



Brandl, Heike; Arslan, Emre; Langelahn, Elke; Riemer, Claudia (Hrsg.):

Mehrsprachig in Wissenschaft und Gesellschaft

Mehrsprachigkeit, Bildungsbeteiligung und Potenziale von Studierenden mit Migrationshintergrund



Interdisziplinäres Symposium
vom 06. bis 07. Februar 2012
im Zentrum für interdisziplinäre Forschung, Bielefeld

Juli 2013

Universität Bielefeld
Universitätsstr. 25
33615 Bielefeld

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://www.d-nb.de> abrufbar.

DOI: 10.2390/biecoll-mehrspr2013_9

Bielefeld 2013

Alle Rechte vorbehalten.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung der Herausgeber unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen

Inhalt

Vorwort der Herausgeberinnen und des Herausgebers

DaZ an der Hochschule oder: wie sich Mehrsprachigkeit für Wissenschaft und Gesellschaft bei Studierenden mit Migrationshintergrund entfalten kann..... 1

Mehrsprachigkeit in der globalisierten Gesellschaft und ihre Relevanz für Forschung und Lehre

Katarina Wagner & Claudia Maria Riehl

Mehrsprachigkeit: gesellschaftliche Wahrnehmung und zukünftige Potenziale..... 1

Yasemin Karakaşoğlu

Interkulturelle Öffnung als Rahmen hochschulpolitischer Maßnahmen (nicht nur) zur Unterstützung des Studienerfolgs von Studierenden mit Migrationshintergrund..... 9

Mark Becker

Bildungsaufstieg unterstützen: ‚Chance hoch 2 – das Programm für Bildungsaufsteiger/-innen‘..... 19

Potenzial und Entfaltung von Migranten im tertiären Bildungsbereich: Forschungsergebnisse aus soziologischer und psychologischer Perspektive

Haci Halil Uslucan

Psychologische Bedingungen des Bildungserfolgs von Migranten..... 29

Hannah Burger, Joanna Pfaff-Czarnecka & Patricia Pielage

Heterogenität an der Universität – Studieren mit Migrationshintergrund. Skizze eines Forschungsprojekts..... 37

Lisa Unger-Fischer

Das Secondos-Programm der Universität Regensburg..... 51

Sprachkompetenzen in der Erst- und Zweitsprache für die Domäne Wissenschaftskommunikation

Kirsten Schindler

Texte beurteilen – Feedback geben. Kompetenzen für Lehramtsstudierende..... 57

Inger Petersen

Entwicklung schriftlicher Argumentationskompetenz bei ein- und mehrsprachigen Oberstufenschüler/-innen und Studierenden..... 69

Sonja Zimmermann & Ellen Rupprecht

Typisch DaZ? – Ein Vergleich schriftlicher Leistungen von Studierenden mit Deutsch als Erst-, Zweit- und Fremdsprache..... 81

Jutta Çıkar

Fordern, Fördern und Zertifizieren. Türkischkompetenzen von Schülerinnen und Schülern mit Migrationshintergrund..... 91

Christoph Schroeder & Meral Dollnick	
Mehrsprachige Gymnasiasten mit türkischem Hintergrund schreiben auf Türkisch.....	101
Elke Langelahn, Heike Brandl & Emre Arslan	
„Profilbildung und Mehrsprachigkeit für Studierende mit Migrationshintergrund“ an der Universität Bielefeld.....	115
Autorinnen und Autoren.....	135

Erschienen in: Brandl, Heike; Arslan, Emre; Langelahn, Elke; Riemer, Claudia (Hrsg.) (2013): *Mehrsprachig in Wissenschaft und Gesellschaft. Mehrsprachigkeit, Bildungsbeteiligung und Potenziale von Studierenden mit Migrationshintergrund*. Bielefeld, 81-89. http://biecoll.ub.uni-bielefeld.de/volltexte/2013/5274/index_de.html.

Typisch DaZ? – Ein Vergleich schriftlicher Leistungen von Studierenden mit Deutsch als Erst-, Zweit- und Fremdsprache

Sonja Zimmermann, Ellen Rupprecht, Bochum

Der Beitrag geht der Frage nach, ob und inwieweit sich schriftliche Leistungen von Studierenden mit Deutsch als Zweitsprache¹ im Vergleich zu den Texten von Studierenden mit Deutsch als Erstsprache einerseits und denen einer fremdsprachlichen Vergleichsgruppe andererseits unterscheiden lassen. Da es sich bei den zur Verfügung stehenden Texten um ein relativ kleines Sample handelt, soll der Beitrag auch dazu dienen, exemplarisch aufzuzeigen, wie man quantitative Analysen von Lernerleistungen² nutzen kann, um Sprachkompetenzen näher zu beschreiben.

1 Ausgangslage

Der Test Deutsch als Fremdsprache (TestDaF) ist ein standardisierter Sprachstandstest, der als Nachweis für die sprachliche Studierfähigkeit ausländischer Studierender von allen deutschen Hochschulen anerkannt wird. Als solcher hat er für die Teilnehmer weitreichende persönliche Konsequenzen, daher orientiert sich die Sicherstellung der Qualität des TestDaF in der Konstruktion, Durchführung und Evaluation an hohen wissenschaftlichen Kriterien (vgl. Eckes 2003). Zu diesen Qualitätssicherungsmaßnahmen gehört u.a. die Erprobung neuer Testaufgaben, die standardmäßig auch mit einer Gruppe von deutschen Muttersprachlern durchgeführt wird. Diese umfasst das Lösen der dichotomen (d.h. entweder als „richtig“ oder „falsch“ kodierten) Items in den Prüfungsteilen Lese- und Hörverstehen sowie ein qualitatives Feedback zu den offenen Aufgaben in den produktiven Prüfungsteilen Schriftlicher und Mündlicher Ausdruck.

Im Rahmen einer Validierungsstudie wurden im April 2009 erstmals auch schriftliche und mündliche Leistungen von muttersprachlichen Studierenden erhoben, die diese unter realen Testbedingungen erbrachten.³

¹ Wir verwenden den Begriff Deutsch als Zweitsprache als übergeordneten „Begriff für alle Formen der Sprachaneignung nach der Erstsprache“ (Ahrenholz 2010: 6).

² Aus Gründen der sprachlichen Vereinfachung werden Ausdrücke wie „Lerner“, „Teilnehmer“, „Beurteiler“ usw. im generischen Sinne verwendet.

³ An dieser Stelle sei noch einmal ausdrücklich den Mitarbeitern des Bereichs Deutsch als Fremdsprache am Zentrum für Fremdsprachenausbildung der Ruhr-Universität Bochum für ihre Unterstützung bei der Durchführung der Studie gedankt.

Muttersprachliche Probanden sind im Kontext des Sprachtestens nicht unumstritten (vgl. Alderson; Clapham; Wall 1995: 97): Einerseits ist es gängige Praxis, die Lösungsraten von Muttersprachlern bei objektiv auswertbaren Aufgaben in die Testerstellung einzubeziehen, indem beispielsweise Items, die sich für diese Probandengruppe als zu schwierig erwiesen, ausgeschlossen werden. Andererseits ist es problematisch, Muttersprachler als Referenz für fremdsprachliche Leistungen anzusetzen, denn diese muttersprachliche Norm ist aufgrund großer Varianz innerhalb der Gruppe der Muttersprachler nur schwer zu beschreiben (Davies 2008: 431). Dennoch kann der Vergleich mit einer muttersprachlichen Referenzgruppe interessante Daten liefern, u.a. zu der Frage, wie valide die Prüfungsaufgaben und ihre Auswertung sind.

Im Schriftlichen Ausdruck des TestDaF beispielsweise sollen bestimmte Merkmale, die für den Hochschulkontext relevant sind, abgeprüft werden. Hierzu zählen v.a. die Versprachlichung eines Sachverhalts und die Diskursivität eines Textes. Ziel des Prüfungsteils ist es festzustellen, inwieweit die Prüfungsteilnehmer in der Lage sind, in einem begrenzten Zeitrahmen und ohne Verwendung von Hilfsmitteln einen kohärenten und strukturierten Text zu schreiben, der zwei für den Hochschulkontext relevante Schreibfertigkeiten beinhaltet: das Beschreiben statistischer Daten und das Argumentieren.⁴

Die zu erhebenden Daten sollten daher im Anschluss unter folgenden Gesichtspunkten analysiert werden: Mit welchen sprachlichen Mitteln strukturieren und gliedern Muttersprachler ihre Texte und welche syntaktischen Variationen setzen sie ein, um einen zusammenhängenden und diskursiven Text zu schreiben? Wie abwechslungsreich ist die Lexik?

Bei der Analyse der Daten soll vor allem der Wortschatz im Mittelpunkt stehen, da dieser ausgehend vom Sprachmodell der kommunikativen Kompetenz nach Bachman; Palmer (1996) im TestDaF in unterschiedlichen Bewertungskriterien eine Rolle spielt: Einerseits wird der Wortschatz im TestDaF explizit durch das Kriterium *Sprachliche Realisierung* bewertet, in dem u.a. beurteilt wird, ob der Wortschatz zur Bewältigung der Aufgabe ausreichend und präzise ist. Andererseits spielt der Wortschatz auch in die Beurteilung des Kriteriums *Gesamteindruck* hinein, in dem erfasst wird, inwieweit ein muttersprachlicher Leser den Text flüssig lesen und den Gedankengang sowie den formalen Aufbau des Textes nachvollziehen kann.⁵

2 Forschungsstand

Das Lexikon eines Fremdsprachenlerner wird allgemein nicht nur als ein guter Prädiktor für seine Sprachfähigkeit in der Fremdsprache verstanden (Vidaković; Barker 2010), sondern es gilt auch als ein entscheidender Faktor für seinen Ausbildungserfolg (vgl. Llach 2007: 2). Als gute Indikatoren für die Sprachkompetenz eines Lerner gelten dabei besonders die Länge seiner produzierten mündlichen und schriftlichen Texte, die Zahl der unterschiedlichen Wörter, die er verwendet, sowie die Type-Token-Relation (vgl. Vidaković; Barker 2010).

Die Begriffe *Type* und *Token* werden in der Sprachwissenschaft benutzt, um die absolute Zahl der Wörter (Token) von der Zahl der unterschiedlichen Wörter (Type) in einem Text abzugrenzen. Die meisten Arbeiten unterscheiden dabei nicht weiter zwischen Wort und Lexem und rechnen daher *gehe*, *gehst*, *gegangen* als drei verschiedene Types, jedes Vorkommen der jeweiligen Wortform wiederum als ein Token.

Die Auswertung der Textlänge und der Zahl der unterschiedlichen verwendeten Wörter gibt jedoch nur Aufschluss über die Varianz des Wortschatzes, Aussagen über die Qualität des verwendeten Wortschatzes sind auf der Grundlage der rein intrinsischen Messeinheiten von Type und Token nicht möglich. Erst ein Vergleich mit einem außenstehenden Kriterium wie beispielsweise einer Häufigkeitsliste kann Aufschluss über die Breite und Komplexität des verwendeten Wortschatzes geben.

Für das Deutsche stehen im Gegensatz zum Englischen (vgl. Taylor; Barker 2008) nur wenige aktuelle Häufigkeitslisten oder Korpora zur Verfügung. Eine Ausnahme bildet hier das Frequenzwörterbuch von

⁴ Zum Konstrukt des Prüfungsteils Schriftlicher Ausdruck im TestDaF vgl. auch Zimmermann (2010).

⁵ Vgl. auch die Beschreibung der Sprachkompetenz im Gemeinsamen europäischen Referenzrahmen für Sprachen (GER). Hier wird die lexikalische Kompetenz nach Breite und Beherrschung des Wortschatzes unterschieden und als zentraler Bestandteil der linguistischen Kompetenzen eines Lerner angeführt (Europarat 2001: 110ff.)

Jones; Tschirner (2006), das auf einem Korpus von 4,2 Millionen Tokens geschriebener und gesprochener Sprache basiert.⁶

Durch den Abgleich der vorliegenden Texte mit einem solchen externen Kriterium können zwar Aussagen hinsichtlich der Wortschatzbreite getroffen werden, allerdings nur auf Einzelwortebene. Neuere Ansätze gehen daher dazu über, in mündlichen wie auch in schriftlichen Texten, Einheiten zu untersuchen, die aus mehreren Wörtern bestehen. Diese werden oft als *Cluster*, *Chunks* oder *lexikalische Bündel* bezeichnet (vgl. Hyland 2008: 4). Vidaković; Barker (2010) kommen in ihrer Studie zu dem Ergebnis, dass sowohl die Anzahl als auch die Art der verwendeten 4-Wort-Bündel ein zuverlässiger Indikator für das Sprachniveau ist. Sie beziehen sich dabei auf Biber; Barbieri (2007), die die lexikalischen Bündel nach ihrer Funktion in drei Gruppen unterteilen: Bündel, die die Meinung des Schreibers bzw. Sprechers wiedergeben (*stance*); Bündel, die den Diskurs strukturieren (*discourse organizer*), und referentielle Bündel (*referential*), die auf ein Objekt oder ein Attribut verweisen.⁷ Insgesamt stellen Vidaković; Barker (2010) fest, dass mit steigendem Sprachniveau mehr lexikalische Bündel verwendet werden. Zum anderen werden niveauübergreifend referentielle Bündel am häufigsten eingesetzt, diskursorganisierende Bündel hingegen werden auf den unteren Sprachniveaus kaum verwendet, erst ab dem Sprachniveau B2/C1 kommen sie vermehrt vor.

3 Methode

3.1 Teilnehmer

Die Gruppe der Muttersprachler (n= 28) setzte sich aus 19 weiblichen und neun männlichen Studenten zusammen, die im Durchschnitt 21 Jahre alt waren. Die Mehrheit der Probanden kam aus den Fachbereichen Sprach- und Kulturwissenschaften sowie Rechts- und Sozialwissenschaften und befand sich zum Zeitpunkt der Durchführung der Studie im zweiten Hochschulsemerster. In einem Begleitfragebogen gaben alle Studierenden an, dass sie bereits schriftliche Leistungsnachweise an der Hochschule in Form von Klausuren, Hausarbeiten und Protokollen erbracht hatten, acht von ihnen hatten auch einen Kurs zum wissenschaftlichen Arbeiten belegt.

Unter den Muttersprachlern befanden sich auch Studierende, die neben dem Deutschen über eine weitere muttersprachliche Kompetenz verfügen, darunter drei Studierende mit türkischsprachigem Hintergrund sowie ein Student mit Französisch als weiterer Muttersprache.

Zusätzlich wurden aus dem Prüfungssatz, den auch die muttersprachlichen Studierenden bearbeitet hatten, eine gleich große Anzahl (n=28) von TestDaF-Prüfungsleistungen zufällig ausgewählt. Bei dieser Gruppe handelt es sich zu einem Großteil um weibliche Teilnehmer, die im Durchschnitt 23 Jahre alt waren und überwiegend aus Mittel- und Osteuropa stammten.⁸

3.2 Beurteilung

Die von den Muttersprachlern erbrachten schriftlichen Leistungen wurden von geschulten TestDaF-Beurteilern zusammen mit den Texten von Prüfungsteilnehmern aus aller Welt entsprechend der vorgegebenen Kriterien beurteilt, ohne dass diese wussten, dass es sich um muttersprachliche Leistungen handelte. Die Texte wurden daher unvoreingenommen von den Beurteilern nach insgesamt neun gleichgewichteten Kriterien im Bereich *Gesamteindruck*, *Behandlung der Aufgabe* und *Sprachliche Realisierung* beurteilt. Die endgültige Einstufung und Zuordnung der Leistung zu einer der drei TestDaF-Niveaustufen (TDN 3, TDN 4, TDN 5)⁹ erfolgte im TestDaF-Institut mittels Multifacetten-Rasch-Analysen¹⁰.

⁶ Das Herder/BYU-Korpus wurde in Zusammenarbeit zwischen dem Herder-Institut der Universität Leipzig und der amerikanischen Brigham Young University erstellt und ist nicht öffentlich zugänglich. Zu einer genaueren Beschreibung siehe Tschirner (2006).

⁷ Daneben führt Nekrasova (2009) noch eine weitere Gruppe von lexikalischen Bündeln an, die sogenannten *special conversational bundles*, die u.a. Höflichkeit ausdrücken (nach Vidaković; Barker 2010: 8).

⁸ Weitere Daten zum Spracherwerb und/oder Studienhintergrund dieser Gruppe liegen nicht vor, da das TestDaF-Institut solche biografischen Daten seiner Prüfungsteilnehmer erst seit 2010 systematisch erfasst.

⁹ Die TestDaF-Niveaustufen 3, 4 und 5 entsprechen den Stufen B2 und C1 des Gemeinsamen europäischen Referenzrahmens für Sprachen, vgl. hierzu auch Kecker (2011).

¹⁰ Das Multifacetten-Verfahren, das routinemäßig vom TestDaF-Institut eingesetzt wird, wird u.a. bei Eckes (2011) näher beschrieben.

3.3 Textanalyse

Um die Leistungen detaillierter analysieren zu können, wurden sowohl die Texte der Muttersprachler als auch die der fremdsprachlichen Vergleichsgruppe anschließend im TestDaF-Institut aufbereitet: Die handschriftlichen Texte wurden digitalisiert und mithilfe eines Part-of-Speech-Taggers (kurz: POS-Tagger) bearbeitet. Der POS-Tagger, in diesem Fall das Programm Morphy¹¹, weist jedem Wort eine Wortart und eine Grundform zu. Verglichen mit anderen gängigen POS-Taggern zeichnet das Programm Morphy sich durch seine Benutzerfreundlichkeit aus. Es hat zudem den großen Vorteil, dass man das Lexikon des Programms unproblematisch erweitern kann. Gerade bei Texten, die sich sprachlich von den Texten der Trainingskorpora (hauptsächlich Zeitungstexte) unterscheiden, kann man so die Korrektheitsrate des Programms erhöhen. Mithilfe von POS-Taggern können größere Mengen von Daten schnell ausgewertet werden: Man kann beispielsweise gezielt nach bestimmten Wortarten oder Kombinationen von Wortarten suchen und verlässliche Angaben dazu machen, wie viele unterschiedliche Wörter in einem Text vorkommen. Problematisch sind bei der Auswertung allerdings homonyme Formen (*sein Haus/es soll so sein*), trennbare Partikelverben (*die Grafik stellt ... dar*), feste Fügungen (*im Großen und Ganzen*), Abkürzungen (z.B.) und Wörter, die nicht im Lexikon des Programms gespeichert sind. Da unser Korpus relativ klein ist und die Texte unterschiedliche Arten von Fehlern enthalten, wurden die Daten von Hand nachkorrigiert.

Die so aufbereiteten Texte wurden abschließend mithilfe der Software WordSmith Tools (Scott 2008) hinsichtlich verschiedener Parameter ausgewertet: So wurden zum einen die Zahl der Token und Types sowie die Type-Token-Relation ermittelt. Dabei haben wir alle Wortformen eines Lexems zusammengefasst und betrachten z.B. *gehe*, *gehst* und *gegangen* als unterschiedliche Realisierungen (Token) des Types *gehen*. Diese Modifikation wurde vorgenommen, da wir uns so präzisere Angaben zu Variation und Breite des Wortschatzes erhoffen. Zudem stammen die meisten Arbeiten, die mit den Begriffen Type und Token arbeiten, aus dem englischsprachigen Raum. Verglichen mit dem Englischen flektiert das Deutsche aber stärker, sodass jedem Lexem mehr Wortformen (Types im traditionellen Sinne) gegenüberstehen.

Des Weiteren wurden Wortlisten erstellt und diese mit einer Liste der 4000 häufigsten Wörter des Deutschen verglichen, um Angaben über die Verteilung von hoch- und niedrigfrequenten Wörtern machen zu können. Außerdem haben wir die Texte mithilfe von WordSmith Tools nach wiederkehrenden Wortkombinationen (lexikalischen Bündeln) durchsucht.

4 Ergebnisse

4.1 Bewertung der Texte

Die Ergebnisermittlung zeigte, dass eine überwiegende Mehrheit (n=23) der muttersprachlichen Texte der höchsten TestDaF-Niveaustufe zugeordnet wurde, es aber auch Leistungen gab, die darunter eingestuft wurden. Für diese niedrigeren Einstufungen war v.a. die inhaltliche Umsetzung der Aufgabe ausschlaggebend, besonders die Beschreibung der Grafik bereitete den Muttersprachlern Schwierigkeiten.¹² Sprachlich dagegen wurden die muttersprachlichen Texte fast durchgängig der höchsten Niveaustufe zugeordnet.

4.2 Wortschatz

In einem ersten Schritt haben wir die Type- und Token-Zahlen der Texte ermittelt, aufgeschlüsselt in Muttersprachler ohne die Gruppe der Bilingualen, die Gruppe der bilingualen Muttersprachler und die Gruppe der Fremdsprachler.

Für die analysierten Texte ergibt sich folgendes Bild:

¹¹ Das von Wolfgang Lezius erarbeitete frei verfügbare Softwarepaket wird nicht mehr weiter entwickelt. Die von uns verwendete Version kann unter <http://www.wolfganglezius.de/doku.php?id=cl:morphy> (15.10.2012) heruntergeladen werden.

¹² Dies deckt sich mit statistischen Auswertungen der Prüfungsleistungen von Fremdsprachlern: Das Kriterium Grafikbeschreibung ist eines der schwersten Kriterien, d.h. die Wahrscheinlichkeit, dass ein Teilnehmer in diesem Kriterium die höchste TDN-Einstufung bekommt, ist im Vergleich zu den anderen Kriterien geringer.

Tab. 1: Durchschnittliche Zahl der Token und Types

	<i>M</i> Token	<i>SD</i> Token	<i>M</i> Type	<i>SD</i> Type
Muttersprachler (ohne Bilinguale)	560,6	113,7	236,3	32,3
Bilinguale Muttersprachler	531	71	226,6	20,5
Fremdsprachler	467,3	79,5	195,2	23,4

Erläuterung: *M*= Durchschnitt, *SD*= Standardabweichung

Es wird deutlich, dass die Muttersprachler im Durchschnitt längere Texte als die Fremdsprachler produzieren und auch mehr unterschiedliche Wörter verwenden. Die Texte der bilingualen Muttersprachler sind etwas kürzer als die der monolingualen und auch die Zahl der Types ist etwas niedriger, die Unterschiede sind aber nicht auffällig. Einzig ein Text einer Studierenden mit Türkisch als weiterer Muttersprache neben dem Deutschen fällt auf: Mit 198 Types liegt er deutlich unter dem Durchschnitt der Muttersprachler und näher an den Werten der Fremdsprachler.

Die Standardabweichung (*SD*) zeigt, dass es bei den Muttersprachlern größere individuelle Unterschiede in der Textlänge gibt als bei den Fremdsprachlern oder den bilingualen Muttersprachlern.

Setzt man die Zahl der Types und Token in Beziehung zueinander, um für die Texte Aussagen darüber machen zu können, wie sich die Wortschatzbreite in Relation zur Textlänge verhält, so zeigen sich keine auffälligen Unterschiede zwischen den einzelnen Gruppen mehr (vgl. Tab. 2).

Tab. 2: Durchschnittliche Type-Token-Relation

	Type-Token-Relation (TTR) ¹³	<i>SD</i>
Muttersprachler (ohne Bilinguale)	42,7	4,0
Bilinguale Muttersprachler	43,1	5,5
Fremdsprachler	42,3	4,9

Erläuterung: *SD*= Standardabweichung

Wie die Standardabweichung zeigt, sind die Unterschiede zwischen den Texten innerhalb einer Gruppe hier deutlich größer als die Unterschiede zwischen den Durchschnittswerten der drei Gruppen.

Problematisch bei der Verwendung der Type-Token-Relation (TTR) ist, dass sie bei zunehmender Textlänge sinkt, da man allgemein von der Annahme ausgeht, dass Schreiber, wenn sie mehr Text produzieren, immer weniger Wörter verwenden, die bisher noch nicht darin vorgekommen sind (vgl. Yu 2009: 236; Read 2000: 201).¹⁴ Dies zeigt sich auch bei unseren relativ kurzen Texten: Die niedrigste TTR in unserem Sample haben die Texte, die deutlich länger als der Durchschnitt sind, obwohl sie, absolut gerechnet, die höchste Type-Zahl aufweisen.

Genaueren Aufschluss über die Varianz des Vokabulars bietet der Vergleich der Texte mit der Liste der häufigsten 4000 Wörter des Deutschen. Tschirner (2006: 1284) gibt an, dass in seinem Korpus akademische Texte (Fachtexte, die aus Fachzeitschriften und Einführungswerken ausgewählt wurden) zu 81 % aus den häufigsten 4000 Wörtern des Korpus bestehen.

¹³ Die Type-Token-Relation ergibt sich aus der Zahl der Types, dividiert durch die Zahl der Token, wobei dieser Wert abschließend mit 100 multipliziert wird. Je weniger Wiederholungen es in einem Text gibt, desto größer ist der Wert der Type-Token-Relation.

¹⁴ Es gibt verschiedene Ansätze, diese Tendenz rechnerisch auszugleichen, eine davon ist die Standardisierte Type-Token-Relation (STTR). Hierbei wird über den Text für jeweils 50 Wörter die TTR ermittelt und aus den Werten schließlich ein Durchschnittswert berechnet.

Da die Aufgabenstellung des TestDaF das Verfassen eines kohärenten und strukturierten Textes fordert, der in einem akademischen Kontext angesiedelt ist, gehen wir davon aus, dass die Texte nicht nur hochfrequente Wörter enthalten, sondern einen ähnlichen Anteil von niedrigfrequenten Wörtern haben wie akademische Texte.

Für die von uns analysierten Texte ergibt sich folgendes Bild:

Tab. 3: Textdeckung durch die häufigsten 2000 und 4000 Wörter des Deutschen

	K 2000	SD	K 4000	SD
Muttersprachler (ohne Bilinguale)	80,6 %	3,18	83,7 %	2,72
Bilinguale Muttersprachler	80,1 %	0,7	82,9 %	1,09
Fremdsprachler	85,4 %	2,29	86,9 %	2,13

Erläuterung: K 2000= die häufigsten 2000 Wörter des Deutschen, K 4000= die häufigsten 4000 Wörter des Deutschen, SD= Standardabweichung

In der Gruppe der Muttersprachler gibt es keinen Unterschied zwischen den einsprachigen und den bilingualen Sprechern, die Fremdsprachler hingegen verwenden deutlich mehr hochfrequente Wörter, vor allem die 2000 häufigsten Wörter machen hier einen deutlich größeren Teil des Textes aus als bei den Muttersprachlern, die nächsten 2000 Wörter tragen kaum mehr zur Textdeckung bei.

Insgesamt entsprechen die Ergebnisse aber unseren Erwartungen, vor allem die Texte der Muttersprachler zeigen eine ähnliche Verteilung von hoch- und niedrigfrequenten Wörtern wie die Texte in Tschirners akademischem Subkorpus.

4.3 Lexikalische Bündel

In Anlehnung an bestehende Forschungen zu lexikalischen Bündeln (Biber; Barbieri 2007; Hyland 2008) wurden die Texte der Muttersprachler und Fremdsprachler nach 4-Wort-Gruppen untersucht, da sich so die größte Menge relevanter Muster finden lässt.¹⁵ Bei der Analyse wurden nur die lexikalischen Bündel berücksichtigt, die im jeweiligen Subkorpus – also in den Texten der Muttersprachler ohne Bilinguale, nur in den Texten der bilingualen Muttersprachler und in den Texten der Fremdsprachler – mindestens dreimal und in drei verschiedenen Texten vorkommen. Wortgruppen, die wortwörtlich oder in leicht abgewandelter Form aus der Aufgabenstellung übernommen wurden, wurden nicht berücksichtigt.

Die gefundenen Bündel haben wir in Anlehnung an Biber; Barbieri (2007) in referentielle, diskursstrukturierende Bündel und Bündel, die die Meinung des Sprechers wiedergeben, unterteilt. Nach Vidaković; Barker (2010) hängt, wie bereits erwähnt, die Art und Menge der verwendeten lexikalischen Bündel vom Sprachniveau ab. Da es sich bei unseren Daten um Texte von fortgeschrittenen Lernern und Muttersprachlern handelt, würden wir erwarten, dass die Texte viele lexikalische Bündel enthalten und neben referentiellen Bündeln, die zu den häufigsten Arten von lexikalischen Bündeln gehören, auch diskursstrukturierende Bündel zu finden sind, die nach Vidaković; Barker (2010) erst ab dem Sprachniveau B2/C1 vorkommen. Weiterhin könnte man erwarten, dass die Texte der Muttersprachler mehr lexikalische Bündel enthalten als die Texte der Fremdsprachler.

Tatsächlich jedoch bestätigen die Ergebnisse der Analyse diese Hypothese nicht in allen Punkten: Die Texte der Fremdsprachler weisen insgesamt deutlich mehr Bündel auf, bei den Muttersprachlern finden sich insgesamt nur vier Bündel, verglichen mit 23 bei den Fremdsprachlern.¹⁶ Betrachtet man die gefundenen lexikalischen Bündel genauer, stellt man fest, dass der überwiegende Teil der Bündel im Kontext der Grafikbeschreibung verwendet wird und der Gruppe der referentiellen Bündel zuzuordnen ist, die nächstgrößere Gruppe bilden Bündel, die die Haltung des Sprechers ausdrücken (vgl. Tab. 4).

¹⁵ Diese Festlegung gilt jedoch hauptsächlich für das Englische, für die deutsche Sprache müsste der Umfang der Bündel noch bestätigt werden.

¹⁶ Für die Daten der Bilingualen lassen sich hier nur sehr bedingt Aussagen machen, weil das Korpus zu klein für eine aussagekräftige Analyse ist.

Tab. 4: Lexikalische Bündel in den Texten der Fremdsprachler

<i>referential</i>	<i>Aus den Jahren 2007 / Aus dem Jahr 2007</i> <i>Stammen aus der HIS-Studienanfängerbefragung / Daten stammen aus der / Die Daten stammen aus / Studienfachs die Daten stammen</i> <i>Mit dem Titel Wahl / Tabelle mit dem Titel / Dem Titel aus welchen / Mit dem Titel aus / Titel aus welchen Gründen</i> <i>Grafik mit dem Titel / Die Grafik mit dem</i>
<i>stance</i>	<i>Bin der Meinung, dass / Ich bin der Meinung</i> <i>Meiner Meinung nach ist</i> <i>Kann man sagen, dass</i> <i>Diskutiert wird die</i> <i>Richtige Entscheidung zu treffen</i>

Lexikalische Bündel, die diskursstrukturierend eingesetzt werden, finden sich nur sehr selten: insgesamt nur ein Bündel dieser Art in den Texten der Muttersprachler und zwei diskursstrukturierende Bündel in den Texten der Fremdsprachler.

5 Zusammenfassung und Diskussion

Die quantitative Auswertung der Texte hat zunächst einmal gezeigt, dass sich die Texte der Muttersprachler mit Deutsch als Zweitsprache in Bezug auf Textlänge und Wortschatzbreite nicht nennenswert von den Texten der einsprachigen Muttersprachler unterscheiden. Auch zwischen der Gruppe der Muttersprachler und den fortgeschrittenen Fremdsprachlern sind die Unterschiede hinsichtlich dieser Untersuchungsparameter eher gering. Dennoch gibt es ein paar Auffälligkeiten:

Es zeigt sich, dass es eine relativ große Streuung der Daten innerhalb der einzelnen Gruppen gibt, besonders die Gruppe der Muttersprachler ist diesbezüglich auffällig. Somit bestätigt sich auch die eingangs erwähnte Skepsis gegenüber einer muttersprachlichen Referenznorm, die auf der Annahme beruht, dass es sich hierbei um eine homogene Gruppe handelt. Die Daten der Fremdsprachler streuen nicht ganz so breit, die Gruppe scheint aufgrund der rein quantitativen Daten homogener, was aber auch daran liegen kann, dass durch die Art der Prüfungsvorbereitung und der zur Verfügung stehenden Vorbereitungsmaterialien sich die Texte sprachlich und inhaltlich ähnlich sind.¹⁷

Hinsichtlich der Wortschatzbreite hat die Analyse der Texte ergeben, dass sich die absoluten Type-Zahlen nur bedingt als Prädiktor für die Einstufung des Kriteriums *Wortschatz* im TestDaF eignen: Zwar haben alle Texte, die in diesem Kriterium nicht auf dem höchsten Niveau eingestuft wurden, eine unterdurchschnittlich niedrige Type-Zahl, umgekehrt hat aber eine unterdurchschnittliche Type-Zahl nicht unbedingt eine niedrigere Einstufung im Kriterium *Wortschatz* zur Folge, denn ein Text mit der niedrigsten Type-Zahl beispielsweise wurde in diesem Kriterium auf dem höchsten Niveau eingestuft.¹⁸

Zudem wurde deutlich, dass sowohl die Muttersprachler als auch die Fremdsprachler weniger hochfrequente Wörter verwenden als erwartet. Die wenigsten hochfrequenten Wörter verwenden die bilingualen Muttersprachler, allerdings ist der Abstand zu den einsprachigen Muttersprachlern gering. Der Unterschied zwischen Muttersprachlern und Fremdsprachlern hingegen ist auffälliger. Insgesamt stützt der Vergleich mit hochfrequenten Wörtern des Deutschen die These, dass Muttersprachler einen größeren Anteil niedrigfrequenter Wörter in ihren Texten verwenden.

Die Annahme, dass Muttersprachler mehr lexikalische Bündel verwenden als die Fremdsprachler, hat sich nicht bestätigt. Dieses unerwartete Ergebnis bedeutet jedoch nicht, dass die muttersprachlichen Texte keine lexikalischen Bündel enthalten, sondern legt eher die Vermutung nahe, dass die Muttersprachler eine so große Bandbreite an Formulierungen verwenden, dass es kaum Wortfolgen gibt,

¹⁷ Zur Prüfungsvorbereitung vgl. Zimmermann (2009).

¹⁸ Ein Zusammenhang zwischen der Einstufung im Kriterium *Wortschatz* und der Einstufung im Kriterium *Gesamteindruck* lässt sich aus den Texten nicht ableiten.

die sich über verschiedene Texte hinweg wiederholen. Die Fremdsprachler hingegen greifen eher auf ein begrenztes Set an vorgefertigten Redemitteln zurück.

Hier wird bereits eines der Probleme einer rein quantitativen Auswertung deutlich: Schon leichte Abweichungen in der Formulierung sorgen dafür, dass eine Wortgruppe nicht als lexikalisches Bündel erkannt wird bzw. als neues Bündel gezählt wird und somit in der Gesamtauswertung unberücksichtigt bleibt. Gleiches gilt für diskontinuierliche Redemittel wie „*einerseits ... andererseits*“.

Problematisch ist außerdem, dass es sich bei den vorliegenden Daten um ein sehr kleines Sample handelt. So werden zwar beim Lesen der Texte Unterschiede zwischen Muttersprachlern und Bilingualen offenbar, sie werden aber statistisch nicht erfasst. Dies betrifft beispielsweise im Bereich der Lexik kreative Wortschöpfungen (*Sprach- und Kulturwissenschaftenliebhaber* anstatt Studierende der Sprach- und Kulturwissenschaften), Schwierigkeiten bei der Wahl des passenden Wortes (*Visualisierung* statt Grafik, *Einstimmung* anstelle von Zustimmung), aber auch grammatikalische Unsicherheiten bei der Wahl von Präpositionen (*Jedoch wird in einigen Bereichen aufgrund des Forschungsgeldmangels geklagt*).

Eine abschließende qualitative Analyse der Texte, die besonders im Bereich der Lexik auch den Textkontext und die richtige Verwendung mit berücksichtigt, könnte hier weitere Aufschlüsse geben.

Um fundierte Aussagen über eine Spezifik des Deutschen als Zweitsprache machen zu können, wäre zudem ein breiteres Lernerkorpus notwendig.

Literatur

- Ahrenholz, Bernt (2010): Erstsprache – Zweitsprache – Fremdsprache. In: Ahrenholz, Bernt; Oomen-Welke, Ingelore (Hrsg.): *Deutsch als Zweitsprache*. (2. korrigierte und überarbeitete Auflage) Baltmannsweiler: Schneider Verlag Hohengehren, 3-16.
- Alderson, J. Charles; Clapham, Caroline; Wall, Diane (1995): *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bachman, Lyle F.; Palmer, Adrian S. (1996): *Language Testing in Practice*. Oxford: University Press.
- Biber, Douglas; Barbieri, Federica (2007): Lexical bundles in university spoken and written registers. *English for Specific Purposes* 26, 263-286.
- Davies, Alan (2008): The Native Speaker in Applied Linguistics. In: Davies, Alan; Elder, Catherine (Hrsg.): *The Handbook of Applied Linguistics*. (5. Auflage) Oxford: Blackwell Publishing, 431-450.
- Eckes, Thomas (2003): Qualitätssicherung beim TestDaF: Konzepte, Methoden, Ergebnisse. *Fremdsprachen und Hochschule* 69, 43-68.
- Eckes, Thomas (2011): *Introduction to many-facet Rasch measurement: Analyzing and evaluating rater-mediated assessments*. Frankfurt: Lang.
- Europarat (2001): *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin u.a.: Langenscheidt.
- Hyland, Ken (2008): As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes* 27, 4-21.
- Jones, Randall L.; Tschirner, Erwin (2006): *A Frequency Dictionary of German. Core vocabulary for learners*. London, New York: Routledge.
- Kecker, Gabriele (2011): *Validierung von Sprachprüfungen. Die Zuordnung des TestDaF zum Gemeinsamen europäischen Referenzrahmen für Sprachen*. Frankfurt: Lang.
- Llach, María Pilar Agustín (2007): Lexical Errors as Writing Quality Predictors. *Studia Linguistica* 61 (1), 1-9.
- Read, John (2000): *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Scott, Mike (2008): WordSmith Tools, version 5. Liverpool: Lexical Analysis Software.
- Taylor, Lynda; Barker, Fiona (2008): Using Corpora for Language Assessment. In: Shohamy, Elena; Hornberger, Nancy H. (Hrsg.): *Encyclopedia of Language and Education. Vol 7: Language Testing and Assessment*. New York: Springer, 241-254.

- Tschirner, Erwin (2006): Häufigkeitsverteilungen im Deutschen und ihr Einfluss auf den Erwerb des Deutschen als Fremdsprache. In: Corino, Elisa; Marelllo, Carla; Onesti, Cristina (Hrsg.): *Atti del XII Congresso Internazionale di Lessicografia*. Alessandria: Edizioni dell' Orso, 1277-1288.
- Vidaković, Ivana; Barker, Fiona (2010): Lexical development across second language proficiency levels: a corpus-informed study. In: Millar, Robert McColl; Durham, Mercedes (Eds.): *Language, Learning & Context: Proceedings of the 42nd Annual Meeting of the British Association for Applied Linguistics*. London: Scitsiugnil Press, 143-146.
- Yu, Guoxing (2009): Lexical Diversity in Writing and Speaking Task Performances. *Applied Linguistics* 31 (2), 236-259.
- Zimmermann, Sonja (2009): Falsche Vorbereitung? – Erkenntnisse aus Teilnehmerleistungen der Prüfungsteile Mündlicher und Schriftlicher Ausdruck im Test Deutsch als Fremdsprache (TestDaF). In: Yu, Xuemei (Hrsg.): *TestDaF-Training und Studienvorbereitung: Beiträge zur chinesisch-deutschen Fachkonferenz: „TestDaF-Training und Studienvorbereitung“ vom 11. bis 12. Oktober 2008 am Deutschkolleg der Tongji-Universität Shanghai*. München: Iudicium, 63-83.
- Zimmermann, Sonja (2010): Schreiben im Test – Schreiben im Studium: Wie können Anforderungen an das akademische Schreiben in Tests abgebildet werden? Das Beispiel TestDaF. In: Brandl, Heike; Duxa, Susanne; Leder, Gabriela & Riemer, Claudia (Hrsg.): *Ansätze zur Förderung akademischer Schreibkompetenz an der Hochschule (Materialien Deutsch als Fremdsprache, Bd. 83)*. Göttingen: Universitätsverlag, 1-10.