deutsch Hochschule und TestDaF abzugleichen, inwieweit die sprachlichen Anforderungen im Studium möglichst genaue Entsprechungen in den jeweiligen Prüfungen finden. Im Fokus steht dabei der Wortschatz, der eine wesentliche Komponente der Fertigkeit Lesen darstellt (vgl. Schmitt et al. 2015).

Im folgenden Abschnitt wird der aktuelle Forschungsstand zu den sprachlichen Anforderungen in der Studieneingangsphase referiert. Nachfolgend wird das Forschungsdesign der vorliegenden Studie vorgestellt, im Rahmen dessen der Wortschatz in den Prüfungsteilen Lesen der DSH, von telc C1 Deutsch Hochschule und von TestDaF zunächst mit verschiedenen, allgemeinsprachlichen, korpusbasierten Worthäufigkeitslisten abgeglichen und mit in der Studieneingangsphase relevanten Lehrwerkstexten verglichen wurde. Um auch Aussagen über den spezifisch studiersprachlichen Wortschatz treffen zu können, wurden die analysierten Texte überdies dem ITT Textkorpus Bildungssprache gegenübergestellt, das vom Leipziger Institut für Testforschung und Testentwicklung auf der Grundlage einer umfassenden Sprachbedarfsanalyse zusammengestellt worden war (vgl. Bärenfänger/Lange/Möhring 2016). Hierbei lag im Rahmen der vorliegenden Untersuchung der Fokus auf korpusbasiert ermittelten, bildungssprachlich bedeutsamen Mehrwortverbindungen und deren Auftreten in den benannten Sprachprüfungen und Lehrwerkstexten. Ein letzter Abschnitt zeigt Schlussfolgerungen und Desiderate aus der empirischen Untersuchung auf.

2 Theoriebildung und empirische Erkenntnisse zu sprachlichen Anforderungen im Studium

Forschungen zu den sprachlichen Anforderungen eines Hochschulstudiums lassen sich einordnen in die Diskussion der Bildungswissenschaften und der Linguistik um das Deutsche als Bildungssprache. Diese erstreckt sich über alle Sektoren des Bildungssystems (vgl. dazu beispielsweise die Beiträge in Becker-Mrotzek et al. 2013; Gogolin et al. 2013; Tschirner/Bärenfänger/Möhring 2016).

Das Konzept der Bildungssprache ist dabei klar abzugrenzen von der *Fachsprache*, die gemäß der kanonischen Definition von Hoffmann (1984: 53) die „Gesamtheit aller sprachlichen Mittel“ umfasst, „die in einem fachlich begrenzten Kommunikationsbereich verwendet werden, um die Verständigung zwischen den in diesem Bereich tätigen Menschen (und die Popularisierung der fachlichen Inhalte sowie den Kontakt zu bestimmten Nicht-Fachleuten) zu gewährleisten“. Im Unterschied dazu kommt der *Bildungssprache* eine dezidiert epistemische Rolle zu, insofern sie nämlich untrennbar mit Lernprozessen verbunden ist (vgl. Feilke 2012). Betrachtet werden kann *Bildungssprache* dabei gemäß einer Klassifikation von Prediger (2013) als *Lernmedium*, innerhalb dessen Lernprozesse stattfinden; demzufolge als *Lernvoraussetzung*; nicht zuletzt aber auch als *Lerngegenstand*, da sich bildungssprachliche Kompetenzen im Rahmen von Lernprozessen sukzessive erweitern (sollen). Diese Perspektive unterscheidet *Bildungssprache* auch vom Konzept der *Wissenschaftssprache*, die gemäß der kanonischen Definition von Benes (1981: 186) gewissermaßen den Querschnitt aller wissenschaftlichen Fachsprachen bildet. Mit ihrem spezifischen Fokus auf dem Lernen beschäftigt sich der überwiegende Teil der Publikationen zur Bildungssprache mit der Modellierung und Beschreibung des Deutschen im Bereich des schulischen Lernens (vgl. z. B. Geist et al. 2017). Obschon für den Bereich der Hochschulbildung noch keine Forschungen zum Zusammenspiel von bildungssprachlicher Kompetenz und Bildungserfolg vorliegen, ist dieser Zusammenhang für den schulischen Bereich gut belegt (Gantefort/Roth 2010; Reich 2013).

Im vorliegenden Beitrag leitend ist die Frage, welche sprachlichen Funktionen und Mittel Studierenden zu Gebote stehen müssen, um erfolgreich an studienbezogenen Lernprozessen teilzunehmen, oder, anders gesagt: um ihr Studium erfolgreich zu absolvieren. Wenn auch eine souveräne Beherrschung des Englischen als Studier- und Wissenschaftssprache für zahlreiche Studiengänge zweifellos von großer Bedeutung ist, steht im Brennpunkt des vorliegenden Beitrags das Deutsche als Studier- und Wissenschaftssprache. Trotz der großen Bedeutung der Frage nach dem Sprachbedarf mit weitreichenden Konsequenzen für sämtliche studienpropädeutischen Sprachkurse, Lehrwerke sowie die hier im

Fokus stehenden Studieneingangsprüfungen steckt die Analyse und Beschreibung des Deutschen als Bildungssprache im Kontext der Hochschulbildung noch in den Anfängen. Nachfolgend werden wichtige Erkenntnisse zum Deutschen als Studiersprache referiert, wobei zunächst auf Ratgeber bzw. Lehr-, Lern- und Übungsbücher eingegangen wird; dann auf linguistisch-empirische Untersuchungen zum Deutschen als Studiersprache; sowie schließlich auf empirische Analysen zum Sprachbedarf von Studienanfängern.

2.1 Ratgeber, Lehr-, Lern- und Übungsbücher

Mit eher praktischen Zielsetzungen steht Studierenden eine Reihe von Ratgebern sowie Lehr-, Lern- und Übungsbüchern zur Verfügung, die verschiedenste konkrete Hilfestellungen in Zusammenhang mit Sprache im Studium leisten sollen (vgl. für eine Übersicht Bärenfänger et al. 2016: 14 ff.). Im Vordergrund dieser Veröffentlichungen stehen dabei Aspekte des wissenschaftlichen Stils (z. B. Nominalstil, Ich-Tabu, Tempuswahl), Besonderheiten studienbezogener Textsorten (z. B. Seminararbeiten, studentische Referate), von den Studierenden zu bewältigende Sprachhandlungssituationen (z. B. Kontakte knüpfen, Beteiligung an Lehrveranstaltungen, Sprechstundengespräche) sowie Strategien und Techniken des wissenschaftlichen Schreibens (z. B. wissenschaftliches Arbeiten, korrektes Zitieren; vgl. zu diesen Aspekten Bünting/Bitterlich/Pospiech 2006; Moll/Thielmann 2017). Wenn sich zwar auch Hinweise zur gesprochenen Studiersprache finden, so überwiegen in den betrachteten Publikationen doch Texte zur geschriebenen Studiersprache. Die in akademischen Sprachverwendungskontexten mindestens genauso wichtige mündliche Studiersprache, wie sie in Lehrveranstaltungen wie Vorlesungen und Seminaren, bei Partner- und Gruppenarbeiten, bei mündlichen Präsentationen oder in Prüfungsgesprächen in Erscheinung tritt, bleibt jedoch mit wenigen Ausnahmen unberücksichtigt (Bärenfänger et al. 2016). Auch die mittlerweile omnipräsenten digitalen Textsorten und Strategien der Informationsbeschaffung oder der Arbeitsorganisation mit Hilfe digitaler Medien werden bestenfalls marginal thematisiert (Bärenfänger et al. 2016). Nicht zuletzt finden für das Studium zentrale Textsorten wie Vorlesungsfolien, Vorlesungsskripte, fachbezogene Lehrbücher und Klausuren keine systematische Erwähnung (Bärenfänger et al. 2016). Ein weiterer Kritikpunkt betrifft den Umstand, dass die Beschreibung der Studiersprache in Ratgebern resp. Lehr-, Lern- und Übungsbüchern fast durchweg auf intuitiver Grundlage erfolgt (Bärenfänger et al. 2016: 33). Auf diese Weise kann es keinesfalls als gesichert gelten, dass entsprechende Beschreibungen des Deutschen als Studiersprache tatsächlich eine Entsprechung in der sprachlichen Realität haben. Zusammenfassend lässt sich für die Ratgeberlitera-

tur sowie die Lehr-, Lern- und Übungsbücher festhalten, dass diese bestenfalls eine sehr lückenhafte Beschreibung des Deutschen als Studiersprache erbringen, zumal ohne verlässliche empirische Basis. Als Grundlage für bedarfsgerechte Kurscurricula, studienvorbereitende Lehr- und Lernmaterialien oder Studienzulassungsprüfungen sind sie mithin ungeeignet.

2.2 Linguistisch-empirische Untersuchungen zum Deutschen als Studiersprache

Empirische Untersuchungen zum Deutschen als Studiersprache sind vergleichsweise rar und beziehen sich in der Regel auf kleine Ausschnitte der sprachlichen Wirklichkeit, ohne dabei ein zusammenhängendes Bild zu vermitteln (vgl. Bärenfänger et al. 2016: 34 ff.). Mit Blick auf die Wissenschaftssprache liegen beispielsweise einige Publikationen zu den Merkmalen von Fachartikeln vor (vgl. z.B. Fandrych 2002; Szurawitzki 2015; Busch-Lauer 2012; Thielmann 2007), die jedoch über dem studiersprachlichen Niveau liegen und von Studienanfängern nicht in allen Fächern rezeptiv oder gar produktiv beherrscht werden müssen.

Die demgegenüber fraglos im Bereich der Studiersprache angesiedelten Haus- bzw. Seminararbeiten sind vergleichsweise wenig empirisch erforscht (Stezano Cotelo 2006). Graefen (2002) hatte korpusbasiert typische Fehler bei der Erstellung von Seminararbeiten identifiziert. Demgegenüber standen bei Stezano Cotelo (2006, 2008) Techniken im Vordergrund, mit Hilfe derer Studierende in Seminararbeiten auf Wissen in anderen Quellen referieren. In der umfänglichen Monographie von Pohl (2007) steht das Verfassen von Seminararbeiten als komplexer, gestufter Entwicklungsprozess im Zentrum. Alle drei genannten Publikationen zeigen auf der Grundlage relativ kleiner Textkorpora auf, in welchen Bereichen Studierende auf Schwierigkeiten beim Verfassen von Seminararbeiten treffen. Die gleichfalls bislang kaum empirisch untersuchte Textsorte Mitschrift (vgl. Steets 2003: 51ff.) wird Untersuchungsgegenstand des seit Juli 2017 an den Universitäten Leipzig und Würzburg angesiedelten Projekts „SpraStu – Sprache und Studienerfolg bei Bildungsausländer/-innen“ (vgl. Wisniewski 2018) sein, in dessen Rahmen neben zahlreichen anderen Erhebungen auch ein (Lerner-)Korpus studentischer Mitschriften angelegt und ausgewertet werden wird. Steets (2003: 52) hebt für die Mitschrift die Eigenschaften „Schnelligkeit des Produktionsprozesses“ und den „Charakter des Provisorischen“ (ebd.) hervor, wodurch nicht zuletzt empirische Untersuchungen im Vergleich zu stärker standardisierten Textsorten deutlich aufwendiger sind. Sicherlich tragen diese und ähnlich gelagerte Hürden auch für die bislang nur unvollständige empirische Beschreibung der gesprochenen Studiersprache bei (vgl. Fandrych/Meißner/Wallner

2017). Eine frühe Studie von Wiesmann (1999) beinhaltet hierzu die Transkription und Analyse von 26 Lehrveranstaltungen unterschiedlichen Typus (Seminar, Übung, Laborpraktikum) aus den Fächern Psychologie, Spanisch, Interkulturelle Kommunikation, Geschichte, Ökonomie, Mathematik, Physik, Biologie, Chemie, die sie noch durch weitere bereits veröffentlichte Transkripte aus der Chemie ergänzte. Auf dieser Datenbasis identifizierte Wiesmann (1999) typische Sprachhandlungsmuster wie Assertieren, Aufgaben-Stellen/Aufgaben-Lösen, Begründen, Einschätzen, Einwenden, Erklären, Erläutern, Frage/Antwort (Lehrerfrage, Nachfragen/Nachhaken) und Vorschläge. Der Frage, wie sich diese Handlungen bildungssprachlich manifestieren, war Wiesmann (1999) jedoch nicht vorrangig nachgegangen. Crawford Camiciottoli (2010) untersuchte korpusbasiert englischsprachige Vorlesungen und identifizierte in diesem Zusammenhang typische studentische Verstehensprobleme. Im Rahmen des *euroWiss*-Projektes waren zudem kontrastiv für das Italienische und das Deutsche Textkorpora aufgebaut worden, die unterschiedliche Aspekte der Wissensvermittlung an der Hochschule zu untersuchen erlauben (vgl. z.B. Redder/Heller/Thielmann 2014). Im Fokus der Untersuchungen stand dabei nicht zuletzt die Frage, wie Forschungswissen in den Lehrveranstaltungen naturwissenschaftlicher Fächer (vgl. z.B. Thielmann 2014a, b), in der Mathematik (Krause 2015; Wagner 2014) und in der Germanistik (Heller 2014) vermittelt wird. Im Rahmen des *GeWiss*-Projekts steht gesprochene Wissenschaftssprache im Vordergrund. Mit dem Fokus auf die beiden prominenten Textsorten Vortrag (Studierender und Dozierender) und Prüfungsgespräch werden kontrastiv für das Deutsche, Englische und Polnische sprachliche Routinen gebunden an die jeweiligen Kommunikationssituationen analysiert (vgl. Fandrych/Meißner/Slavcheva 2014). Im Rahmen des Projekts wurden neben den sprachwissenschaftlichen Analysen im engeren Sinne auch methodische Fragen systematisch bearbeitet sowie die eingangs erwähnten Hürden benannt und mögliche Lösungsansätze, etwa zur orthografischen Normalisierung oder der Wortartenannotation, vorgestellt (vgl. Fandrych/Meißner/Wallner 2017).

In ihrer Würdigung des empirischen Forschungsstands bemängeln Bärenfänger et al. (2016: 51) die in den meisten Fällen unzureichende Datenbasis. Oft handele es sich um relativ kleine, in der Regel nicht frei zugängliche Textkorpora. Dies macht es unmöglich, die Schlussfolgerungen der Autoren zu überprüfen und die Ergebnisse auf allgemeinere Zusammenhänge hin zu übertragen. Da die vorhandenen empirischen Untersuchungen zudem nur einzelne Facetten der studien sprachlichen Realität beleuchten und kein zusammenhängendes Gesamtbild ergeben, sind sie als Grundlage für die Planung von studienpropädeutischen Curricula oder Lehr- und Lernmaterialien als noch nicht ausreichend zu betrachten. Da sie die Studienwirklichkeit nicht repräsentativ erfassen, lassen sie sich

ebenfalls nicht für die Erstellung von Prüfungen heranziehen, die die für die Aufnahme eines Studiums erforderlichen Kompetenzen erheben sollen.

2.3 Bedarfsanalysen zu den sprachlichen Anforderungen eines Hochschulstudiums

Unter der Fragestellung, inwieweit Hochschulzugangsprüfungen wie die DSH die sprachliche Wirklichkeit eines Hochschulstudiums abbilden, hatten Wollert/Zschill (2017) exemplarisch einen Lesetext aus einer DSH-Prüfung untersucht. Sie kamen diesbezüglich für die untersuchte DSH zu der Schlussfolgerung, dass „die vorgefundenen sprachlichen Mittel nicht repräsentativ für sprachliche Mittel und Handlungen eines wissenschaftlichen Textes“ seien und deshalb aus dessen Bearbeitung „auch keine spezifischen Rückschlüsse auf eine sprachliche Studierfähigkeit abgeleitet werden“ könnten (10). Auch wenn hier lediglich ein einziger Prüfungstext aus einer einzigen DSH Gegenstand der Betrachtung war, so ergeben sich doch Zweifel an der Konstruktvalidität der untersuchten DSH. Vor diesem Hintergrund fordern Wollert/Zschill nachdrücklich ein empirisch fundiertes Konzept der sprachlichen Studierfähigkeit (16).

Auch hinsichtlich umfassenderer Bedarfsanalysen zu den sprachlichen Anforderungen eines Hochschulstudiums liegen bislang nur wenige belastbare Befunde vor. Eine Studie fand im Umfeld des TestDaF statt. So wurden beispielsweise systematisch die Webseiten von 25 deutschen Hochschulen darauf hin analysiert, welche Prüfungsformate verbreitet sind und welche Hinweise sich auf sprachliche Anforderungen finden. Bei der Betrachtung von 100 Studiengängen zeigte sich, dass Klausuren, Seminararbeiten, mündliche Prüfungen sowie mündliche Referate bzw. Präsentationen die gängigsten Prüfungsformate an Hochschulen darstellen (Arras 2012: 141). Sprachbezogene Anforderungen an Studierende seien „analysieren, erläutern, präsentieren, kommentieren und argumentieren“ (ebd.). Diese Ergebnisse wurden für die Entwicklung eines Interviewleitfadens verwendet, der bei Gesprächen mit 38 deutschen und ausländischen Studierenden sowie mit Hochschullehrenden angewendet wurde (Marks 2015: 23). Diese Daten wiederum wurden zur Erstellung eines Fragebogens zu den sprachlichen Anforderungen eines Hochschulstudiums herangezogen, den 1327 ausländische Studierende beantworteten. Von Seiten der Hochschullehrenden ergab sich ein Rücklauf von 120 Fragebögen. Von den Studierenden wurden folgende Aktivitäten als für ein Hochschulstudium nötige Sprachhandlungen genannt: Vorlesung hören; in einer Lehrveranstaltung Notizen machen; Material für eine Lehrveranstaltung vor- bzw. nachbereiten (lesen); Fachkommunikation mit Mitstudierenden; private schriftliche Kommunikation (SMS, E-Mail mit Mitstudierenden); All-

tagskommunikation mit Mitstudierenden; Flyer/Aushänge lesen; Fachliteratur lesen und im Seminar besprechen; während einer Vorlesung Vorlesungsskript ergänzen; Kommunikation in der Hochschulverwaltung (Marks 2015: 24). In teilweiser Übereinstimmung hiermit erachteten die befragten Hochschullehrenden die folgenden Aktivitäten für zentral: Fachvorträge/Vorlesungen verfolgen/zuhören; Material zur Vorbereitung einer Veranstaltung lesen; in Vorlesung/Seminar mitschreiben, Notizen machen; einem Vortrag folgen anhand eines Handouts, einer PowerPoint-Präsentation o. Ä.; sich im Seminar aktiv mündlich beteiligen, an Fachdiskussionen teilnehmen; Aufgaben in einer Lern- oder Arbeitsgruppe bearbeiten und präsentieren; Fachliteratur lesen und im Seminar besprechen; in einer Vorlesung zuhören, dabei Skript lesen und ergänzen (Marks 2015: 24). Die Studien von Arras (2012) und Marks (2015) nähern sich dem Ziel, die sprachlichen Anforderungen eines Hochschulstudiums empirisch zu beschreiben, auf eindrückliche Weise. Erstmalig liegen belastbare Informationen über typische Leistungsnachweise und Gruppen von sprachlichen Aktivitäten vor. Wollert/Zschill (2017) bemerken allerdings völlig zu Recht, dass es sich bei den aufgezählten sprachlichen Aktivitäten nicht um linguistische Kategorien im pragmatisch-linguistischen Sinne handelt. Die genannten Aktivitäten erlauben lediglich Auskunft über *Klassen* von Sprachhandlungen, aber nicht über den realen Sprachbedarf von Hochschulstudierenden. Auch handelt es sich letzten Endes um subjektive Einschätzungen von Studierenden und Hochschullehrenden, die oftmals die sprachliche Realität nicht erfassen (vgl. zu diesem Problem Long 2005). Hierfür müssten detailliert konkrete Sprachverwendungssituationen beschrieben werden, die sich über sprachliche Domänen (privat, öffentlich, Beruf, Bildung), Orte und Zeiten, Institutionen und Organisationen, beteiligte Personen, Objekte, Ereignisse, Operationen (vulgo: Handlungen) und Texte beschreiben lassen (zu diesem Ansatz vgl. Kapitel 4.1.2 im Gemeinsamen europäischen Referenzrahmen, Europarat 2001: 53–55). In jedem Fall wäre es für eine fundierte Beschreibung des sprachlichen Bedarfs erforderlich, konkrete sprachliche Texte zu identifizieren und einer linguistischen Analyse zugänglich zu machen.

Den Weg, für den Zweck einer umfassenden empirischen Beschreibung der sprachlichen Anforderungen von Studienanfängern typische Kommunikationssituationen zu erfassen, gingen Bärenfänger et al. (2016) im Rahmen einer mehrmethodisch angelegten Sprachbedarfsanalyse in den Fächern Chemie, Medizin und Wirtschaftswissenschaften. Hierzu kamen sechs leitfadenbasierte Fokusgruppeninterviews mit 16 Dozierenden und 13 Studierenden der drei genannten Fächer an zwei großen deutschen Universitäten zum Einsatz. Diese Daten wurden durch ein Shadowing, bei dem Mitarbeitende des Forschungsteams insgesamt elf Studierende der drei Fächer über einen längeren Zeitraum in Echtzeit beobachteten, ergänzt. Als dritte Datenquelle kamen Logbücher zur Verwendung, in denen

insgesamt elf Studierende der drei Fächer ihre sprachbezogenen Studienaktivitäten nach einem vorgegebenen Schema dokumentierten. Aus der vergleichenden Interpretation der drei Datenquellen ging zum einen hervor, dass Studierende ausgeprägte Schwierigkeiten beim simultanen Zuhören in Lehrveranstaltungen, dem Mitlesen von Skripten oder PowerPoint-Folien sowie dem Anfertigen von Mitschriften geltend machten. Das kombinierte Hör-Lese-Verstehen und Verschriftlichen des Rezipierten ist indessen – insbesondere mit Blick auf unterschiedliche Anforderungen in den einzelnen Fächern – praktisch nicht erforscht. In Zusammenhang hiermit fällt zum anderen auf, dass auch für den Studienerfolg hoch relevante und ubiquitäre Textsorten wie Vorlesungsskripte, PowerPoint-Präsentationen, Lehrbuchtexte und unterschiedliche Klausurentypen linguistisch gesehen empirisches Neuland sind. Des Weiteren wurde deutlich, dass in den untersuchten Fächern Mathematik in vielen Teilbereichen essenziell und mithin für den Studienerfolg unabdingbar ist. Gerade für nicht-muttersprachliche Studierende stellt der Bezug zwischen in Deutsch als Nicht-Muttersprache gestellten Aufgaben oder Beschreibungen und der mathematischen Formelsprache eine besondere Herausforderung dar. Es erscheint dringend notwendig, die Verwendung mathematischer Formelsprache in natürlich-, jedoch nicht-muttersprachlichen studienbezogenen Kommunikationssituationen gezielt empirisch zu erforschen und in studienpropädeutische Materialien zu überführen respektive in Hochschulzugangsprüfungen zu berücksichtigen. Schließlich wurde in der Bedarfsanalyse von Bärenfänger et. al. (2016) deutlich, dass ein großer Teil des Lehrens und Lernens an Hochschulen mittlerweile über digitale Kanäle (z.B. Fachportale, Lernplattformen, soziale Medien) und Textsorten (z.B. elektronische Vorlesungsskripte, PowerPoint-Folien, wissenschaftsbezogene Youtube-Videos, Vorlesungsmitschnitte, E-Klausuren usw.) erfolgt. Sprachliche Spezifika dieser elektronischen Textsorten ebenso wie die Besonderheiten der digitalen Kommunikation mit dem Ziel des Wissens- und Kompetenzerwerbs sind – obwohl sie mittlerweile konstitutiver Teil des Studiums sind – so gut wie unerforscht. Insgesamt konnten Bärenfänger et al. (2016) für die Fächer Chemie, Medizin und Wirtschaftswissenschaften zeigen, welche Kommunikationssituationen und damit verbundene typische Textsorten jeweils konstitutiv sind. Eine Beschreibung der sprachlichen Mittel, die für die Kommunikation an der Hochschule notwendig sind, konnte im Rahmen dieser Studie jedoch nicht erbracht werden. Gleichwohl konnte die Identifikation der sprachlichen Anforderungen in der Studieneingangsphase als Grundlage dafür genutzt werden, um für die drei Fächer repräsentative Textkorpora mit studientypischen Texten/Textsorten aufzubauen. Das in der Folge aufgebaute Akademische Textkorpus Bildungssprache des Instituts für Testforschung und Testentwicklung (vgl. Kap. 4.4) wurde, neben anderen Korpora, für die Analysen im vorliegenden Beitrag herangezogen.

3 Zwischenfazit

Zwar sind für das Deutsche als Studiersprache eine Reihe von nicht-empirischen Beschreibungen vorhanden, jedoch zielen diese auf häufig nicht relevante Textsorten und sind mit dem Makel behaftet, lediglich durch die Intuition ihrer Autorinnen und Autoren validiert zu sein. Durch die Empirieferne sehen sich diese Publikationen mit dem Vorwurf konfrontiert, gegenüber den tatsächlichen Anforderungen eines Hochschulstudiums blind zu sein. Demgegenüber sind die echt empirischen Studien zum Deutschen als Studier- und Wissenschaftssprache noch so lückenhaft und parzelliert, dass auch sie kaum als verlässliche Grundlage für die Entwicklung von Curricula, Lehr- und Lernmaterialien und Prüfungen dienen können. Hinsichtlich der wenigen Sprachbedarfsanalysen ist festzuhalten, dass diese entweder den tatsächlichen Sprachbedarf nicht in der nötigen Tiefe beschreiben oder die Studienwirklichkeit nicht in der gesamten Breite der Fächer erfassen. Zusammenfassend ist also mit Blick auf die theoretische Modellierung und empirische Beschreibung des Deutschen als Studiersprache noch immer Demmigs (2012: 38) Einschätzung zutreffend, dass der Forschungsstand nicht als gesicherte Grundlage für die Entwicklung von Hochschulzugangsprüfungen ausreicht. Indessen stellt die möglichst lückenlose Erforschung und Beschreibung von Wissenschaftssprache auf empirischer Basis eine notwendige Voraussetzung dafür dar, dass gezielt Maßnahmen zur Förderung bildungssprachlicher Kompetenz konzipiert und durchgeführt werden können (Fandrych/Graefen 2010: 509). Wie eingangs beschrieben, unterstreichen die hohen Studienabbruchquoten die Notwendigkeit einer optimierten Praxis der Studienvorbereitung bzw. -begleitung sowie der Studienzulassung.

Hierzu soll die vorliegende Studie beitragen, indem hier Lesetexte untersucht werden, die Studierende einerseits unmittelbar vor einem Studium im Rahmen einer Hochschulzugangsprüfung bewältigen können müssen, und andererseits Lehrbuchtexte, die in der Studieneingangsphase in großer Zahl präsent sind. In der Studie werden sowohl die Hochschulzulassungsprüfungen DSH, telc C1 Hochschule und TestDaF als auch die Lehrbuchtexte hinsichtlich ihrer Wortschatzschwierigkeit betrachtet. Hierfür werden Häufigkeitswortlisten großer allgemeinsprachlicher Korpora des Deutschen zum Vergleich herangezogen. Zudem wird exemplarisch auf die mögliche, empirische Ermittlung bildungssprachlich relevanter Mehrwortverbindungen und deren Vorkommen in den besagten Texten eingegangen.

4 Datengrundlage

4.1 Lesetexte Hochschulzugangsprüfungen

Für die vorliegende Analyse wurden Lesetexte aus den das Leseverstehen überprüfenden Testteilen der Sprachtests telc Deutsch C1 Hochschule (Übungstest 1) und TestDaF (Modellsatz 2 und 3, Modellsatz 1 gilt als veraltet und wurde daher von der Analyse ausgenommen) sowie Modellsätze der DSH untersucht. In diesem Kontext muss freilich darauf hingewiesen werden, dass die DSH keinen einheitlichen homogenen Test darstellt. Geleitet von einer gemeinsamen Rahmenordnung erstellt eine Vielzahl von Hochschulen eigene lokale DSH-Prüfungen, die nur begrenzt miteinander vergleichbar sind. Aus diesem Grund wurden für die vorliegende Untersuchung exemplarisch die DSH der Universitäten Leipzig und München ausgewählt.

Die jeweils acht Lesetexte A bis H, die Teil 1 des TestDaF konstituieren, wurden zusammengefasst, da sie ansonsten zu kurz gewesen wären, um die im Anschluss vorgenommenen Vergleiche zuverlässig durchführen zu können. Alle Texte werden in Form von Modellsätzen, Handbüchern o.a. von den Testanbietern bzw. Universitäten selbst online zur Verfügung gestellt. Tabelle 1 zeigt einen Überblick zu den einzelnen Texten sowie einige deskriptive Angaben zur Anzahl der laufenden Wörter (Token) und unterschiedlicher Wortformen (Types). Die Erfassung von Types und Tokens erfolgte mit der Software AntWord-Profiler1.4.0w.

Tab. 1: Übersicht analysierter Lesetexte aus Hochschulzugangsprüfungen

Nr.	Testsatz	Lesetext	Tokens	Types
1	DSH Leipzig	Lesetext 1	694	368
2	DSH München	Lesetext 1	603	353
3	Telc C1 Übungstest 1	Lesetext 1	373	230
4		Lesetext 2	634	349
5		Lesetext 3	1002	463
6	TestDaF Modellsatz 02	Lesetexte 1A-H	432	274
7		Lesetext 2	598	338
8		Lesetext 3	580	338
9	TestDaF Modellsatz 02	Lesetexte 1A – H	392	238
10		Lesetext 2	567	323
11		Lesetext 3	591	331

4.2 Lesetexte Lehrwerke

Zudem wurden für die drei Fächer Chemie, Wirtschaftswissenschaften (Wiwi) und Medizin jeweils zehn Textauszüge aus in der Studieneingangsphase relevanten Lehrbüchern extrahiert, wobei darauf geachtet wurde, dass alle in beteiligten Subdisziplinen der Fächer vertreten sind (vgl. Bärenfänger et al. 2016). Die einzelnen Auszüge sollten 1000 laufende Wortformen umfassen; die exakten Werte sind Tabelle 2 zu entnehmen.

Tab. 2: Übersicht analysierter Lesetexte aus Lehrwerken

Nr.	Disziplin	Teildisziplin	Tokens	Types
1	Chemie	Analytische Chemie	957	401
2	Chemie	Anorganische Chemie	976	488
3	Chemie	Experimentalphysik	1014	468
4	Chemie	Informatikrecht	1055	486
5	Chemie	Mathematik für Chemiker	1031	390
6	Chemie	Molekülspektroskopie	964	465
7	Chemie	Organische Chemie	1020	448
8	Chemie	Physikalische Chemie	992	402
9	Chemie	Technische Chemie	1027	456
10	Chemie	Theoretische Chemie	981	488
11	Medizin	Biochemie	1087	424
12	Medizin	Anatomie	1016	508
13	Medizin	Anatomie	934	458
14	Medizin	Biologie	1018	489
15	Medizin	Biologie	1065	481
16	Medizin	Biochemie	969	459
17	Medizin	Physiologie	1099	436
18	Medizin	Physik	1091	492
19	Medizin	Soziologie/Psychologie	1047	479
20	Medizin	Terminologie	979	506
21	Wiwi	BWL/VWL	978	452
22	Wiwi	BWL/VWL	995	531
23	Wiwi	Wirtschaftsinformatik	921	473
24	Wiwi	Recht	993	448
25	Wiwi	Mathematik	1014	485

Tab. 2: (fortgesetzt)

Nr.	Disziplin	Teildisziplin	Tokens	Types
26	Wiwi	Makroökonomie	943	397
27	Wiwi	Makroökonomie	979	413
28	Wiwi	Mikroökonomie	1000	402
29	Wiwi	Marketing	1079	476
30	Wiwi	Statistik	962	513

4.3 Wortschatzlisten

Für die Überprüfung der lexikalischen Komplexität der oben beschriebenen Lesetexte wurden u.a. ebenfalls mit verschiedene, auf großen Korpora basierende Wortlisten eingesetzt.

Zum einen wurde eine lemmatisierte Wortliste der 5000 häufigsten Wörter des Deutschen basierend auf dem 4,2 Millionen Tokens großen Herder/BYU-Korpus genutzt (vgl. Tschirner/Jones 2005; Tschirner 2006; Tschirner 2008). Diese Liste enthält die 5000 häufigsten Grundwortformen, denen jeweils alle flektierten Formen zugeordnet sind. Die Liste enthält insgesamt ca. 25 000 Wortformen, durchschnittlich sind jeder Grundform somit etwa fünf flektierte Formen zugeordnet. Für die durchgeführten Analysen wurde die Wortliste in die häufigsten 1000, 1001–3000 und 3001–5000 Wörter aufgeteilt.

Zur Validierung der Ergebnisse wurde zudem auf eine auf dem Deutschen Referenzkorpus (DeReKo-2014-II, zum Zeitpunkt der Erstellung ca. 7 Milliarden Tokens) des IDS Mannheim basierende Wortformenliste (Institut für Deutsche Sprache 2014) zurückgegriffen. Diese DeReWo-Wortliste wurde insofern aufbereitet, als dass sie auf Wortformen bzw. Types im engeren Sinne, also Buchstabenkombinationen, reduziert wurde. Kombinationen mit Sonderzeichen wie etwa „Pressestimmen/“ wurden ebenso gelöscht wie alleinstehende Sonderzeichen und Zahlen. Doppelte Einträge, die etwa auf das separate Erfassen der Groß- und Kleinschreibung in der DeReWo-Liste zurückzuführen waren (*die* erscheint etwa auf Rang zwei und als *DIE* auf Platz 1377), wurden entfernt. Die ursprünglich 100 000 Einträge umfassende Wortformenliste enthielt nach der Bereinigung noch 80 843 Einträge, wobei für die weiteren Analysen nur Wortformen bis zum Rang 80 000 berücksichtigt wurden. Diese 80 000 Wortformen umfassende Liste wurde aufgeteilt in 16 Teillisten zu je 5000 Einträgen, welche die 1–5000, 5001–10 000 ... 75 001–80 000 häufigsten Wortformen des Deutschen enthalten.

Es sei vorweggenommen, dass die Textdeckungsuntersuchung mit beiden Häufigkeitswortlisten zu sehr ähnlichen Ergebnissen führt, wodurch gleichfalls gezeigt wird, dass es im Bereich des hochfrequenten Wortschatzes unabhängig von der Zusammensetzung der zugrundeliegenden Korpora große Schnittmengen vorliegen und die hier vorgestellten Ergebnisse eine sehr hohe Reliabilität aufweisen.

4.4 ITT Textkorpus Bildungssprache

Das am Institut für Testforschung und Testentwicklung e.V. entwickelte Textkorpus Bildungssprache entstand im Rahmen des von der telc gGmbH geförderten Forschungsprojekts zur „Skalierung von Deutsch als Bildungssprache“. Die Zusammensetzung des Korpus erfolgte kriteriengeleitet auf der Grundlage einer umfangreichen Bedarfsanalyse zu den bildungssprachlichen Leseanforderungen (internationaler) Studierender in der Studieneingangsphase in den Fächern Chemie, Medizin und Wirtschaftswissenschaften (vgl. Bärenfänger et al. 2016; vgl. Kap. 2.3). Die auf der Bedarfsanalyse basierenden Kriterien betreffen beteiligte Textsorten, relevante Subdisziplinen sowie die Relevanz der einzelnen Textsorte/Subdisziplin und die damit verbundene Menge an Tokens, welche pro Textsorte je Subdisziplin aufgenommen wurden. Für jedes der Fächer wurde ein Fachkorpus im Umfang von ca. 1 Million Tokens zusammengestellt, insgesamt somit ca. 3 Millionen Tokens, welche sich auf insgesamt 1629 Einzeltexte verteilen. Zu den in das Korpus aufgenommenen Textsorten gehören beispielsweise Lehrbuchtexte, Skripte, PowerPoint Folien, Aufgaben- und Übungsblätter, (Alt-) Klausuren, Mitschriften, Protokolle und Versuchsanleitungen sowie Fachtexte wie wissenschaftliche Artikel oder Monographien, die beiden letztgenannten jedoch deutlich geringeren Anteils als gemeinhin angenommen werden mag, da die benannte Bedarfsanalyse gezeigt hat, dass sie, wenn überhaupt, dann nur in einigen wenigen Subdisziplinen im Lesekanon der Studieneingangsphase in den benannten Fächern eine Rolle spielen. Im vorliegenden Beitrag wurde das Korpus hinsichtlich bildungssprachlich relevanter Mehrwortverbindungen untersucht und deren Vorkommen in den oben beschriebenen Test- und Lehrbuchtexten analysiert. Es sei hierbei darauf hingewiesen, dass die Lehrbuchtexte dem Korpus selbst entstammen.

5 Analyse/Methode

Die Untersuchung der Textdeckung der Lesetexte ist durch die Ergebnisse empirischer Studien motiviert, welche belegen, dass die Kenntnis von ca. 98 % der laufenden Wörter eines Textes notwendig ist, um angemessenes Textverständnis zu erzielen (vgl. Nation/Chung 2009: 543; Schmitt/Cobb/Horst/Schmitt 2015; Tschirner 2006: 1279) und dass im Deutschen Wortschatzkenntnisse weit über den 4000 häufigsten Wörtern erforderlich sind, um Zeitungstexte und Fachtexte problemlos zu verstehen (vgl. Tschirner 2006: 1284). Die Häufigkeit der Wörter in Lesetexten ist außerdem ein wichtiges Kriterium zur Bestimmung der Textschwierigkeit. Sie spielt daher u. a. beim fremdsprachlichen Lesen eine wichtige Rolle. Die häufigsten 4000 Wörter des Deutschen decken ca. 93,1 % der Tokens in Texten der gesprochenen Sprache, 88,6 % der Tokens in belletristischen Texten, 86,9 % der Tokens in Zeitungstexten und nur 82,8 % der Tokens in Fachtexten ab (Tschirner 2006: 1284 ff.; siehe auch Tschirner 2008: 195) – Ergebnisse, welche in der vorliegenden Studie sehr ähnlich ausfallen.

Im Rahmen der hier vorgestellten Untersuchung wurden die elf benannten Lesetexte aus Hochschulzugangsprüfungen sowie die 30 Lehrbuchtexte auf deren Textdeckung mit den häufigsten 1000, 1001–3000 und 3001–5000 Wörtern des Deutschen basierend auf der Herder/BYU Wortliste und mit den 16 Einzellisten der DeReWo-Wortliste überprüft. Hierfür wurde die Software AntWordProfiler genutzt. Die Wortlisten wurden hierbei als „Level Lists“ definiert und mit den 41 Lesetexten verglichen. Auch die akkumulierten Textdeckungsraten für die gesamten Wortlisten wurden erfasst. Alle Textdeckungswerte sind in Prozent angegeben. Ob es zwischen den Textdeckungsraten für die verschiedenen Textgruppen Unterschiede gibt, wurde mittels verschiedener Varianzanalysen überprüft. Des Weiteren wurde das ITT Textkorpus Bildungssprache auf häufige Mehrwortverbindungen hin untersucht, wobei N-Gramme und Funktionsverbgefüge berücksichtigt wurden. Deren Vorkommen wurde wiederum in den oben beschriebenen Texten überprüft.

5.1 Ergebnisse – Textdeckung

Mit den 1000 häufigsten Wörtern der Herder/BYU Liste erreichte man bei den elf Prüfungstexten eine Textdeckung zwischen 63,4 % und 75,3 % ($M = 68.77$, $SD = 3.86$). Die 1001–3000 häufigsten Wörter tragen mit 9 % bis 18,1 % zur Textdeckung bei ($M = 13$, $SD = 2.82$), die 3001–5000 häufigsten Wörter tragen weitere 2,1 % bis 5,9 % bei ($M = 3.96$, $SD = 1.27$). Insgesamt wurde mit den häufigsten 1–5000 Wörtern bei den Prüfungstexten eine Textdeckung zwischen 79,1 % und 90,5 % erreicht ($M = 85.75$, $SD = 4$).

Insgesamt ergibt sich folgendes Bild (vgl. Abb. 1) für den Abgleich der Texte aus Test/Hochschulzugangsprüfung mit den Lehrbuchtexten aus Medizin, Chemie und Wirtschaftswissenschaft mit der Herder/BYU Häufigkeitsliste:

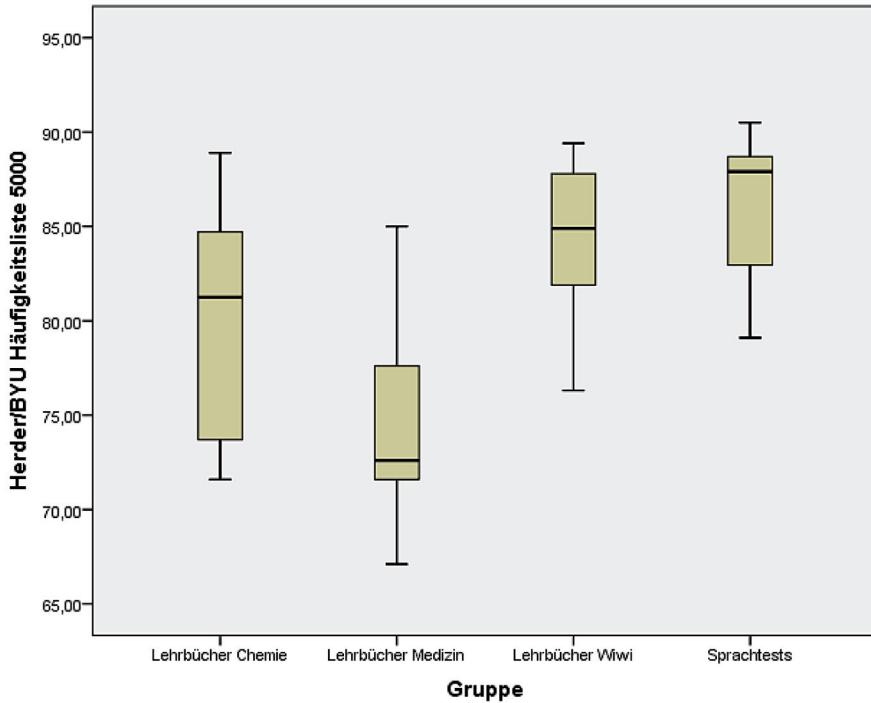


Abb. 1: Abgleich Texte und Wortliste Herder/BYU 1 bis 5000

Die Unterschiede beim Grad der Textdeckung der vier Textgruppen (Hochschulzugangsprüfung ($n = 11$), Lehrbuch Chemie ($n = 10$), Lehrbuch Medizin ($n = 10$) und Lehrbuch Wirtschaftswissenschaften ($n = 10$)) mit den häufigsten 5000 Wörtern (gesamt) der Herder/BYU Häufigkeitsliste wurden mittels einer einfaktoriellen ANOVA untersucht. Varianzhomogenität war gemäß Levene-Test gegeben ($p > .05$). Der Grad der Textdeckung unterschied sich statistisch signifikant für die verschiedenen Textgruppen, $F(3, 37) = 11.014$, $p < .001$. Ein Post Hoc Tukey Test ergab, dass sich die Lehrbuchtexte Medizin mit ihrer niedrigen Textdeckung signifikant von den anderen drei Textgruppen unterscheiden ($p < .05$). Zwischen den Gruppen Lehrbuch Chemie, Lehrbuch Wirtschaftswissenschaften und Hochschulzugangsprüfung traten keine signifikanten Unterschiede im Frequenzbereich 1–5000 auf.

Ähnlich fällt das Ergebnis aus, wenn man die analog untersuchte Textdeckung mit den hochfrequenten 1000 Wörtern des Deutschen betrachtet ($F(3, 37) = 6.318$, $p = .001$). Auch hier zeigt der Post Hoc Tukey Test Gruppenunterschiede zwischen Medizin und den Textgruppen Hochschulzugangsprüfung ($p < .05$) und Wirtschaftswissenschaften ($p < .05$), im Gegensatz zur oben genannten Analyse mit den häufigsten 5000 Wörtern aber nicht gegenüber der Chemie (vgl. Abb. 2).

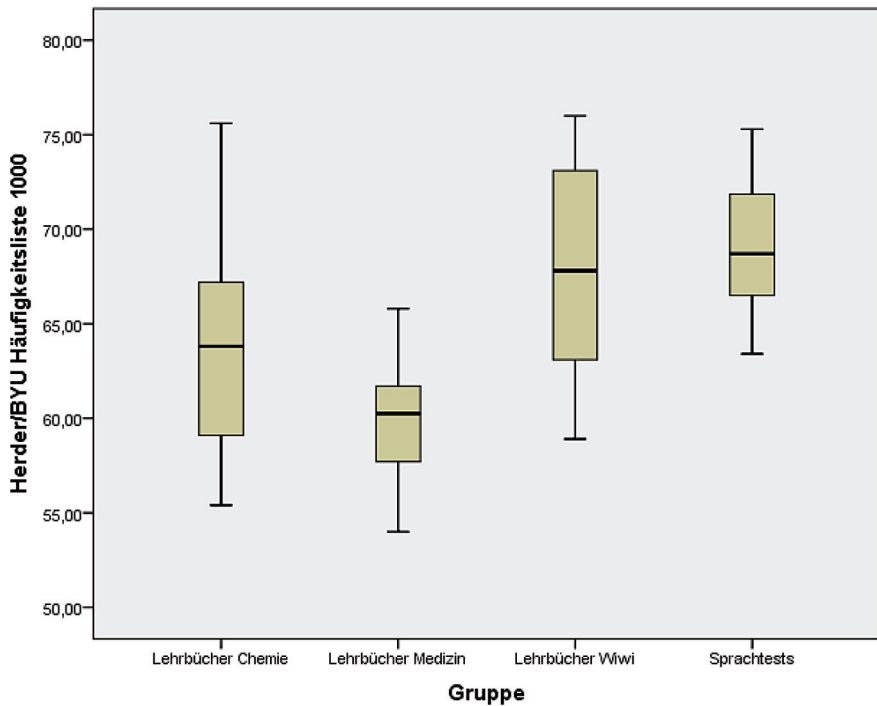


Abb. 2: Abgleich Texte und Wortliste Herder/BYU 1 bis 1.000

Zur Validierung dieser Ergebnisse wurde die Textdeckungsuntersuchung auch mit der Wortformenliste des IDS durchgeführt. Die wie oben beschrieben bereinigte Liste enthielt noch 80 000 eingetragene Wortformen. Diese wurde in Teillisten zu 5000 Wortformen der Häufigkeit nach aufgeteilt, um die Zuwachsraten nach Frequenzbereich abbilden zu können (vgl. Abb. 3).

Die Überblicksdarstellung (rechts) der Zuwachsraten entlang der 16 untersuchten Frequenzbereiche zeigt deutlich, dass die Prüfungstexte, gefolgt von den Texten im Bereich Wirtschaft, die größte Progression der Textdeckung aufweisen,

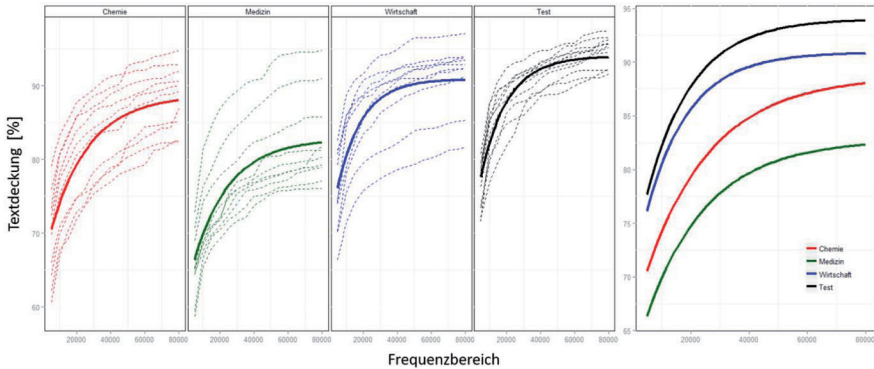


Abb. 3: Zuwachsraten der Textdeckung mit der DeReWo-Liste

während die Texte der naturwissenschaftlichen Fächer bis zum Schlusspunkt im Bereich der 75 000 bis 80 000 häufigsten Wortformen durchweg geringere Textdeckungsraten zeigen. Gleichfalls wird deutlich, dass sich die einzelnen Texte in den untersuchten Gruppen und Frequenzbereichen hinsichtlich ihrer Varianz stark unterscheiden, wobei sich die Prüfungstexte (vgl. Abb. 3) am engsten um die Gruppentrendkurve bewegen. Zudem wird deutlich, dass der Zuwachs an Textdeckung pro Frequenzbereich (mit einer gleichen Menge an Wortformen) erwartungsgemäß sehr schnell abnimmt – wie auch schon die differenzierten Werte für Textdeckung mit den häufigsten 1000, 3000 und 5000 Wörter der Herder/BYU Liste gezeigt haben.

Der Grad der Textdeckung der vier Textgruppen (Hochschulzugangsprüfung ($n = 11$), Lehrbuch Chemie ($n = 10$), Lehrbuch Medizin ($n = 10$) und Lehrbuch Wirtschaftswissenschaften ($n = 10$)) mit den häufigsten 5.000 Wortformen der DeReWo Häufigkeitsliste wurde ebenfalls mittels einer einfaktoriellen ANOVA untersucht. Varianzhomogenität war gemäß dem Levene-Test gegeben ($p > .05$). Der Grad der Textdeckung unterschied sich statistisch signifikant für die verschiedenen Textgruppen, $F(3, 37) = 11.40$, $p < .001$. Ein Post Hoc Tukey Test ergab, dass sich die Lehrbuchtexte Medizin und Chemie signifikant von den Textgruppen Wiwi und Hochschulzugangsprüfung unterscheiden (je $p < .05$), zwischen den Gruppen Lehrbuch Chemie und Medizin traten keine signifikanten Unterschiede auf, analog verhält es sich für Wiwi und Hochschulzugangsprüfung (vgl. Abb. 4).

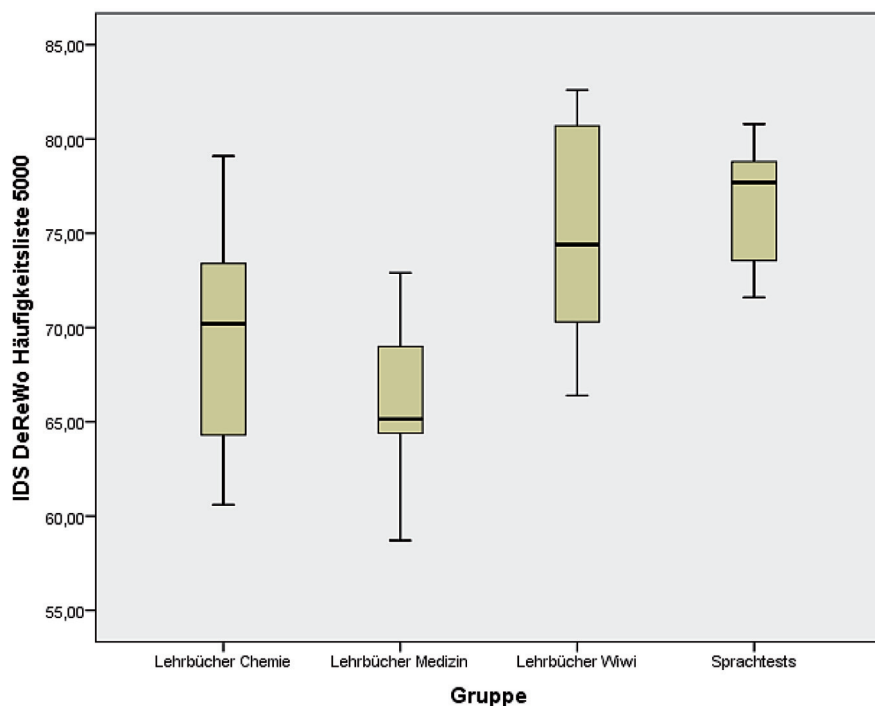


Abb. 4: Abgleich Texte und Wortliste DeReWo 1 bis 5.000

Auch für alle 80 000 in der bereinigten DeReWo Liste eingetragenen Wortformen wurde eine Varianzanalyse durchgeführt, $F(3, 37) = 14.09$, $p < .001$. Ein Post Hoc Tukey Test zeigt hier ein noch differenzierteres Bild von drei Subsets. Die Medizintexte ($M = 82.63$) zeigen sich verschieden zu allen anderen untersuchten Textgruppen ($p < .05$), Chemie ($M = 88.67$) und Wiwi ($M = 91.42$) unterschieden sich nicht voneinander, ebenso wenig wie Hochschulzugangsprüfung ($M = 94.77$) und Wiwi, signifikante Unterschiede traten jedoch zwischen Chemie und Hochschulzugangsprüfung auf ($p < .05$) (vgl. Abb. 5).

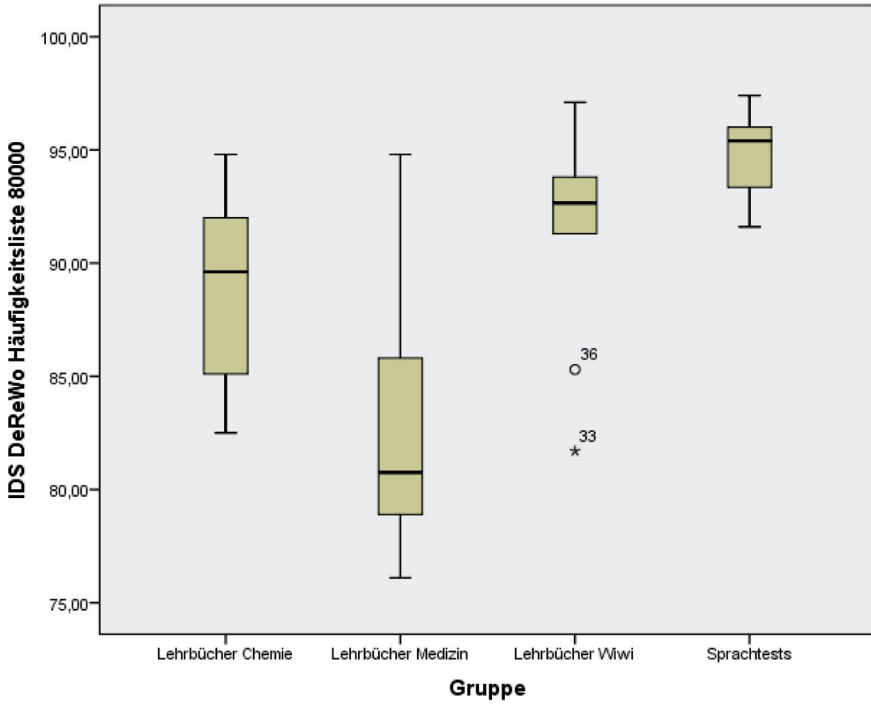


Abb. 5: Abgleich Texte und Wortliste DeReWo 1 bis 80 000

Die erreichten Mittelwerte der Textdeckung für alle Texte zusammen betragen $M = 89.51$ für die häufigsten 80 000 Wortformen der DeReWo-Liste und $M = 81.11$ für die häufigsten 5000 Wörter der Herder/BYU Liste. Die Korrelation der Ergebnisse beträgt $r = .936$ ($df(39)$, $p < .01$). Es hat sich somit gezeigt, dass der Grad der Textdeckung bei beiden Wortlisten stark korreliert (vgl. Abb. 6).

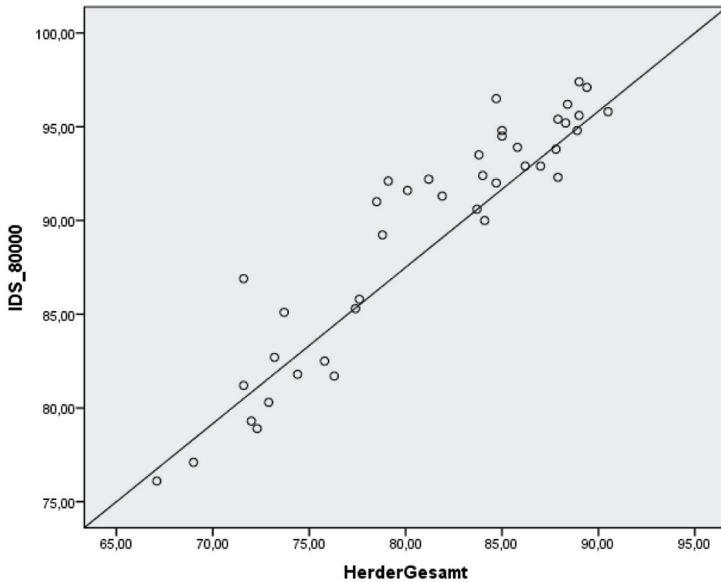


Abb. 6: Korrelation der Textdeckungswerte zwischen Herder/BYU und DeReWo Wortliste

5.2 Ergebnisse – ITT Textkorpus Bildungssprache | Mehrwortverbindungen

Mittels der korpuslinguistischen Software AntConc (Anthony 2016) wurde das *ITT Textkorpus Bildungssprache* auf häufige, feststehende Mehrwortkombinationen (N-Gramme) und statistisch signifikante Funktionsverbgefüge hin untersucht. In einer noch nicht veröffentlichten, umfassenderen N-Gramm-Analyse wurde für die Definition einer bildungssprachlich relevanten Wortkombination festgelegt, dass Verbindungen insgesamt mindestens 70-mal im Korpus belegt sind, sich diese Belege auf mindestens 20 Einzeltexte erstrecken und sie in allen drei Fachbereichen vorkommen. Die beiden letztgenannten Kriterien dienen dazu, eine möglichst große Streubreite der Wortverbindungen und damit die Relevanz für möglichst viele Studierende in der Studieneingangsphase zu gewährleisten und Wortverbindungen, die nur in sehr wenigen Texten oder nur in einem oder zwei der Fachbereiche gehäuft vorkommen und damit sehr spezifisch sind, auszuschließen.

Fünf der häufigsten N-Gramme seien an dieser Stelle eingehender betrachtet: *handelt (es) sich um* ($f = 514$), *in der Regel* ($f = 446$), *im Rahmen der/des* ($f = 255$), *in diesem Fall* ($f = 255$), *im Bereich der/des* ($f = 198$).

Ergänzend dazu wurde untersucht, mit welchen nominalen Partnern prototypische Funktionsverben im *ITT Textkorpus Bildungssprache* gemeinsam auftreten. Hierfür wurden die von Heine (2006) benannten Funktionsverben *bringen, finden, gehen, halten, kommen, nehmen, setzen, stehen, stellen, treten* und *ziehen* mittels einer Kollokationsanalyse im Suchhorizont von fünf Wörtern links und rechts des Basisverbs analysiert. Diese Untersuchung ergab einige prototypische und im Korpus häufig belegte Funktionsverbgefüge (vgl. zu deren Eigenschaften u. a. Riebling 2013: 146), z. B.: (zum) *Ausdruck bringen*, *Anwendung finden*, (zum) *Einsatz kommen*, (in) *Anspruch nehmen*, (zur) *Verfügung stehen*, oder (in) *Frage stellen*. Andere häufige, nominale Partner ergaben in Kombination zwar keine prototypischen Funktionsverbgefüge, wohl aber bildungssprachlich spezifisch gebrauchte Lesarten wie *etwas (ins) Gleichgewicht bringen*, *Temperatur halten*, *Preis setzen*, (im) *Vordergrund/Mittelpunkt stehen* oder (in) *Wechselwirkung treten*. Die folgende Analyse konzentriert sich auf die oben benannte Auswahl von Funktionsverbgefügen und N-Grammen und überprüft deren Vorkommen in den benannten Hochschulzugangsprüfungen und Lehrbuchtexten (vgl. Tab. 3).

Tab. 3: Mehrworteinheiten in Hochschulzugangsprüfung und Lehrbuch

Wortverbindung	Hochschulzu- gangsprüfung	Lehrbuch Chemie	Lehrbuch Wiwi	Lehrbuch Medizin
<i>handelt (es) sich um</i>	1	4	2	3
<i>in der Regel</i>	4	0	4	2
<i>im Rahmen der/des</i>	0	3	3	0
<i>in diesem Fall</i>	0	1	1	0
<i>im Bereich der</i>	0	0	0	2
<i>(zum) Ausdruck bringen</i>	1	0	0	0
<i>Anwendung finden</i>	0	0	0	0
<i>(zum) Einsatz kommen</i>	0	0	0	0
<i>(in) Anspruch nehmen</i>	0	0	1	0
<i>(zur) Verfügung stehen</i>	1	1	2	1
<i>(in) Frage stellen</i>	0	0	0	0

6 Diskussion der Ergebnisse

Die Textdeckungsuntersuchungen, bei denen Lehrbuchtexte und Lesetexte aus Hochschulzugangsprüfungen auf ihren Anteil hochfrequenter Wörter auf Grundlage des Herder/BYU-Korpus und des DeReWo hin überprüft wurde, erlauben zahlreiche Rückschlüsse.

Zum einen untermauern sie die Relevanz empirisch basierter Wortschatzbestände, auch für das Fremdsprachenlernen. Schon mit den 1000 wichtigsten Grundformen (inklusive flektierter Formen) der Herder-Liste bzw. den 5000 häufigsten Wortformen des DeReWo erreicht man, auch bei diesen durchweg als bildungssprachlich zu beschreibenden Texten, bereits Textdeckungsraten von über 60 %. Mit den häufigsten 5000 Wörtern (Herder/BYU) bzw. 80 000 Wortformen (DeReWo) werden fast ausnahmslos Werte von über 80 % erreicht, wobei die Werte für die Wortformenliste des IDS etwas höher ausfallen, was durch die insgesamt mehr als dreimal so große Menge von in dieser Liste enthaltenen Types im Vergleich zur 5000er Liste (Herder/BYU) begründet ist.

Insgesamt entwickeln sich die Textdeckungswerte mit zunehmendem Umfang der Wortlisten jedoch vergleichbar, was für große Schnittmengen der enthaltenen Lexikinventare und deren Relevanz spricht, obgleich die Zusammenstellung der zugrunde liegenden Korpora und die Art der Aufbereitung verschieden erfolgte (zur Vergleichbarkeit dieser Listen auch Wallner/Möhring 2013). Ein weiteres Indiz für die Vergleichbarkeit ist die sehr starke Korrelation der Textdeckungswerte mit beiden Listen. Sie illustriert, dass ein untersuchter Text, der eine eher geringe Textdeckungsrate mit der einen Liste aufweist, auch mit der anderen Liste einen eher geringen Wert erreichen wird und umgekehrt. Gehen wir davon aus, dass eine hohe Textdeckung mit diesen hochfrequenten, allgemeinsprachlichen Lexikinventaren für einen leichten Text spricht und eine geringe Textdeckung für einen fach- und bildungssprachlich anspruchsvolleren Text mit einem hohen Anteil an niederfrequentem Wortschatz, so würde dies vereinfacht bedeuten: Man kann mit sehr großer Wahrscheinlichkeit davon ausgehen, dass ein mittels der einen Liste als eher leicht eingestuftes Text auch mit der anderen Liste ein solches Ergebnis erreichen wird (hohe Textdeckung), womit man von einer sehr hohen Validität der Ergebnisse sprechen darf.

Eine der wichtigsten Fragen dieser Studie war: Sind Lesetexte in relevanten Textsorten der Studieneingangsphase (hier: Lehrbücher) hinsichtlich des Wortschatzes vergleichbar zu Lesetexten in den Hochschulzugangsprüfungen? Die durchgeführten Varianzanalysen haben gezeigt, dass die naturwissenschaftlichen Fächer, vor allem die Medizin, signifikant schwierigeren Wortschatz aufweisen als die Prüfungen, es aber keine Unterschiede zu den Wirtschaftswissenschaften gab. Dieses Ergebnis darf nicht zu dem Rückschluss führen, dass die

untersuchten Hochschulzugangsprüfungen für einige Fächer in Bezug auf die untersuchten Aspekte „zu leicht“ wären, denn der hohe Grad an Fachsprachlichkeit kann und darf nicht von einer fächerübergreifenden Hochschulzugangsprüfung erwartet werden. Zu den häufigsten Wörtern, die in den Lehrbuchtexten Medizin vorkommen, aber in keiner der beiden Grundwortschatzlisten enthalten sind, gehören beispielsweise *Spektroskopie*, *Wellenfunktion*, *Orbital* und *Zustandsgleichung* – fachsprachliche Ausdrücke, die vor Antritt eines Studiums weder von muttersprachlichen noch von internationalen Studierenden zwingend erworben werden müssen. Tendenziell allgemeine, bildungssprachliche Ausdrücke wie *kohärent* stellen nur eine Minderheit in dieser Ausdifferenzierung dar. Unter diesem Gesichtspunkt ist es nicht verwunderlich, dass die Lesetexte der Sprachtests die im Durchschnitt höchste Textdeckung enthalten, es fehlt weitgehend der in den Lehrbuchtexten enthaltene Fachwortschatz. Es sei jedoch angeregt, Analysen wie die hier vorgestellten konsequent zu nutzen, um lexikalische Mindeststandards für Hochschulzugangsprüfungen festzulegen.

Die in dieser Untersuchung genutzten, allgemeinsprachlichen Häufigkeitslisten, die auf dem Herder/BYU-Korpus und dem DeReKo basieren, können dabei nur einen Startpunkt bedeuten. Zusätzlich braucht es empirisch belastbare Zusammenstellungen von in der Studieneingangsphase, vor allem der Bachelorstudiengänge, relevantem Wortschatz – eben jenem Teil des Vokabulars, der jenseits der 5000 häufigsten Wörter des Deutschen und damit der Schwelle von Niveau B2 zu C1 angesiedelt ist, aber in verschiedenen Studiengängen in den tatsächlich rezipierten Textsorten bildungssprachlich relevant, jedoch nicht fachspezifisch ist. Es ist anzunehmen, dass eben dieses Inventar fachübergreifend maßgeblich zu einem besseren Textverständnis beitragen kann (vgl. zu einer korpusbasierten Wortliste für Akademisches Englisch Coxhead 2000).

Hierzu gehören auch Mehrwortverbindungen in ihrer unterschiedlichen Ausprägung als (lexikalische) Kollokation (vgl. für den Bereich bildungssprachlicher Kollokationen Wallner 2014), N-Gramm/Chunk, Funktionsverbgefüge und allen sonstigen Sprachroutinen, die in der Studienwirklichkeit wichtig sind. Wie dies für einen kleinen Ausschnitt der möglichen Sprachroutinen in den Fächern bewerkstelligt werden kann, wurde in der beschriebenen Mehrwortextraktion aus dem *ITT Textkorpus Bildungssprache* gezeigt. Der Abgleich des Vorkommens einiger der in diesem Korpus wichtigsten N-Gramme und Funktionsverbgefüge mit den 41 untersuchten Lesetexten hat gezeigt, dass (fast) alle Verbindungen sich auch in dieser relativ kleinen Stichprobe wiederfinden lassen. Die Darstellung des Auftretens dieser Wortkombinationen erfolgte bewusst rein deskriptiv, da die Datenbasis für einen statistischen Vergleich der Auftretenshäufigkeit von Mehrwortkombinationen keine Aussagekraft besäße. Es wäre jedoch wünschenswert, auch diesen Ansatz weiter zu verfolgen und auf dieser Grundlage empirisch

ermittelte Chunks für Studienvorbereitung und Hochschulzugangsprüfung nutzbar zu machen und nicht zuletzt bestehende Materialien auf Angemessenheit und Vollständigkeit hin zu untersuchen (vgl. Kap. 2.1).

7 Fazit und Ausblick

Die vorliegende Studie konnte zeigen, dass sich vor allem Lehrbuchtexte der Medizin deutlich von den Lesetexten der Hochschulzugangsprüfungen unterscheiden, sich die Textdeckungen mit häufigem, allgemeinsprachlichen Wortschatz jedoch ähneln. Die Unterschiede basieren zum einen auf fachsprachlichen, zum anderen aber auch auf allgemein bildungssprachlichen Ausdrücken, die nicht in den beiden Wortlisten Herder/BYU und DeReWo enthalten sind. Eine genauere Analyse eben dieser Ausdrücke wäre wünschenswert, um Studienvorbereitung und Hochschulzugangsprüfungen auf empirischer Basis zu verbessern.

Potenzial für diesbezügliche Studien bieten beispielsweise das *ITT Textkorpus Bildungssprache*, Forschungsprojekte wie *GeSIG – Das gemeinsame sprachliche Inventar der Geisteswissenschaften* (vgl. Meißner/Wallner 2016) oder das 2017 beginnende DFG-Forschungsprojekt *Zugang zu multimodalen Korpora gesprochener Sprache*, dessen Ziel es ist „die Nutzungsmöglichkeiten von multimodalen Korpora (darunter auch *GeWiss*, JM/OB) gesprochener Sprache zu vernetzen, zu professionalisieren und nutzergruppenbezogen auszudifferenzieren“ (http://herder.philol.uni-leipzig.de/hi.site,postext,projekte,a_id,1646.html, 15.08.2017), vgl. auch <http://zumult.org>, 18.05.2018). Es gilt hierbei, alle relevanten Fächer bzw. Fachbereiche zu berücksichtigen und vornehmlich die relevanten Textsorten, etwa der Studieneingangsphase, für korpuslinguistische Studien heranzuziehen. Nur so kann verlässlich benannt werden, welche bildungssprachlichen Wörter und Wortverbindungen zukünftige Studierende (insbesondere mit nicht-deutscher Muttersprache) beherrschen sollten, um möglichst keine sprachlichen Schwierigkeiten im Studium selbst bewerkstelligen zu müssen.

Ein gut ausgebauter allgemeinsprachlicher Wortschatz genügt, wie die vorliegenden Analysen gezeigt haben, nicht für eine ausreichende Lesekompetenz im Studium. Zudem sei hinterfragt, ob es in Anbetracht der großen lexikalischen Unterschiede zwischen verschiedenen Fachbereichen zielführend ist, mit allgemeingültigen Hochschulzugangsprüfungen zu agieren, oder ob fachbereichsspezifische Prüfungen, etwa für technische und geisteswissenschaftliche Studiengänge, nicht eine zuverlässigere Gatekeeper-Funktion erfüllen könnten.

Zum Verhältnis von erfolgreich absolvierter Hochschulzugangsprüfung, sprachlicher Kompetenz und Studienerfolg sind bspw. vom Forschungsprojekt *SPRASTU – Sprache und Studienerfolg bei Bildungsausländer/-innen* (Wisniewski

2018) aufschlussreiche Ergebnisse zu erwarten. Ein Studium bedeutet eine enorme (fremd)sprachliche Hürde, und auch eine erfolgreich absolvierte Hochschulzugangsprüfung entbindet nicht von Bemühungen der Studierenden aber auch Hochschulen, den Sprachlernprozess auch studienbegleitend weiterzuführen (vgl. Althaus 2004; Plassmann/Zeidler 2014; Zimmermann 2010), wie dies auch in zahlreichen außercurricularen Angeboten, wie etwa dem Academic Lab | UniSpracheDeutsch der Universität Leipzig (<http://home.uni-leipzig.de/academiclab>, 29.05.2017) erfolgt. Breit angelegte, empirische Forschungsanstrengungen sind jedoch unverzichtbar, um Sprachförderung und Sprachtests vor und während des Studiums möglichst effizient zu gestalten, und dies natürlich über die hier fokussierte Komponente des Wortschatzes und die hier exemplarisch analysierten Textsorten hinaus.

Danksagung: Die Autoren bedanken sich herzlich bei Prof. Dr. Erwin Tschirner für die Bereitstellung der Herder/BYU Häufigkeitsliste und bei Dr. Georg Frenck für die grafische Darstellung der Zuwachsraten (Abb. 3).

Literatur

- Althaus, Hans Joachim (2004): „Der TestDaF“. In: DAAD (Hrsg.): *Die internationale Hochschule: Ein Handbuch für Politik und Praxis* (Bd. 8). Bielefeld: Bertelsmann, 80–87.
- Apolinarski, Beate; Poskowsky, Jonas (2013): *Ausländische Studierende in Deutschland 2012: Ergebnisse der 20. Sozialerhebung des Deutschen Studentenwerks, durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW)*. Online: http://www.sozialerhebung.de/download/20/soz20_auslaenderbericht.pdf. (24.07.2017).
- Arras, Ulrike (2012): „Im Rahmen eines Hochschulstudiums in Deutschland erforderliche sprachliche Kompetenzen: Ergebnisse einer empirischen Bedarfsanalyse“. In: Tinnefeld, Thomas (Hrsg.): *Hochschulischer Fremdsprachenunterricht: Anforderungen, Ausrichtung, Spezifik*. Saarbrücken: HTW, 137–148.
- Bärenfänger, Olaf; Lange, Daisy; Möhring, Jupp (2016): „Sprache und Bildungserfolg: Sprachliche Anforderungen in der Studieneingangsphase“. In: Tschirner, Erwin; Bärenfänger, Olaf (Hrsg.): *Research Papers in Assessment*, Band 1. Leipzig: Institut für Testforschung und Testentwicklung. Online: <http://nbn-resolving.de/urn:nbn:de:bsz:15-qucosa-188820> (28.05.2017).
- Becker-Mrotzek, Michael; Schramm, Karen; Thürmann, Eike; Vollmer, Helmut Johannes (2013): *Sprache im Fach: Sprachlichkeit und fachliches Lernen*. Münster: Waxmann.
- Benes, Eduard (1981): „Die formale Struktur der wissenschaftlichen Fachsprachen in syntaktischer Hinsicht“. In: Bungarten, Theo (Hrsg.): *Wissenschaftssprache: Beiträge zur Methodologie, theoretischen Fundierung und Deskription*. München: Wilhelm Fink Verlag, 185–212.
- Bünting, Karl-Dieter; Bitterlich, Axel; Pospiech, Ulrike (2006): *Schreiben im Studium: mit Erfolg*. 6. Auflage. Berlin: Cornelsen.

- Busch-Lauer, Ines (2012): „Abstracts – eine facettenreiche Textsorte der Wissenschaft“. In: *Linguistik Online*, 52 (2), 5–22.
- Chehab, Nabila; Grützmacher, Judith; Heublein, Ulrich; Kercher, Jan (2017): *Wissenschaft weltweit offen kompakt. Daten und Fakten zur Internationalität von Studium und Forschung in Deutschland*. Bielefeld: Bertelsmann.
- Coxhead, Averil (2000): „A New Academic Word List“. In: *TESOL Quarterly*, 34, (2), 213–238.
- Crawford Camiciottoli, Belinda (2010): „Meeting the Challenges of European Student Mobility: Preparing Italian Erasmus Students for Business Lectures in English“. In: *English for Specific Purposes*, 29, 268–280.
- Demmig, Silvia (2012): *Sprachkompetenzen prüfen – aber wie? Zur Konstruktvalidität von standardisierten Sprachtests am Beispiel der DSH: Theoretische Grundlegung und empirische Longitudinalstudie*. Jena: Habilitationsschrift Friedrich-Schiller-Universität Jena.
- Europarat (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: Lernen, Lehren, Beurteilen*. München: Langenscheidt.
- Fandrych, Christian (2002): „Herausarbeiten vs. illustrieren: Kontraste bei der Versprachlichung von Sprechhandlungen in der englischen und deutschen Wissenschaftssprache“. In: Ehlich, Konrad (Hrsg.): *Wissenschaftskommunikation: Vol. 2. Mehrsprachige Wissenschaft – europäische Perspektiven. Eine Konferenz zum Europäischen Jahr der Sprachen 2001*. München: Institut für Deutsch als Fremdsprache / Transnationale Germanistik. Heidelberg: Synchron, 1–28.
- Fandrych, Christian; Graefen, Gabriele (2002): „Text-commenting Devices in German and English Academic Articles“. In: *Multilingua* 21, 17–43.
- Fandrych, Christian; Graefen, Gabriele (2010): „Wissenschafts- und Studiensprache Deutsch“. In: Krumm, Hans-Jürgen; Fandrych, Christian; Hufeisen, Britta; Riemer, Claudia (Hrsg.): *Deutsch als Fremd- und Zweitsprache*. Vollst. Neubearbeitung, Handbücher zur Sprach- und Kommunikationswissenschaft 19, Berlin: De Gruyter, 509–517.
- Fandrych, Christian; Meißner, Cordula; Slavcheva, Adriana (2014): „Das Korpusprojekt ‚Gesprochene Wissenschaftssprache kontrastiv‘ und seine Relevanz für die Vermittlung des Deutschen als Wissenschaftssprache“. In: Mackus, Nicole; Möhring, Jupp (Hrsg.): *Wege für Bildung, Beruf und Gesellschaft – mit Deutsch als Fremd- und Zweitsprache. 38. Jahrestagung des Fachverbandes Deutsch als Fremdsprache an der Universität Leipzig 2011*. Göttingen: Universitätsverlag, 141–160.
- Fandrych, Christian; Meißner, Cordula; Wallner, Franziska (2017) (Hrsg.): *Gesprochene Wissenschaftssprache – digital. Verfahren zur Annotation und Analyse mündlicher Korpora*. Tübingen: Stauffenburg Verlag.
- Feilke, Helmuth (2012): „Bildungssprachliche Kompetenzen – fördern und entwickeln“. In: *Praxis Deutsch* (39), 4–13.
- Gantefort, Christoph; Roth, Hans-Joachim (2010): „Sprachdiagnostische Grundlagen für die Förderung bildungssprachlicher Fähigkeiten“. In: *Zeitschrift für Erziehungswissenschaft* 13 (4), 573–591.
- Gogolin, Ingrid; Lange, Imke; Michel, Ute; Reich, Hans H. (Hrsg.) (2013): *Herausforderung Bildungssprache – und wie man sie meistert*. Münster: Waxmann.
- Graefen, Gabriele (2002): „Probleme mit der Alltäglichen Wissenschaftssprache in Hausarbeiten ausländischer StudentInnen“. In: Rehbein, Jochen; Riedel, Sabine (Hrsg.): *Lernen in der zweiten Sprache*. Münster: Waxmann, 2–20. Online: www.wissenschaftssprache.de/Texte/GG_Probleme_mit_AWS_2002.pdf (30.05.2017).

- Geist, Barbar; Dietze, Kim; Lange-Schubert, Kim (2017): „Bildungssprachliche Merkmale im Sachunterricht der Grundschule: Theoretische und empirische Annäherungen“. In: Tschirner, Erwin; Möhring, Jupp; Cothrun, Keith (Hrsg.): *Deutsch als zweite Bildungssprache in MINT-Fächern*. Tübingen: Stauffenburg Verlag.
- Heine, Antje (2006): Funktionsverbgefüge in System, Text und korpusbasierter (Lerner-)Lexikografie. Frankfurt am Main et. al.: Peter Lang. [Finnische Beiträge zur Germanistik 18].
- Heller, Dorothee (2014): „Dozentenseitige Beiträge zum Verständigungshandeln und Hörereinbezug – Exemplarische Analysen einer italienischen Germanistikvorlesung“. In: Hornung, Antonie; Carobbio, Gabriella; Sorrentino, Daniela (Hrsg.): *Diskursive und textuelle Strukturen in der Hochschuldidaktik: Deutsch und Italienisch im Vergleich*. Münster: Waxmann.
- Heublein, Ulrich; Richter, Johanna; Schmelzer, Robert; Sommer, Dieter (2014): *Die Entwicklung der Studienabbruchquoten an den deutschen Hochschulen: Statistische Berechnungen auf der Basis des Absolventenjahrgangs 2012*. Hannover: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung.
- Hoffmann, Lothar (1984): *Kommunikationsmittel Fachsprache*. Berlin: Akademie-Verlag.
- Institut für Deutsche Sprache (2014): *Korpusbasierte Wortformenliste DeReWo, DeReKo-2014-II-MainArchive-STT.100000*. Mannheim: Institut für Deutsche Sprache. Online: <http://www.id-s-mannheim.de/derewo> (30.05.2017).
- HRK/KMK (2015): Rahmenordnung über Deutsche Sprachprüfungen für das Studium an deutschen Hochschulen (RO-DT). Online: http://www.kmk.org/fileadmin/Dateien/veroeffentlichtungen_beschlusse/2004/2004_06_25_RO_DT.pdf (24.03.2017).
- Kliegl, Christina; Müller, Ursula M. (2012): „Diversity und Studienabbruch im Zeitalter von Bologna. Bedingungsfaktoren für Studienabbruchsgedanken in den alten und neuen Studiengängen an der Universität Duisburg-Essen“. In: *Die Hochschule*, 21 (1), 73–90.
- Krause, Arne (2015): „Sprachliche Verfahren zur Vermittlung mathematischen Problemlösungswissens in der Hochschule: Exemplarische Analysen mathematischer Übungen“. In: Ferraresi, Gisela; Liebner, Sarah (Hrsg.): *SprachBrückenBauen. 40. Jahrestagung des Fachverbandes Deutsch als Fremd- und Zweitsprache an der Universität Bamberg 2013. Materialien Deutsch als Fremdsprache*, 29. Göttingen: Universitätsverlag, 203–218.
- Long, Michael H. (2005): „Methodological Issues in Learner Needs Analysis“. In: Long, Michael H. (Ed.): *Second Language Needs Analysis*. Cambridge: Cambridge University Press, 19–78.
- Marks, Daniela (2015): „Prüfen sprachlicher Kompetenzen internationaler Studienanfänger an deutschen Hochschulen: Was leistet der TestDaF?“ In: *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 20 (1), 21–39.
- Mehlhorn, Grit (2005): *Studienbegleitung für ausländische Studierende an deutschen Hochschulen*. München: iudicium.
- Meißner, Cordula; Wallner, Franziska (2016): „Die datengeleitete Ermittlung des gemeinsamen sprachlichen Inventars der Geisteswissenschaften“. *DHD 2016 Modellierung – Vernetzung – Visualisierung. Die Digital Humanities als fächerübergreifendes Forschungsparadigma. Universität Leipzig 7.-12.03.2016, Konferenzabstractband*, 169–171.
- Moll, Melanie; Thielmann, Winfried (2017): *Wissenschaftliches Deutsch*. Konstanz: Huter & Roth.
- Nation, Paul; Chung, Teresa (2009): “Teaching and Testing Vocabulary”. In: Long, Michael H.; Doughty, Catherine (Eds.): *The Handbook of Language Teaching*. Malden, MA: Wiley-Blackwell, 543–559.
- Plassmann, Sibylle; Zeidler, Beate (2014): “Taking Decisions: Assessment for University Entry”. In: *CercleS* 4 (1), 237–255.

- Pohl, Thorsten (2007): *Studien zur Ontogenese wissenschaftlichen Schreibens*. Tübingen: Niemeyer.
- Prediger, Susanne (2013): „Darstellungen, Register und mentale Konstruktion von Bedeutungen und Beziehungen – mathematikspezifische sprachliche Herausforderungen identifizieren und bearbeiten“. In: Becker-Mrotzek, Michael; Schramm, Karen; Thürmann, Eike; Vollmer, Helmut Johannes (Hrsg.): *Sprache im Fach: Sprachliches und fachliches Lernen*. Münster: Waxmann, 167–183.
- Reich, Hans H. (2013): „Durchgängige Sprachbildung“. In: Gogolin, Ingrid; Lange, Imke; Michel, Ute; Reich, Hans H. (Hrsg.): *Herausforderung Bildungssprache – und wie man sie meistert*. Münster: Waxmann, 55–70.
- Riebling, Linda (2013): „Heuristik der Bildungssprache“. In: Gogolin, Ingrid; Lange, Imke; Michel, Ute; Reich, Hans H. (Hrsg.): *Herausforderung Bildungssprache – und wie man sie meistert*. Münster: Waxmann, 106–153.
- Redder, Angelika; Heller, Dorothee; Thielmann, Winfried (Hrsg.) (2014): *Eristische Strukturen in Vorlesungen und Seminaren deutscher und italienischer Universitäten: Analysen und Transkripte*. Heidelberg: Synchron.
- Schmitt, Norbert; Cobb, Tom; Horst, Marlise; Schmitt, Diane (2015): “How Much Vocabulary is Needed to Use English? Replication of van Zeeland and Schmitt (2012), Nation (2006) and Cobb (2007)”. In: *Language Teaching* 50 (2), 221–226.
- Steets, Angelika (2003): „Die Mitschrift als universitäre Textart – schwieriger als gedacht, wichtiger als vermutet.“ In: Ehlich, Konrad; Steets, Angelika (Hrsg.): *Wissenschaftlich schreiben – lehren und lernen*. Berlin: De Gruyter, 51–64.
- Stezano Cotelo, Kristin (2006): „Die studentische Seminararbeit – studentische Wissensverarbeitung zwischen Alltagswissen und wissenschaftlichem Wissen“. In: Ehlich, Konrad; Heller, Dorothee (Eds.), *Linguistic Insights: Vol. 52. Die Wissenschaft und ihre Sprachen*. Bern: Lang, 87–114.
- Stezano Cotelo, Kristin (2008): *Verarbeitung wissenschaftlichen Wissens in Seminararbeiten ausländischer Studierender: Eine empirische Sprachanalyse*. München: iudicium Verlag.
- Szurawitzki, Michael (2015): „Makro- und mikrostrukturelle Analysen wissenschaftlicher Artikel II: Untersuchungen soziologischer Einleitungen und Vergleich mit medizinwissenschaftlichen Text- und Sprachstrukturen“. In: *Glottology*, 6 (1), 113–138.
- Thielmann, Winfried (2007): *Zur Wissensbearbeitung beim Leser in deutschen und englischen Wissenschaftstexten: Hinführen, Verknüpfen, Benennen. Wissenschaftskommunikation: Vol. 3*. Heidelberg: Synchron.
- Thielmann, Winfried (2014a) : „Marie, das wird nichts‘ – sprachliche Verfahren der Wissensbearbeitung in einer Vorlesung im Fach Maschinenbau“. In: Fandrych, Christian et al. (Hrsg.), *Gesprochene Wissenschaftssprache. Korpusmethodische Fragen und empirische Analysen*. Heidelberg: Synchron, 193–206.
- Thielmann, Winfried (2014b) (Hrsg.): „Verfahren diskursiver Wissensvermittlung im Fach Maschinenbau: Wenn einmal der Wert eingeführt ist, kriegst du ihn nicht mehr weg“. In: Hornung, Antonie; Carobbio, Gabriella; Sorrentino, Daniela (Hrsg.): *Diskursive und textuelle Strukturen in der Hochschuldidaktik. Deutsch und Italienisch im Vergleich*. Münster: Waxmann, 55–67.
- Tschirner, Erwin; Bärenfänger, Olaf; Möhring, Jupp (Hrsg.) (2016): *Deutsch als fremde Bildungssprache. Das Spannungsfeld von Fachwissen, sprachlicher Kompetenz, Diagnostik und Didaktik*. Tübingen: Stauffenburg Verlag.

- Tschirner, Erwin; Jones, Randall L. (2005): *The Herder-BYU electronic corpus of contemporary German*. Leipzig: Herder-Institut.
- Tschirner, Erwin (2006): „Häufigkeitsverteilungen im Deutschen und ihr Einfluss auf den Erwerb des Deutschen als Fremdsprache“. In: Corino, Elisa; Marella, Carla; Onesti, Christina (Eds.): *Atti del XII Congresso Internazionale di Lessicografia. Proceedings – XII Euralex International Congress. Torino, Italia, September 6th–9th 2006*. Alessandria: Edizioni dell’Orso, 1277–1288.
- Tschirner, Erwin (2008): „Das professionelle Wortschatzminimum im Deutschen als Fremdsprache“. In: *Deutsch als Fremdsprache* 45 (4), 195–207.
- Wagner, Jonas (2014): „Wissensvermittelnde Verfahren im Fach Mathematik“. In: Horning, Antonie; Carobbio, Gabriella; Sorrentino, Daniela (Hrsg.): *Diskursive und textuelle Strukturen in der Hochschuldidaktik. Deutsch und Italienisch im Vergleich*. Münster: Waxmann, 69–92.
- Wallner, Franziska (2014): *Kollokationen in Wissenschaftssprachen: Zur lernerlexikographischen Relevanz ihrer wissenschaftssprachlichen Gebrauchsspezifika*. Tübingen: Stauffenburg Verlag.
- Wallner, Franziska; Möhring, Jupp (2013): „Wortschatzlisten auf dem Prüfstand. In: Bergerová, Hana; Schmidt, Marek und Schuppener, Georg (Hrsg.): *Aussiger Beiträge: Germanistische Schriftenreihe aus Forschung und Lehre* 7, 119–133.
- Wisniewski, Katrin (2018): „Studienerfolg und sprachliche Kompetenzen bei Bildungsausländer/-innen: Das Forschungsprojekt SPRASTU“. In: *Info DaF* in diesem Band, 573–597.
- Wiesmann, Bettina (1999): *Mündliche Kommunikation im Studium: Diskursanalysen von Lehrveranstaltungen und Konzeptualisierung der Sprachqualifizierung ausländischer Studienbewerber. Studien Deutsch: Vol. 27*. München: iudicium.
- Wollert, Mattheus; Zschill, Stefanie (2017): „Sprachliche Studierfähigkeit: Ein Konstrukt auf dem Prüfstand“. In: *Info DaF* 44 (2), 2–17.
- Zimmermann, Sonja (2010): „Schreiben im Test – Schreiben im Studium: Wie können Anforderungen an das akademische Schreiben in Tests abgebildet werden? Das Beispiel Test-DaF“. In: Brandl, Heike; Duxa, Susanne; Leder, Gabriele; Riemer, Claudia (Hrsg.): *Ansätze zur Förderung akademischer Schreibkompetenz an der Hochschule*. Göttingen: Universitätsverlag Göttingen (Materialien Deutsch als Fremdsprache 83), 1–10.

Software

- Anthony, Laurence (2016): AntConc (Version 3.4.4) [Computer Software]. Tokyo, Japan: Waseda University. Online: <http://www.antlab.sci.waseda.ac.jp> (30.05.2017).
- Anthony, Laurence (2013): AntWordProfiler (Version 1.4.0) [Computer Software]. Tokyo, Japan: Waseda University. Online: <http://www.antlab.sci.waseda.ac.jp> (30.05.2017).

Biographische Angaben

Jupp Möhring

ist Geschäftsführer am Institut für Testforschung und Testentwicklung e.V. und wissenschaftlicher Mitarbeiter am Herder-Institut der Universität Leipzig.

Olaf Bärenfänger

ist Direktor des Sprachenzentrums an der Universität Leipzig und Vorstand am Institut für Testforschung und Testentwicklung e. V. Darüber hinaus lehrt er am Herder-Institut der Universität Leipzig als Privatdozent mit dem Schwerpunkt Angewandte Linguistik.