

Zarah Weiss, Kim Lange-Schubert, Barbara Geist und  
Detmar Meurers\*

# Sprachliche Komplexität im Unterricht

## Linguistic complexity in the classroom

### Eine computerlinguistische Analyse der gesprochenen Sprache von Lehrenden und Lernenden im naturwissenschaftlichen Unterricht in der Primar- und Sekundarstufe

A computational-linguistic analysis of the spoken language of teachers and learners in science education in primary and secondary schools

<https://doi.org/10.1515/zgl-2022-2052>

**Abstract:** We investigate the linguistic complexity of oral classroom interactions in late primary and early secondary school across German school types. The goal is to explore whether teachers and students align in terms of their use of the academic language register. We empirically base this investigation on transcriptions of teacher and student contributions during content matter lessons on the vaporisation and condensation of water. Across school types and grade levels, we compare the extent to which teachers offer language that is adaptively rich in linguistic constructs commonly associated with academic language, such as deagentivation, nominal style, and cohesive devices. Putting this in relation to the developing academic language competence of the students, we then compare the language offered by the teachers to the use of these academic language constructs in the students' spoken language contributions. We discuss the methodological challenges arising from analyzing oral classroom interactions and from applying automatic linguistic complexity analyses to such data.

---

\***Kontaktpersonen:** Zarah Weiss: Universität Tübingen, Seminar für Sprachwissenschaft, Wilhelmstraße 19, D-72074 Tübingen, E-Mail: [zweiss@sfs.uni-tuebingen.de](mailto:zweiss@sfs.uni-tuebingen.de)

**Prof. Dr. Kim Lange-Schubert:** Universität Leipzig, Erziehungswissenschaftliche Fakultät, Marschnerstraße 31, D-04109 Leipzig, E-Mail: [kim.lange-schubert@uni-leipzig.de](mailto:kim.lange-schubert@uni-leipzig.de)

**Dr. Barbara Geist:** Pädagogische Hochschule Freiburg, Institut für deutsche Sprache und Literatur, Kunzenweg 21, D-79117 Freiburg, E-Mail: [barbara.geist@ph-freiburg.de](mailto:barbara.geist@ph-freiburg.de)

**Prof. Dr. Detmar Meurers:** Universität Tübingen, Seminar für Sprachwissenschaft, Wilhelmstraße 19, D-72074 Tübingen, E-Mail: [dm@sfs.uni-tuebingen.de](mailto:dm@sfs.uni-tuebingen.de)

1	Einleitung
2	Hintergrund
2.1	Bildungsspracherwerb im sprachsensiblen Fachunterricht
2.2	Bildungssprache und sprachliche Komplexität
3	Daten
4	Automatische Messung sprachlicher Komplexität
5	Studie 1: Vergleich der Lehrkräfte nach Schulformen
5.1	Aufbau von Studie 1
5.2	Ergebnisse von Studie 1
6	Studie 2: Vergleich von Lehrkräften und Schüler*innen
6.1	Aufbau von Studie 2
6.2	Ergebnisse von Studie 2
7	Diskussion
8	Ausblick
	Danksagung
	Fördermittel
	Literatur

## 1 Einleitung

Bildungssprachliche Kompetenzen sind laut der Kultusministerkonferenz ein obligatorisches Bildungsziel, das in der Schule vermittelt werden soll (KMK, 2013, S. 5). Dieser Erwerb soll fachübergreifend über alle Schulstufen gefördert werden, da Schüler\*innen bildungssprachliche Kompetenzen nur durch gezielte Heranführung (Cummins, 2000) und unterstützendes *Scaffolding* (Gibbons, 2002) erwerben. Sprachliche Bildung sollte somit auch im Fachunterricht stattfinden, sowohl in der Sekundar- als auch in der Primarstufe (Feilke, 2012, S. 5), und adaptiv auf den sprachlichen Fähigkeiten der Schüler\*innen aufbauen (Kleinschmidt-Schinke, 2018). Dies stellt Lehrkräfte vor besondere Herausforderungen, da sie dazu auf die individuellen Bedürfnisse von Schüler\*innen mit heterogenen sprachlichen Lernvoraussetzungen eingehen sollten. Inwieweit ihnen dies im Sprach- und Fachunterricht gelingt, ist eine zentrale Frage. Studien zu diskursiven und bildungssprachlichen Praktiken im Fachunterricht liegen inzwischen für verschiedene Schulfächer und Praktiken wie Erklären und Argumentieren vor (u. a. Erath, Prediger, Quasthoff & Heller, 2018; Heller & Morek, 2015). In gesprächsanalytischen Studien wird zudem der Frage nachgegangen, wie Schüler\*innen im Unterricht zu diesen diskursiven Praktiken aufgefordert werden, wie sie diese realisieren, und ausgehend davon, wie sie im Erwerb unterstützt werden (z. B. Heller, 2017). Berendes et al. (2018) beschäftigen sich mit der sprachlichen Komplexität des Inputs für Schüler\*innen und dessen Entwicklung im Verlauf

der Schulzeit und zeigen in einer computerlinguistischen Analyse der aktuell zugelassenen Geographiebücher, dass deren sprachliche Komplexität nicht systematisch nach Klassenstufe zunimmt und sich Schulbücher für verschiedene Schulformen nicht systematisch in ihrer sprachlichen Komplexität unterscheiden. Kleinschmidt-Schinke (2018) untersucht die sprachliche Komplexität der Äußerungen von acht Lehrkräften in Grundschule (Klasse 3) und Gymnasium im Deutsch- oder im Sach- bzw. Biologieunterricht. Die Studie zeigt eine Veränderung der an die Schüler\*innen gerichteten Sprache „über Jahrgangsstufen hin zu stärkerer konzeptioneller Schriftlichkeit“ (Kleinschmidt-Schinke, 2018, S. 555). Eine Analyse des gesamten gesprochenen Angebots von Lehrkräften für Schüler\*innen im Unterricht zwischen Klassenstufen und im Vergleich zwischen Schulformen der Sekundarstufe 1 fehlt jedoch. So benennt Kleinschmidt-Schinke (2018, S. 561) bspw. die Analyse der Sprache der Lehrpersonen in derselben Jahrgangsstufe in unterschiedlichen Schulformen als Desiderat.

Wir adressieren diese Forschungslücke mittels einer computerlinguistischen Analyse des Lerner\*innen-Korpus „Naturwissenschaftlicher Unterricht“, das im Rahmen der PLUS-Studie (DFG-Forschergruppe, Kennzeichen MO 942/3-1 und FI 477/21-1) elizitiert wurde. Es handelt sich hierbei nicht um ein Korpus für Zweitsprache, sondern um ein Lehr-Lern-Korpus, welches primär den Erwerb des deutschen Bildungssprachregisters in der Erst- oder Zweitsprache (L1/L2) in authentischen Lehr-Lern-Kontexten abbildet (vgl. Wisniewski, i. d. H.). Auf dieser Basis untersuchen wir folgende Forschungsfragen:

1. Unterscheiden sich Lehrkräfte über Schulformen hinweg systematisch im Hinblick auf die sprachliche Komplexität ihrer Äußerungen?
2. Besteht ein Zusammenhang zwischen der sprachlichen Komplexität von Lehrkraft- und Schüler\*innen-Äußerungen in Abhängigkeit von der Schulform?

Zur Beantwortung dieser Fragen untersuchen wir die sprachliche Komplexität der transkribierten Äußerungen von Lehrkräften und Schüler\*innen im Sach- und Physikunterricht der 4. und 6. Klassen in den Schulformen Grund- und Hauptschule sowie Gymnasium, die im Rahmen von Unterrichtseinheiten zum Thema „Aggregatzustände“ gesammelt wurden und in denen das Unterrichtsthema über alle Klassenstufen und Schulformen hinweg konstant gehalten wurde. In diesem Zusammenhang diskutieren wir auch den Einfluss von Gesprächsthemen, Diskursformen und diskursiven Funktionen in dieser Art von Daten in den Abschnitten 3 und 8.

Die Untersuchung leistet ebenfalls einen wichtigen methodischen Beitrag zur linguistischen Analyse von Korpora gesprochener Interaktion und der Lerner\*innen-Korpusforschung. Die Zahl der einschlägigen (Zweit-)Spracherwerbsstudien, die sich mit gesprochener Sprache beschäftigen, ist begrenzt und kon-

zentriert sich in der Regel auf die Analyse spezifischer sprachlicher Konstrukte, welche unter (quasi-)experimentellen Bedingungen eliziert wurden (z. B. De Clercq & Housen, 2019; Lintunen & Mäkilä, 2014; Pallotti, 2019). Eine Vielzahl bildungssprachlicher Adaptionstrategien sind im Unterricht denkbar, bspw. im Bezug auf die Variation und Häufigkeit spezifischen Fach-Vokabulars, die Einbettung von Nebensätzen oder die Nutzung von Nominalisierungen, welche auch bei Muttersprachler\*innen in der frühen Adoleszenz die Informationsverarbeitung beeinflussen können (Daroczy, Artemenko, Meurers, Wolska & Nuerk, 2020). Jedoch wurde bisher kaum empirisch etabliert, auf welche Konstruktionen es bei der sprachlichen Adaption tatsächlich ankommt. Es fehlt eine linguistisch breit aufgestellte empirische Untersuchung des frühen Bildungsspracherwerbs in ökologisch validen Lehr-Lern-Kontexten. Wir tragen hierzu mit unserer explorativen Analyse authentischer Unterrichtsgespräche bei, indem wir eine große Bandbreite von Merkmalen sprachlicher Komplexität aus den Bereichen Syntax, Lexikon, Morphologie, Diskurs, Sprachverarbeitung und Sprachnutzung untersuchen.

Die vorliegende Arbeit illustriert dabei auch, wie gesprochensprachliche Lehr-Lern-Korpora, die bislang konversationsanalytisch untersucht wurden, für quantitative Analysen zur sprachlichen Komplexität erschlossen und genutzt werden können. Dabei gehen wir insbesondere auf die spezifischen Anforderungen von Lehr-Lern-Korpora ein und legen dar, inwiefern diese durch solche Datenquellen bedient werden können, bspw. im Hinblick auf die Differenzierung von möglichst lautgetreuen Verschriftlichungen des Sprachsignals und normalisierten Transkriptionen und Zielhypothesen. Wir diskutieren, welche Chancen und Herausforderungen dabei für linguistische Analysen zur Beantwortung interdisziplinärer Fragestellungen bestehen, und stellen computerlinguistische und statistische Analyseverfahren vor, die sich für solche Untersuchungen besonders eignen.

Im folgenden Abschnitt wird die fach- und sprachdidaktische Diskussion zu Bildungssprache im Fachunterricht dargestellt und auf die bildungspolitischen Forderungen eingegangen (2.1). Anschließend wird der Forschungsstand zu sprachlicher Komplexität im Bildungsspracherwerb im Kontext der Zweitspracherwerbsforschung skizziert (2.2). Das Korpus wird in Abschnitt 3 beschrieben und unser Vorgehen für die automatische Messung sprachlicher Komplexitätsmerkmale in Abschnitt 4. Daraufhin adressieren wir unsere Forschungsfragen in zwei separaten Studien (5 und 6). Es folgt eine Diskussion der Ergebnisse (7) sowie ein Ausblick mit einem Fokus auf methodische Aspekte (8).

## 2 Hintergrund

### 2.1 Bildungsspracherwerb im sprachsensiblen Fachunterricht

Der Bildungsspracherwerb ist eine zentrale Voraussetzung zur gesellschaftlichen Teilhabe (Hüttis-Graff, Merklinger, Klenz & Speck-Hamdan, 2010) und für schulischen Lernerfolg (Bos, Valtin, Hußmann, Wendt & Goy, 2017; Kempert, Schalk & Saalbach, 2019). Dies gilt insbesondere auch im naturwissenschaftlichen Unterricht (Holtappels & Heerdegen, 2005). Während in den PISA-Studien für die Fächer Deutsch und Mathematik ein Rückgang des Einflusses von Bildungssprachkompetenz auf den Bildungserfolg über die letzten Jahre zu erkennen ist, bleibt er in den Naturwissenschaften konstant hoch (Reiss, Sälzer, Schiepe-Tiska, Klieme & Köller, 2015). Diese Leistungsunterschiede sind bereits im Elementarbereich zu erkennen (Bürgermeister, Große, Leuchter, Studhalter & Saalbach, 2018) und bleiben bis in die Sekundarstufe bestehen (Reiss et al., 2015), was die besondere Bedeutung von Sprache für den naturwissenschaftlichen Unterricht in allen Schulstufen und die Notwendigkeit der Untersuchung der im Unterricht gesprochenen Sprache unterstreicht.

Die Kultusministerkonferenz (KMK, 2013, S. 5) erklärt bildungssprachliche Kompetenz zu einem obligatorischen Bildungsziel. Dem Konstrukt der durchgängigen Sprachbildung unterliegt die Annahme einer horizontalen Dimension im Sinne eines fächerübergreifenden Sprachlernens und einer vertikalen Dimension, die Sprachbildung als Kontinuum beschreibt (Gogolin & Lange, 2010; Riebling, 2013; Schmölzer-Eibinger, 2013; Vollmer & Thürmann, 2013): Man geht von einer bildungsbiographischen Entwicklung aus, also dem Zuwachs an sprachlichem Wissen und Können mit zunehmendem Lernalter, welche einhergeht mit einem sukzessiven Ausbau des bildungssprachlichen Registers in den Institutionen der allgemeinen Bildung (Riebling, 2013). Dabei wird das bildungsbiographische Kontinuum durchschnitten von Zäsuren, die sich aus den institutionellen Grenzen im Bildungssystem, also den Übergängen zwischen den Bildungsstufen, ergeben. Diese Zäsuren können Brüche im Bildungsprozess darstellen oder als Ausbau der eigenen Fähigkeiten erlebt werden. Den sprachlichen Aspekten der Übergänge sollte daher im Modell der durchgängigen Sprachbildung besondere Aufmerksamkeit geschenkt werden (Riebling, 2013). Auch im fachdidaktischen Diskurs wird Sprache als bedeutender Faktor für das Lernen und als „Konstruktionsmittel für fachliche Verstehensprozesse“ (Leisen, 2015a, S. 135) bezeichnet. Der Sprache wird beim fachlichen Unterricht sowohl eine kommunikative Funktion bei der Vermittlung von Informationen als auch eine kognitive Funktion bei der Konstruktion neuen Wissens zugesprochen (Kempert et al., 2019). Man geht davon

aus, dass sich Bildungssprache im Fach „langsam und kumulativ über mehrere Stufen“ (Leisen, 2015b, S. 117) entwickelt und demnach mit fortschreitender Klassenstufe ein zunehmend hohes Maß an fachsprachlichen Elementen im Unterricht zu erwarten ist. Allerdings wird in der Fachliteratur häufig angemahnt, dass zu erwerbende Sprachkompetenzen vonseiten der Bildungsinstitutionen oftmals fälschlicherweise vorausgesetzt und eben nicht gelehrt werden (Feilke, 2012). Ein reines Hineinwachsen in die Bildungssprache wird als nicht ausreichend zurückgewiesen. Vielmehr wird eine Heranführung (Cummins, 2000) oder explizite Lehre nach dem *Scaffolding*-Prinzip (Gibbons, 2002) im Rahmen des sprachsensiblen Fachunterrichts eingefordert – auch in der Primarstufe (Feilke, 2012, S. 5). Dieses sprachliche Lerngerüst (*scaffold*) soll Schüler\*innen in ihrer *Zone of Proximal Development* (ZPD; Vygotsky, 1979) zum weiteren Denken anregen und zur Bewältigung einer Aufgabe befähigen (Wood, Bruner & Ross, 1976), mit zunehmendem Fortschritt jedoch reduziert werden (Hardy, Kloetzer, Moeller & Sodian, 2010). Empirische Befunde zeigen die Wirksamkeit verbalen Scaffoldings für die Lernleistung der Schüler\*innen (van de Pol, Volman, Oort & Beishuizen, 2015), insbesondere im naturwissenschaftlichen Unterricht (Leuchter & Saalbach, 2014). Durch die Notwendigkeit expliziter Heranführung und Instruktion ähnelt der Bildungsspracherwerb auch in der L1 mehr dem L2-Erwerb als anderen L1-Erwerbsformen. Im Unterschied zum regulären L2-Erwerb können Lerner\*innen beim L1-Bildungsspracherwerb jedoch zuverlässig auf ihre Alltagsprachkenntnisse aufbauen. Zudem wird in der Forschung zum Bildungsspracherwerb davon ausgegangen, dass die bewusste Verzahnung sprachlichen und fachinhaltlichen Lernens eine zunehmende Ausdifferenzierung der Alltagssprache hin zu Bildungssprache positiv beeinflusst (Heintze, 2010; Reich, 2013). Für eine Diskussion der Unterschiede zwischen fortgeschrittenem L1- und L2-Bildungsspracherwerb des Deutschen, siehe auch Fandrych und Wallner (i. d. H.).

Wie der Erwerb bildungssprachlicher Fähigkeiten im Detail verläuft, ist bislang kaum erforscht (vgl. aber Wildemann, Rank, Hartinger & Sutter, 2016). Bislang fehlen auch empirische Befunde über das Konstrukt Bildungssprache im realen naturwissenschaftlichen Unterricht. Rincke und Markic (2018) stellen dazu fest, dass die empirische und theoretische Auseinandersetzung mit der im Unterricht beobachteten Sprache auch unter Berücksichtigung von typischen Diskursverläufen bereits seit dem vergangenen Jahrhundert Thema der pädagogischen und didaktischen Literatur gewesen ist, dabei aber hauptsächlich zwei Analyserichtungen ausschlaggebend waren: Eine, die sich der Analyse und Problematisierung von Diskursverläufen im Unterricht zugewandt hat und eine, die Spannungen zwischen alltäglichem und fachlichem Sprachgebrauch im Blick hat (S. 31–32). Sie stellen heraus, dass in der Naturwissenschaftsdidaktik bislang kaum systematisch untersucht wurde, wie sich Merkmale von Bildungssprache

(sie meinen Merkmale der Mündlichkeit oder Schriftlichkeit) konzeptionell auf die Textmuster im naturwissenschaftlichen Unterricht durchschlagen und damit Anforderungen im naturwissenschaftlichen Unterricht definieren. Unsere Studie knüpft zwar an Modellierungen und Studien zu „Sprache im Fach“ (vgl. Beiträge in Becker-Mrotzek, Schramm, Thürmann & Vollmer, 2013) und Fachsprache im naturwissenschaftlichen Unterricht an, siehe u. a. Härtig, Bernholt, Prechtel und Retelsdorf (2015) zum Textverständnis und Rincke (2007, 2010) für Studien zum Umgang mit Fachsprache im Unterricht. Sie geht jedoch explorativ(er) an das Konstrukt Bildungssprache heran (kein Fokus bspw. auf Fachtermini) und untersucht ein Korpus aus dem tatsächlichen Unterricht – d. h. kein spezifisch für Forschungszwecke konzipiertes Unterrichtsgeschehen wie bei Rincke (2007) – sowohl in der Primar- als auch in der Sekundarstufe. Zwar gibt es einige Arbeiten zum schriftlichen Bildungssprachangebot im Fachunterricht, bspw. in Schulbüchern für die Fächer Biologie, Sachunterricht und Geschichte (Berendes et al., 2018; Drumm, 2019a, 2019b; Pohl, 2017). Jedoch wurde bislang weder die Passung der sprachlichen Angebote der Lehrpersonen zu den sprachlichen Kompetenzen der Schüler\*innen noch die implizit angenommene Steigerung der sprachlichen Komplexität über die Bildungsbiographie für die tatsächlich im Unterricht gesprochene Sprache hinreichend untersucht.

Eine Ausnahme stellt hier die Fallstudie von Kleinschmidt-Schinke (2018) dar, die anhand von acht Lehrkräften aus Grundschule und Gymnasium zeigt, wie komplex die an die Schüler\*innen gerichtete Sprache ist. Außerdem belegt ihre Studie erstmals nicht nur interindividuelle Unterschiede zwischen den Lehrpersonen, sondern auch intraindividuelle Veränderungen ein und derselben Lehrperson im Unterricht in verschiedenen Jahrgangsstufen des Gymnasiums. Es gelang ihr, konzeptionelle Schriftlichkeit in Transkripten aus Plenumsunterrichtsinteraktionen schüler\*innen- wie lehrkraftseitig zu operationalisieren sowie die Adaptivität der lehrkraftseitigen Äußerungen, die auf Schüler\*innen-Äußerungen folgten, zu analysieren. Sie zeigt, dass mehr als 56 % der Expansionen, Re- und Umformulierungen schülerseitiger Äußerungen einen „höheren Grad konzeptioneller Schriftlichkeit“ (Kleinschmidt-Schinke, 2018, S. 558) aufwiesen. Demnach können die Lehrpersonen sprachförderlich auf Schüler\*innen-Äußerungen reagieren, steigern jedoch nicht in jeder Formulierung die Komplexität. Ein Unterschied zwischen Grundschule und Gymnasium ist z. B., dass im Grundschuldiskurs ein höheres Maß geschlossener oder halboffener Aufgabenstellungen eingesetzt wird, während in Mittel- und Oberstufe vermehrt explikativ-erläuternde Aufgabenstellungen und Aufgabenstellungen, die diskursiv-argumentierende Antworten erfordern, zu verzeichnen sind (Kleinschmidt-Schinke, 2018). Dies führt dazu, dass die Studie nicht nur Veränderungen der im Unterricht genutzten Strukturen „im Sinne einer verstärkten Nutzung konzep-

tionell schriftlicher Formen [nachzeichnet] [...], sondern auch eine Veränderung der Strukturierung des gesamten Unterrichtsdiskurses hin zur Oberstufe“ (Kleinschmidt-Schinke, 2018, S. 559). Lehrpersonen – das zeigt Kleinschmidt-Schinke ebenfalls – können als konzeptionell schriftliches Modell für die Schüler\*innen wirken und so die konzeptionelle Schriftlichkeit fördern und im Sinne einer Entwicklung veränderungssensitiv einsetzen (Kleinschmidt-Schinke, 2018, S. 560).

## 2.2 Bildungssprache und sprachliche Komplexität

Zur Erschließung des linguistisch wie gesellschaftlich zentralen Konzepts „Bildungssprache“ wurden diverse Ansätze vorgeschlagen. Im deutschsprachigen Raum erwies sich die Erschließung von Bildungssprache durch die Unterscheidung von Nähe- und Distanzsprache (Ágel & Hennig, 2006, 2007; Koch & Oesterreicher, 1985) als äußerst produktiv. Hierbei wird konzeptionelle Schriftlichkeit (Generalisierung, Objektivierung, Abstraktion) als Wesensmerkmal von Bildungssprache fokussiert, welche sich unabhängig von ihrer Medialität (mündlich oder schriftlich) an Konventionen der Schriftsprache orientiert (Koch & Oesterreicher, 1985), mit dem Ziel einer sach- bzw. lernprozessbezogenen Verständigung (Feilke, 2012; Gogolin & Lange, 2010; Riebling, 2013). So verstehen auch Gogolin und Kolleg\*innen Bildungssprache als funktionales Sprachregister (Gogolin, 2007; Gogolin & Lange, 2011; Gogolin & Roth, 2007). Nach Brandt und Gogolin (2016) ist der Zweck dieses Registers, „hoch verdichtete, kognitiv anspruchsvolle Informationen in kontextarmen, formellen Konstellationen“ (Brandt & Gogolin, 2016, S. 8) in wechselseitiger Abhängigkeit von Sender\*in und Empfänger\*in konzeptionell schriftlich auszutauschen. Cummins (2000) hingegen unterscheidet bei Sprachkompetenz zwischen *Basic Interpersonal Communicative Skills* (BICS) und *Cognitiv Academic Language Proficiency* (CALP). Der Erwerb der BICS beginnt bereits vor dem Schuleintritt, während CALP bildungssprachliche Kompetenzen darstellen, die ein hohes Maß an kognitiven und sprachlichen Fähigkeiten erfordern und ab dem Schuleintritt erworben werden müssen, um Bildungserfolge zu erzielen. Dieser Ansatz hat seinen Ursprung in der L2-Forschung, kann jedoch auch für den Bildungsspracherwerb in der L1 fruchtbar gemacht werden, da Bildungsspracherwerb zum Teil explizite Instruktion erfordert (siehe Abschnitt 2.1).

Verschiedene sprachliche Mittel wurden in Anlehnung an diese Perspektiven zur Charakterisierung von Bildungssprache vorgeschlagen. In ihrer Synopse zur Kategorisierung von Bildungssprache zeigen Geist, Lange-Schubert und Dietze (2017), dass über unterschiedliche Autor\*innen hinweg lexikalische, morphosyntaktische und diskursive Merkmale als charakteristisch für Bildungssprache herausgestellt werden. Allerdings mangelt es insbesondere im deutschsprachi-



gen Raum an empirischen Studien dazu, welche linguistischen bildungssprachlichen Mittel für verschiedene Lehr-Lern-Kontexte, bspw. Schulfächer, charakteristisch sind (Brandt & Gogolin, 2016, S. 19). In diesem Beitrag verzichten wir daher darauf, uns auf einen bestimmten Ansatz zur Charakterisierung von Bildungssprache festzulegen. Stattdessen stützen wir unsere explorative, empirische Untersuchung auf eine breite Erfassung von sprachlichen Mitteln, die in der Forschungsliteratur mit Bildungssprache und Sprachkompetenz in Zusammenhang gebracht wurden. Hierbei erweist sich das Konzept der *sprachlichen Komplexität* als besonders produktiv. Dieses wurde in der Vergangenheit nicht nur zur Messung konzeptioneller Schriftlichkeit herangezogen (Kleinschmidt-Schinke, 2018; Pohl, 2017), sondern ist auch mit einer erwerbsbezogenen Perspektive kompatibel, wie im Folgenden ausgeführt wird.

Im Rahmen der L2-Erwerbsforschung (engl. *Second Language Acquisition research*, SLA) wird sprachliche Performanz oft entlang der CAF-Triade durch linguistische Komplexität (engl. *complexity*), normative Richtigkeit (engl. *accuracy*) und Produktionsflüssigkeit (engl. *fluency*) gemessen (Housen, Kuiken & Vedder, 2012). In diesem Zusammenhang wird sprachliche Komplexität gemäß Ellis (2003, S. 340) typischerweise definiert als die Elaboriertheit und Variabilität von Sprache. Rescher (1998) betrachtet ebenfalls die Verknüpfung von Elementen in einem System als komplexitätsstiftend. Diese Maße werden in SLA als deskriptive Größe verstanden. Höhere Komplexität ist also nicht automatisch mit einer höheren Performanz gleichzusetzen (vgl. auch Bulté & Housen, 2014; Pallotti, 2009), bspw. kann höhere Komplexität in informellen Kontexten indikativ für geringere Sprachkompetenz sein, da sie auf eine mangelnde Registerkompetenz hinweisen kann.

Sprachliche Komplexität kann entlang verschiedener Ebenen des theoretisch-linguistischen Systems gemessen werden, bspw. Lexikon, Morphologie und Syntax. Solche Maße mit Bezug auf das linguistische Sprachsystem erfassen, was Housen et al. (2012) als „absolute Komplexität“ bezeichnen. Im Kontrast hierzu steht die sogenannte „relative Komplexität“, welche die Komplexität der psycho-linguistischen Verarbeitung und des Erwerbs abbildet. Im Bezug auf Bildungssprache ist dabei festzustellen, dass die höhere absolute Komplexität von Bildungssprache bei an Bildungssprache gewöhnten Personen nicht zwingend zu einer Steigerung der relativen Komplexität führt, da automatisierte Prozesse keine eigenen Ressourcen verlangen (DeKeyser, 2001).

Die Komplexitätsforschung im Zweitspracherwerb konzentriert sich vornehmlich auf die Charakterisierung von L2-Kompetenz und L2-Entwicklung (siehe Housen, De Clercq, Kuiken und Vedder (2019) für einen aktuellen Überblick über den Forschungsstand). Auffällig ist hierbei die von Housen et al. (2019) als reduktionistisch kritisierte Fokussierung der Forschung auf lexikalische und

syntaktische Komplexität und auf Englisch als untersuchte Sprache. Neuere Beiträge der L2-Forschung tragen zu einer Erweiterung des Forschungsgegenstandes bei. Hierbei sind insbesondere die zunehmende Untersuchung morphologischer Komplexität zu nennen, welche vor allem in morphologisch reicheren Sprachen eine wichtige Rolle spielt (Brezina & Pallotti, 2019; De Clercq & Housen, 2019) sowie Ansätze zur Messung phraseologischer Komplexität (Garner, Crossley & Kyle, 2019; Paquot, 2019).

Maße sprachlicher Komplexität werden auch außerhalb der Zweitspracherwerbsforschung verwendet, bspw. in den Bereichen *Writing Quality Assessment* und zur Erfassung von Textlesbarkeit. In diesen Bereichen spielt neben grammatischen Komplexitätsmaßen die Erfassung von Diskurskohäsion traditionell eine wichtige Rolle (Bachmann & Becker-Mrotzek, 2017; Becker-Mrotzek, Grabowski, Jost, Knopp & Linnemann, 2014; Crossley, 2020). Die Methodenübertragung in den Bereich der Bildungssprachenanalyse basiert in Teilen auf der klaren Verbindung zwischen Merkmalen sprachlicher Komplexität und Merkmalen von Bildungssprache. Ein prominentes Beispiel hierfür ist die Komplexität der Nominalphrasen. Diese ist sowohl ein Merkmal von Bildungssprache (Fang, Schleppegrell & Cox, 2006; Hennig & Niemann, 2013) als auch von phrasaler Elaboriertheit, welche zur Erfassung von L2-Kompetenz gemessen wird (Crossley, Kyle, Allen, Guo & McNamara, 2014).

Unabhängig vom Anwendungskontext spielen Aufgabeneffekte eine zentrale Rolle bei der Erfassung sprachlicher Komplexität. Forschung im Bereich des *Task-based Language Learning* (TBLL) hat wiederholt gezeigt, dass Aufgabeneffekte (Alexopoulou, Michel, Murakami & Meurers, 2017; Vajjala, 2018) ebenso wie Register (Biber, Gray & Staples, 2016), Thema (Yoon, 2017) und Genre (Yoon & Polio, 2016) sprachliche Komplexität maßgeblich beeinflussen. Dies wurde bisher nicht nur für den L2-Erwerb gezeigt, sondern auch für die Evaluation von Bildungssprache (Weiss, Riemenschneider, Schröter & Meurers, 2019) und gesprochener Lerner\*innen-Sprache (Karges, Studer & Müller, i. d. H.; Pallotti, 2019; Pallotti & Ferrari, 2008). Obwohl die Faktoren Register, Genre und Thema in TBLL gemeinhin auf Schriftsprache angewandt werden, lassen sich die Kategorien auch auf gesprochene Lehr-Lern-Interaktionen im (Fach-)Unterricht anwenden. Beispielsweise wird das Thema maßgeblich durch den Unterrichtsinhalt vorgegeben. Überlegungen zu Aufgabeneffekten in Lehr-Lern-Interaktionen knüpfen direkt an fachdidaktische Fragestellungen an. Die sprachlichen Anforderungen der Beschreibung konkreter Beobachtungen in einer Experimentssituation weichen zum Beispiel von denen evidenzbasierten Argumentierens ab und es ist unklar, wie in diesen verschiedenen Unterrichtssituationen der parallele Erwerb von Bildungssprache und Fachkompetenzen gefördert werden kann. Während wir uns in Abschnitt 3 auf Aspekte von Aufgabeneffekten konzentrie-

ren, die für unsere empirische Untersuchung zentral sind, diskutieren wir diese interdisziplinäre Perspektive in Abschnitt 8 ausführlicher.

### 3 Daten

Der vorliegende Beitrag basiert auf Daten aus der an die DFG-Forschergruppe „Naturwissenschaftlicher Unterricht“ angebundenen Querschnittsstudie PLUS (Professionswissen von Lehrkräften und naturwissenschaftlicher Unterricht im Übergang von der Primar- zur Sekundarstufe), welche den physikbezogenen Sachunterricht am Ende der Grundschulzeit und den physikbezogenen Anfangsunterricht zu Beginn der Sekundarstufe vergleichend untersucht (Möller, Kleickmann & Lange, 2013). Die Studie wurde von 2007 bis 2010 in Nordrhein-Westfalen (NRW) im Raum Essen und Münster durchgeführt (Stadt- und Landschulen). Im Rahmen der Studie wurden 114 Lehrpersonen gebeten, das Thema „Aggregatzustände und ihre Übergänge am Beispiel Wasser“ (kurz: Aggregatzustände) in ihren 4. bzw. 6. Klassen im Umfang von sechs Unterrichtsstunden zu unterrichten. Der Unterricht wurde von den Lehrkräften selbständig geplant und im regulären Schulalltag in ihren Schulen durchgeführt.

Die erste dieser Doppelstunden wurde videographiert. Hierbei wurden zwei digitale Kameras eingesetzt, eine zur Aufnahme der Lehrkraft und ihrer Interaktionen mit den Schüler\*innen und eine für den Klassenraumüberblick. Eine dritte, bewegliche Kamera wurde für Unterrichtsaktivitäten außerhalb des Klassenraums eingesetzt. Die Aufnahmen wurden mit Hilfe der Software „Videograph“ (Rimmele, 2003) synchronisiert und transkribiert. Die Transkripte wurden auf Basis des „PLUS Transkriptionsmanuals“ (Ewerhardy, 2010, Anhang D) von geschulten Transkriptor\*innen angefertigt. Im Folgenden führen wir im Detail aus, welche sprachlichen Phänomene in der Transkription berücksichtigt wurden und welche Limitationen für die linguistische Analyse die Nutzung eines Korpus mit sich bringt, das in gesprächsanalytischen Studien ursprünglich ohne linguistische Fragestellung erstellt wurde. Hierbei orientieren wir uns an den von Hirschmann und Schmidt (i. d. H.) diskutierten Charakteristiken für gesprochene Lerner\*innen-Korpora.

Bei der Transkription wurden die Sprachsignale der Lehrkräfte und Schüler\*innen mittels normaler Alphabetschrift verschriftet und dabei gemäß der deutschen Standardorthographie segmentiert und normalisiert, inklusive Interpunktion und Groß- und Kleinschreibung. Dabei wurden individuelle Lautungen ebenso normalisiert wie klitisierte und synkopierte Wortformen. Eine nicht-normalisierte Transkription liegt nicht vor, was eine Einschränkung für die Beant-

wortung linguistischer Fragestellungen mit diesen Daten darstellt. Folgende für gesprochene (Lerner\*innen-)Korpora relevante Phänomene wurden in der Transkription in einem für die PLUS-Studie erstellten Annotationsformat vermerkt: Interjektionen, Wort- und Satzabbrüche, Stottern, Wiederholungen, Filler und nicht-verbale Phänomene wie Räuspern und Lachen, unverständliche Äußerungen, Dehnungen und Betonungen, Unterbrechungen durch andere Gesprächsteilnehmer\*innen, nicht-verbale Informationen zum Gesprächskontext (bspw. Gesten oder direkte Ansprache einzelner Personen) und Sprechpausen von mehr als vier Sekunden. Die Namen der Gesprächsteilnehmer\*innen wurden nur teilweise maskiert. Eine zeitliche Alignierung der Transkripte mit den Primärdaten liegt vor, erfolgte jedoch nur grob in Zeiträumen von 15 Sekunden, sodass eine Identifikation zeitlich paralleler Äußerungen nur begrenzt möglich ist. Diese Transkription und Annotation unterstützt zwar keine phonetischen Studien (siehe z. B. Belz & Odebrecht, i. d. H.; Trouvain, i. d. H.), sie erlaubt jedoch die Beantwortung einer Vielzahl anderer linguistischer Fragestellungen zu authentischen Lehr-Lern-Interaktionen im Schulunterricht. Die Meta-Informationen beinhalten Alter, Geschlecht und berufliche Angaben zu den Lehrkräften ebenso wie Alter, Geschlecht und Sprachbiographie der Schüler\*innen. Der Online-Appendix A illustriert die Transkription und Annotation anhand eines Beispiel-ausschnitts.<sup>1</sup> Das Korpus ist auf Anfrage zu wissenschaftlichen Zwecken frei verfügbar, die im Zusammenhang mit den Fragestellungen der PLUS-Studie stehen.

**Tab. 1:** Zahl der Turns, Sätze und Worte nach Schulform und Sprecher\*innen-Rolle (L = Lehrkräfte; S = Schüler\*innen)

Schulform	#-Videos	Rolle	#-Sprecher*innen	#-Turns	#-Sätze	#-Worte
Grundschule	58	L	65	17.907	51.513	308.878
		S	1.070	10.762	13.880	78.787
Hauptschule	22	L	23	6.716	4.397	107.660
		S	323	3.496	19.465	22.546
Gymnasium	19	L	20	5.721	15.894	108.244
		S	342	2.309	2.978	19.634
Σ	99	L	108	30.344	86.872	524.782
		S	1.735	16.567	21.255	120.967

Für den vorliegenden Beitrag wird auf die Daten von 58 Grundschulklassen sowie 41 Sekundarschulklassen aus Hauptschule und Gymnasium zurückgegriffen. Diese Stichprobe ergibt sich daraus, dass nur für diese Klassen komplette Daten-

<sup>1</sup> <https://doi.org/10.1515/zgl-2022-2052>

reihen inklusive Transkriptionen der Unterrichtsstunden vorliegen. Tabelle 1 gibt einen Überblick über die Verteilung von Turns, Sätzen und Worten zwischen Lehrkräften und Schüler\*innen pro Schulform in dem Korpus, wobei wir Turns als ununterbrochene Äußerungseinheiten definieren. Die untersuchten Schulklassen setzen sich aus Schüler\*innen mit Deutsch als Erst- (64,6 %) und als Zweitsprache (7,4 %) zusammen (28,0 % keine Angabe). Die Meta-Informationen zu Alter, Geschlecht sowie Familien- und Muttersprachen der Schüler\*innen ebenso wie zu Alter, Geschlecht und Lehrerfahrung der Lehrkräfte können Tabellen B.1, B.2 und B.3 im Online-Appendix entnommen werden.<sup>2</sup> Im Hinblick auf Aufgabeneffekte sind die Daten der Stichprobe so vergleichbar, wie die authentische Unterrichtssituation zulässt. Das Thema wird über alle Klassenstufen und Schulformen hinweg konstant gehalten, zudem wird das Register zumindest teilweise durch den schulischen und fachlichen Kontext fixiert. Weitere potentiell relevante Aufgabenfaktoren, die in diesem Beitrag nicht berücksichtigt werden konnten, werden in Abschnitt 8 diskutiert.

## 4 Automatische Messung sprachlicher Komplexität

Wir erfassen 290 automatisch berechnete Merkmale sprachlicher Komplexität mit Hilfe des Systems von Weiss und Meurers (2018).<sup>3,4</sup> Das System erfasst absolute sprachliche Komplexität in den Domänen Lexikon, Morphologie und Syntax, Maße von Diskurskohäsion sowie Merkmale relativer Komplexität aus der menschlichen Sprachverarbeitung und Sprachnutzung.<sup>5</sup> Es stellt damit die umfassendste Sammlung automatisch extrahierbarer Komplexitätsmerkmale des Deutschen bereit und erlaubt somit eine linguistisch breit aufgestellte Analyse der Unterrichtssprache in unseren Daten. Die parallele Betrachtung so vieler sprachlicher Maße hat den Nachteil, dass eine Vielzahl von Faktoren

---

<sup>2</sup> <https://doi.org/10.1515/zgl-2022-2052>

<sup>3</sup> Die unseren Studien zugrunde liegenden Analyseskripte sowie die Datentabelle mit den Komplexitätsanalysen sind über die Plattform der Open Science Foundation (OSF) öffentlich zugänglich: [https://osf.io/fusye/?view\\_only=964f965111bc4514aad80e4dd8967fd2](https://osf.io/fusye/?view_only=964f965111bc4514aad80e4dd8967fd2)

<sup>4</sup> Das System steht zur freien Nutzung in der *Common Text Analysis Platform* (X. Chen & Meurers, 2016) zur Verfügung ([www.ctapweb.com](http://www.ctapweb.com)), welche von Weiss, Chen & Meurers (2021) für mehrsprachige Komplexitätsanalysen erweitert wurde. Die Daten müssen zur Analyse als einfache Textdateien (.txt) ohne Annotationen vorliegen. Die Berücksichtigung bestehender Annotationen wird derzeit nicht unterstützt.

<sup>5</sup> Siehe Diskussion relativer und absoluter Komplexität in Abschnitt 2.2.

zueinander in Bezug gesetzt werden müssen und nicht immer die Interpretation aller Beobachtungen im Detail möglich ist. Die Methode erlaubt jedoch eine linguistisch breit aufgestellte explorative Untersuchung des frühen Bildungsspracherwerbs, welche einen wertvollen Gesamteindruck in die Sprachnutzung in ökologisch validen Lehr-Lern-Kontexten vermitteln kann. Dieses Verfahren erlaubt im Anschluss eine genauere Betrachtung jener Merkmale, die in den Daten relevante Kontraste sichtbar machen. Frühere Studien konnten zeigen, dass ein solch breit aufgestelltes Analyseverfahren äußerst erfolgreiche Einblicke erlaubt, beispielsweise in Unterschiede zwischen Textproduktionen in verschiedenen Klassenstufen und L2-Kompetenzniveaus (Weiss & Meurers, 2019a, 2019b, 2021), in die Lesbarkeit von Texten für junge Leser\*innen (Weiss & Meurers, 2018) und von Schulbüchern (Berendes et al., 2018; Kühberger, Bramann, Weiss & Meurers, 2019), und für L2-Lernende (Weiss, Chen & Meurers, 2021), oder in die Bewertungsobjektivität von Lehrkräften im Deutsch-Abitur (Weiss et al., 2019).

Im Folgenden motivieren wir die Erfassung der einzelnen Domänen für die vorliegende Untersuchung authentischer Interaktionen in Lehr-Lern-Kontexten anhand einer Auswahl konkreter Merkmale. Weiss und Meurers (2019b, 2021) enthalten eine detaillierte Beschreibung des gesamten Systems sowie Definitionen aller Merkmale.<sup>6</sup> Bei der folgenden Kategorisierung ist festzuhalten, dass sich Komplexitätsmaße nicht immer eindeutig einer einzigen Domäne zuweisen lassen. Beispielsweise tragen attributive Partizipien zur phrasalen Komplexität von Nominalphrasen bei, es handelt sich bei ihnen jedoch auch um ein Deagentivierungsmittel. Konnektoren erhöhen die explizite Kohäsion von Texten, gehen jedoch auch immer mit einer erhöhten klausalen Komplexität einher. Im Folgenden schlagen wir eine mögliche Kategorisierung von Merkmalen vor, die mit unserer späteren Ergebnisinterpretation kongruent ist.

**Klausale Komplexität** Die Komplexität klausaler Strukturen wird in der SLA-Komplexitätsforschung traditionell über die Länge, Anzahl und Art der Einbettung von (Teil-)Sätzen erfasst (bspw. die maximale Satzlänge in Worten oder die Anzahl eingeleiteter Nebensätze pro Satz), siehe Wolfe-Quintero, Inagaki und Kim (1998). Auch die Häufigkeit von Phrasen pro Satz (bspw. Nominalphrasen pro Satz) wurde zur Erfassung der Elaboriertheit syntaktischer Strukturen vorgeschlagen. Wir zählen diese Maße zum Bereich klausaler Komplexität, da sie keine Auskunft über die Elaboriertheit oder Variabilität der gemessenen Phrasen

---

<sup>6</sup> Appendix A und B aus Weiss und Meurers (2021) mit den Definitionen aller Merkmale finden sich auch unter [https://osf.io/pu9hy/?view\\_only=588b4fb0c26e48e58d7f06219a566713](https://osf.io/pu9hy/?view_only=588b4fb0c26e48e58d7f06219a566713)

geben. Da die Verwendung para- und hypotaktischer Satzstrukturen auch häufig zur Charakterisierung von Bildungssprache (Feilke, 2012; Schleppegrell, 2004) und gesprochener Sprache (Biber & Gray, 2010) herangezogen wird, sind diese Maße nicht nur für den Zweitspracherwerbskontext sondern auch für die vorliegende Studie direkt relevant.<sup>7</sup>

**Phrasale Komplexität** Innerhalb der phrasalen Domäne wird Komplexität primär in Form der Länge und Art der Modifikation einzelner Phrasen gemessen (bspw. Worte pro Präpositionalphrase oder Anzahl pränominaler Modifikatoren), siehe Bulté und Housen (2014); Kyle (2016). Seltener wird auch die Variabilität der Nutzung verschiedener Modifikationstypen erfasst (bspw. Variation nominaler Modifikationstypen). Diese Maße stehen in direktem Zusammenhang mit sprachlichen Konstruktionen, wie wir in Abschnitt 2.2 am Beispiel der komplexen Nominalphrase erörtert haben.

**Komplexe grammatische Strukturen** Wir erfassen zudem die Häufigkeit und Variabilität von Strukturen, welche sich nicht klar der klausalen oder phrasalen Domäne zuordnen lassen. In diesen Bereich gehören insbesondere Deagentivierungsmaße, welche direkt aus der Bildungssprachforschung stammen (siehe bspw. Bailey, 2007; Hennig & Niemann, 2013; Snow & Uccelli, 2009), deren Realisierung jedoch ebenfalls Einfluss auf die Elaboriertheit und Variabilität der syntaktischen Domäne haben (bspw. wie häufig der Agens durch die Bildung von Partizipialattributen oder die Nutzung von (Quasi-)Passiven ausgelassen wird oder wie viele verschiedene Deagentivierungsstrategien in einer Sprachproduktion vorkommen).

In diesen Bereich fallen nach unserer Kategorisierung auch Maße, welche die Besetzung verschiedener Felder im topologischen Modell erfassen. Diese spielen insbesondere für den L2-Erwerb eine wichtige Rolle. So stellen Vorfelder, welche nicht durch ein Subjekt besetzt werden, eine besondere Herausforderung für Deutschlernende da (siehe bspw. Ballestracci, 2010). Dies gilt auch für elaborierte Mittelfelder, insbesondere wenn die rechte Verbklammer besetzt ist. Inwiefern diese Maße auch für den sprachsensiblen Fachunterricht eine Rolle spielen, bleibt empirisch zu bestimmen.

---

<sup>7</sup> Aufgrund der zentralen Bedeutung des Satzbegriffes für die Erfassung dieser Domäne, ist die Berechnung dieser Maße für gesprochener Sprache alles andere als trivial. In der vorliegenden Arbeit nutzen wir für die Identifikation von Sätzen die graphematischen Markierungen, welche in den Transkripten nach dem „PLUS Transkriptionsmanual“ gesetzt wurden, siehe Ende dieses Abschnitts für eine Diskussion der Validität dieses Vorgehens.

**Lexikalische Komplexität** Die Komplexität in der lexikalischen Domäne wird traditionell ebenfalls über die Länge von Worten gemessen (bspw. Anzahl der Buchstaben pro Wort oder maximale Wortlänge in Silben). Zur Erfassung der Variabilität der lexikalischen Domäne haben sich vor allem Maße lexikalischer Diversität etabliert, welche die Zahl der Wort-Token in Verhältnis zu zur Anzahl der vorkommenden Wort-Typen setzen (bspw. die Type-Token Ratio), siehe bspw. McCarthy und Jarvis (2010). Auch wortartspezifische Varianten dieses Maßes wurden vorgeschlagen, um die lexikalische Variation insbesondere im Bereich der Autosemantika zu erfassen. Für die vorliegende Studie spielt hier insbesondere die Variation im Bereich der Verben eine Rolle (bspw. Voll- und Modalverb-Typen pro Voll- und Modalverb-Token),<sup>8</sup> da sich Bildungs- und Fachsprache unter anderem durch die fachspezifische Nutzung von Verben abgrenzen (vgl. bspw. Uessler, Runge & Redder, 2013).

**Morphologische Komplexität** Die Erfassung morphologischer Komplexität wurde im Rahmen der SLA-Komplexitätsforschung lange vernachlässigt (Housen et al., 2019; Pallotti, 2015). Häufig konzentrieren sich Maße morphologischer Komplexität auch auf den Bereich der Verbflexion (Bulté & Housen, 2012), bspw. auf die Flexion nach Personenmerkmalen oder Tempora. Hierbei können Flexionsänderungen einerseits als formal komplexer betrachtet werden, aufgrund der Zunahme diskreter morphologischer Komponenten durch ihre Realisierung. Andererseits können Maße wie der Anteil der Flexion nach der zweiten Person im L2-Erwerb oder das Vorkommen eines bestimmten Kasus Rückschlüsse auf die strukturelle Komplexität des sich entwickelnden Sprachsystem der Lernenden erlauben. Obwohl diese Überlegungen für den Bildungsspracherwerb weniger zentral sind, ist die Erfassung der Verbflexion für die vorliegende Untersuchung aus anderen Gründen interessant, da sie Rückschlüsse auf die Adressierung der Schüler\*innen im Unterricht erlaubt: So lässt sich durch ihre Erfassung untersuchen, ob Lehrkräfte Schüler\*innen tendenziell direkt ansprechen (also die zweite Person nutzen), ob sie sich mit einbeziehen (Nutzung der ersten Person) oder ob sie sowohl sich als auch die Schüler\*innen ausschließen, bspw. im Rahmen von Deagentivierungsstrategien (Nutzung der dritten Person).

Andere Merkmale, die primär für den L2-Erwerb relevant sind (bspw. die Nutzung von Infinitiven oder Dativen), können als zusätzliche Evidenz für die

---

<sup>8</sup> Es handelt sich hierbei um ein Maß für die lexikalische Varianz bei autosemantischen Verben, bei dem Modalverben zu den Autosemantika gezählt werden. Da Modalverben in manchen Varianten dieses Maßes ebenfalls ausgeschlossen werden, nutzen wir hier und im Folgenden die komplette Nennung aller gemessener Verbarten.



Interpretation von Beobachtungen aus anderen Domänen dienen. Daher schließen wir diese Maße nicht *a priori* von unserer Analyse aus. Neben Flexionsmaßen enthält diese Domäne auch Maße für die Elaboriertheit von Komposita (bspw. durchschnittliche Kompositalänge) und die Erfassung von Derivationsprozessen (bspw. derivierte Nomen pro Nomen), welche die formale Komplexität im morphologischen Bereich durch die Erhöhung morphologischer Komponenten beeinflussen.

**Diskurskohäsion** Anders als die zuvor besprochenen Domänen wird Diskurskohäsion häufiger im Bereich des *Writing Quality Assessment* erfasst, welcher eng mit Bildungssprache verbunden ist, da sich die Untersuchungen von Textqualität zumeist auf akademische Essays beziehen (Crossley, 2020). Dabei spielt insbesondere die Nutzung expliziter Kohäsionsmarker eine Rolle, bspw. kausale oder temporale Konnektoren, wobei wir zwischen Mehrwortkonnektoren (z. B. *sowohl ... als auch ...*) und Einwortkonnektoren (z. B. *aber*) unterscheiden. Zur Identifikation spezifischer Konnektorenarten nutzen wir Listen von Konnektorentypen (bspw. temporale Konnektoren oder kausale Konnektoren), welche entweder auf Breindl, Volodina und Waßner (2015) – im Folgenden Breindl-Konnektorenliste genannt – oder Eisenberg et al. (2009) – im Folgenden Eisenberg-Konnektorenliste genannt – basieren.

Kohäsion kann jedoch auch implizit durch die Wiederaufnahme sprachlichen Materials erzeugt werden, wie die Wiederholung von Phrasen oder durch koreferenzielle Ausdrücke. Sowohl implizite als auch explizite Kohäsionsmarker erhöhen die Verknüpfung von diskursiven Einheiten und somit die diskursive Komplexität. In unserer Untersuchung spielen Maße für die Nutzung koreferenzieller Ausdrücke wie Pronomen mit verschiedenen Personenmerkmalen (bspw. Personal- oder Possessivpronomen der zweiten Person) zusätzlich eine wichtige Rolle, da sie zusätzliche Evidenz für den Adressatenbezug im Unterrichtsgeschehen bereitstellen. Aus einer interdisziplinären Perspektive ist ebenso die Nutzung von Eigennamen als Indikator für *Classroom Management* interessant.

**Sprachnutzung** Wortfrequenzen wurden seit den Anfängen der Untersuchung sprachlicher Komplexität in SLA als Komponente lexikalischer Komplexität erfasst (Wolfe-Quintero et al., 1998), da sie einen starken Einfluss auf die menschliche Sprachverarbeitung haben (vgl. Brysbaert et al., 2011). Obwohl die SLA-Komplexitätsforschung im Bestreben um größere terminologische Klarheit Abstand von der Erfassung solcher entwicklungsbezogener Maße genommen hat (siehe bspw. Housen et al., 2012; Pallotti, 2015), ist die Relevanz von Wortfrequenzen sowohl für den Spracherwerb und insbesondere für den Bildungsspracherwerb nicht von der Hand zu weisen. Neben der Erfassung durchschnittlicher Frequen-

zen hat sich auch die Messung sogenannter logarithmischer Frequenzbänder als nützlich erwiesen, für die alle Worte innerhalb eines bestimmten Frequenzintervalls dem gleichen Frequenzband zugeordnet werden (bspw. alle Worte mit einer logarithmierten Frequenz zwischen 2,0 und 2,9). Diese sind unabhängiger von der Größe der Frequenzdatenbank (vgl. Van Heuven, Mandera, Keuleers & Brysbaert, 2014) und können eine differenziertere Ergebnisinterpretation erlauben als einfache Frequenzdurchschnitte. So sind zwei Texte mit sehr ähnlicher durchschnittlicher Wortfrequenz denkbar, bei denen einer nahezu ausschließlich mittel-frequentes Vokabular enthält während der andere sowohl hoch- als auch niedrig-frequentes Vokabular nutzt. Der Unterschied zwischen diesen Texten kann durch Frequenzbänder erfasst werden.

Eine zentrale Rolle für solche Maße von Sprachnutzung ist die Referenzquelle, aus der die Wortfrequenzen bezogen werden. Für unsere Untersuchung nutzen wir vier verschiedene Datenquellen: Wir nutzen die Frequenzdatenbank Subtlex-DE (Brysbaert et al., 2011), welche die Frequenzen von Worten in Filmen und Fernsehserien enthält und somit möglichst repräsentativ für den Sprachgebrauch außerhalb der Schule ist. Zur Erfassung von geschriebener Bildungssprache erfassen wir Wortfrequenzen aus Google Books 2000 und der dlexDB Datenbank (Heister et al., 2011), welche den Sprachgebrauch in Zeitungsartikeln abbildet. Zuletzt erfassen wir den Sprachgebrauch von deutschen Schulkindern in der Schriftsprache zwischen der ersten bis achten Klasse über Wortfrequenzen, die wir aus dem Karlsruhe Children's Texts Korpus (KCT; Lavalley, Berkling & Stüker, 2015) extrahiert haben. Diese Datenquelle erlaubt uns auch die Verknüpfung von Vokabelvorkommen mit dem Alter von Schüler\*innen, indem wir das durchschnittliche und minimale Alter der aktiven Nutzung für Worte im KCT Korpus erfassen.

**Sprachverarbeitung** Maße menschlicher Sprachverarbeitung betrachten sprachliche Komplexität ebenfalls aus einer psycho-linguistischen Perspektive, in der höhere kognitive Verarbeitungskosten mit höherer Komplexität assoziiert werden. Wir nutzen für die Erfassung von Sprachverarbeitungskosten die Dependency Locality Theory von Gibson (2000, DLT). Es handelt sich hierbei um eine prominente Satzverarbeitungstheorie, welche auf zwei Kernannahmen beruht. Erstens, die menschliche Sprachverarbeitung besteht aus zwei Prozessen, die begrenzte Aufmerksamkeitsressourcen verbrauchen: die Zwischenspeicherung unvollständiger Diskursstrukturen (operationalisiert als nicht vollständig saturierte finite Verben) und die Integration von Diskursreferenten (operationalisiert als Verben und Nomen) in unvollständige Diskursstrukturen. Zweitens, je größer die Distanz zwischen Diskursreferenten und ihrer Diskursstruktur, desto höher sind die Integrationskosten. Die Integrationskosten werden basierend auf der

Anzahl der Diskursreferenten zwischen dem zu integrierenden Diskursreferenten und der Diskursstruktur gemessen.

Es handelt sich bei Integrationskosten in der DLT um ein prozedurales Maß. Wir aggregieren dieses, indem wir die durchschnittlichen und maximalen Integrationskosten pro finitem Verb berechnen (Weiss, 2017). Shain, van Schijndel, Futrell, Gibson und Schuler (2016) schlagen unterschiedliche Gewichtungen für koordinierte Diskursreferenten, finite Verben und Modifizierer vor, welche wir als verschiedenen Varianten dieses Merkmals implementiert haben. Obwohl sich Untersuchungen zur menschlichen Sprachverarbeitung in der Regel mit geschriebener Sprache beschäftigen, sollten die der Theorie zugrundeliegenden Annahmen unabhängig vom Rezeptionsmodus Anwendung finden. Inwiefern diese Maße eine Rolle in der gesprochenen Unterrichtssprache spielen, ist daher eine relevante Fragestellung.

**Textoberflächenmaße** Neben den zuvor beschriebenen Merkmalen berechnen wir ebenfalls die Gesamtzahl der Sätze und Worte als globales Maß für den Äußerungsumfang der Gesprächsteilnehmer\*innen.

Zwar wurde das deutsche System bisher nicht für gesprochene Sprache eingesetzt, allerdings ließ sich unsere Analysemethode erfolgreich auf englische und deutsche Fernsehuntertitel anwenden (Vajjala & Meurers, 2014; Weiss & Meurers, 2018) und ermöglichte auch die Analyse kurzer Antworten auf Leseverständnisfragen von beginnenden Deutschlerner\*innen (Weiss & Meurers, 2021), sodass die Extraktion von Komplexitätsmerkmalen augenscheinlich auch auf sehr kurzen Äußerungen funktioniert und trotz der auf Schriftsprache trainierten Modelle für automatische Sprachverarbeitung (engl. *Natural Language Processing*) auch für nicht-standardisierte Sprache relativ robuste verwertbare Ergebnisse liefert. Zudem kann die Definition schriftsprachlicher linguistischer Konstrukte, wie das des (graphematischen) Satzes (Schmidt, 2016), aus den PLUS-Transkription übernommen werden. Die Modelle können wie in standardisierter Schriftsprache auf diese in der Transkription eingeführten graphematischen Markierungen zurückgreifen. Auch wenn in der Anwendung schriftsprachlicher Kategorien auf gesprochene Sprache eine gewisse *comparative fallacy* (Bley-Vroman, 1983) durch den Skriptizismus (Linell, 2005, engl. *written language bias*) auftritt, so begünstigt die konzeptionelle Schriftlichkeit von Bildungssprache doch die valide Anwendbarkeit schriftsprachlicher Kategorien.

## 4.1 Allgemeines Vorgehen zur Merkmalsextraktion und -auswahl

Grundlage aller im Folgenden berichteten Studien ist die Komplexitätsanalyse der Unterrichtsäußerungen von Lehrkräften und Schüler\*innen in den in Abschnitt 3 beschriebenen Daten (46.911 Turns). Wir extrahieren zunächst 290 Merkmale sprachlicher Komplexität mit dem in Abschnitt 4 eingeführten System. Nach Ausschluss der Merkmale, die keine oder nur geringe Varianz über die Lehrkraft-Turns ( $N=30.344$ ) hinweg zeigen, bleiben 169 Merkmale übrig. Wir definieren hier geringe Varianz als das Vorkommen des gleichen Merkmalswertes in mindestens 80 % der Datenpunkte.

Schließlich standardisieren wir alle Werte aller Merkmale in den Äußerungen der Lehrkräfte in Form von z-Werten, welche den Merkmalswert in Form von Standardabweichungen vom Durchschnitt angeben und identifizieren alle Merkmale, die mit anderen Merkmalen der gleichen Merkmalsdomäne eine absolute Pearson Korrelation  $r \geq 0,8$  aufweisen. Aus den so identifizierten Merkmalen wählen wir für die Analyse immer dasjenige Merkmal aus, das die spezifischere linguistische Interpretation bietet. Beispielsweise beobachten wir eine Korrelation von  $r = 0,99$  zwischen den Diskursmerkmalen *Konnektoren pro Satz* und *Einwortkonnektoren pro Satz*. Wir behalten Letzteres als das spezifischere Maß, da die Zahl der Einwortkonnektoren eine Untermenge der Gesamtzahl der Konnektoren ist. Während es prinzipiell denkbar ist, dass sich die Gesamtzahl der Konnektoren und die Zahl der Einwortkonnektoren unterschiedlich entwickelt, ist das in unseren Daten nur sehr selten der Fall und eine separate Erhebung somit nicht informativ. Mit diesem Vorgehen reduzieren wir die Fälle, in denen das gleiche linguistische Phänomen durch mehrere Maße abgebildet wird. Bei dem erwähnten Beispiel wird so vermieden, den Einfluss von Konnektoren auf Unterschiede zwischen Schulformen mehrfach abzubilden und so womöglich zu überschätzen. Dieses Vorgehen eliminiert 54 Merkmale.

Eine Zusammenfassung der Merkmalsanzahl pro Merkmalsdomäne und ihre Reduktion auf 111 variable, unkorrelierte Merkmale ist in Tabelle 2 zu finden. Insbesondere in den morpho-syntaktischen Domänen sowie im Diskursbereich beobachten wir eine relativ hohe Anzahl invariabler Merkmale in den Äußerungen der Lehrkräfte, während die Domänen Lexikon, Sprachnutzung, Sprachverarbeitung und Oberflächenmaße mehr variable Maße enthalten. Die Bereiche klausale Syntax, Sprachnutzung, Sprachverarbeitung, und Oberflächenmaße zeigen zudem eine hohe Zahl miteinander korrelierter Maße.

Tab. 2: Übersicht zur Merkmalsreduktion aggregiert über Merkmalsdomänen

Domäne	Gesamt	Variabel	Unkorreliert
Klausale Komplexität	22	14	8
Phrasale Komplexität	27	16	15
Grammatische Strukturen	20	7	7
Lexikalische Komplexität	34	33	24
Morphologische Komplexität	41	10	9
Diskurskohäsion	64	22	16
Sprachnutzung	54	44	28
Sprachverarbeitung	25	17	3
Oberflächenmaß	3	2	1
$\Sigma$	290	165	111

## 5 Studie 1: Vergleich der Lehrkräfte nach Schulformen

### 5.1 Aufbau von Studie 1

Um herauszufinden, ob die Komplexität der Sprache der Lehrkraft von der Schulform abhängt, berechnen wir für jedes der 111 Komplexitätsmerkmale ein Mehrebenenmodell, welches sich für die hierarchischen Datenstrukturen in den hier untersuchten Daten anbietet. Hierfür nutzen wir eine gemischte Regression auf den Lehrkraft-Turns ( $N=30.344$ ). Das jeweilige Komplexitätsmerkmal ist hier die abhängige Variable und die Schulform der Prädiktor. Um individuelle Lehrkraft-Unterschiede soweit wie möglich zu neutralisieren, fügen wir auch die IDs der Lehrkräfte als *random intercept* hinzu, sodass signifikante Unterschiede zwischen Schulformen unabhängig von inter-individuellen Lehrkraftunterschieden festgestellt werden können. Formel 1 zeigt das resultierende Modellmuster:

$$\text{Merkmal} \sim \beta_0 + \beta_1 * \text{Schulform} + (1|\text{Lehrkraft}) + \epsilon \quad (1)$$

Hier repräsentiert  $\beta_0$  den *intercept*,  $\beta_1$  den Merkmalskoeffizienten für die Schulform,  $(1|\text{Lehrkraft})$  den *random intercept* für die individuellen Unterschiede der Lehrkräfte, und  $\epsilon$  den *residual error*. Für die Parameterschätzung nutzen wir *Restricted Maximum Likelihood Estimation*, wie sie im R-Paket *lme4* (Bates, Mächler, Bolker & Walker, 2015) umgesetzt ist. Um eine Normalverteilung und Homoske-

dastizität (also die homogene Varianz) der Störgröße zu gewährleisten, transformieren wir alle Komplexitätsmaße mit der *Ordered Quantile Normalization Transformation* aus dem R-Paket *bestNormalize* (Peterson & Cavanaugh, 2019).

Die Durchführung vieler Signifikanztests, hier einer pro Merkmal, erhöht das Risiko falsch-positiver Resultate (siehe z. B. S.-Y. Chen, Feng & Yi, 2017). Um einer Inflation signifikant scheinender Schulformunterschiede vorzubeugen, wenden wir die Benjamini-Hochberg-Korrektur an (Benjamini & Hochberg, 1995). Dieses Verfahren dient der Korrektur von Signifikanzwerten unter Berücksichtigung der Anzahl von durchgeführten Tests. Es handelt sich hierbei um eine Korrekturmethode, die sich besonders für die Durchführung einer großen Zahl von Signifikanztests anbietet, siehe S.-Y. Chen et al. (2017) für weitere Erläuterungen und einen Vergleich dieser mit anderen Signifikanzkorrekturmethoden. Wir verwenden einen kritischen Wert von 0,10; es werden in dieser explorativen Untersuchung also bis zu 10 % falsch positive Resultate akzeptiert. Wir extrahieren p-Werte zur Signifikanzbestimmung für unsere Haupteffekte mit Hilfe des R-Packets *lmerTest* (Kuznetsova, Brockhoff & Christensen, 2017), sowie *MuMIn* (Barton, 2020) zur Bestimmung der aufgeklärten Varianz  $R^2$ .

## 5.2 Ergebnisse von Studie 1

Tabelle 3 zeigt die Verteilung der 49 Komplexitätsmerkmale, die mindestens in einer Sekundarschulform (Gymnasium oder Hauptschule) signifikant von der Grundschule abweichen, gruppiert nach Merkmalsdomänen.<sup>9</sup> Spalte  $\Sigma$  zeigt die Gesamtzahl der Merkmale der jeweiligen Domäne und Spalte Sign., wie viele davon bei mindestens einer Sekundarschulform signifikante Unterschiede zur Grundschule zeigen. Die folgenden Spalten zeigen, wie viele der signifikanten Unterschiede bei beiden Sekundarschulformen zunehmen ( $\uparrow$ HG) oder abnehmen ( $\downarrow$ HG), wie viele in der Hauptschule ab- und im Gymnasium zunehmen ( $\downarrow$ H $\uparrow$ G) und wie viele jeweils nur bei einer Schulform signifikante Unterschiede zur Grundschule zeigen ( $\uparrow$ G,  $\downarrow$ G,  $\uparrow$ H,  $\downarrow$ H). So lässt sich an den ersten Spalten der ersten Zeile ablesen, dass sich die Äußerungen von Lehrkräften in mindestens einer der beiden Sekundarschulformen signifikant von den Äußerungen von Grundschullehrkräften im Hinblick auf sechs der acht getesteten Merkmale aus

<sup>9</sup> Hier und im Folgenden muss aus Platzgründen auf explizite Modellzusammenfassungen verzichtet werden. Eine detaillierte Darstellung für alle Modelle ist im Online-Appendix zu finden, in Tabellen C.1 bis C.9 unter <https://doi.org/10.1515/zgl-2022-2052>. Dort finden sich auch Angaben zu den *random effects* der Lehrkräfte, welche Aufschluss über die Varianz zwischen individuellen Lehrkräften geben.

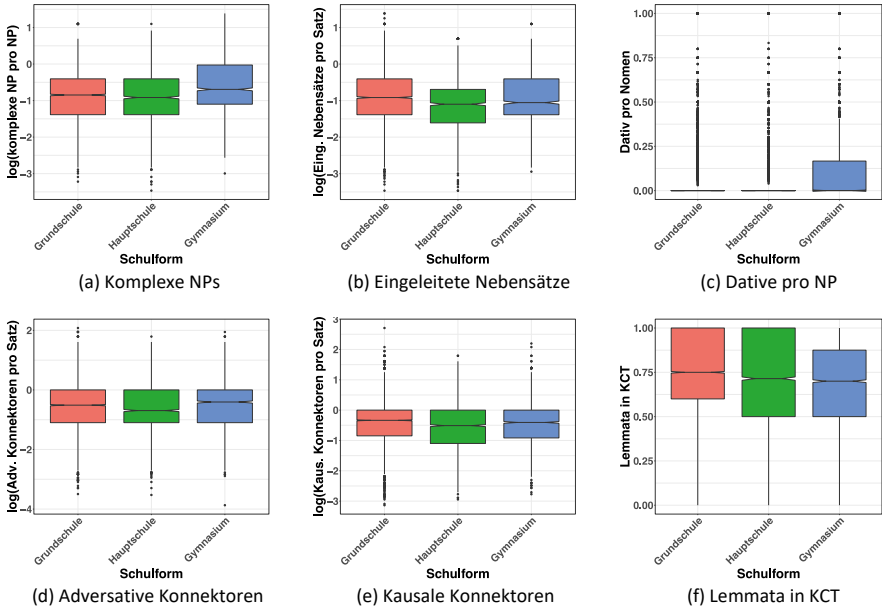
dem Bereich der klausalen Komplexität unterscheiden. Die folgenden Spalten (abgetrennt durch eine vertikale Linie) schlüsseln dieses Ergebnis im Detail auf: von den sechs signifikanten Merkmalen ist eines in den Äußerungen von Hauptschullehrkräften signifikant niedriger als in den Äußerungen von Grundschullehrkräften und zugleich in Äußerungen von Gymnasiallehrkräften signifikant höher als in den Äußerungen von Grundschullehrkräften ( $\downarrow H \uparrow G$ ).<sup>10</sup> Drei Merkmale sind im Gymnasium signifikant höher als in der Grundschule, zeigen jedoch keinen signifikanten Unterschied zwischen Grund- und Hauptschule ( $\uparrow G$ ). Zwei Merkmale klausaler Komplexität sind in der Hauptschule signifikant niedriger als in der Grundschule, während kein signifikanter Unterschied zwischen Gymnasium und Grundschule besteht ( $\downarrow H$ ). Andere denkbare Muster ( $\uparrow HG$ ,  $\downarrow HG$ ,  $\downarrow G$ ,  $\uparrow H$ ) sind bei den getesteten Merkmalen klausaler Komplexität nicht vertreten.

**Tab. 3:** Anzahl von Merkmalen mit signifikanter Zu- ( $\uparrow$ ) und Abnahme ( $\downarrow$ ) zwischen Grund- und Sekundarschulen aufgeschlüsselt nach Domänen (G: nur zum Gymnasium; H: nur zur Hauptschule; HG: zu Hauptschule und Gymnasium) sowie die Summe dieser signifikanten Unterschiede (Sign.) und die Gesamtsumme der getesteten Merkmale ( $\Sigma$ ).

Domäne	$\Sigma$	Sign.	$\uparrow HG$	$\downarrow HG$	$\downarrow H \uparrow G$	$\uparrow G$	$\downarrow G$	$\uparrow H$	$\downarrow H$
Klausale Komplexität	8	6	0	0	1	3	0	0	2
Phrasale Komplexität	15	6	0	0	0	5	0	0	1
Grammatische Strukturen	7	5	0	0	0	1	0	0	4
Lexikalische Komplexität	24	8	2	1	0	1	2	0	2
Morphologische Komplexität	9	4	0	1	0	2	0	0	1
Diskurskohäsion	16	6	0	1	0	2	1	0	2
Sprachnutzung	28	14	1	5	0	3	3	0	2
Sprachverarbeitung	3	0	0	0	0	0	0	0	0
Oberflächenmaß	1	0	0	0	0	0	0	0	0
$\Sigma$	111	49	3	8	1	17	6	0	14

Welche konkreten Merkmale in den jeweiligen Merkmalsdomänen die signifikanten Unterschiede ausmachen, besprechen wir im Folgenden im Detail. Hierfür listen wir zunächst explizit auf, welche Merkmale in der jeweiligen Domäne den in Tabelle 3 ausdifferenzierten Mustern zuzuordnen sind, bevor wir die Beobachtungen konkret diskutieren. Abbildung 1 visualisiert einige der prominenteren Unterschiede, welche im Folgenden an gegebener Stelle im Zusammenhang mit Tabelle 3 diskutiert werden.

<sup>10</sup> Im Folgenden werden wir die Formulierung „Äußerungen von Grundschul-/Hauptschul-/Gymnasiallehrkräften“ zu „in der Grund-/Hauptschule“ oder „im Gymnasium“ vereinfachen.



**Abb. 1:** Boxplots teilweise log-transformierter Merkmalswerte nach Schulform

## Klausale und phrasale Komplexität und komplexe grammatische Strukturen

Viele dieser Merkmale zeigen signifikante Unterschiede zwischen Grundschule und den Sekundarschulformen. Allerdings unterscheiden sich Gymnasium und Hauptschule stark im Bezug darauf, welche syntaktischen Charakteristiken sich verändern.

### Klausale Komplexität

- ↓**H**↑**G**    längster Satz in Worten
- ↑**G**        NPs pro Satz, komplexe NPs pro Satz, PPs pro Satz
- ↓**H**        VPs pro Satz, eingeleitete Nebensätze pro Satz



### Phrasale Komplexität

- ↑**G** pränominaler Modifizierer pro Nomen, mittlere NP-Länge in Worten, NP-Modifizierer pro NP, mittlere PP-Länge in Worten, NP-Dependenten pro Nomen mit Dependenten
- ↓**H** postnominale Modifizierer pro NP

### Grammatische Strukturen

- ↑**G** Abdeckung von Deagentivierungsmustern
- ↓**H** Vorfelder ohne Subjektbesetzung pro Vorfeld, Infinitivkonstruktionen pro VP, mittlerer Silbenabstand von 1. Argument und Vollverb, mittlere Mittelfeldlänge in Silben

Im Gymnasium nimmt die Elaboriertheit von Präpositionalphrasen ebenso zu wie die Elaboriertheit und Variation von komplexen Nominalphrasen (Abbildung 1a) und die Abdeckung von Deagentivierungsmustern. In der Hauptschule hingegen nutzen Lehrkräfte weniger eingeleitete Nebensätze (Konjunktionalsätze und Relativsätze; siehe Abbildung 1b) und insgesamt weniger Verbalphrasen pro Satz. Beides zeigt, dass die Subordination gegenüber der Grundschule abnimmt. Zudem nehmen die postnominale Modifikation und die Distanz zwischen Verben und ihren Argumenten gegenüber der Grundschule ab. Diese unterschiedliche Entwicklung spiegelt sich auch in der konträren Entwicklung der maximalen Satzlänge.

### Lexikalische Komplexität

Im lexikalischen Bereich finden sich sowohl einheitliche Entwicklungen beider Sekundarschulformen im Vergleich zur Grundschule als auch individuelle Unterschiede.

- ↑**HG** Wortlänge in Buchstaben, Wortlänge in Silben
- ↓**HG** Voll- und Modalverben pro Token
- ↑**G** Maximale Wortlänge in Silben
- ↓**G** lexikalische Typen pro Token, lexikalische Typen pro lexikalisches Token
- ↓**H** Voll- und Modalverb-Typen pro Voll- und Modalverb-Token, Modalverben pro Verb

In beiden Sekundarschulformen nimmt die durchschnittliche Wortlänge zu. Im Gymnasium nimmt zudem die maximale Wortlänge zu. In beiden Schulformen nimmt die Verbvarianz ab. Im Gymnasium ist dieser Rückgang einer Reduktion der lexikalischen Varianz im Allgemeinen zuzuschreiben. In der Hauptschule hingegen nimmt spezifisch die Verbvarianz ab, was zu den Beobachtungen in den syntaktischen Domänen passt, in welcher der Rückgang der Subordination mit einem Rückgang der Verbnutzung einhergeht. Insofern unterscheidet sich die Hauptschule von der Grundschule systematisch im Hinblick auf die Verbnutzung mit Blick auf die Frequenz und Varianz von Verben.

### **Morphologische Komplexität und Diskurskohäsion**

Wir berichten die Ergebnisse zu morphologischer Komplexität und Diskurskohäsion zusammen, da sie teils die gleichen Phänomene aus verschiedenen Perspektiven illustrieren.

#### **Morphologische Komplexität**

- ↓**HG** Verbflexion nach der zweiten Person pro finites Verb
- ↑**G** Verbflexion nach der ersten Person pro finites Verb, Dative pro Nomen
- ↓**H** Infinitive pro Verb

#### **Diskurskohäsion**

- ↓**HG** zweite Person Personal- und Possessivpronomen pro Nomen
- ↑**G** adversative Konnektoren pro Satz (nach Eisenberg-Konnektorenliste), erste Person Personalpronomen pro Nomen
- ↓**G** Eigennamen pro Nomen
- ↓**H** temporale Konnektoren pro Satz (nach Breindl-Konnektorenliste), kausale Konnektoren pro Satz (nach Eisenberg-Konnektorenliste)

In beiden Schulformen nimmt im Vergleich zur Grundschule der Anteil der zweiten Person bei den Verbflexionen ab. Dies korrespondiert mit einer reduzierten Nutzung der Possessiv- und Personalpronomen der zweiten Person in den Diskurskohäsionsmaßen. In der Hauptschule nimmt zudem die Nutzung der Infinitivkonstruktionen ab, korrespondierend zu den beobachtbaren Veränderungen in den syntaktischen Domänen. Im Gymnasium nimmt im Gegenzug die Verb-

flexion der ersten Person zu (wieder mit einem Zuwachs der Nutzung der ersten Person Possessiv- und Personalpronomen im Diskursbereich). Zudem nimmt die Nutzung von Dativen signifikant zu (Abbildung 1c).

Im Diskursbereich sind zusätzlich auch Unterschiede in der Nutzung von Konnektoren zu sehen. Im Gymnasium nimmt die Nutzung adversativer Konnektoren zu (Abbildung 1d). In der Hauptschule hingegen nimmt die Nutzung temporaler und kausaler Konnektoren ab (Abbildung 1e).

## Sprachnutzung

Im Bereich der Sprachnutzung finden sich die meisten signifikanten Unterschiede, wobei zu beachten ist, dass es sich auch um die Domäne mit den meisten Merkmalen handelt.

- ↑**HG** Prozentsatz der Worte im zweitniedrigsten Log-Frequenzband (LFB) in Subtlex-DE
- ↓**HG** Prozentsatz der Lemmata im KCT-Korpus, Prozentsatz der Typen in Subtlex-DE, Prozentsatz der Worte im drittniedrigsten LFB in Subtlex-DE, Prozentsatz der Worte im niedrigsten LFB im KCT-Korpus, Prozentsatz der Worte im vierten LFB in Google Books 2000
- ↑**G** mittleres Alter aktiver Wortproduktion pro im KCT-Korpus vorkommendem Lemma, Log Lemmafrequenz pro im KCT-Korpus vorkommendem Lemma, Prozentsatz der Worte im zweithöchsten LFB in Google Books 2000
- ↓**G** Log Typenfrequenz pro in Subtlex-DE vorkommendem Typ, Typenfrequenz pro in Subtlex-DE vorkommendem Typ, maximales Alter aktiver Wortproduktion pro im KCT-Korpus vorkommendem Lemma
- ↓**H** Prozentsatz der Worte im dritten LFB in dlexDB, Prozentsatz der Worte im fünften LFB in Subtlex-DE

In beiden Sekundarschulformen nimmt die Nutzung niedrig-frequenter Worte in der Subtlex-DE Frequenzdatenbank zu. Diese basiert auf Film- und Serienuntertiteln und kommt einer Repräsentation gesprochener Sprache von allen in unserem System genutzten Frequenzdatenbanken am nächsten (Brysbaert et al., 2011). Zugleich nimmt in beiden Schulformen die Nutzung mittel-frequenter Worte in Subtlex-DE und der auf Google Books 2000 basierenden Frequenzdatenbank ab, ebenso wie die Übereinstimmungen mit dem im KCT-Korpus genutzten Vokabular (Abbildung 1f). Das KCT-Korpus besteht aus von Kindern der 1. bis 8. Klasse geschriebenen Texten und dient uns zur Approximation kindlicher Sprachnutzung (Lavalley et al., 2015). Im Gymnasium nimmt zudem die Nutzung

von Worten im höchsten Log-Frequenzband in Google Books 2000 zu und die Frequenz von Worten nach dem KCT-Korpus ab. Außerdem nimmt das durchschnittliche Alter der ersten aktiven Nutzung von Lemmata nach dem KCT-Korpus zu. Jedoch nimmt zugleich das maximale Alter der aktiven Nutzung ab. In der Hauptschule nimmt die Nutzung hochfrequenter Worte in Subtlex-DE ab, ebenso wie die Nutzung mittel-frequenter Worte in der dlexDB Frequenzdatenbank, welche auf Zeitungstexten basiert und damit ein akademisches schriftsprachliches Register repräsentiert (Heister et al., 2011).

### **Sprachverarbeitung und Oberflächenmaße**

Weder bei der Sprachverarbeitung noch den Oberflächenmaßen wurden signifikante Unterschiede zwischen der Grundschule und den Sekundarschulformen festgestellt.

## **6 Studie 2: Vergleich von Lehrkräften und Schüler\*innen**

### **6.1 Aufbau von Studie 2**

Die zweite Studie untersucht, inwiefern die Unterschiede zwischen Lehrkräften der unterschiedlichen Schulformen Unterschieden zwischen den Schüler\*innen der jeweiligen Schulform entsprechen. Hierfür nutzen wir den Gesamtdatensatz bestehend aus den Äußerungen der Lehrkräfte und der Schüler\*innen (N=46.911). Wir untersuchen die 111 Komplexitätsmerkmale, die in den Lehrkraftäußerungen variabel und unkorreliert waren. Da wir keine longitudinalen Daten für die selben Individuen in der Grund- und einer Sekundarschule haben, können wir die Unterschiede nur auf der Gruppenebene betrachten: Wir berechnen pro Domäne sprachlicher Komplexitätsmerkmale die Korrelationen von Merkmalsunterschieden zwischen Schulformen bei Schüler\*innen und Lehrkräften. Dies bietet einen gewissen Einblick darin, inwiefern die zuvor diskutierten Unterschiede zwischen Lehrkräften durch etwaige Unterschiede zwischen Schüler\*innen motiviert sein können.

Als Grundlage für die Berechnung der Korrelation wird jedes Merkmal zunächst z-transformiert, sodass alle Merkmale auf der gleichen Skala liegen und somit vergleichbar sind. Dann wird für jedes Merkmal pro Schulform der

Mittelwert für die Lehrkräfte und für die Schüler\*innen berechnet. Auf den Mittelwerten werden dann jeweils für Lehrkräfte und Schüler\*innen die Differenzen zwischen Schulformen berechnet, um je drei Maße zu erhalten: die Differenz ( $\Delta$ ) der durchschnittlichen Merkmalsausprägung zwischen Grundschule (hier: E für „Elementary School“) und Hauptschule ( $\Delta_{EH}$ ), zwischen Grundschule und Gymnasium ( $\Delta_{EG}$ ), und zwischen Hauptschule und Gymnasium ( $\Delta_{HG}$ ). Schließlich wird für jede der neun Komplexitätsdomänen die Pearson-Korrelation zwischen den Schüler\*innen- und den Lehrkraft-Merkmalisdifferenzen zwischen den Schulformen berechnet.

## 6.2 Ergebnisse von Studie 2

Tabelle 4 zeigt die Korrelationsergebnisse für die neun linguistischen Domänen. Die erste Spalte zeigt die Korrelation der Unterschiede zwischen Grund- und Hauptschule  $c(\Delta_{EH}^S, \Delta_{EH}^L)$ , die zweite Spalte die Korrelation der Unterschiede zwischen Grundschule und Gymnasium  $c(\Delta_{EG}^S, \Delta_{EG}^L)$  und die dritte Spalte die Korrelation der Unterschiede zwischen Hauptschule und Gymnasium  $c(\Delta_{HG}^S, \Delta_{HG}^L)$ . Da die Domäne der Oberflächenmaße nur ein Maß enthält (Anzahl der Worte) kann hierfür keine Korrelation berechnet werden.

Wir beobachten in den meisten Domänen eine mittlere bis starke Korrelation zwischen den Komplexitätsunterschieden bei den Lehrkräften unterschiedlicher Schulformen und denen der Schüler\*innen. Insbesondere im Bereich der Sprachverarbeitung und der klausalen Komplexität, aber auch in dem der lexikalischen Komplexität liegt eine starke Korrelation vor – die Komplexität der Äußerungen der Lehrkräfte unterscheidet sich also zwischen den Schulformen in einer sehr ähnlichen Art und Weise, wie das bei den Äußerungen der Schüler\*innen der Fall ist.

Im phrasalen Bereich hingegen beobachten wir eine moderate Korrelation für die Unterschiede zwischen Grundschule und Gymnasium, und nur eine schwache Korrelation für die Unterschiede zwischen Grund- und Hauptschule und Hauptschule und Gymnasium. Dies bedeutet, dass die Unterschiede in der phrasalen Komplexität von Lehrkraftäußerungen nur in Teilen der von Schüler\*innen-Äußerungen entsprechen.

**Tab. 4:** Domänenweise Korrelation der Durchschnittsdifferenzen  $c(\Delta, \Delta)$  zwischen Schulformen für Schüler\*innen und Lehrkräften (kursiv: absolute Korrelationen  $\leq 0,2$ ; fett:  $\leq 0,2$ ; E = Grundschule; H = Hauptschule; G = Gymnasium; S = Schüler\*innen; L = Lehrkräfte)

Domäne	$c\left(\overset{S}{\Delta}_{EH}, \overset{L}{\Delta}_{EH}\right)$	$c\left(\overset{S}{\Delta}_{EG}, \overset{L}{\Delta}_{EG}\right)$	$\left(\overset{S}{\Delta}_{HG}, \overset{L}{\Delta}_{HG}\right)$
Klausale Komplexität	0,819	0,863	0,917
Phrasale Komplexität	0,311	0,489	0,275
Grammatische Strukturen	0,360	0,770	0,799
Lexikalische Komplexität	0,751	0,760	0,772
Morphologische Komplexität	<b>-0,312</b>	-0,021	0,619
Diskurskohäsion	-0,054	0,405	0,676
Sprachnutzung	0,238	0,202	0,257
Sprachverarbeitung	0,900	0,908	0,765
Oberflächenmaß	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>

Im Bereich der komplexen grammatischen Konstruktionen beobachten wir für den Unterschied zwischen Grundschule und Gymnasium und Hauptschule und Gymnasium eine starke Korrelation, jedoch nur eine bedeutend schwächere Korrelation für den Unterschied zwischen Grund- und Hauptschule.

Bei der morphologischen Komplexität liegt eine schwache negative Korrelation für die Unterschiede zwischen Grund- und Hauptschule vor. Die Unterschiede in der morphologischen Komplexität von Lehrkraftäußerungen zwischen diesen Schulformen entsprechen also nicht dem, was bei den Schüler\*innen zu beobachten ist. Für die Unterschiede zwischen Grundschule und Gymnasium ist praktisch keine Korrelation zu beobachten.

Im Bereich Diskurskohäsion sehen wir keine Korrelation für die Unterschiede zwischen Grund- und Hauptschule und eine moderate Korrelation für die Unterschiede zwischen Grundschule und Gymnasium. Im Bereich Sprachnutzung sehen wir nur schwache Korrelationen über Schulvergleiche hinweg.

## 7 Diskussion

Die erste Studie hat gezeigt, dass sich Lehrkräfte in der Tat systematisch über Schulformen und Klassenstufen hinweg im Hinblick auf die sprachliche Komplexität ihrer Äußerungen unterscheiden und zwar für eine große Bandbreite linguistischer Merkmale. Dies gilt für alle von uns erfassten Bereiche (außer der Sprachverarbeitung und den Oberflächenmaßen, die allerdings nur mit wenigen Merkmalen repräsentiert waren). Auch wenn vielfältige signifikante Unterschiede zwischen der Grundschule und den Sekundarschulformen zu beobachten sind,

gibt es nur wenige Fälle, in denen sich beide Sekundarschulformen in gleicher Weise von der Grundschule abheben. Sekundarschullehrkräfte nutzen signifikant längere Worte als Grundschullehrkräfte und haben eine geringere Verbvariation. Letzteres scheint jedoch im Fall der Hauptschule auf einen spezifischen Rückgang der Nutzung von Modalverben zurückzuführen zu sein, während im Gymnasium lexikalische Varianz im Allgemeinen abzunehmen scheint. In beiden Schulformen ist ein Rückgang der Nutzung der zweiten Person zu beobachten. Dies steht im Einklang mit dem zu erwartenden Wechsel in der Ansprache der Schüler\*innen durch die Lehrkräfte. Aus didaktischen Gründen wird in der Grundschule stärker auf eine direkte Anrede der Schüler\*innen gesetzt („Was siehst du?“), während in der Sekundarstufe dieser Bezug reduziert wird („Was kann man sehen?“). Allerdings ist nur im Gymnasium ein Rückgang der Nutzung von Eigennamen zu beobachten, welcher sich als reduziertes oder weniger explizites Klassenzimmermanagement interpretieren lässt. Die Sprachnutzung in beiden Schulformen entfernt sich von der Sprachnutzung im KCT-Korpus (Lavalley et al., 2015), welches kindlichen Sprachgebrauch repräsentiert. Seltenes gesprochen-sprachliches Vokabular nimmt zu und mittelfrequentes Vokabular gesprochener und geschriebener Sprache ab. Es überwiegen jedoch auch hier die Unterschiede. Beispielsweise scheint sich das Gymnasium stärker von der in SUBTLEX-DE (Brysaert et al., 2011) abgebildeten gesprochenen Sprache zu entfernen als die Hauptschule. Es zeigt auch mehr Annäherungen an auf Schriftsprache basierenden Frequenzdatenbanken, allerdings sind die Ergebnisse sehr heterogen und so nicht einfach interpretierbar.

Besonders auffällig sind die Unterschiede im Bereich der syntaktischen Komplexität. Nur im Gymnasium beobachten wir einen Zuwachs komplexer Nominalphrasen, welche ein bedeutendes Merkmal von Bildungssprache sind. Auch Kleinschmidt-Schinke (2018) zeigt eine Zunahme der lexikalischen Dichte und damit der Integration auf nominalphrasaler Ebene über die Jahrgangsstufen (Kleinschmidt-Schinke, 2018, S. 652f.). In der Hauptschule hingegen scheint die klausale Komplexität im Vergleich zur Grundschule sogar abzunehmen. Der entgegengesetzte Trend in der syntaktischen Entwicklung spiegelt sich insbesondere in der maximalen Satzlänge wider, welche im Gymnasium signifikant zu- (vgl. auch Kleinschmidt-Schinke, 2018, S. 362) und in der Hauptschule signifikant abnimmt. Hierzu passt der signifikante Rückgang der Nutzung temporaler und klausaler Konnektoren in der Hauptschule. Im Gymnasium hingegen ist kein signifikanter Rückgang in der Konnektorennutzung zu beobachten und die Nutzung konzessiver Konnektoren nimmt sogar zu. Dies könnte inhaltlich als Erweiterung der Argumentationsformen im gymnasialen Unterricht gewertet werden, welche in der Hauptschule nicht nur ausbleibt, sondern sogar reduziert wird.

Zusammenfassend scheint die Analyse der Äußerungen der Lehrpersonen im authentischen naturwissenschaftlichen Unterricht einen sehr detaillierten Einblick in bildungssprachlich relevante Eigenschaften und Unterschiede zu bieten. Einige der Unterschiede der sprachlichen Komplexität der Lehrkraftäußerungen zwischen den Schulformen entsprechen den Erwartungen – allerdings bieten sich auch detaillierte Einsichten in Bereiche, in denen die von Bildungspolitik, Fach- und Sprachdidaktik geforderte und/oder erwartete Steigerung sprachlicher Komplexität im Sinne einer durchgängigen Sprachbildung derzeit nicht systematisch umgesetzt werden.

In unserer zweiten Studie haben wir untersucht, ob die Unterschiede in der Komplexität der Lehrkraftäußerungen zwischen den Schulformen den demonstrierten sprachlichen Fähigkeiten der Schüler\*innen in den jeweiligen Schulformen entsprechen, sich also adaptiv an dem jeweiligen Schüler\*innen-Niveau ausrichten. Wir beobachten bis auf wenige Ausnahmen weitestgehend und insbesondere im Bereich der Sprachverarbeitung sowie der klausalen und lexikalischen Komplexität moderate bis starke Korrelationen. Einzig im Bereich der morphologischen Komplexität beobachten wir eine negative Korrelation beim Unterschied von Grund- und Hauptschule. Hier liegt demnach keine Entsprechung zwischen den Unterschieden zwischen Grund- und Hauptschullehrkräften und Unterschieden zwischen Grundschul- und Hauptschüler\*innen vor. Dies legt weitere Untersuchungen nahe, wobei die in Studie 1 festgestellten Unterschiede in der Nutzung von Personenmerkmalen seitens der Lehrkräfte zumindest in Teilen eine potentielle Erklärung darstellen.

Die mehrheitlich positiven Korrelationen zwischen den Unterschieden zwischen Lehrkräften und Schüler\*innen können mit Blick auf die sinkende sprachliche Komplexität zwischen Grund- und Hauptschule ein Beleg für eine Anpassung der lehrkraftseitigen Äußerungen an die sprachlichen Fähigkeiten der Schüler\*innen sein. Es stellt sich jedoch dann die Frage, ob die im Verhältnis zur Hauptschule erstaunlich hohe sprachliche Komplexität der Lehrkraftäußerungen in der Grundschule allen Schüler\*innen gerecht wird oder bereits so komplex ist, dass sie einigen Kindern nicht mehr als Gerüst im Sinne des *Scaffoldings* in der individuellen Entwicklungsdomäne (*ZPD*) dient, sondern weit über deren sprachlichen Fähigkeiten steht. Dies könnte ein Grund dafür sein, dass es laut der Studie von Holtappels und Heerdegen (2005) in der Grundschule nicht gelingt, fachliches Lernen für alle Kinder gleichermaßen zu ermöglichen.

Im konzeptuellen Kontext der Anforderungen an eine durchgängige Sprachbildung interpretiert, zeigen unsere Ergebnisse, dass nicht grundsätzlich von einer steigenden sprachlichen Komplexität mit Voranschreiten der Bildungsbiografie (hier abgebildet über die Klassenstufen) auszugehen ist (im Unterschied



zu Kleinschmidt-Schinke, 2018, für den Vergleich Grundschule und Gymnasium). Was bedeutet das insbesondere im Hinblick auf die sprachliche Komplexität, der Schüler\*innen in den Hauptschulen begegnen? Es ist möglich, dass sich die Lehrkräfte in der Hauptschule bewusst an den Sprachstand der Schüler\*innen anpassen, um diese in Ihrer ZPD zu fördern. Dies hätte Implikationen für die Grundschule, da in diesem Fall zur Diskussion stünde, inwiefern die Lehrkraftäußerungen in der Grundschule dem heterogenen Sprachstand der Schüler\*innen gerecht werden. Ob diese bewusste oder unbewusste Anpassung an den Sprachstand der Schüler\*innen in der Hauptschule fachliches und sprachliches Lernen unterstützt oder ob die Lehrkräfte so nicht hinreichend als Sprachvorbild für die weitere Entwicklung fungieren, ergibt sich als zusätzliche wichtige Frage für die weitere Forschung, mit direkter praktischer Relevanz für die Lehrkraft(fort)bildung.

## 8 Ausblick

In dem vorliegenden Beitrag haben wir die sprachliche Komplexität der Äußerungen von Lehrkräften im Fachunterricht über Schulformen und Klassenstufen hinweg verglichen, um zu identifizieren, inwiefern sie Schüler\*innen mit ihrer Sprache gezielt an zunehmend bildungssprachliche Konstruktionen heranführen. Unsere erste Studie zeigt, dass sich Lehrkräfte unterschiedlicher Schulformen in der Tat im Hinblick auf die sprachliche Komplexität ihrer Äußerungen unterscheiden – bei einer großen Bandbreite linguistischer Merkmale in fast allen repräsentierten Komplexitätsdomänen. Allerdings zeigt eine detaillierte Analyse der Unterschiede, dass sich Äußerungen von Lehrkräften primär im Gymnasium durch die vermehrte Nutzung bildungssprachlicher Konstruktionen im Vergleich zur Grundschule auszeichnen. Lehrkraftäußerungen in der Hauptschule hingegen zeigen in erster Linie eine Reduktion sprachlicher Komplexität über alle Domänen hinweg. Dies widerspricht der fachdidaktischen Annahme, dass sich Bildungssprache im Fachunterricht „langsam und kumulativ über mehrere Stufen“ (Leisen, 2015a, S. 135) entwickelt und somit im Unterricht höherer Klassenstufen mit komplexerer Sprache zu rechnen ist. Hiermit leistet unsere Untersuchung eine wertvolle Ergänzung zu den Erkenntnissen von Kleinschmidt-Schinke (2018), indem wir die Perspektive auf weitere Schulformen erweitern. Zugleich eröffnet unsere Studie neue Forschungsdesiderate: Mit den hier angewandten Methoden lässt sich nicht klären, wie die beobachteten Unterschiede zwischen Gymnasium und Hauptschule zu bewerten sind und welche Konsequenzen für die Lehrkraft(fort)bildung sich hieraus ergeben.

In unserer zweiten Studie haben wir untersucht, inwiefern ein Zusammenhang zwischen der sprachlichen Komplexität der Lehrkräfte und der der Schüler\*innen in Abhängigkeit von der Schulform besteht. Insbesondere in den Bereichen Lexikon, klausale Syntax und Sprachverarbeitung beobachten wir für beide Sekundarschulformen starke positive Korrelationen für die sprachlichen Unterschiede zwischen Schulformen bei Schüler\*innen und Lehrkräften, allerdings ebenfalls teilweise negative Korrelationen für den Vergleich von Grund- und Hauptschule im Bereich Morphologie. Bei der Interpretation dieser Ergebnisse sind jedoch die explorative Natur unserer Untersuchung sowie die methodischen Limitationen unserer zweiten Studie zu bedenken, die nur auf Mittelwerten pro Schulformen basiert. Essentiell für eine adaptive sprachliche Förderung ist jedoch die individuelle Interaktion und entsprechende Passung von sprachlichem Angebot und Sprachkompetenz. In einem nächsten Analyseschritt möchten wir daher in den hier untersuchten Daten annotieren, an wen sich die Äußerungen von Lehrkräften richten, um sprachliche Alignmentprozesse in individuellen Interaktionen analysieren zu können. Dann können auch minimale diskursive Einheiten identifiziert werden, sodass der Einfluss des Kontextes und der aktuellen sprachlichen Aufgabe methodisch dingfest und separierbar gemacht werden kann. Die notwendigen Methoden für die linguistische Analyse individueller Alignmentprozesse in authentischen Unterrichtskontexten werden derzeit im Rahmen des COLD Verbundprojekt (<https://www.die-bonn.de/cold/>) entwickelt, in dem die professionellen Kompetenzen von Lehrkräften in Schule und Erwachsenenbildung beim Unterrichten von Deutsch als Zweitsprache in sprachlich heterogenen Lerngruppen erfasst werden. Mittelfristig erlaubt eine solche Aufschlüsselung auch die Berücksichtigung weiterer interdisziplinärer Perspektiven aus der Gesprächsanalyse, da sie den interaktiven Charakter der Daten zugänglich machen. Dies würde es auch erlauben, den Einfluss von verschiedenen Diskursformen (z. B. Diskussion im Plenum vs. Partner\*innen- oder Kleingruppenarbeit) und diskursiven Funktionen (z. B. Inhaltsvermittlung vs. Organisationsphasen) auf die sprachliche Komplexität des Unterrichtsgesprächs zu untersuchen. Von der Datenbasis her wäre langfristig für eine Analyse individueller bildungssprachlicher Entwicklungsverläufe im Kontext des von Lehrkräften und Lehrwerken gebotenen sprachlichen Inputs eine longitudinale Erhebung von Daten wünschenswert, welche die sprachliche Entwicklung von Schüler\*innen am Übergang von der Grundschule auf die weiterführende Schulform intra-individuell analysierbar macht.

Unser Beitrag stellt insgesamt einen wertvollen Ausgangspunkt für zukünftige interdisziplinäre Forschungsarbeiten dar, die den Aufgaben- und Kontexteffekten aus unterschiedlichen Perspektiven weiter nachgehen. Aus Sicht der fachdidaktischen Forschung bleibt zu klären, inwiefern die Erkenntnisse zu den

linguistischen Eigenschaften von Bildungssprachangeboten in naturwissenschaftlichen Schulbüchern auf gesprochene Lehr-Lern-Situationen im naturwissenschaftlichen Unterricht anwendbar sind (für einen Überblick, siehe Härtig et al., 2015). Zugleich verbleibt empirisch zu bestimmen, ob bestimmte fachspezifische Aufgabenschemata, bspw. nach Feilke (2014), in Übertragung auf den naturwissenschaftlichen Unterricht (z. B. beschreiben, experimentieren, sachgemäß und evidenzbasiert argumentieren) mit bestimmten sprachlichen Komplexitätsmerkmalen einhergehen. So ist der Grad der (De-)Kontextualisierung von naturwissenschaftlichen Aufgaben (konkrete Beschreibung in einer Experimentsituation vs. allgemeingültige Erklärung) mit deren sprachlichen Anforderungen in Verbindung gebracht worden (Quehl & Trapp, 2013). Auch wurden der Umgang mit unpersönlichen Ausdrücken, Fachverben mit festen Präpositionen oder fachspezifischen Abkürzungen als bedeutsam für naturwissenschaftliches Lernen deklariert (Beese et al., 2014). Die in diesem Beitrag zusammengeführten interdisziplinären Perspektiven der (Computer-)Linguistik, Sprachdidaktik und Fachdidaktik können zur Addressierung dieser Forschungsdesiderate beitragen und somit sowohl einen wichtigen Beitrag zur Grundlagenforschung als auch zu der Gestaltung von Maßnahmen für die Unterrichtsgestaltung und Lehrkraft(fort-)bildung leisten.

Eine methodische Frage für die linguistische Forschung betrifft die Analyse gesprochener (Lerner\*innen-)Sprache mit computerlinguistischen Modellen, die auf geschriebener Standardsprache trainiert wurden. Dies ist auch für die Lerner\*innen-Korpusforschung eine wichtige Fragestellung. Da die computerlinguistischen Methoden zur Berechnung der sprachlichen Komplexität erfolgreich auf kurze schriftlichen Antworten von Deutsch L2-Lerner\*innen auf Anfängerniveau angewandt werden können (Weiss & Meurers, 2021), ist davon auszugehen, dass die Methoden auch hinreichend robust für die standardisierte Transkription der PLUS-Studie sind, für die eine Wort- und Satzsegmentierung vorliegt. Eine zukünftige genauere Untersuchung der Frage, welche Komplexitätsmaße für gesprochene Sprache effektiv automatisiert berechnet werden können, wäre interessant um herauszuarbeiten, welche Unterschiede derzeit aufgrund unzuverlässiger Merkmalsberechnung nicht gefunden werden können – die bereits identifizierten signifikanten Unterschiede haben jedoch davon unabhängig Bestand.

Trotz der erwähnten Limitationen und weiterzuverfolgenden Fragen leisten unsere Studien einen empirisch expliziten Beitrag zur Erfassung des bildungssprachlichen Inputs, den Lehrkräfte Schüler\*innen im Fachunterricht bieten. Ein wichtiger methodischer Beitrag ist, dass wir weitgehende Analysemöglichkeiten für Daten aufzeigen, die bisher kaum Annotationen für eine linguistische Analyse beinhalten. Dies ist insbesondere deshalb relevant, da zahlreiche Korpora von authentischen Lehr-Lern-Situationen in gesprächsanalytischen Studien kom-

piliert werden, die so für eine Vielzahl von Forschungsfragen nutzbar gemacht werden können. Unser Beitrag demonstriert, dass solche Daten sinnvoll für eine Vielzahl linguistischer Studien eingesetzt werden können. Neben der Erschließung neuer Datenquellen bietet auch für die Lerner\*innen-Korpusforschung an sich die Erweiterung der Analysemethoden um ein großes Repertoire von relativ robusten Komplexitätsmerkmalen wichtige neue Perspektiven.

**Danksagung:** Wir möchten allen Teilnehmer\*innen des ZGL-Kolloquiums zu diesem Themenheft für die intensive Diskussion und die vielfältigen Perspektiven danken, die diesen Beitrag bereichert haben.

---

**Fördermittel:** Diese Arbeit wurde von der Leibniz-Gemeinschaft gefördert (Leibniz-Wettbewerb Kooperative Exzellenz, Förderkennzeichen: K113/2018).

---

**Zusatzmaterial:** Die Onlineversion dieses Artikels bietet Zusatzmaterial (<https://doi.org/10.1515/zgl-2022-2052>).

## Literatur

- Ágel, Vilmos. & Hennig, Mathilde (2006). Theorie des Nähe- und Distanzsprechens. In V. Ágel & M. Hennig (Hrsg.), *Grammatik aus Nähe und Distanz. Theorie und Praxis am Beispiel von Nahetexten 1650–2000* (S. 3–31). Tübingen: Niemeyer.
- Ágel, Vilmos & Hennig, Mathilde. (2007). Überlegung zur Theorie und Praxis des Nähe- und Distanzsprechens. In V. Ágel & M. Hennig (Hrsg.), *Zugänge zur Grammatik der gesprochenen Sprache* (S. 179–214). Tübingen: Niemeyer.
- Alexopoulou, Theodora, Michel, Marije, Murakami, Akira & Meurers, Detmar. (2017). Task effects on linguistic complexity and accuracy: A large-scale learner corpus analysis employing natural language processing techniques. *Language Learning*, 67, 181–209.
- Bachmann, Thomas & Becker-Mrotzek, Michael. (2017). Schreibkompetenz und Textproduktion modellieren. *Forschungshandbuch empirische Schreibdidaktik*, 25, 53.
- Bailey, Alison L. (Hrsg.). (2007). *The language demands of school. Putting academic English to the test*. New Haven and London: Yale University Press.
- Ballestracci, Sabrina. (2010). Der Erwerb von Verbzweitsätzen mit Subjekt im Mittelfeld bei italophonen DaF-Studierenden. Erwerbsphasen, Lernschwierigkeiten und didaktische Implikationen. *Linguistik online*, 41 (1).
- Bartón, Kamil. (2020). Mumin: Multi-model inference [Software-Handbuch]. (R package version 1.43.17)
- Bates, Douglas, Mächler, Martin, Bolker, Ben & Walker, Steve. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67 (1), 1–48.
- Becker-Mrotzek, Michael, Grabowski, Joachim, Jost, Jörg, Knopp, Matthias & Linnemann, Markus. (2014). Adressatenorientierung und Kohärenzherstellung im Text. Zum

- Zusammenhang kognitiver und sprachlicher realisierter Teilkompetenzen von Schreibkompetenz. *Didaktik Deutsch: Halbjahresschrift für die Didaktik der deutschen Sprache und Literatur*, 19 (37), 21–43.
- Becker-Mrotzek, Michael, Schramm, Karin, Thürmann, Eike & Vollmer, Helmut J. (2013). *Sprache im Fach: Sprachlichkeit und fachliches Lernen*. Münster: Waxmann Verlag.
- Beese, Melanie, Benholz, Claudia, Chlosta, Christoph, Gürsoy, Erkan, Hinrichs, Beatrix, Niederhaus, Constanze & Oleschko, Sven. (2014). *Sprachbildung in allen Fächern* (Bd. 16). Stuttgart: Ernst Klett Sprachen.
- Belz, Malte & Odebrecht, Carolin. (2022). Abschnittsweise Analyse sprachlicher Flüssigkeit in der Lernaltersprache: Das Ganze ist weniger informativ als seine Teile. *Zeitschrift für germanistische Linguistik*, 50 (1), 131–158.
- Benjamini, Yoav & Hochberg, Yosef. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society (B)*, 57 (1), 289–300.
- Berendes, Karin, Vajjala, Sowmya, Meurers, Detmar, Bryant, Doreen, Wagner, Wolfgang, Chinkina, Maria & Trautwein, Ulrich. (2018). Reading demands in secondary school: Does the linguistic complexity of textbooks increase with grade level and the academic orientation of the school track? *Journal of Educational Psychology*, 110 (4), 518–543.
- Biber, Douglas & Gray, Bethany. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, 9 (1), 2–20.
- Biber, Douglas, Gray, Bethany & Staples, Shelley. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics*, 37 (5), 639–668.
- Bley-Vroman, Robert. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning*, 33 (1), 1–17.
- Bos, Wilfried, Valtin, Renate, Hußmann, Anke, Wendt, Heike & Goy, Martin. (2017). IGLU 2016: Wichtige Ergebnisse im Überblick. *IGLU*, 13–28.
- Brandt, Hanne & Gogolin, Ingrid. (2016). *Sprachförderlicher Fachunterricht: Erfahrungen und Beispiele. Unter Mitarbeit von Margit Maronde-Heyl und Heidi Scheinhardt-Stettner* (Bd. 8). Münster: Waxmann.
- Breindl, Eva, Volodina, Anna & Waßner, Ulrich H. (2015). *Handbuch der deutschen Konnektoren 2. Semantik der deutschen Satzverknüpfers* (Bd. 13). De Gruyter.
- Brezina, Vaclav & Pallotti, Gabriele. (2019). Morphological complexity in written L2 texts. *Second language research*, 35 (1), 99–119.
- Brysbaert, Marc, Buchmeier, Matthias, Conrad, Markus, Jacobs, Arthur M., Bölte, Jens & Böhl, Andreas. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, 58, 412–424.
- Bulté, Bram & Housen, Alex. (2012). Defining and operationalising L2 complexity. In A. Hou-sen, F. Kuiken & I. Vedder (Hrsg.), *Dimensions of L2 performance and proficiency* (S. 21–46). John Benjamins. Zugriff auf <https://doi.org/10.1075/llt.32.02bul>
- Bulté, Bram & Housen, Alex. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26 (0), 42–65. (Comparing perspectives on L2 writing: Multiple analyses of a common corpus)
- Bürgermeister, Anika, Große, Gerlind, Leuchter, Miriam, Studhalter, Ueli & Saalbach, Henrik. (2018). Interaktion von pädagogischen Fachkräften und Kindern in naturwissenschaftlichen Lerngelegenheiten im Kindergarten. *Frühe Bildung*, 8, 13–21.

- Chen, Shi-Yi, Feng, Zhe & Yi, Xiaolian. (2017). A general introduction to adjustment for multiple comparisons. *Journal of thoracic disease*, 9 (6), 1725.
- Chen, Xiaobin & Meurers, Detmar. (2016). CTAP: A web-based tool supporting automatic complexity analysis. In *Proceedings of the workshop on computational linguistics for linguistic complexity* (S. 113–119). Osaka, Japan: COLING.
- Crossley, Scott A. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, 11 (3), 415–443.
- Crossley, Scott A., Kyle, Kristopher, Allen, Laura K., Guo, Liang & McNamara, Danielle S. (2014). Linguistic microfeatures to predict L2 writing proficiency: A case study in automated writing evaluation. *Journal of Writing Assessment*, 7 (1), 1–15.
- Cummins, Jim. (2000). Language, power and pedagogy: Bilingual children in the crossfire. *Multilingual Matters (Bilingual education and bilingualism)*, 23.
- Daroczy, Gabriella, Artemenko, Christina, Meurers, Detmar, Wolska, Magdalena & Nuerk, Hans-Christoph. (2020). Influence of task characteristics on eye-movement patterns related to numerical and textual information in arithmetic word problems. *Neuroscience*, 9, 1–8.
- De Clercq, Barbara & Housen, Alex. (2019). The development of morphological complexity: A cross-linguistic study of L2 French and English. *Second Language Research*, 35 (1), 71–97.
- DeKeyser, Robert M. (2001). Automaticity and automatization. *Cognition and second language instruction*, 125–151.
- Drumm, Sandra. (2019a). Analyse von schulischen Vermittlungstexten in Biologielehrwerken. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 24 (1).
- Drumm, Sandra. (2019b). Umgang mit Sachtexten im CLIL-Unterricht. Ein Beispiel anhand des Faches Biologie. In *IDT 2017*. Berlin: Erich Schmidt Verlag.
- Eisenberg, Peter, Peters, Jörg, Gallmann, Peter, Fabricius-Hansen, Cathrine, Nübling, Damaris, Barz, Irmhild, ... Fiehler, Reinhard. (2009). *Duden. Deutsche Grammatik* (8. überarbeitete Aufl., Bd. 4). Mannheim: Bibliographisches Institut & F. A. Brockhaus AG. Dudenredaktion.
- Ellis, Rod. (2003). *Task-based language learning and teaching*. Oxford, UK: Oxford University Press.
- Erath, Kirstin, Prediger, Susanne, Quasthoff, Uta & Heller, Vivian. (2018). Discourse competence as important part of academic language proficiency in mathematics classrooms: The case of explaining to learn and learning to explain. *Educational Studies in Mathematics*, 99 (2), 161–179.
- Ewerhardy, Anne. (2010). *Zusammenhänge zwischen Verständnisorientierung von naturwissenschaftsbezogenem Sachunterricht und Fortschritten im Verständnis naturwissenschaftlicher Konzepte bei Lernenden der Grundschule* (Inaugural-Dissertation). Westfälische Wilhelms-Universität, Münster.
- Fandrych, Christian & Wallner, Franziska. (2022). Funktionale und stilistische Merkmale fortgeschrittener Lerner:innensprache: Methodische und konzeptionelle Überlegungen am Beispiel von GeWiss. *Zeitschrift für germanistische Linguistik*, 50 (1), 202–239.
- Fang, Zhihui, Schleppegrell, Mary J. & Cox, Beverly E. (2006). Understanding the language demands of schooling: Nouns in academic registers. *Journal of literacy research*, 38 (3), 247–273.
- Feilke, Helmuth. (2012). Bildungssprachliche Kompetenzen – fördern und entwickeln. *Praxis Deutsch*, 39 (233), 3–13.
- Feilke, Helmuth. (2014). Argumente für eine Didaktik der Textprozeduren. In T. B. und Helmuth Feilke (Hrsg.), *Werkzeuge des Schreibens. Beiträge zu einer Didaktik der Textprozeduren* (1. Aufl., S. 11–34). Fillibach bei Klett.

- Garner, James, Crossley, Scott & Kyle, Kristopher. (2019). N-gram measures and L2 writing proficiency. *System*, 80, 176–187.
- Geist, Barbara, Lange-Schubert, Kim & Dietze, Stefan. (2017). Bildungssprachliche Merkmale im Sachunterricht der Grundschule: Theoretische und empirische Annäherungen. In E. Tschirner, J. Möhring & K. Cothrun (Hrsg.), *Deutsch als zweite Bildungssprache in MINT-Fächer*.
- Gibbons, Pauline. (2002). *Scaffolding language, scaffolding learning*. Portsmouth, NH: Heinemann.
- Gibson, Edward. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita & W. O'Neil (Hrsg.), *Image, language, brain: papers from the first mind articulation project symposium* (S. 95–126). MIT.
- Gogolin, Ingrid. (2007). *Wem nützt oder schadet Zweisprachigkeit?*
- Gogolin, Ingrid & Lange, Imke. (2010). *Durchgängige Sprachbildung: Eine Handreichung* (Bd. 2). Münster: Waxmann.
- Gogolin, Ingrid & Lange, Imke. (2011). Bildungssprache und Durchgängige Sprachbildung. In S. Fürstenau & M. Gomolla (Hrsg.), *Migration und schulischer Wandel: Mehrsprachigkeit. Lehrbuch* (S. 107–127). VS Verlag für Sozialwissenschaften.
- Gogolin, Ingrid & Roth, Hans-Joachim. (2007). Bilinguale Grundschule. Ein Beitrag zur Förderung der Mehrsprachigkeit. In *Mehrsprachigkeit bei Kindern und Erwachsenen* (S. 31–46). Tübingen: Attempto.
- Hardy, Ilonca, Kloetzer, Birgit, Moeller, Kornelia & Sodian, Beate. (2010). The analysis of classroom discourse: Elementary school science curricula advancing reasoning with evidence. *Educational Assessment*, 15 (3–4), 197–221.
- Härtig, Hendrik, Bernholt, Sascha, Prechtel, Helmut & Retelsdorf, Jan. (2015). Unterrichtssprache im Fachunterricht – Stand der Forschung und Forschungsperspektiven am Beispiel des Textverständnisses. *Zeitschrift für Didaktik der Naturwissenschaften*, 21, 55–67.
- Heintze, Andreas. (2010). Durchgängige Sprachbildung. eine gemeinsame Aufgabe für jede schule. *Grundschulunterricht Deutsch*, 57 (4).
- Heister, Julian, Würzner, Kay-Michael, Bubbenzer, Johannes, Pohl, Edmund, Hanneforth, Thomas, Geyken, Alexander & Kliegl, Reinhold. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, 62, 10–20.
- Heller, Vivian. (2017). Lerngelegenheiten für bildungssprachliche Kompetenzen: Wie partizipieren DaZ-Lerner am Erklären und Argumentieren im Unterricht? In S. J. u. W. K. I. Fuchs (Hrsg.), *Mehrsprachigkeit: Spracherwerb, Unterrichtsprozesse, Schulentwicklung* (S. 162–185). Stuttgart: Fillibach bei Klett.
- Heller, Vivian & Morek, Miriam. (2015). Unterrichtsgespräche als Erwerbskontext: Kommunikative Gelegenheiten für bildungssprachliche Praktiken erkennen und nutzen. *Les forum.ch*, 3, 1–23.
- Hennig, Mathilde & Niemann, Robert. (2013). Unpersönliches Schreiben in der Wissenschaft. *Informationen Deutsch als Fremdsprache*, 4, 439–455.
- Hirschmann, Hagen & Schmidt, Thomas. (2022). Mündliche Lernerkorpora: Methodisch-technische Aspekte der Erhebung, Erschließung und Nutzung. *Zeitschrift für germanistische Linguistik*, 50 (1), 36–81.
- Holtappels, Heinz G. & Heerdegen, Marco. (2005). Schülerleistungen in unterschiedlichen Lernumwelten im Vergleich zweier Grundschulmodelle in Bremen. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (S. 361–397). Münster: Waxmann.

- Housen, Alex, De Clercq, Barbara, Kuiken, Folkert & Vedder, Ineke. (2019). Multiple approaches to complexity in second language research. *Second Language Research*, 35 (1), 2–31.
- Housen, Alex, Kuiken, Folkert & Vedder, Ineke. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken & I. Vedder (Hrsg.), *Dimensions of L2 performance and proficiency* (S. 1–20). John Benjamins.
- Hüttis-Graff, Petra, Merklinger, Daniela, Klens, Stefanie & Speck-Hamdan, Angelika. (2010). Bildungssprache als Bedingung für erfolgreiches Lernen. In H. Bartnitzky & U. Hecker (Hrsg.), *Allen Kindern gerecht werden. Aufgabe und Wege* (Bd. 129, S. 238–265). Frankfurt am Main: Grundschulverband.
- Karges, Katharina, Studer, Thomas & Hicks, Nina S. (2022). Lernalternativen, Aufgaben und Modalität: Beobachtungen zu Texten aus dem Schweizer Lernerkorpus SWIKO. *Zeitschrift für germanistische Linguistik*, 50 (1), 104–130.
- Kempert, Sebastian, Schalk, Lennart & Saalbach, Henrik. (2019). Übersichtsartikel: Sprache als Werkzeug des Lernens: Ein Überblick zu den kommunikativen und kognitiven Funktionen der Sprache und deren Bedeutung für den fachlichen Wissenserwerb. *Psychologie in Erziehung und Unterricht*, 66 (3), 176–195.
- Kleinschmidt-Schinke, Katrin. (2018). *Die an die Schüler/-innen gerichtete Sprache (SgS): Studien zur Veränderung der Lehrer/-innensprache von der Grundschule bis zur Oberstufe* (Bd. 310). Berlin: De Gruyter.
- KMK. (2013). *Interkulturelle Bildung und Erziehung in der Schule: Beschluss der Kultusministerkonferenz vom 25.10.1996*.
- Koch, Peter & Oesterreicher, Wulf. (1985). Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. In *Romanisches Jahrbuch* (Bd. 36, S. 15–43). Berlin: De Gruyter.
- Kühberger, Christoph, Bramann, Christoph, Weiss, Zarah & Meurers, Detmar. (2019). Task complexity in history textbooks. a multidisciplinary case study on triangulation in history education research. *History Education International Research Journal (HEIRJ)*, 16 (1).
- Kuznetsova, Alexandra, Brockhoff, Per B. & Christensen, Rune H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82 (13), 1–26.
- Kyle, Kristopher. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication* (Dissertation, Georgia State University). Zugriff auf [http://scholarworks.gsu.edu/ales\\_diss/35](http://scholarworks.gsu.edu/ales_diss/35)
- Lavalley, Rémi, Berkling, Kay & Stüker, Sebastian. (2015). Preparing children's writing database for automated processing. In *Ltlt@ slate* (S. 9–15).
- Leisen, Josef. (2015a). Fachlernen und Sprachlernen! Bringt zusammen, was zusammen gehört! *Der mathematische und naturwissenschaftliche Unterricht (MNU)*, 68 (3), 132–137.
- Leisen, Josef. (2015b). Sprachlehren und Sprachlernen im Fach – Bildungssprache im CLIL-Unterricht. In A. Dittmann, B. Giblak & M. Witt (Hrsg.), *Bildungsziel: Mehrsprachigkeit* (S. 241–257). Leipzig: Leipziger Universitätsverlag.
- Leuchter, Miriam & Saalbach, Henrik. (2014). Verbale Unterstützungsmaßnahmen im Rahmen eines naturwissenschaftlichen Lernangebots in Kindergarten und Grundschule. *Unterrichtswissenschaft*, 42, 117–131.
- Linell, Per. (2005). *The written language bias in linguistics: Its nature, origins and transformations*. London/New York: Routledge.
- Lintunen, Pekka & Mäkilä, Mari. (2014). Measuring syntactic complexity in spoken and written learner language: comparing the incomparable? *Research in Language*, 12 (4), 377–399.



- McCarthy, Philip M. & Jarvis, Scott. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42 (2), 381–392.
- Möller, Kornelia, Kleickmann, Thilo & Lange, Kim. (2013). Naturwissenschaftliches Lernen im Übergang von der Grundschule zur Sekundarstufe. In H. E. Fischer & E. Sumfleth (Hrsg.), *nwu-essen – 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht* (S. 57–120). Berlin: Fischer.
- Pallotti, Gabriele. (2009). Caf: Defining, refining and differentiating constructs. *Applied Linguistics*, 30 (4), 590–601.
- Pallotti, Gabriele. (2015). A simple view of linguistic complexity. *Second Language Research*, 31 (1), 117–134.
- Pallotti, Gabriele. (2019). An approach of assessing the linguistic difficulty of tasks. *Journal of the European Second Language Association*, 3 (1), 58–70.
- Pallotti, Gabriele & Ferrari, Stefania. (2008). La variabilità situazionale dell'interlingua: implicazioni per la ricerca acquisizionale e il testing linguistico [Interlanguage situational variability: implications for SLA research and language testing]. In G. Bernini, L. Spreafico & A. Valentini (Hrsg.), *Competenze lessicali e discorsive nell'acquisizione di lingue seconde* (S. 437–462). Bergamo: Guerra Perugia.
- Paquot, Magali. (2019). The phraseological dimension in interlanguage complexity research. *Second Language Research*, 35 (1), 121–145.
- Peterson, Ryan A. & Cavanaugh, Joseph E. (2019). Ordered quantile normalization: a semi-parametric transformation built for the cross-validation era. *Journal of Applied Statistics*, 1–16.
- Pohl, Thorsten. (2017). Komplexität als Operationalisierungsdimension konzeptioneller Schriftlichkeit in Untersuchungen zum Unterrichtsdiskurs. In M. Hennig (Hrsg.), *Linguistische Komplexität – ein Phantom?* (Bd. 94, S. 253–280). Tübingen: Stauffenberg.
- Quehl, Thomas & Trapp, Ulrike. (2013). *Sprachbildung im Sachunterricht der Grundschule: Mit dem Scaffolding-Konzept unterwegs zur Bildungssprache: Lehrprogramm gemäß § 14 JuSchG* (Bd. 4). Münster: Waxmann.
- Reich, Hans H. (2013). Durchgängige Sprachbildung. In I. Gogolin, I. Lange, U. Michel & H. H. Reich (Hrsg.), *Herausforderung Bildungssprache – und wie man sie meistert* (S. 55–70). Münster: Waxmann.
- Reiss, Kristina, Sälzer, Christine, Schiepe-Tiska, Anja, Klieme, Eckhard & Köller, Olaf. (2015). *PISA 2015. Eine Studie zwischen Kontinuität und Innovation*. Münster: Waxmann.
- Rescher, Nicholas. (1998). *Complexity: A philosophical overview*. London: Transaction Publishers.
- Riebling, Linda. (2013). Heuristik der Bildungssprache. In I. Gogolin, I. Lange, U. Michel & H. H. Reich (Hrsg.), *Herausforderung Bildungssprache – und wie man sie meistert* (S. 106–153). Münster: Waxmann.
- Rimmele, Rolf. (2003). *Videograph. Multimedia Player zur Kodierung von Videos*. Kiel: Institut für die Pädagogik der Naturwissenschaften.
- Rincke, Karsten. (2007). *Sprachentwicklung und Fachlernen im Mechanikunterricht* (Bd. 66; H. Niedderer, H. Fischler & E. Sumfleth, Hrsg.). Berlin: Logos. Zugriff auf <https://kobra.bibliothek.uni-kassel.de/handle/urn:nbn:de:hebis:34-2007101519358>
- Rincke, Karsten. (2010). Von der Alltagssprache zur Fachsprache – Bruch oder schrittweiser Übergang? In G. Fenkhar, E. Zeitlinger & A. Lembens (Hrsg.), *Sprache, Mathematik und Naturwissenschaften* (S. 47–62). Innsbruck: Studien Verlag.

- Rincke, Karsten & Markić, Silvija. (2018). Sprache und das Lernen von Naturwissenschaften. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (S. 31–48). Berlin, Heidelberg: Springer.
- Schleppegrell, Mary J. (2004). *The language of schooling. A functional linguistics perspective*. Mahwah, New Jersey and London: Lawrence Erlbaum Associates Publishers.
- Schmidt, Karsten. (2016). Der graphematische Satz. *Zeitschrift für germanistische Linguistik*, 44 (2), 215–256.
- Schmölzer-Eibinger, Sabine. (2013). Sprache als Medium des Lernens im Fach. In M. Becker-Mrotzek, K. Schramm, E. Thürmann & H. J. Vollmer (Hrsg.), *Sprache im Fach: Sprachlichkeit und fachliches Lernen* (S. 25–40). Münster: Waxmann Verlag.
- Shain, Cory, van Schijndel, Marten, Futrell, Richard, Gibson, Edward & Schuler, William. (2016). Memory access during incremental sentence processing causes reading time latency. In *Proceedings of the workshop on computational linguistics for linguistic complexity (CL4LC)* (S. 49–58). Osaka.
- Snow, Catherine E. & Uccelli, Paola. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Hrsg.), *The Cambridge handbook of literacy* (S. 112–133). Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511609664.008
- Trouvain, Jürgen. (2022). Das IFCASL-Korpus als phonetisches Lernerkorpus. *Zeitschrift für germanistische Linguistik*, 50 (1), 82–103.
- Uessler, Stella, Runge, Anna & Redder, Angelika. (2013). „Bildungssprache“ diagnostizieren. Entwicklung eines Instruments zur Erfassung von bildungssprachlichen Fähigkeiten bei Viert- und Fünftklässlern. In A. Redder & S. Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik: Interdisziplinäre Perspektiven* (S. 42–67). Münster: Waxmann Verlag.
- Vajjala, Sowmya. (2018). Automated assessment of non-native learner essays: Investigating the role of linguistic features. *International Journal of Artificial Intelligence in Education*, 28 (1), 79–105.
- Vajjala, Sowmya & Meurers, Detmar. (2014). Exploring measures of „readability“ for spoken language: Analyzing linguistic features of subtitles to identify age-specific tv programs. In *Proceedings of the 3rd workshop on predicting and improving text readability for target reader populations (PITR)* (S. 21–29).
- van de Pol, Janneke, Volman, Monique, Oort, Frans. & Beishuizen, Jos. (2015). The effects of scaffolding in the classroom: support contingency and student independent working time in relation to student achievement, task effort and appreciation of support. *Instructional Science*, 43 (5), 615–641.
- van Heuven, Walter J., Mandera, Pawel, Keuleers, Emmanuel & Brysbaert, Marc. (2014). Subtlex-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, 1176–1190. Zugriff auf <http://dx.doi.org/10.1080/17470218.2013.850521>
- Vollmer, Helmut & Thürmann, Eike. (2013). Sprachbildung und Bildungssprache als Aufgabe aller Fächer der Regelschule. In M. Becker-Mrotzek, K. Schramm, E. Thürmann & H. J. Vollmer (Hrsg.), *Sprache im Fach. Sprachlichkeit und fachliches Lernen* (Bd. 3, S. 41–58). Münster: Waxmann.
- Vygotsky, Lev S. (1979). *Mind in society: The development of higher psychological processes* (2nd Aufl.). Cambridge, MA: Harvard University Press.
- Weiss, Zarah. (2017). *Using measures of linguistic complexity to assess German L2 proficiency in learner corpora under consideration of task-effects* (Masterarbeit, University of

- Tübingen, Germany). Zugriff auf <http://www.sfs.uni-tuebingen.de/~zweiss/ma-thesis/weiss2017-distr.pdf>
- Weiss, Zarah, Chen, Xiaobin & Meurers, Detmar. (2021). Using broad linguistic complexity modeling for cross-lingual readability assessment. In *Proceedings of the joint 10th workshop on NLP for computer assisted language learning* (S. 38–54). Linköping Electronic Conference Proceedings.
- Weiss, Zarah & Meurers, Detmar. (2018). Modeling the readability of German targeting adults and children: An empirically broad analysis and its cross-corpus validation. In *Proceedings of the 27th international conference on computational linguistics (COLING)*. Santa Fe, New Mexico, USA.
- Weiss, Zarah & Meurers, Detmar. (2019a). Analyzing linguistic complexity and accuracy in academic language development of German across elementary and secondary school. In *Proceedings of the 14th workshop on innovative use of NLP for building educational applications (BEA)*. Florence, Italy.
- Weiss, Zarah & Meurers, Detmar. (2019b). Broad linguistic modeling is beneficial for German L2 proficiency assessment. In A. Abel, A. Glaznieks, V. Lyding & L. Nicolas (Hrsg.), *Widening the scope of learner corpus research. selected papers from the fourth learner corpus research conference*. Louvain-La-Neuve: Presses Universitaires de Louvain.
- Weiss, Zarah & Meurers, Detmar. (2021). Analyzing the linguistic complexity of German learner language in a reading comprehension task: Using proficiency classification to investigate short answer data, cross-data generalizability, and the impact of linguistic analysis quality. *International Journal of Learner Corpus Research*, 83–130.
- Weiss, Zarah, Riemenschneider, Anja, Schröter, Pauline & Meurers, Detmar. (2019). Computationally modeling the impact of task-appropriate language complexity and accuracy on human grading of German essays. In *Proceedings of the 14th workshop on innovative use of NLP for building educational applications (BEA)*. Florence, Italy.
- Wildemann, Anja, Rank, Astrid, Hartinger, Andreas & Sutter, Sabrina. (2016). Bildungssprache im Kontext kindlicher Entwicklung: Eine Studie zur Erfassung früher bildungssprachlicher Fähigkeiten. In B. Koch-Priewe & M. Krüger-Potratz (Hrsg.), *Qualifizierung für sprachliche Bildung* (S. 65–81). Münster: Waxmann.
- Wisniewski, Katrin. (2022). Gesprochene Lernerkorpora des Deutschen. *Zeitschrift für germanistische Linguistik*, 50 (1), 1–35.
- Wolfe-Quintero, Kate, Inagaki, Shunji & Kim, Hae-Young. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (Bericht). Manoa, Hawaii: Second Language Teaching & Curriculum Center, University of Hawaii at Manoa. Zugriff auf <https://doi.org/10.2307/3587656>
- Wood, David J., Bruner, Jerome S. & Ross, Gail. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology*, 17 (2), 89–100.
- Yoon, Hyung-Jo. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System*, 66, 130–141.
- Yoon, Hyung-Jo & Polio, Charlene. (2016). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 275–301.