Institut für Schulentwicklungsforschung

Nele McElvany | Sabrina König |
Rahim Schaufelberger | Michael Becker |
Hanna Gaspard | Birgit Heppt |
Alexander Naumann (Hrsg.)

Jahrbuch der Schulentwicklung
Band 24

Bildungsprozesse und Kompetenzentwicklungen im Kontext sprachlicher und sozialer Heterogenität

BELTZ JUVENTA

Dieses Buch ist erhältlich als:
ISBN 978-3-7799-9322-3 Print
ISBN 978-3-7799-9323-0 E-Book (PDF)
DOI 10.3262/978-3-7799-9323-0

1. Auflage 2025

Verlagsgruppe Beltz
Werderstraße 10, 69469 Weinheim
service@beltz.de

Satz: Helmut Rohde, Euskirchen
Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza
Beltz Grafische Betriebe ist ein Unternehmen mit finanziellem Klimabeitrag
(ID 15985-2104-1001)
Printed in Germany

Weitere Informationen zu unseren Autor:innen und Titeln finden Sie unter: www.beltz.de

Inhalt

Grußwort der Herausgeber*innen

Der 24. Band des Jahrbuchs der Schulentwicklung umfasst in bewährter Tradition der Reihe Bestandsaufnahmen, Analysen und Forschungsbefunde und stellt zudem innovative Interventionskonzepte zu aktuellen und bedeutsamen Themen der schul- und unterrichtsbezogenen Bildungsforschung vor. Thematischer Fokus dieses Jahrbuchs sind *Bildungsprozesse und Kompetenzentwicklungen im Kontext sprachlicher und sozialer Heterogenität*. Vor dem Hintergrund gesunkener Schüler*innenkompetenzen in zentralen Bereichen wie der deutschen Sprache, des Fortbestehens sozialer Disparitäten und der Relevanz von Bildungsgerechtigkeit steht das Schulsystem vor großen Herausforderungen, die es zu adressieren gilt. Zentrale Fragen schul- und unterrichtsbezogener Bildungsforschung sind daher unter anderem, wie Merkmale sprachlicher und sozialer Heterogenität mit Unterschieden in Kompetenzentwicklungen zusammenhängen und wie Bildungsprozesse zur Reduktion heterogenitätsbedingter Bildungsungerechtigkeit optimiert werden können. Zusätzlich werden zwei Interventionskonzepte zur systematischen Förderung im Rahmen einer digitalen Schreib- bzw. Wortschatz- und Grammatikförderung vorgestellt. Die Beiträge des vorliegenden Bandes des Jahrbuchs der Schulentwicklung nähern sich der Thematik aus unterschiedlichen Perspektiven. In dem ersten Bereich liegt der Fokus neben einer umfassenden Bestandsaufnahme auf aktuellen Forschungsergebnissen:

Vogel und *Stang-Rabrig* fokussieren auf ein Maß lexikalischer Distanz zwischen der Herkunftssprache multilingualer Viertklässler*innen und deren Unterrichtssprache Deutsch. Unter Berücksichtigung des Alters, genereller kognitiver Fähigkeiten und dem sozioökonomischen Status werden Zusammenhänge zwischen der lexikalischen Distanz und der Lesekompetenz sowie zwischen der lexikalischen Distanz und dem Wortschatz regressionsanalytisch untersucht.

Heppt, *Löfflad*, *Henschel*, *Gabler*, *Hardy* und *Meurers* erforschen das Potenzial computerlinguistischer Analyseverfahren zur Evaluation der Effekte von Lehrkräftefortbildungen. Es wird untersucht, ob Lehrkräfte, die für die Umsetzung fachintegrierter Sprachbildung im Sachunterricht der Grundschule fortgebildet wurden, ihren Schüler*innen einen variableren und anspruchsvolleren Sprachinput anboten als Lehrkräfte der Kontrollgruppe.

Kosubek, *Gaspard*, *Schröter* und *Lauermann* befassen sich in ihrem Beitrag mit Zusammenhängen zwischen dem Unterrichtsenthusiasmus von Lehrkräften und der intrinsischen Motivation von Lernenden. Die Zusammenhänge werden für sprachlich heterogene Gruppen von Lernenden in Vorbereitungsklassen an Regelschulen der Sekundarstufe I sowie in Integrationskursen für Erwachsene analysiert.

Kleinkorres, *Ludewig* und *Lorenz* gehen der Frage nach, inwiefern Gymnasialempfehlungen von Lehrkräften vom familiären Hintergrund von Schüler*innen abhängig sind. Unter Berücksichtigung leistungsbezogener Kriterien für den Übergang am Ende der vierten Klassenstufe wie Zeugnisnoten und Arbeitsverhalten werden in den Analysen der Zuwanderungshintergrund und der sozioökonomische Status fokussiert.

Molitor, *Fabian* und *Stang-Rabrig* untersuchen die Bedingungen eines erfolgreichen Gymnasialbesuchs von Schüler*innen mit und ohne Gymnasialempfehlung am Ende der Grundschulzeit. Dabei prüfen sie anhand von Strukturgleichungsmodellen, ob sich diese beiden Gruppen systematisch bezüglich des sozioökonomischen Status, der Opportunitätskosten und der Motivation unterscheiden.

Grecu, *Neuhaus*, *Yotyodying* und *McElvany* befassen sich mit Schulentfremdung und nehmen dabei die Heterogenität der Einstellungen von Siebtklässler*innen sowie von Zehntklässler*innen zur Schule in den Blick. Sie stellen latente Schulentfremdungsprofile vor und analysieren, wie diese mit Merkmalen der Schüler*innen wie dem Geschlecht und einem Migrationshintergrund zusammenhängen.

Becks, *Czaja* und *Klein* beschäftigen sich in ihrem Beitrag mit kulturresponsivem Handeln als Bedingung erfolgreicher Bildungsprozesse von Schüler*innen aus marginalisierten Herkunftsmilieus. Auf Basis eines systematischen Literaturreviews bieten sie einen Überblick über die Forschung zu kulturresponsivem Schulleitungshandeln und diskutieren Adaptionsmöglichkeiten für den deutschsprachigen Kontext.

In einem weiteren Bereich werden zwei aktuelle Interventionskonzepte vorgestellt:

Gade und *Busse* stellen eine zielgruppenspezifische Intervention vor, die im Ganztag an Grundschulen mit schreibschwachen Schüler*innen in der dritten und vierten Klassenstufe durchgeführt wurde, um prozessorientiertes Schreiben und kooperatives Lernen zu fördern. Sie diskutieren Herausforderungen und Potenziale einer digitalen Schreibförderung für Schulentwicklungsprozesse.

Dargiewicz, *Wehrhöfer*, *Ohle-Peters*, *Lauermann* und *McElvany* präsentieren die Konzeption eines Interventionsprojektes. Sie beschreiben, wie

im Rahmen des Projektes eine digitale Sprachförderung in den Bereichen Wortschatz und Grammatik für den Unterricht in der ersten Klassenstufe entwickelt wird, um schulisch relevanten Wortschatz zu fördern sowie das Genus- und Kasussystem zu trainieren.

Die Beiträge des 24. Bandes des Jahrbuchs der Schulentwicklung geben einen breiten Überblick über aktuelle schul- und unterrichtsbezogene Themen im Kontext sprachlicher und sozialer Heterogenität. Sie leisten damit einen wertvollen Beitrag für die Bildungsforschung.

Besonderer Dank gilt allen Gutachter*innen, die die interne und externe Begutachtung übernommen haben und mit ihrer Expertise zur Sicherung der Qualität der Beiträge des vorliegenden Jahrbuchs beigetragen haben. Neben den Kolleg*innen am IFS danken wir Birgit Becker, Michaela Gläser-Zikuda, Markus Hasselhorn, Jörg Jost, Harm Kuper, Katrin Lintorf, Elena Makarova, Camilla Rjosk, Nadine Spörer und Tobias C. Stubbe herzlich für ihre Mitwirkung.

Dortmund, im Juli 2025
Nele McElvany
Sabrina König
Rahim Schaufelberger
Michael Becker
Hanna Gaspard
Birgit Heppt
Alexander Naumann

Herausgeber*innen

Teil A:
Bestandsaufnahmen, Analysen und Forschungsbefunde

Reading Competence and Vocabulary of Students From Diverse Language Backgrounds: Employing a Lexical Distance Measure

Sebastian Nicolas Thomas Vogel & Justine Stang-Rabrig

Zusammenfassung

Mit Blick auf die wachsende sprachliche Heterogenität fokussiert dieser Beitrag auf ein Maß lexikalischer Distanz zwischen der Herkunftssprache von multilingualen Lernenden und der Zielsprache Deutsch, welche Verkehrs- und Unterrichtssprache zugleich ist. Lexikalische Distanz wird als Prädiktor von Lesekompetenz und Wortschatz, zentrale Indikatoren des Spracherwerbs im Deutschen, genutzt. Darüber hinaus wird das Alter zugewanderter Lernender bei der Ankunft in Deutschland als möglicher Moderator von Effekten lexikalischer Distanz auf den Spracherwerb im Deutschen berücksichtigt. Anhand einer sprachlich heterogenen Stichprobe von $N = 193$ multilingualen Vierklässler*innen mit einem hohen Anteil zugewanderter Lernender zeigten Regressionsanalysen signifikant negative Zusammenhänge der lexikalischen Distanz mit Lesekompetenz, aber nicht Wortschatz, wenn für die Sprache, die die Lernenden primär zuhause sprechen (Deutsch vs. Herkunftssprache), die generellen kognitiven Fähigkeiten und den sozioökonomischen Status kontrolliert wurde. Das Alter zugewanderter Lernender bei der Ankunft war weder als Moderator der lexikalischen Distanz noch direkt mit Lesekompetenz oder Wortschatz verbunden. Die Ergebnisse konnten also zeigen, dass lexikalische Distanz unabhängig von Zuwanderungserfahrungen der Lernenden negativ mit der Entwicklung von Lesekompetenz im Deutschen assoziiert war, während es keinen Zusammenhang mit der Herausbildung des Wortschatzes über den Einfluss der primär gesprochenen Sprache und der allgemeinen kognitiven Fähigkeiten hinaus gab.

Schlagworte: Lexikalische Distanz, sprachliche Heterogenität, Lesekompetenz, Wortschatz, zugewanderte Lernende, Grundschule

Abstract

In light of growing linguistic heterogeneity, this contribution focuses on a measure of lexical distance between multilingual students' heritage language and German as a predictor of reading competence and vocabulary, indicators of the acquisition of language skills in the target language German, the common language and language of instruction in German schools. Additionally, first-generation immigrant students' age at arrival in Germany is considered as a potential moderator of lexical distance effects for language acquisition in the target language. Using a sample of $N = 193$ multilingual fourth-grade students with a large share of first-generation immigrant students and high linguistic diversity, regression analyses showed that lexical distance was significantly negatively related to reading competence, but not vocabulary when considering whether students primarily spoke German or the heritage language at home, general cognitive abilities, and socioeconomic status as important control variables. Including immigrant students' age at arrival and its interaction with lexical distance to check moderation effects revealed no additional significant associations. Therefore, independent of immigration experiences, lexical distance was detrimental to developing reading comprehension in German but did not affect vocabulary learning beyond the effects of general cognitive abilities and primarily speaking the heritage language rather than German at home.

Keywords: Lexical distance, linguistic diversity, reading competence, vocabulary, immigrant students, primary school

1. Introduction

Global migration movements cause classrooms to become increasingly heterogeneous in terms of cultures and languages. This especially applies to Germany as well, a popular destination for immigration, where 21.0% of fourth grade students now primarily speak another language than the common language German at home (Stubbe et al., 2023). Studies have shown that these students are often disadvantaged regarding the development of reading competence and acquisition of vocabulary in German (e.g., Henschel et al., 2022; Novita et al., 2022). However, most studies compared students based on

dichotomized measures of primarily speaking German at home or not, but did not consider the large heterogeneity of languages spoken by the students who do not exclusively speak German (e.g., Stubbe et al., 2023). Although such approaches have advantages, they artificially reduce variance in the group of multilingual students, despite theoretical frameworks suggesting that specific aspects of language learning might be helped or hindered when students' first language is more or less similar to the target language (e.g., Chung et al., 2019). Furthermore, another crucial individual aspect when learning a new language is age, which might be of particular importance for immigrant students learning the common language in their new country of residence as it directly relates to their exposure to the new language. Nevertheless, the role of immigrant students' age at arrival has rarely been studied in regard of its potential moderating role for language learning when considering the linguistic variety of heritage languages (e.g., Schepens et al., 2013a). To address these research desiderates, we employed a measurement of lexical distance to assess the heterogeneity in multilingual students' backgrounds and its role for reading competence and vocabulary in the common language German as two indicators of language learning. We investigated these questions in a multilingual sample surveyed in the fourth year of primary school which is not only an important time in the development of central language competences, especially reading (Chall, 1983), but a central point for shaping educational trajectories in the German school system as well (e.g., Maaz & Nagy, 2010). Finally, we also considered the role of immigrant students' age at arrival in Germany as a potential moderator of linguistic distance's effects on language learning.

2. Theory

2.1. Immigrant-Origin Students and Language Diversity in German Schools

As one of the most sought-out destinations for immigration worldwide (International Organization for Migration [IOM], 2021), society and by extension schools in Germany are continuously becoming more diverse. The share of immigrant-origin students in Grade 4 throughout Germany was 38.3% in 2021, marking an increase of more than 50% in ten years, with the proportion of first-generation immigrant students growing disproportionately, from making up less than one tenth to now representing more than one quarter of all immigrant-origin students (Henschel et al., 2022). Since immigrant fam-

ilies move to Germany from a large variety of different regions and countries of origin (Statistical Federal Office, 2022) and the vast majority of immigrant-origin students grows up speaking at least one heritage language next to German at home (Henschel et al., 2022), the language diversity among German students is at an unprecedented high, especially in primary schools.

2.2. Language-Related Outcomes in Primary School

Theoretical models regarding the stages of reading development (Chall, 1983) and central developmental tasks (McCormick et al., 2011) as well as international educational guidelines (Organization for Economic Co-operation and Development [OECD] et al., 2015) emphasize the importance of students developing adequate reading competence during the first years of primary school that allows them to use reading as a means to learn new information in later years of education and life. For this shift from learning to read to learning through reading (see Chall, 1983) to happen by the end of primary school, the development of good *reading competence*, describing students' ability to engage with texts in a way that allows them to generate meaning (McElvany et al., 2009), is crucial. Other closely related measures like the formation of an adequate *vocabulary*, which allows students to understand the lexical meaning of words and the concepts related to them (Aarnoutse et al., 2001), can also be understood to be indicators of achieving good language skills.

For multilingual students, who in Germany usually are of immigrant origin (Henschel et al., 2022), acquiring language skills in the country's primary language is of special importance as it represents a core acculturative task for these students (Suárez-Orozco et al., 2018). Accordingly, reading competence and vocabulary in the language of instruction have been understood and studied frequently as important indicators of language proficiency of multilingual and language-minority students. Studies focusing on these constructs tend to find disadvantages of language-minority students compared to students speaking German as their primary language in terms of both reading competence (e.g., Segerer et al., 2021; Seuring et al., 2020; Stubbe et al., 2023) and vocabulary (e.g., Kigel et al., 2015; Marx et al., 2015; Novita et al., 2022). However, empirically investigating language-minority students inevitably means that the large heterogeneity of language backgrounds has to be operationalized in a way that still allows for meaningful statistical analyses. In consequence, studies frequently do not distinguish between different language backgrounds at all (e.g., Kigel et al., 2015; Stubbe et al., 2023) or only include large language-minority groups on their own (in the German context, often

Turkish-language background students, e.g., Marx et al., 2015; Segerer et al., 2021) and subsume all other language-minority students in a single group. While the advantages of this method are apparent, it makes a compromise of sacrificing variance in students' language backgrounds to enable statistical comparison of the groups. Crucially, analyzing language effects in this way implicitly assumes that differences in multilingual students' reading competence or vocabulary when compared to other students will only be found if German is not their primary language, and that this is the only explanatory factor. That is not to say that meaningful results cannot still be attained with this method; however, it does by design largely or entirely negate the role of individual characteristics of multilingual students' heritage language for language learning and often excludes multilingual students who speak German as the primary language.

2.3. Language Diversity and Language Learning

Theories of cross-language transfer have recognized the role of the similarity of the target language that is to be learned and the language – or languages – that speakers have already achieved some degree of fluency in for language learning. This aspect of language similarity, reversely framed as language distance (i.e., the dissimilarity of two languages' linguistic features), is identified as one key factor of the interactive language transfer framework by Chung et al. (2019) which combines features of four influential cross-language transfer frameworks. While some skills, like phonological awareness, appear largely unaffected by language distance, the authors emphasize that shared linguistic features can help with the transfer of others, for example morphological and cognate awareness as well as orthographic processing to the second language. Similarly, the established model of Chiswick and Miller (1995, 2007), which focuses on the economics of language learning among immigrants and other language minorities, includes linguistic distance as part of the efficiency of language acquisition and one of the core determinants of language proficiency in second language learning.

Different studies have investigated the role of linguistic distance for language learning, and more specifically reading competence and vocabulary acquisition as central indicators of language proficiency. Common ways to operationalize linguistic distance measures utilize either morphological features, which are a result of the languages' systems of forming words and reflected in their internal structure (Aronoff & Fudeman, 2023), or lexical distance between languages, describing dissimilarities based on languages' basic

vocabularies and shared lexical forms (Schepens et al., 2016). Between these two attempts, lexical distance appears to be better suited to explain differences in language proficiency (e.g., Schepens et al., 2013b, 2016). Therefore, we will focus on lexical distance as an indicator for linguistic distance in this study.

Investigating lexical distance, a series of studies using large databases of second language learners' Dutch proficiency tests has shown negative associations with language proficiency for learners with Indo-European (Schepens et al., 2013a) as well as non-Indo-European first languages (Schepens et al., 2013b). Additionally, lexical distance of both the first and second language was found to explain differences in language proficiency when Dutch was learned as a third language (Schepens et al., 2016). Similarly, using a German dataset, a negative association of lexical distance with German proficiency could be shown in a sample of immigrants over the age of 17, even when controlling for multiple other relevant constructs (Isphording & Otten, 2011). Investigating reading proficiency specifically rather than language proficiency more broadly, Borgonovi and Ferrara (2020) showed a significant negative association with lexical distance in adolescent language-minority students even when multiple control variables were considered. Finally, a study focused on Dutch adolescents learning English as a second language showed that vocabulary gains were higher when the lexical distance of English target words and their Dutch counterparts was small (Mulder et al., 2019), implying that building a vocabulary should be relatively harder the more distant the target language is to other languages the learner is already speaking.

2.4. Language Learning and Age of Immigration

Another aspect that theoretical models deem crucial for learning a new language is exposure to that language (e.g., Chiswick & Miller, 2007; Esser, 2006). This factor can be considered especially important for first-generation immigrants. Due to the strong prevalence of German as a language in public spaces in Germany, native-born multilingual speakers will usually have some exposure to German in their every-day life even if it is not the primary language spoken in their family. In contrast, first-generation immigrants will usually have little to no exposure to German before immigration, as indicated by low average self-reported German skills upon immigration among different immigrant groups (e.g., Kristen & Seuring, 2021) and the low number of immigrants from German-speaking countries (Statistical Federal Office, 2022). Therefore, meaningful exposure to the German language often begins

upon the arrival in Germany, especially for young immigrants who lack the resources to begin learning German on their own prior to migration.

The role of lexical distance for language learning can arguably be more important for students who have less exposure to the language they are striving to learn. If language-learners have had sufficient opportunities to gain a core understanding of the language, lexical distance to their heritage language might be less relevant to learning the new language as they can draw from the knowledge about said language they already have developed through exposure. However, if they have had little exposure to the new language, they can only relate it to their heritage language in their learning process, meaning that students with a more dissimilar heritage language would be additionally disadvantaged in such scenarios. Therefore, theoretical models of second language learning include the efficiency of language learning, which is influenced by linguistic distance, and the exposure to the target language not only as relevant factors in their own right, but also their interaction as another important contributor (e.g., Esser, 2006). Extant research appears to back these theoretical considerations: Schepens et al. (2013a) found a significant negative interaction of lexical distance and age at arrival in the host country, implying that lexical distance hindered language acquisition more the older subjects were upon their immigration. A similar interaction was present among adolescent immigrant students who had participated in the PISA study between 2006 and 2015, where a positive interaction of lexical distance and early arrival (i.e., before the age of 12) emerged (Borgonovi & Ferrara, 2020). This again implied a stronger negative impact of lexical distance for language learning when students' later arrival in the host country had largely prevented earlier exposure to the new language.

2.5. Aims of the Study and Research Questions

The reported theoretical considerations and extant research clearly show that the individual features of students' heritage language, indicated by lexical distance, have an influence on the acquisition of a second language, and reading competence and vocabulary in that language specifically. The role of lexical distance might be even more pronounced for first-generation immigrant students with little exposure to the target language before migration. However, previous studies were focused on either adolescent students (Borgonovi & Ferrara, 2020; Mulder et al., 2019) or largely adult immigrants (Isphording & Otten, 2011; Schepens et al., 2013a, 2013b, 2016) whereas students in primary school have not been focused, despite its status as a critical phase in language

learning and the development of reading competence. Additionally, to our knowledge no study has attempted to investigate lexical distance in the context of German primary schools specifically, where students speaking a large variety of first languages face the task of becoming proficient in German. Therefore, we investigated the following research questions:

Research Question 1a: How is lexical distance of multilingual students' heritage language to German related to their reading competence and vocabulary in German?

Hypothesis 1a: The lexical distance of multilingual students' heritage language to German is significantly negatively related to students' reading competence and vocabulary in German.

Research Question 1b: Do the associations identified in Research Question 1a remain when controlling for which language (German vs. heritage language) students primarily speak at home?

Hypothesis 1b: The negative association between lexical distance of students' heritage language to German and students' reading competence and vocabulary in German is lowered but remains significant when controlling for which language students primarily speak at home.

Research Question 1c: Do the relations identified in Research Question 1b change when controlling for other significant contributors to students' reading competence and vocabulary (general cognitive abilities, socioeconomic status)?

Hypothesis 1c: The negative association between lexical distance of students' heritage language to German and students' reading competence and vocabulary in German remains stable when additionally controlling for general cognitive abilities and socioeconomic status.

Research Question 2: Is the relation of lexical distance with reading competence and vocabulary moderated by first-generation immigrant students' age at arrival in Germany?

Hypothesis 2: Age at arrival in Germany significantly negatively moderates the association of lexical distance with reading competence and vocabulary.

3. Methods

3.1. Participants and Procedure

The analysis sample consisted of $N = 193$ multilingual fourth-grade students from 47 classes in the Ruhr metropolitan area of North Rhine-Westphalia, Germany, after all students who exclusively spoke German at home ($n = 74$),

for whom insufficient information on the language spoken at home was available (n = 10), or whose data were unfit for analyses (n = 2) had been excluded. Students were 10.49 years old on average (SD = 0.58), and the sample was skewed towards female students (58.5% female, 39.9% male). Of the participants, 50.8% primarily spoke a heritage language and never or only sometimes German at home, whereas 49.2% mostly or almost always spoke German at home, and only sometimes a heritage language. Students reported a total of 42 different languages, detailed information on the languages is depicted in Table 1. The majority of the sample were immigrant-origin students with at least one parent born abroad (89.1%), including a comparably large amount of first-generation immigrant students born outside of Germany (n = 85), who on average were 10.63 years old (SD = 0.63) at the time of the study and had been 4.74 years old (SD = 2.50) upon their arrival in Germany. The share of students primarily speaking a heritage language at home was significantly higher in first-generation immigrant students (74.1%) than in the native-born student subsample (32.4%), $\chi^2(1) = 31.5$, $p < .001$. Data was collected towards the end of the 2021/2022 school year in the context of the MERCUR-funded study *School Integration of Newly Immigrated Children (SIGN)*, a joint research project by the University of Duisburg-Essen, Ruhr University Bochum, and the Center for Research on Education and School Development at TU Dortmund University. During the study, which took part in the classroom context or small groups during regular instruction time, students filled in well-established, standardized tests and self-report questionnaires in paper-pencil form. The study was conducted by two or three trained research administrators per class, depending on class size, and took 90 minutes. Participation in the study required parents' consent and was voluntary, with students who did not participate being given alternative tasks by their teachers. The study met ethical standards and was approved by the ethics committee of the Department of Psychology, University of Duisburg-Essen.

Table 1: Number of Times Languages and Language Families Other Than German Were Reported as Being Spoken at Home in the Sample

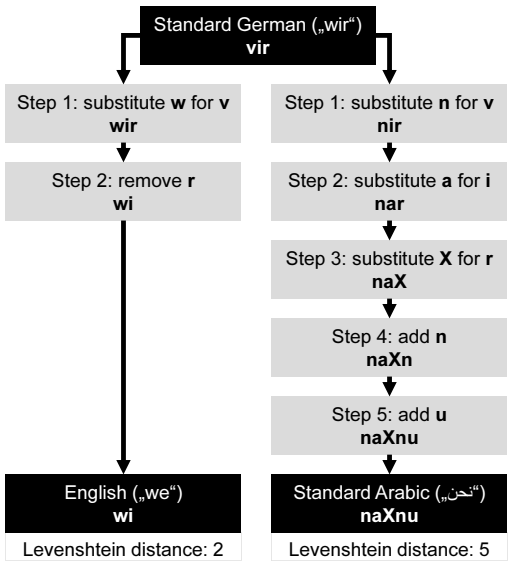
Language	*n*	%
Afro-Asiatic languages	**59**	**30.6**
Arabic (incl. different dialects)	53	27.5
Other Afro-Asiatic languages	6	3.1
Dravidian languages	**2**	**1.0**
Indo-European languages	**123**	**63.7**
English	28	14.5
Kurdish	14	7.3
Polish	17	8.8
Romanian	7	3.6
Russian	6	3.1
Spanish	8	4.1
Indo-Aryan languages	8	4.1
Italo-Western Romance languages	9	4.6
Persian languages	7	3.6
South Slavic languages	10	5.2
Other Indo-European languages	9	4.6
Niger-Congo languages	**12**	**6.2**
Volta-Congo languages	10	5.2
Other Niger-Congo languages	2	1.0
Turkic languages	**35**	**18.1**
Turkish	34	17.6
Other Turkic languages	1	0.5

Note. Students could name multiple languages spoken at home. Only languages reported by more than five children are depicted separately. All other languages were grouped into the next higher subgroup (indicated by *italics*) following classification by Eberhard et al. (2024) until a group comprising more than five students was reached.

3.2. Measures

An overview of descriptive information for all measures included in the study is given in Table 2. To assess *lexical distance* between students' heritage languages and German, we used the Automated Similarity Judgment Program (ASJP; Wichmann et al., 2022), which encompasses information on 5590 distinct languages following the ISO 639-3 classification (International Organization for Standardization, 2023). ASJP makes use of wordlists of these different languages containing 40 words (Holman et al., 2008), a subset of a 100-word list developed by Swadesh (1955) for lexicostatistic dating. These words are transcribed into a standardized orthography featuring seven vowel and 34 consonant symbols (Holman et al., 2008). With these wordlists, a Levenshtein distance is calculated (Levenshtein, 1966), an indicator of the smallest possible number of substitutions, additions, or deletions of symbols necessary to transform the target word in one language into the word with the same meaning in the other (for an example, see Figure 1).

Figure 1: Example for the Calculation of a Levenshtein Distance for the Target Word "We" for Standard German With English and Standard Arabic, Respectively



Note. ASJP standardized orthography transcriptions of the words that are used in the calculation of the Levenshtein distance are printed bold.

The distance index is then modified in two steps to take into account coincidental similarities, for example introduced by a general overlap of the phoneme inventories of two languages, and better reflect actual relation between languages (for a detailed description, see Wichmann et al., 2010). We calculated the lexical distance to German[1] for each language reported by students, then transformed values so they would fall on a scale ranging from 1 (lowest distance, in our sample: English) to 10 (highest lexical distance to German, in our sample: Fula). Students who reported speaking multiple languages other than German at home were appointed the value of the language with the smallest lexical distance to German. The distribution of lexical distance in the sample, differentiated by the language students primarily spoke at home, is depicted in Figure 2. The language families depicted in Table 1 are roughly reflected in the distribution of linguistic distance: While Germanic languages, mostly English in our sample, have a value equal or close to 1, other Indo-European languages tend to fall in the range between 6 and 9,

1 ASJP offers two wordlists for Standard German, the results reported here use wordlist "Standard German". We also replicated analyses using wordlist "Standard German 2", and results remained stable.

Effectiveness of Professional Development for Language-Supportive Teaching: Insights From Employing Computational Linguistic Analysis Methods

Birgit Heppt, Denise Löfflad, Sofie Henschel, Katrin Gabler, Ilonca Hardy & Detmar Meurers

Zusammenfassung

Die vorliegende Studie untersucht das Potenzial computerlinguistischer Analyseverfahren zur Evaluation der Effekte von Lehrkräftefortbildungen. Entsprechende Methoden ermöglichen die automatisierte Analyse einer Vielzahl sprachlicher Merkmale und dürften daher eine vertiefte Analyse der mündlichen Unterrichtssprache von Lehrkräften erleichtern. Auf Grundlage eines positiv evaluierten Professionalisierungsansatzes (z. B. Heppt et al., 2022) wurde untersucht, ob Lehrkräfte, die für die Umsetzung fachintegrierter Sprachbildung im Sachunterricht der Grundschule fortgebildet wurden (Interventionsgruppe [IG], $n = 9$), ihren Schüler*innen einen variableren und anspruchsvolleren Sprachinput anboten als Lehrkräfte der Kontrollgruppe (KG; $n = 19$). Anhand von Transkripten von zwei videografierten Sachunterrichtsstunden je Lehrkraft (Thema 1: Schwimmen und Sinken, Thema 2: Verdunstung und Kondensation) fanden wir nur wenige Merkmale, die für eine elaboriertere und komplexere Sprachverwendung der IG im Vergleich zur KG sprechen. Die Gruppenunterschiede waren überdies bei Thema 2 weniger stark ausgeprägt als bei Thema 1, was darauf hindeutet, dass sich die Fortbildungsteilnahme nur geringfügig auf den mündlichen Sprachgebrauch der Lehrkräfte während des Sachunterrichts auswirkte.

Schlagworte: Lehrkräftefortbildung, Sprachbildung, Grundschule, Sachunterricht, sprachliche Komplexität, computerlinguistische Analyseverfahren

Abstract

The present study aims at exploring the potential of using computational linguistic analysis methods for examining professional development (PD) effects. As respective methods allow for an automatic analysis of a wide range of linguistic features, they may facilitate an in-depth analysis of teachers' oral language use in classroom discourse. Building on a positively evaluated PD approach (e.g., Heppt et al., 2022), we investigated whether in-service teachers who were trained for language-supportive teaching in elementary school science classes (intervention group [IG], $n = 9$) provide their students with a more varied and more elaborate language input than teachers from the control group (CG; $n = 19$). Based on transcripts of video recordings of two science lessons per teacher (Topic 1: floating and sinking, Topic 2: evaporation and condensation), we found only a small number of features that point to more sophisticated and complex language use in the IG than in the CG. Moreover, group differences were less pronounced for Topic 2 than for Topic 1, suggesting that PD participation had only small effects on teachers' oral language use during science instruction.

Keywords: Teacher professional development, language support, elementary school, science education, linguistic complexity, computational linguistic analysis methods

1. Introduction

Language-supportive teaching that fosters both domain-specific learning and language development is increasingly seen as a general teaching principle across subjects (e.g., Becker-Mrotzek & Woerfel, 2020; Prediger & Hardy, 2023). Although considered beneficial for all students, such classroom instruction is particularly aimed at students who are at risk of falling behind. Among them are students with an immigrant background, often growing up as dual language learners, who have repeatedly been shown to lag behind their peers without immigrant background in academic achievement and (academic) language proficiency (Henschel et al., 2022; Ludewig et al., 2022). Similarly, students from families with low socioeconomic status (SES) consistently perform, on average, below their peers from high-SES families (e.g., Sachse et al.,

2022). Both student groups are likely to have only limited opportunities for acquiring the register of schooling within their families. Children's language development depends on the amount and quality of language input at home, which, in turn, varies with family SES. Children from high-SES families, thus, typically outperform their counterparts from low-SES families in their mastery of academic language, helping them to benefit from classroom discourse, accomplish written assignments, or understand specialized texts (e.g., Volodina et al., 2021). Students with an immigrant background face similar obstacles as they often grow up in families with low SES (e.g., Henschel et al., 2022). Moreover, many of them have limited access to the language of instruction within their homes, thus further hampering their mastery of the language of schooling (for an overview, see Heppt & Schröter, 2023). With the aim of adapting to these students' needs, language-supportive teaching is considered a potential remedy for tackling educational inequalities.

The need for integrating language support into regular classroom teaching is widely accepted, resulting in an expansion of course offerings on language support and second language acquisition in university teacher training in Germany (Paetsch & Heppt, 2021). While an increasing number of German elementary school teachers participates in such courses during university teacher training, many of them are still ill-prepared for providing subject-integrated language support in their daily classroom instruction (Henschel & Heppt, 2024). This highlights the importance of effective professional development (PD).

In investigating the effectiveness of PD for language-supportive teaching, research frequently focuses on teachers' classroom practice (e.g., Gabler et al., 2024; Heppt et al., 2022; van Dijk et al., 2019). Based on laborious coding of selected language-support strategies (e.g., use of language-supportive questions) or highly inferential ratings of the overall quality of language support, studies have pointed to the general effectiveness of PD for developing teachers' classroom practice (for a meta-analysis, see Kalinowski et al., 2020).

With the present study, we aim at exploring the potential of using computational linguistic analysis methods for examining PD effects. As respective methods allow for an automatic analysis of a wide range of linguistic features, they may facilitate an in-depth analysis of teachers' oral language use in classroom discourse (cf. Weiss et al., 2022). Building on a PD approach that has been positively evaluated regarding participants' knowledge on language support (Heppt et al., 2022) and their use of specific language-support strategies in science instruction (e.g., the use of language-supportive questions; Gabler et al., 2024), this study focuses on teachers' oral language input. Specifically, we investigate whether, upon completing the PD, German elementary school

teachers use more stimulating and sophisticated oral language in their science instruction than teachers who did not take part in the PD.

2. Theoretical and Empirical Background

2.1. Language-Supportive Classroom Instruction

Effective language-supportive classroom instruction helps students attain domain-specific learning goals, while developing the necessary academic language skills. In elementary school science classes, for instance, experiments on the floating and sinking of objects are clearly aimed at developing students' conceptual knowledge on phenomena like water displacement or water pressure. Yet, in order to construct this type of knowledge in a co-constructive process, students need to be able to formulate and justify assumptions and to describe and explain their observations (e.g., Vorholzer & Aufschnaiter, 2019). These language functions form a core part of the academic language register (e.g., Bailey et al., 2007; Prediger & Hardy, 2023). Performing language functions such as "hypothesizing" or "justifying" requires the precise understanding of their respective meaning. In addition, students need knowledge of adequate linguistic structures (e.g., the correct use of causal connectives such as "therefore" or "due to") and domain-specific academic vocabulary (e.g., "to displace", "wax", "Styrofoam"). As an important prerequisite for learning, language functions need to be systematically developed in classroom instruction, along with the underlying lexical and syntactical skills.

This can best be achieved by cognitively activating instruction that engages students in higher-order thinking and connects concepts with activities (e.g., by conducting experiments), as such learning environments typically offer multiple opportunities for using language in meaningful contexts (e.g., Bravo & Cervetti, 2014). In using language functions and expanding their vocabulary knowledge, students should be assisted through language-support strategies (e.g., open-ended questions, rich and elaborate language input, linguistic feedback that adequately expands the students' utterances; e.g., Gabler et al., 2020; Heppt et al., 2022; Mahan, 2020). The linguistic scaffolding approach (Gibbons, 2002), which builds on the theory of social learning (cf. Wood et al., 1976), considers these linguistic aids as a scaffold that helps students master linguistically demanding tasks. These aids are adapted to students' linguistic needs and gradually reduced as the students increasingly gain proficiency in academic language. The linguistic scaffolding approach has been proven effective in promoting students' domain-specific knowledge with no pronounced

differences across student groups (e.g., multilingual and monolingual learners; Prediger & Neugebauer, 2021; Prediger & Wessel, 2017). Research on its effectiveness for (academic) language development has mostly been conducted in the United States (US) with a focus on students with limited language proficiency (English Language Learners; ELLs). Overall, results indicate that integrating inquiry-based science instruction with the adaptive use of a range of language-support strategies benefits ELLs' academic language proficiency (e.g., Bravo & Cervetti, 2014; Llosa et al., 2016).

2.2. Language Input and its Associations with Content- and Language Learning

The quality and amount of oral language input are among the fundamental drivers of children's language development. Since the seminal study by Hart and Risely (1995) on the huge socioeconomic differences in the number of child-directed words, amounting to a "30-million-word gap" for children from high vs. low-SES families by the age of four, numerous studies have highlighted the relation between caregivers' language input and their child's language proficiency. In line with these findings, the amount and quality of a (preschool) teacher's oral language use are also considered an essential language-support strategy in instructional settings (e.g., Gabler et al., 2020; Kane et al., 2023).

With the aim of acting as language role models, this language-support strategy requires (preschool) teachers to provide frequent, rich, and elaborate language input in instruction. This can be achieved by using thinking-aloud techniques or by mapping one's own or students' actions with language (e.g., "I have a wooden dice in my left hand and a metal dice in my right hand. Now, I put the wooden dice into the water basin."). In doing so, teachers should try and use important general and domain-specific vocabulary, ideally in multiple contexts and by contrasting them with other words (e.g., "Did we hypothesize this or did we establish this? We established this in the end. So this is for sure."), contributing to students' enhanced and differentiated vocabulary knowledge.

Prior research on teacher's oral language use has been conducted in preschool settings and, to a lesser extent, in elementary school. Overall, this research showed that preschool teachers' use of elaborate language can increase students' domain-specific learning and language development (e.g., Kane et al., 2023; Studhalter et al., 2021). However, (preschool) teachers tend to use this high-quality language input rather infrequently (see however,

Weiss et al., 2022). Studhalter et al. (2021), for instance, found that preschool teachers' use of domain-specific vocabulary (e.g., "wax", "clay", "iron") during a 4-week learning unit on the topic "floating and sinking" significantly predicted children's conceptual learning gains. Focusing on the interplay between the quality of preschool teachers' talk and students' language development, Dickinson and Porche (2011) conducted a longitudinal study from preschool to fourth grade. Among the indicators used for evaluating the quality of preschool teachers' oral language were sophisticated vocabulary (i.e., low-frequency words) and utterances aimed at focusing children's attention, correcting or expanding their oral expressions. The authors found that each of these indicators of preschool teachers' language quality contributed to students' reading skills in Grade 4 (Dickinson & Porche, 2011).

Despite its pivotal role for student learning, children seem to receive relatively low amounts of high-quality language input during regular classroom teaching in elementary school. Based on videotaped classroom observations of five upper elementary classrooms in the US, Ernst-Slavit and Mason (2011) found that less than 12% of teachers' oral language input across classrooms and subjects can be classified as "academic language", as reflected in vocabulary (e.g., general and domain-specific vocabulary), grammar (e.g., syntactically long and complex sentences with clause connectives) and discourse (e.g., factual and information-dense style). In a similar vein, a Dutch study aimed at assessing elementary school teachers' use of academic language in whole-classroom discourse in mathematics classrooms in Grades 1 and 2 (Dokter et al., 2017). Drawing on transcripts of two eight-minute sequences of classroom instruction per teacher, the authors found that all teachers used math-specific language to some extent. However, the overall lexical and grammatical complexity of teacher talk was rather low. In sum, prior research suggests that teachers do not deliberately use their oral language input in classroom discourse as a means for modeling students' (academic) language development, pointing to the need for effective teacher PD.

2.3. Effectiveness of Teacher PD for Language-Supportive Teaching

Research on teacher PD in general (e.g., Darling-Hammond et al., 2009; Lipowsky & Rzejak, 2015) as well as the emerging literature on PD for language-supportive teaching (for an overview, see Kalinowski et al., 2020) have identified key characteristics of effective teacher PD. Importantly, teacher PD that helps teachers gain knowledge and skills for integrating language-support strategies into regular classroom teaching combines phases of input with

opportunities for actively using the newly acquired knowledge (e.g., in role plays or in classroom instruction). During implementation phases, teachers should receive feedback and have the chance to reflect upon their experiences, for instance by analyzing and discussing video-recordings of their own classroom teaching (Piwowar et al., 2013; van Dijk et al., 2019).

PD programs that consider these core principles have, indeed, been proven effective in developing teachers' knowledge and skills for subject-integrated language support (e.g., Babinski et al., 2018; van Dijk et al., 2019). A meta-analysis incorporating ten studies, showed small, albeit statistically non-significant gains in teachers' self-efficacy and self-assessed knowledge regarding language-supportive teaching, whereas larger PD effects occurred for teachers' language-supportive classroom behavior (Kalinowski et al., 2020). Most of the studies included in the meta-analysis were conducted in the US with a focus on ELLs and only few studies used (quasi-)experimental designs with intervention group (IG) and control group (CG).

A couple of more recent (quasi-)experimental studies from Germany and the Netherlands aimed at teachers in mainstream classrooms, helping them to provide language-supportive science instruction for *all* students (Gabler et al., 2020; Henrichs & Leseman, 2014; van Dijk et al., 2019). In one such PD program based on the linguistic scaffolding approach (Gabler et al., 2020), teachers were familiarized with and actively used core language-support strategies in group work and classroom teaching. The PD covered the following language-support strategies: (1) language modeling by providing elaborate and targeted language input, (2) asking language-stimulating questions, (3) giving language-supportive feedback (e.g., by elaborating on or rephrasing students' answers), and (4) shifting students' attention to important vocabulary and sentence structures (e.g., by using visual aids). This PD program, which also forms the basis of the present study, has been proven effective in advancing teachers' knowledge on subject-integrated language support (Heppt et al., 2022). Moreover, after dealing intensively with a science curriculum on "floating and sinking" and the implementation of possible language scaffolds, trained teachers showed a higher overall quality and amount of language-supportive behavior and provided better learning-related feedback in classroom teaching (Heppt et al., 2022). They also used language-stimulating questions more frequently than their counterparts who did not participate in the PD (Gabler et al., 2024). Regarding the overall quality of instructional support, including language-supportive teaching, no group differences were observed when participants were required to transfer their newly acquired knowledge and skills for language-supportive teaching to a different science curriculum (i.e., on "evaporation and condensation"). Interestingly, though, instructional

quality of both IG- and CG-teachers was higher when teaching this second topic, possibly reflecting that it was easier to implement with regard to both conceptual demands and classroom organization. At the same time, it should be noted that the two science topics require the use of different science vocabulary and focus on partly different language functions and structures. Hence, the topics may also differ in their potential for modeling sophisticated vocabulary, syntactic structures, and language functions. However, group differences in terms of the amount and quality of teachers' language input have not yet been investigated. Consequently, it remains an open question whether teachers' oral language use differs across topics.

2.4. The Potential of Using Computational Linguistic Analysis Methods for Investigating Teachers' Oral Language Input

In general, research on PD for language-supportive teaching has rarely considered the amount and linguistic sophistication (e.g., the use of low-frequency vocabulary) of teachers' oral language in classroom discourse as an outcome (for exceptions, see Henrichs & Leseman, 2014; van Dijk et al., 2019). At least in part, this is probably due to the enormous amount of time and effort required to assess the frequency of selected features in oral language. In their study on the effectiveness of a short intervention on academic language use in early science instruction, Henrichs and Leseman (2014), for instance, focused on two lexical features: lexical diversity, as measured in the number of word types (i.e., number of different words), and lexical sophistication, as measured in the number of general and domain-specific academic words. The number of word types was determined by counting the number of different words used across all (videotaped and transcribed) lessons. Assessing the number of general and domain-specific vocabulary required the authors to (1) create lists of all words used in the teacher-student conversations, (2) manually double-code all words that were classified as cross-disciplinary vocabulary (e.g., "experiment", "describe") or domain-specific vocabulary (e.g., "air pressure", "force"), and (3) determine interrater-reliability and discuss divergent ratings (Henrichs & Leseman, 2014).

Using computational linguistic analysis methods, by contrast, allows for assessing a wider range of linguistic features automatically in a timely manner. This enables a very fine-grained analysis of teacher talk, specifically incorporating a variety of features that are focused upon in the respective intervention. Moreover, computational linguistic analysis methods come along with maximum accuracy regarding the detection of specific linguistic features in