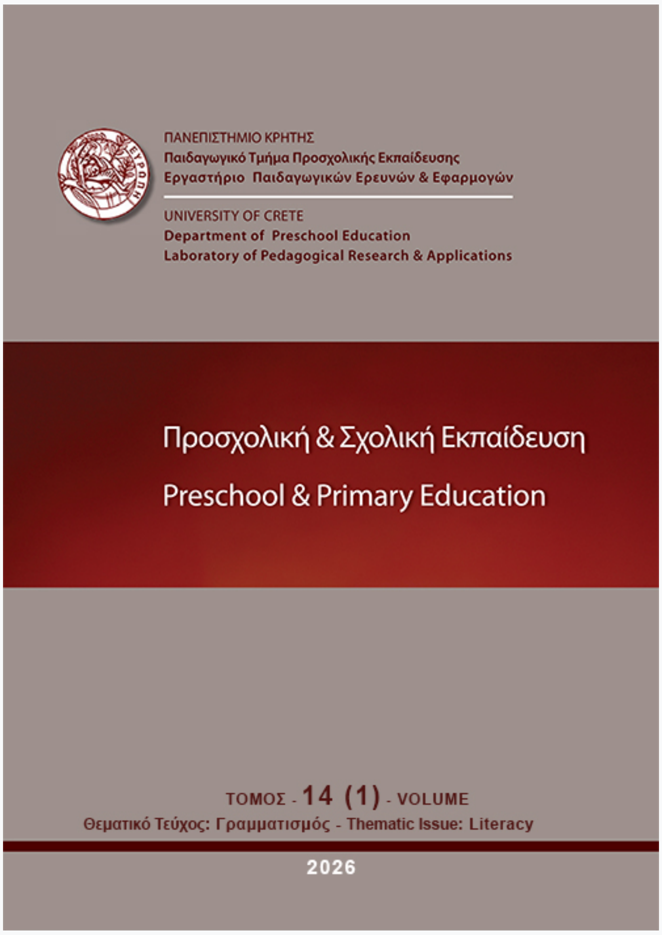
Preschool and Primary Education

Τόμ. 14, Αρ. 1 (2026)

Θεματικό τεύχος: Γραμματισμός



Investigating the predictors of academic language competences in primary school children: A machine learning approach

Astrid Rank, Elisabeth Kraus

doi: [10.12681/ppej.40307](https://doi.org/10.12681/ppej.40307)

Βιβλιογραφική αναφορά:

Rank, A., & Kraus, E. (2026). Investigating the predictors of academic language competences in primary school children: A machine learning approach. *Preschool and Primary Education*, *14*(1), 144–162. https://doi.org/10.12681/ppej.40307

Preschool & Primary Education
2026 Volume 14, Issue 1, pp. 144-162
e-ISSN: 2241-7206

Laboratory of Pedagogical Research & Applications
Department of Preschool Education, University of Crete
doi: https://doi.org//10.12681/ppej.40307

Investigating the predictors of academic language competences in primary school children: A machine learning approach

Astrid Rank
University Regensburg

Elisabeth Barbara Kraus
University Tübingen

Abstract. Academic language plays a critical role in students' ability to succeed in school. Since multilingual and socioeconomically disadvantaged students often face challenges in acquiring the necessary language register, academic language proficiency is widely recognized as a key factor in promoting educational equity. To investigate how academic language skills develop – and what factors influence this process over time, the Eva-Prim study (Rank et al., 2021) analysed longitudinal data from 570 German primary school students. Academic language comprehension and production in mathematical contexts were assessed. Drawing on over 1,000 student-, family-, and school-related variables, a machine learning approach (Random Forests) (Breiman, 2001) was employed to identify the most relevant predictors. The findings show that at the beginning of primary school, distal factors such as intelligence, parental education, and socioeconomic status were associated with academic language comprehension. Over time, however, the influence of these background characteristics declined, while trainable skills – particularly vocabulary, reading fluency, and mathematical competence, became increasingly important. Academic language production in grade four was predicted mainly by these proximal, trainable skills. Demographic factors – including migration background, played only a minor role in predicting academic language performance. These results suggest that academic language development is shaped less by static background variables and more by dynamic, educationally influenceable skills. Thus, supporting vocabulary and reading fluency within subject teaching may be key to fostering academic success for all students, regardless of their background.

Keywords*:* Academic language, primary education, mathematics education, machine learning, language development

Introduction

Understanding the predictors of students' academic achievement has long been a central concern in educational research (Lavelle-Hill et al., 2024). While previous studies have examined a variety of individual, familial, and institutional factors, many have focused on isolated predictors, offering only limited insight into how these variables jointly influence learning outcomes over time. Addressing this gap, this study applies a machine learning approach to analyse a comprehensive set of over 1,000 variables drawn from longitudinal data. This method allows us to examine how different predictors interact and change in terms of their relevance for primary school children's academic language competences. In this study,

Correspondent Author: *Astrid Rank*, Regensburger Str. 15, 93309 Kelheim Germany, e-mail: astrid.rank@ur.de

e-publisher: *National Documentation Centre, National Hellenic Research Foundation*
URL: http://ejournals.epublishing.ekt.gr/index.php/education

academic language skills were assessed specifically within mathematical contexts. To contextualize our analysis, we begin by defining academic language and outlining its role in mathematics learning and educational equity. The discussion is based on a sociocultural understanding of language (e.g. Cummins, 2017).

academic language: foundations and educational relevance

Academic language proficiency is a crucial prerequisite for academic achievement across all subjects. Unlike conversational language (Basic Interpersonal Communicative Skills BICS), academic language (Cognitive Academic Language Proficiency CALP) is more complex, usually requiring years to master. Academic language is characterized by context reduction, which occurs when language is used in a way that relies less on the immediate physical, social, or situational context to be understood. (Cummins, 2008). Its mastery is essential for understanding subject content, solving problems, and participating in classroom discourse. The characteristics of academic language include factual, concise and dense communication (Snow, 2010), as well as specialised lexical, grammatical and functional elements (e.g. reasoning, justifying, hypothesising) that are essential for classroom participation and learning (Morek & Heller, 2012). Further, academic language functions as a bridge between everyday communication and the cognitive demands of schooling, thus necessitating targeted interventions (Lange & Gogolin, 2010).

Development of academic language

It is widely accepted that academic language is a crucial prerequisite for success in school, beginning as early as elementary education (Bailey, 2007; Cummins, 2008; Rank et al., 2019; Schleppegrell, 2004). Rank et al. (2019) show that academic language is not a static construct but evolves alongside children's linguistic development. Children begin acquiring features of academic language, such as passive constructions and complex noun phrases, as early as preschool, while full mastery develops throughout their years of formal schooling (Rank et al., 2019; Volodina et al., 2020). Academic language is not merely a set of static features but rather the result of a dynamic interplay, as highlighted by the heuristic model proposed by Steinhoff (2019). This model emphasizes that the components of academic language are deeply interconnected. It identifies three core dimensions: conditions (e.g., socioeconomic status or institutional contexts), actions (e.g., linguistic practices such as explaining, arguing, or hypothesizing), and resources (e.g., vocabulary, syntactic knowledge, or metalinguistic awareness). Individuals engage in linguistic actions under specific conditions while making use of available resources. Language development occurs as these conditions, actions, and resources evolve continuously, mutually influencing one another, and being shaped by learners' lived experiences in educational settings. Also, in line with Snow and Uccelli (2009), academic language should not be understood solely as a set of linguistic features but rather as a complex, multifunctional register that is shaped by the communicative demands of schooling. Their pragmatics-based approach highlights the fact that mastering academic language requires students to represent abstract content, adopt a detached and authoritative stance, and communicate with an often "intangible" academic audience. These demands place considerable cognitive and linguistic pressure on learners, especially in subject areas such as mathematics and science. Understanding these underlying communicative challenges is essential in order to identify and support the predictors of academic language development. Building on Schleppegrell's (2012) work, academic language can be understood as a set of functionally specialized registers that are central to learning, across all subjects. It goes beyond vocabulary and syntax, encompassing the ability to use language appropriately for a range of academic purposes, genres, and audiences. Importantly, academic language varies across content areas and tasks. For instance, the language of narrating a story differs significantly from explaining a scientific process or solving a mathematics problem. Many students enter school without prior exposure to these registers, which are often shaped by socio-cultural and class-specific language socialization processes. If teachers are unaware of these differences, the mismatch can lead to miscommunication or underestimation of students' abilities. Schleppegrell (2012) emphasizes the need for teachers to make language expectations explicit, and to offer structured opportunities for all students to engage with academic registers – especially in linguistically diverse classrooms.

Features of academic language

Academic language proficiency is essential for communication and knowledge acquisition in school contexts and is a key predictor of academic achievement, yet it often constitutes a significant barrier for multilingual learners. The academic progress of multilingual students is influenced by both socioeconomic status (SES) and linguistic background (Heppt et al., 2015; Schleppegrell, 2004). In Germany, many multilingual students grow up in socially disadvantaged families with limited access to rich linguistic resources – a challenge also faced by native-speaking peers from low-SES backgrounds (Statista Research Department, 2024). As a result, these students often experience reduced exposure to diverse language input and written cultural resources. Consequently, similar to native-speaking peers from low-SES backgrounds, multilingual students frequently lack sufficient opportunities within their families to develop the academic language registers necessary for educational success (Heppt et al., 2015). Furthermore, while multilingual students bring diverse linguistic resources to the classroom, their proficiency in the academic language register at school may lag behind that of their monolingual peers, necessitating tailored support. Therefore, it is crucial to foster academic language development, particularly in multilingual and multicultural classrooms, in order to promote equity in educational outcomes. Addressing these challenges through targeted support not only strengthens language competence, but also contributes to closing achievement gaps and promoting inclusive education (Chamot, 2009).

A range of individual, familial, and institutional factors serve as predictors of academic language proficiency (Heppt & Stanat, 2020). Among the individual predictors, verbal intelligence has consistently shown strong correlations with students' ability to comprehend and produce complex linguistic structures. This is most likely due to its role in abstract thinking and the flexible manipulation of language that is required for academic tasks. Age is another significant predictor: as children mature cognitively and experience more years of formal education, their academic language skills typically become more sophisticated, aligning with developmental progressions in syntax, vocabulary, and the ability to use language in decontextualized contexts (Volodina et al., 2020). Gender may also influence academic language development. Research indicates that girls often outperform boys in linguistic tasks during early school years, possibly due to differences in verbal fluency, attention to detail, and classroom behaviour (McElvany et al., 2023; Reilly et al., 2019). However, these differences tend to diminish over time as they become shaped by broader sociocultural dynamics.

Teacher-related factors also play a central predictive role. Teachers with strong pedagogical content knowledge and a solid understanding of linguistic structures are better equipped to integrate academic language instruction into their lessons effectively (Brandt et al., 2024; Corvacho del Toro, 2013). Additionally, teachers' attitudes toward language diversity are predictive of how inclusively and supportively they approach academic language learning. For example, valuing students' home languages can foster linguistically rich and inclusive classrooms (Lange & Pohlmann-Rother, 2020).

At the institutional level, predictors include school language policies, the availability of professional development opportunities, and classroom resources. A systematic review of 136 studies (Rakesh et al., 2024) identified school climate, teacher–student relationships, support structures, and academic expectations as factors that mediate the effects of socioeconomic status (SES) on language ability and achievement. The socioeconomic status and migration background of students remain two of the most influential contextual predictors (Henschel et al., 2022). Children from high-SES families are typically exposed to a broader range of linguistic input and cognitively stimulating interactions, while children from lower-SES or immigrant backgrounds may have fewer opportunities to engage with complex language at home (Heppt et al., 2016; Heppt & Stanat, 2020; Morales-Reyes et al., 2024).

These predictors mentioned above highlight the multi-layered and dynamic nature of academic language development. They show that academic language proficiency is shaped not only by cognitive and linguistic abilities but also by socio-cultural contexts, instructional quality, and institutional support systems. Accordingly, educational strategies need to adopt a holistic and differentiated approach to strengthen academic language proficiency – particularly in linguistically diverse and socioeconomically heterogeneous classrooms.

Academic language in primary school mathematics

An overview of research studies search shows that academic language skills have a stronger impact on academic performance in subjects such as mathematics and reading comprehension than everyday language skills (Heppt et al., 2016). Although mathematics is often considered a subject with minimal language demands, it relies heavily on language as a prerequisite, instructional medium and learning objective. However, language skills can also pose obstacles in learning mathematics (Prediger et al., 2015; Rank et al., 2020). Research shows that primary school students' mathematical performance correlates with their language proficiency (e.g., Ufer et al., 2013). In 2020, an American research team comprising Peng, Lin, Ünal, Lee, Namkung, Chow, and Sales published a meta-analysis of 344 studies based on 393 independent samples and approximately 360,000 participants. The aim was to provide empirical evidence for the relationship between language and mathematics. The studies included in the analysis covered an age range from roughly 3 to 14 years, thereby encompassing the primary school period. The authors present the results in line with the sequence of their research questions. Overall, a moderate, yet significant relationship was found between language skills and mathematics performance, with stronger effects for complex mathematical skills than for basic ones. Age moderated the link between language and basic mathematics, leading to a weaker association over time. Further, the relationship was stronger for first-language learners compared to second-language learners. Working memory and intelligence together explained 50% of the variance in the language-math relationship. Over the long term, language and mathematic skills were shown to mutually influence each other, predicting each other’s development. The meta-analysis thus highlights that language plays a *developmental function* in learning mathematics, meaning that the relationship is dynamic and can be shaped through development and intervention (Balk 2024; Peng et al., 2020). Additional studies confirm the relationship between language and mathematics during early childhood and primary school. Purpura and Ganley (2014) found that language skills in kindergarten were more strongly associated with mathematical abilities than working memory performance. Viesel-Nordmeyer et al. (2020a, 2020b) demonstrated that language is not only directly related to mathematical skills but also acts as a mediator via working memory. This highlights the importance of accounting for working memory when assessing predictors of mathematical performance. Particularly noteworthy is the role of specific working memory components, especially the phonological loop, which, according to their findings, is linked to mathematical achievement in primary school.

Findings demonstrate that the relationship between language awareness and mathematical modelling is not only correlational but can also be actively strengthened through targeted interventions. Balk (2024) showed that a language-focused intervention, specifically the promotion of language awareness, can enhance students' mathematical modelling competence. Overall, the results indicate that, regardless of the support setting, the intervention led to significant improvements, particularly among initially low-performing students.

The BiSS-Initiative

A key example of how language development has been addressed at the institutional level in Germany was the nationwide "BiSS" (Bildung durch Sprache und Schrift) [Education through Language and Writing] initiative, which sought to systematically improve language support across educational stages. Starting in 2013, the Federal Republic of Germany funded the BiSS initiative – a five-year research and development programme, as a joint initiative across various ministries. The programme aimed to improve language teaching, language diagnostics and reading promotion in early childhood, primary and lower secondary education. BiSS involved more than 400 schools (primary and secondary), more than 200 kindergartens and day-care centres, and 180 partners, including universities, foundations, associations, adult education centres and libraries. These participants formed a total of 101 BiSS networks, working closely together to exchange experiences and to coordinate language education measures. Additionally, BiSS was accompanied by scientific evaluation, during which evaluation teams analysed the performance of these networks, based on specific research questions. This article introduces the Eva-Prim evaluation group (Rank et al., 2021), which focused specifically on the academic language skills of primary school children in the context of mathematics education, recognizing that linguistic competence is crucial for successful mathematics learning. The Eva-Prim study was conducted within the framework of the nationwide BiSS initiative and funded by the German Federal Ministry of Education and Research (01 Jl 1503A). A team of ten researchers from two universities participated in this project[1].

Research question

Academic language plays a pivotal role in ensuring student success in school, serving as a critical medium in terms of both knowledge transfer and cognitive development. However, its acquisition is shaped by a complex interplay of individual, familial, and school-related factors. Individual characteristics such as intelligence, age, and gender shape how children develop and use academic language. Verbal intelligence supports abstract thinking, while age aligns with cognitive and linguistic development. SES and migration background further affect students' exposure to and proficiency in academic language, highlighting disparities in linguistic resources and opportunities. As key agents in the educational process, teachers play a crucial role in their students' learning, through their linguistic knowledge, instructional strategies, and attitudes toward language diversity. However, the specific relationships between all these factors and academic language development

[1] Universität Landau: Prof. Dr. Anja Wildemann; Prof. Dr. Gerlinde Lenske; Dr. Alexandra Merkert, Dr. Lena Bien-Miller
Universität Regensburg: Prof. Dr. Astrid Rank; Prof. Dr. Anita Schilcher, Prof. Dr. Stephan Krauss, Dr. Isabell Deml, Prof. Dr. Karin Binder, Dr. Magdalena Golonka (née Schulte)

remain underexplored. This article seeks to address this gap by investigating the research question: Which student-, family-, and school-related variables are associated with primary school students' academic language development, and how do these associations strengthen or weaken over time?

By identifying the most important predictors of academic language at different time points, this study aims to provide insights for educators, policymakers, and researchers seeking to foster equitable academic outcomes.

Methods

Sample and procedure

The Eva-Prim study collected extensive data over the course of three academic years (2015/16, 2016/17, 2017/18) at three time points (time of measurement (ToM)1, ToM2, ToM3), in grades two to four of the German primary school system. The research team collected data from 984 students, 82 teachers, and 27 schools, encompassing a broad spectrum of constructs (presented in the Measures section). The original dataset comprised 1,090 observations and 2,173 variables (see Rank et al. (2021) for a detailed description of all the study variables). Among the variables, 617 were numerical and 1,556 were categorical. A total of 47.1% of all values were missing. For 517 individuals, more than 50% of values were missing and were therefore non-imputable. Regarding the variables, 1,142 variables had more than 50% missing values. For the two main variables of interest (see section on 'Dependent variables') – academic listening comprehension and academic language production – the extent of missing values was: time of measurement (ToM)1: ToM1: 140 and ToM3: 337 for the former; and ToM3: 336 for the latter. Therefore, the dataset was sorted according to missing values and reduced correspondingly. The resulting dataset contained 8.9% missing values overall, covering 570 individuals and 1,172 variables. Table 1 provides an overview over the number of variables in the dataset, grouped by measurement context.

Table 1 Overview of variables in the dataset

Variables grouped by context	Number in original dataset	Number in reduced dataset
Pedagogical Content Knowledge (PCK) in Language Instruction of Teachers	1,230	682
Pedagogical Content Knowledge in Mathematics Instruction of Teachers	317	153
Sociographic Data of Teachers	304	44
Students' Assessement of Instruction	19	19
Student Variables (Performance and Demographics)	292	268
School Context Variables (e.g. size, financial resources)	2	2
Miscellaneous	9	4
Total	2,173	1,172

Finally, missing values were imputed by state-of-the-art multiple value imputation with using the MissForests algorithm (Stekhoven & Bühlmann, 2012). According to a suggestion by Bodner (2008), the percentage of missing values should guide the number of imputed datasets. Since the reduced dataset contained 8.9% missing values, nine MissForests values were calculated and used to create nine imputed datasets. The multiple imputed datasets were then examined. Distributions of imputations were compared with distributions in the dataset with missing values to assess the imputation quality. Imputed distributions closely resembled non-imputed distributions, thus indicating good imputation quality.

Measures

Dependent variables

In this study, we were interested in two theoretical target constructs of academic language skill: academic listening comprehension and academic language production. Both constructs were measured using standardised and established ability tests, as described further below. Academic listening comprehension was measured using the BiSpra test (Heppt et al., 2024; Heppt & Volodina, 2024), which specifically measures students' abilities to understand language in school-relevant contexts. The test duration (including instructions) is approximately 40 minutes. It evaluates the use of specialized vocabulary, complex grammatical structures, and the capacity to interpret abstract, decontextualized content. Texts and items are played from an audio recording. A sample item consists of listening to a short text, such as: "Sambelo excitedly switches on his Rela. The Rela is a kind of magic ball. It shows the Kumubilan children when they will next write a class test. If they are about to write a test, it turns red. If there is no test ahead, it turns blue. Sambelo has to wait for a moment. He looks at the ball intently. Suddenly the Rela starts to glow red." Then questions need to be answered such as: "Is Sambelo going to write a class test soon?". The scores for this variable ranged between 3.24 and 27. It has proven acceptable to very good psychometric properties, with internal consistencies of Cronbach's α ranging from $\alpha = .75$ to $\alpha = .90$ (Heppt et al., 2024), and convergent validity with vocabulary measured by the PPVT (Peabody Picture Vocabulary Test) (Bulheller & Häcker, 2003) of r (404) = .60.

Academic language production was measured by the SAMT (Sprachliche Ausdrucksfähigkeit in Mathematik Test) (Merkert, 2022), which evaluates linguistic expressive abilities in mathematics of third and fourth-grade students. The test comprises tasks requiring written responses to assess students' proficiency in articulating mathematical concepts and reasoning (e.g. describing the way a fictitious character should take through a maze consisting of geometrical shapes such as hexagons, prisms, or cones). The scores for this variable ranged between 1.0 and 4.57. The SAMT also has proven acceptable psychometric properties, with an internal consistency of Cronbach's $\alpha = .71$ (Merkert & Lenske, 2023); and convergent validity with the following skills: academic listening measured by the BiSpra r (427) = .59*, reading fluency measured by the SLS (Wimmer & Mayringer, 2014) r (443) = .49*, and mathematical ability measured by the DEMAT test (Krajewski et al., 2004) r (400) = .44*.

Predictors

The predictors comprised variables at the school, teacher and student levels. For students, the dataset included standardized test data on reading (SLS; Wimmer & Mayringer, 2014), vocabulary (PPVT; Lenhard et al., 2015), mathematical competence (DEMAT; Krajewski et al., 2004), intelligence (CFT 1-R; Weiß & Osterland, 2013), grades, as well as demographic information such as gender, parental education level, SES, languages spoken at home and migration status. In this study, migration status was defined as the child not being born in

Germany. Test data were scored according to the test manuals and interpreted according to norming tables for the respective age groups.

At the teacher level, teachers provided data on their content knowledge (i.e., their understanding of subject-specific concepts and facts), pedagogical content knowledge (i.e., "that special amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding" (Shulman 1987)), and general pedagogical knowledge (i.e., classroom management, lesson planning, and instructional strategies). They also reported on their professional beliefs (e.g., views on language diversity and inclusion), professional training (e.g., participation in workshops or in-service education), and classroom performance (e.g., self-assessed effectiveness in supporting students' academic language development). At the school levels, data were collected on various aspects, including facilities (e.g., spatial resources) and school-wide educational policies.

Data analyses

All analyses were conducted using the statistical software R (R Core Team, 2024). Initially, the dataset was reduced based on substantive considerations and a sorting algorithm that organized the dataset by completeness in columns and rows. Subsequently, a multiple imputation of missing values was performed using the MissForest algorithm in the missForest package (Stekhoven & Stekhoven, 2013). Bivariate, linear relationships were investigated using a compact graphical approach. The main analyses were conducted using a machine learning (ML) model called Random Forests, estimated by means of the randomForest package (Liaw & Wiener, 2002), and visualized using interpretable ML methods (Molnar, 2020). Post-hoc correlation analyses were performed with the predictors identified as being particularly predictive by means of variable importance analyses.

Linear Relationship Analysis

We used correlational analysis to investigate bivariate, linear relationships between the predictor variables and the dependent variables, and among the predictor variables. Since our dataset was vast, we used a visual approach and plotted the correlational matrix after grouping the variables into teacher variables, student achievement variables, and student demographic information.

Complex relationship analysis

We chose ML as our central analytical strategy because of its ability to combine numerous small, localized statistical models into a single, overarching model (Hilbert et al., 2021). This approach facilitates the efficient analysis of large datasets and enables the identification of complex relationships. Since ML investigates the relationships between all predictor variables and the target variables simultaneously, another of its key advantages is its capability to uncover entirely new relationships that traditional methods might overlook when manually selecting the relationships to be investigated. More specifically, we chose the Random Forests model because it can automatically detect and model interactive and nonlinear effects, as well as subgroup effects within the data. In addition, Random Forests have been found not to overfit to a specific sample (Breiman, 2001), and they perform well even without (extensive) hyper parameter tuning (Probst et al., 2019).

Random forests

A Random Forest is created by combining the results of numerous decision trees (e.g., 500). This combination can be calculated, for example, by averaging the predictions for an observation across all trees. The decision tree algorithm is a nonparametric technique that divides a dataset into subgroups by making binary splits based on predictor variables, thus clustering observations (here: participants) with similar characteristics on the target (dependent) variable. The goal of this process is to maximize the reduction in impurity from the parent node to its sub-nodes. In regression trees, impurity is typically measured by the reduction in target variance after introduction of the sub-node. The construction of a decision tree concludes when a node reaches a predefined minimum number of observations, or when the impurity reduction from a split falls below a specified threshold (Probst et al., 2019).

Each decision tree within a Random Forest selects only a subset of variables and observations for each node decision, thus introducing a random component to individual trees and ensuring greater diversity among them. The size of this subsample is expressed as a proportion of the total number of predictors. It must be chosen by the modeler and is called the mtry hyperparameter. The model's quality is determined using the so called out-of-bag error. The out-of-bag error is an internal validation metric computed by averaging the prediction errors for each observation using only the subset of trees that did not include that observation (Breiman, 2001). For continuous dependent/target variables, it can be used to derive a R^2 measure, similar to the explained variance in a regression analysis. Yet – contrary to the R^2 measure in linear regression, the R^2 measure of a Random Forest is a measure of the generalizability of a model and not susceptible to overfitting with an increasing number of predictors (Breiman, 2001).

In this study, Random Forests constituted the main method of analysis. For each imputed dataset, a Random Forest was calculated with the BiSpra values (at both measurement points) as the target variable, and the parameters were then pooled. Additionally, SAMT values at measurement point 3 were predicted, following a similar procedure to the BiSpra predictions. The Random Forests were tuned with respect to the mtry parameter; and R^2 was estimated.

Variable importance

To determine variable importance (Wei et al., 2015), we calculated permutation variable importance. The underlying idea is to shuffle all values of a predictor and then observe the change in prediction accuracy in the out of bag sample (Hilbert et al., 2025). We used the change in mean squared error (MSE) as a measure of change. Variables were then ranked by their mean variable importance across the nine Random Forests, and the variable importance scores were visualized using boxplots. For the BiSpra predictions, results across the three measurement points were plotted. This was possible because all dependent variables were measured on the same scale and did not differ in their variance. Thus, changes in the importance of predictors could be observed.

Post-hoc analyses

Finally, descriptive correlations of the most important predictors with the target variables were calculated.

Results

Linear relationships between teacher variables, student achievement, and student demographics

To precede our central ML analysis, we investigated linear relationships between the different groups of variables. The visual display of the correlation matrix revealed that variables belonging to the same variable grouping (teacher variables, student achievement, student demographics) were more strongly associated with each other than with variables belonging to the other groups. This finding is displayed in Figure 1, with intensive colouring indicating high absolute associations, and pallid colouring indicating weak associations.

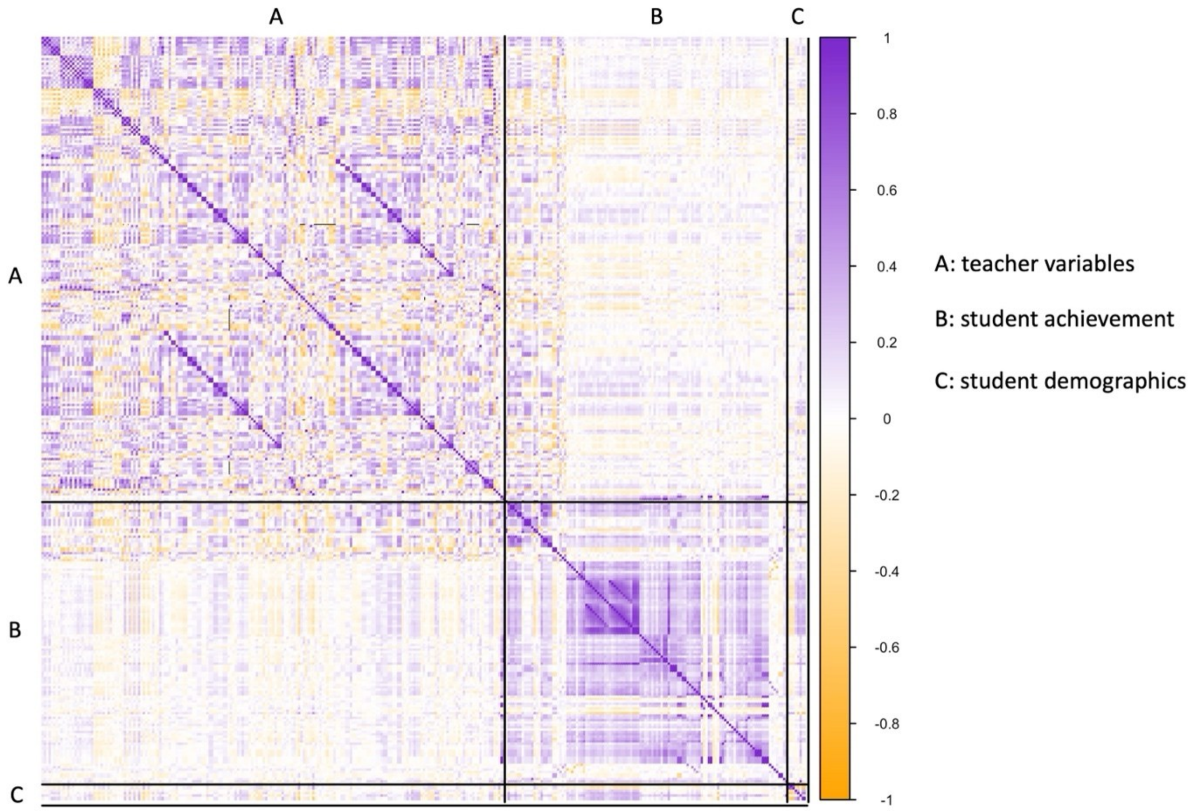


Figure 1 Visual display of variable correlations grouped by origin of variable

Complex relationship analysis using random forests

Academic listening comprehension (BiSpra)

The dependent variable at the first measurement point (grade two) was academic listening comprehension, assessed using the BiSpra test (Heppt et al., 2024). Untuned Random Forests explained between 48.5% and 65.0% of the test set variance in the BiSpra scores; tuning increased this only marginally to 48.9% and 65.5%, depending on the version of the imputed dataset. Variable importance analysis that compared the contributions of over 1,000 variables, indicated that the following variables were particularly predictive for academic listening comprehension at the first measurement point: distal student variables and demographic information – in particular, vocabulary, intelligence, and socioeconomic status had high variable importance in these results. Additional predictive factors included: parental qualifications, teachers' profiles during their university studies, and the number of books in the child's household. Figure 2 displays these results graphically.

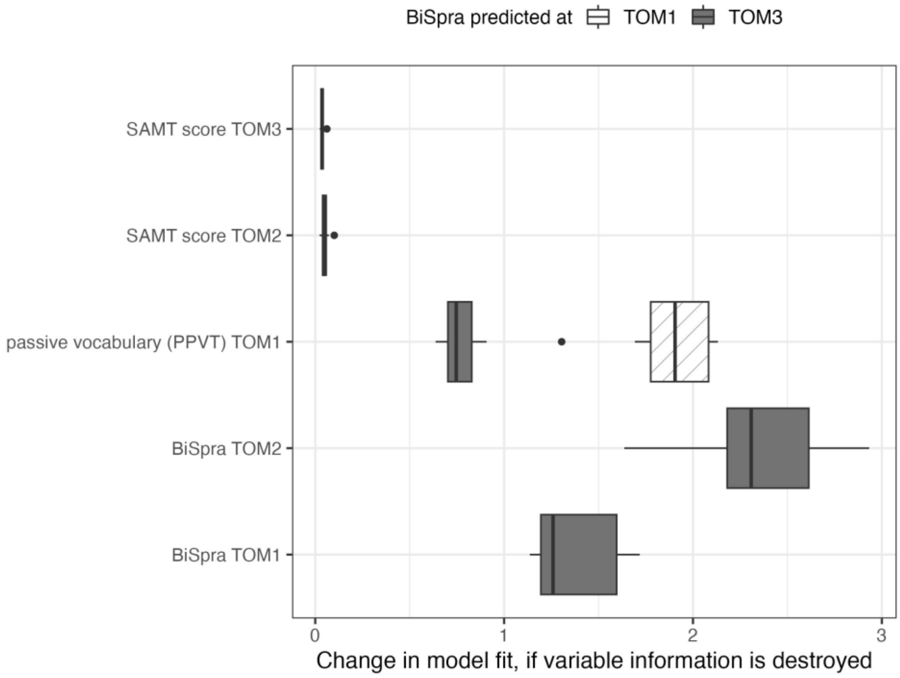


Figure 1 Variable importance scores for language related variables in BiSpra predictions at ToM1 and ToM3
Note. TOM = Time of measurement.

By the third measurement point (grade four), the variable importance ranking changed. While the importance of demographic variables decreased drastically, the following variables were highly predictive of academic listening comprehension in the fourth grade: academic listening comprehension (BiSpra scores) from time points 1 and 2, mathematical competence (DEMAT scores), and vocabulary (PPVT scores).

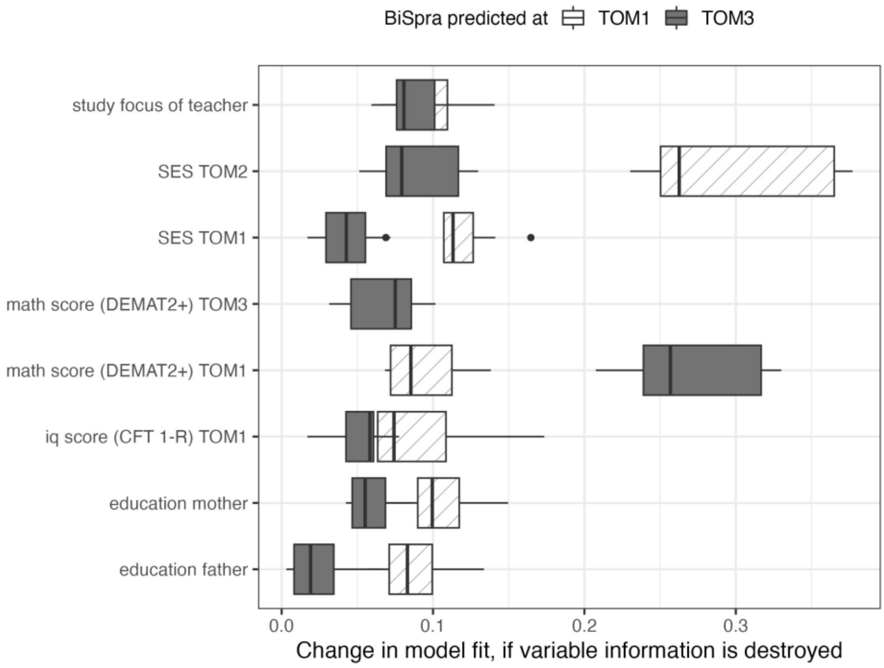


Figure 2 Variable importance scores for non-language related variables in BiSpra prediction at ToM1 and ToM3
Note. TOM = Time of measurement.

Interestingly, the importance of intelligence on academic listening comprehension decreased over time, while the importance of other trainable variables, such as vocabulary and mathematical competence, increased. Figure 3 shows the exact changes in variable importance between TOM1 and TOM3 for the specific predictors. This finding suggests that it could be worthwhile to specifically target and develop these trainable factors throughout primary education.

Academic language production (SAMT)

The second dependent variable was academic language production, assessed by SAMT (Merkert, 2022). SAMT was administered at the third measurement point (grade four). Machine learning analysis could explain between 50.5% and 53.0% of the test set variance in SAMT scores (50.6% and 53.2% when tuned). Variable importance analysis suggested that, again, trainable abilities such as mathematical competence (DEMAT), reading fluency (SLS) and academic language production (SAMT) from earlier measurement time points were most predictive among the possible predictors. A subset of the most important predictors is presented in Figure 4.

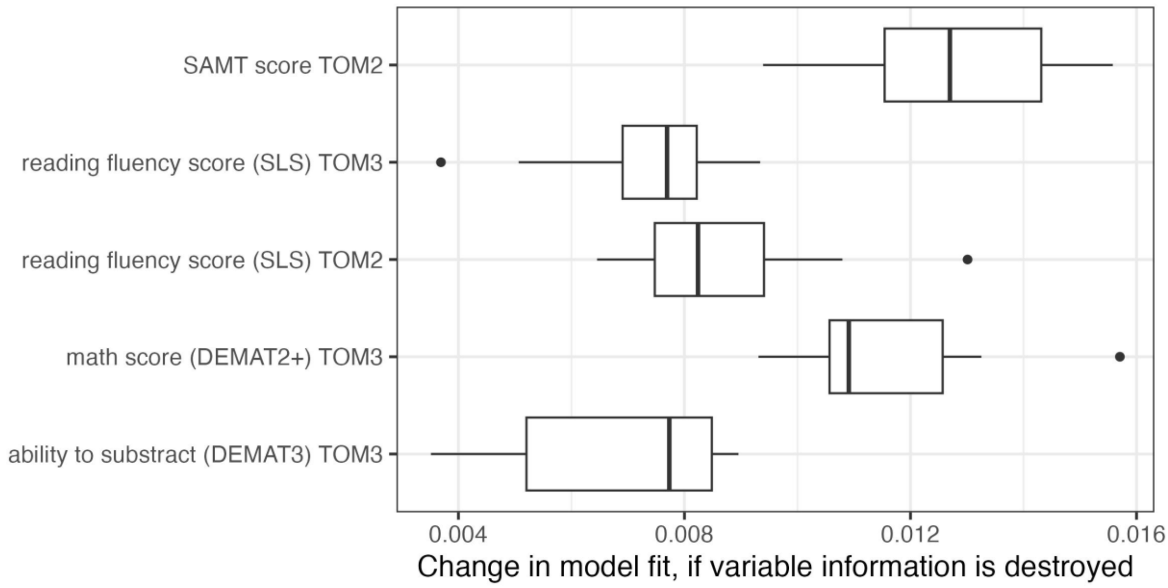


Figure 3 Variable importance scores for most important variables in SAMT predictions at ToM3
Note. TOM = Time of measurement.

Post-hoc correlations with the children’s demographic data revealed no relationship between migration status and academic language production (*r* (370) = 0.03) and only a small positive relationship with gender reported as male (*r* (390) = .18).

In summary, the following can be said: The school and teacher variables were negligibly associated with primary children's academic language competences. The distal individual variables – such as intelligence and socioeconomic background – were important at the beginning of their school career. However, ongoing progression through their school career decoupled this effect. Academic listening comprehension at the third time point, as well as production of academic language in a mathematical setting, were explained mostly through mathematical competences, vocabulary, and reading fluency from earlier time points.

Discussion

This study employed a machine learning approach to identify and understand the most influential predictors of academic language development in primary school students, based on comprehensive longitudinal data. The analysis revealed that academic language skills – in both comprehension and production – can be predicted with a high degree of accuracy, particularly through trainable variables such as vocabulary, reading fluency, and mathematical competence. Importantly, by comparing the predictive power of different groups of variables over time, we gained valuable insights into how the relevance of cognitive, demographic, and instructional factors shifts throughout the primary years. These findings align with sociocultural and interactionist theories of language development, which emphasize the dynamic interplay between cognitive, linguistic, and contextual influences. The decreasing role of distal predictors such as intelligence and SES over time supports the notion that academic language is a learnable competence, shaped by educational input and student experience.

Notably, the growing importance of vocabulary and subject-specific skills points to the potential of targeted interventions. In particular, vocabulary emerged consistently as a strong predictor across both academic listening and language production. This underscores the need for instructional practices that promote academic vocabulary in context. Especially in mathematics, scaffolding language through explicit modelling, sentence starters, and vocabulary support could simultaneously strengthen both subject understanding and language skills. The findings of our study underscore this perspective by showing that trainable skills such as vocabulary and mathematical competence are important predictors of academic language development. These skills can be viewed as access points to the language registers of schooling, as described by Schleppegrell (2012). Supporting students in developing these competencies is not merely a matter of individual cognitive ability, but of guided participation in the discursive practices employed at school.

The strong link between mathematical competence and academic language also suggests that language and content learning are mutually reinforcing. Integrating language-sensitive instruction into mathematics classes may yield dual benefits by supporting conceptual understanding while fostering language development. The lack of substantial effects for migration background challenges deficit-based assumptions and instead highlights the central role played by varied learning opportunities. This, in turn, reinforces the importance of equitable access to high-quality instruction and learning environments for all students. The minor gender effect observed in language production – slightly favouring boys – invites further investigation into gender-specific pathways in academic language development and possible differentiated support strategies.

Interestingly, teacher-related variables showed limited associations in our analyses. This finding may reflect the limitations of self-reported data; or it might be due to the nested data structure, where (by design) predicting individual level data carries more information than class level data provided by the teacher (Lavelle-Hill et al., 2024). However, this finding also points to the need for further research that explores classroom discourse and instructional practices. Professional development that enables teachers to embed academic language support into subject instruction remains a promising area for further investigation.

The use of machine learning in this study proved particularly effective in uncovering complex, non-linear patterns that traditional statistical models might overlook. It provided a more nuanced understanding of how diverse factors interact over time in shaping the academic language skills of primary students.

When interpreting our findings, it is, it is important to note that our machine learning models are data-driven and not based on a top-down, causal inference framework. As with other regression-based approaches, the results should not be interpreted as evidence of causality. While we included variables (based on theoretical grounds) that suggest potential causal relevance, machine learning models do not allow for explicit specification of how these variables are used – such as defining mediation or moderation pathways. This limitation is especially relevant given the large number of predictors in our model. It is possible that some variables with genuine causal effects may appear less influential because their effects were mediated by other variables, thereby diminishing their apparent direct impact. For example, the importance of migration background and socioeconomic status (SES) appeared to be relatively low in this study, which contradicts findings from other studies such as the IQB Trends in Student Achievement, which has been conducted every five years since 2011. That study has consistently reported a strong and stable association between student competences and their SES and migration background, which amounted, for instance, to a difference equivalent of two school years in listening comprehension in the 2021 assessment (Henschel et al., 2022). It may be that, in the present study, the effects of migration background and SES were absorbed into the student achievement scores at measurement timepoint 1 – which are more proximal to the dependent student achievement variables at measurement timepoint 3.

While our methodology offers several advantages, it also comes with limitations. Although we aimed to include the most relevant predictors of academic language competence in this study, we were bound to the scope of the EvaPrim dataset and might have overlooked – or not had available – other important variables. For instance, no student variables on motivation or working memory were available (Viesel-Nordmeyer et al., 2020a; 2020b); and teacher variables were based solely on self-reporting. As a result, it is possible that, if these omitted variables were to be considered, some findings may differ from those reported here.

In conclusion, our findings suggest that early, trainable skills– especially vocabulary, reading fluency, and mathematical competence – should be key targets of educational interventions that aim to foster academic language skills in primary school students. A systematic, language-integrated instructional approach in all subjects, paired with support for teachers and equitable learning conditions, has the potential to bridge linguistic gaps and promote academic success.

Future research should aim to explore the causal mechanisms underlying these relationships, particularly through longitudinal intervention studies. A focus on teacher practices, peer interaction, and the micro-level dynamics of classroom language use could further enhance our understanding of academic language development in diverse educational settings.

Ethics statement

This study was conducted in accordance with the data protection regulations of the University of Regensburg and was approved by the Bavarian State Ministry for Education and Cultural Affairs (under the reference number X.7-BO7106/85/12). All procedures complied with relevant legal and ethical standards for research involving human participants. Informed consent was obtained from all participating teachers as well as from the parents or legal guardians of the children. Participation was voluntary, and participants were informed about their right to withdraw at any time without negative consequences. All data were pseudonymized prior to analysis, in order to ensure confidentiality and data protection.

References

Bailey, A. L. (Eds.). (2007). *The language demands of school: Putting academic English to the test*. Yale University Press.

Balk, D. (2024). *Mathematische Modellierungskompetenz von Grundschulkindern sprachbewusst fördern*. [Promoting mathematical modelling competence of primary school children through language awareness]. Julius Klinkhardt Verlag.

Bodner, T. E. (2008). What improves with increased missing data imputations?. *Structural equation modeling: a multidisciplinary journal*, *15*(4), 651-675. https://doi.org/10.1080/10705510802339072

Brandt, H., Menzel, K. N., Neumann, A., & Weinhold, S. (2024). Gut vorbereitet auf den Umgang mit sprachlicher Diversität im Unterricht? [Well prepared for dealing with linguistic diversity in the classroom?] *DDS – Die Deutsche Schule, 2024(2)*, 163–184. https://doi.org/10.31244/dds.2024.02.05

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

Bulheller, S., & Häcker, H. O. (2003). *Peabody Picture Vocabulary Test (PPVT)-Deutschsprachige Fassung des PPVT-III für Jugendliche und Erwachsene von LM Dunn & LM Dunn*. Swets Test Services.

Chamot, A. U. (2009). *The CALLA handbook: Implementing the cognitive academic language learning approach* (Second edition). Pearson Longman.

Corvacho del Toro, I. M. (2013). *Fachwissen von Grundschullehrkräften: Effekt auf die Rechtschreibleistung von Grundschülern* [Subject knowledge of primary school teachers: Effect on students' spelling performance]. University of Bamberg Press.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In N. H. Hornberger (Eds.), *Encyclopedia of language and education* (pp. 487–499). Springer US. https://doi.org/10.1007/978-0-387-30424-3_36

Cummins, J. (2017). Teaching for transfer in multilingual school contexts. In O. García, A. M. Y. Lin, & S. May (Eds.), *Bilingual and multilingual education* (pp. 103–115). Springer International Publishing. https://doi.org/10.1007/978-3-319-02258-1_8

Henschel, S., Heppt, B., Rjosk, C., & Weirich, S. (2022). Zuwanderungsbezogene Disparitäten [Migration-related disparities]. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich, & S. Henschel (Eds.), *IQB-Bildungstrend 2021 Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (pp. 181-219). Waxmann.

Heppt, B., Haag, N., Böhme, K., & Stanat, P. (2015). The role of academic-language features for reading comprehension of language-minority students and students from low- SES families. *Reading Research Quarterly*, *50*(1), 61–82. https://doi.org/10.1002/rrq.83

Heppt, B., Henschel, S., & Haag, N. (2016). Everyday and academic language comprehension: Investigating their relationships with school success and challenges for language minority learners. *Learning and Individual Differences*, *47*, 244–251. https://doi.org/10.1016/j.lindif.2016.01.004

Heppt, B., Köhne-Fuetterer, J., Eglinsky, J., Volodina, A., Stanat, P., & Weinert, S. (2024). *BiSpra 2–4: Test zur Erfassung bildungssprachlicher Kompetenzen bei Grundschulkindern der Jahrgangsstufen 2 bis 4* [BiSpra 2–4: Test for assessing academic language competences in primary school children in grades 2 to 4](2. überarbeitete und aktualisierte Auflage). Waxmann.

Heppt, B., & Stanat, P. (2020). Development of academic language comprehension of German monolinguals and dual language learners. *Contemporary Educational Psychology*, *62*, 101868. https://doi.org/10.1016/j.cedpsych.2020.101868

Heppt, B., & Volodina, A. (2024). Diagnostik bildungssprachlicher Kompetenzen mit BiSpra 2–4: Grundlagen der Testentwicklung und empirische Befunde zu einsprachigen und mehrsprachigen Lernenden [Diagnosis of academic language competences with BiSpra 2–4: Foundations of test development and empirical findings on monolingual and multilingual learners]. In J. Goschler, P. Rosenberg, & T. Woerfel (Eds.), *Empirische Zugänge zu Bildungssprache und bildungssprachlichen Kompetenzen* (pp. 239–271). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-43737-4_10

Hilbert, S., Coors, S., Kraus, E., Bischl, B., Lindl, A., Frei, M., Wild, J., Krauss, S., Goretzko, D., & Stachl, C. (2021). Machine learning for the educational sciences. *Review of Education*, *9*(3), e3310. https://doi.org/10.1002/rev3.3310

Hilbert, S., Kraus, E., & Lindl, A. (2025). *Machine Learning: Eine Einführung für Psychologie, Geistes-und Sozialwissenschaften*. [Machine learning: An introduction for psychology, the humanities, and social sciences]. Springer Wissensmedien Wiesbaden GmbH.

Krajewski, K., Liehm, S., & Schneider, W. (2004). *Deutscher mathematiktest für zweite klassen: DEMAT 2*. [German mathematics test for second grade: DEMAT 2]. Beltz Test.

Lavelle-Hill, R., Frenzel, A. C., Goetz, T., Lichtenfeld, S., Marsh, H. W., Pekrun, R., Sakaki, M., Smith, G., & Murayama, K.. (2024). How the predictors of math achievement change over time: A longitudinal machine learning approach. *Journal of Educational Psychology*, *116*(8), 1383–1403. https://doi.org/10.1037/edu0000863

Lange, I., & Gogolin, I. (2010). *Durchgängige sprachbildung: Eine handreichung*. [Continuous language education: A handbook]. Waxmann.

Lange, S. D., & Pohlmann-Rother, S. (2020). Überzeugungen von Grundschullehrkräften zum Umgang mit nicht-deutschen Erstsprachen im Unterricht. [Primary school teachers' beliefs about dealing with non-German first languages in the classroom]. *Zeitschrift für Bildungsforschung*, *10*(1), 43–60. https://doi.org/10.1007/s35834-020-00265-4

Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody Picture Vocabulary Test - Revision 4 (PPVT-4), German Version.* Pearson Assessment.

Liaw, A., & Wiener, M., (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

McElvany, N., Lorenz, R., Frey, A., Goldhammer, F., Schilcher, A., & Stubbe, T. C. (2023). *IGLU 2021 Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* [PIRLS 2021: Reading competence of primary school children in international comparison and in the trend over 20 years]. Waxmann.

Merkert, A. (2022). *Sprachdiagnostik im Mathematikunterricht der Grundschule: Konzeption eines Testinstruments.* [Language diagnostics in primary mathematics education: Conceptualization of a test instrument]. Waxmann.

Merkert, A., & Lenske, G. (2023). Validierung eines diagnostischen Instruments zur Erfassung der sprachlichen Ausdrucksfähigkeit in Mathematik für die dritte und vierte Klassenstufe. [Validation of a diagnostic instrument for assessing language production skills in mathematics for third and fourth grade students]. *Diagnostica*, *69*(4), 194–206.

Molnar, C. (2020). *Interpretable machine learning*. https://christophm.github.io/interpretable-ml-book/

Morales-Reyes, A., Pérez-Vargas, J. C., Siberón, J., & Wolfgang, R. (2024). English skills of Puerto Rican students: The effect of gender, SES and school socioeconomic composition.

International Journal of the Sociology of Language, *2024*(286), 23–51. https://doi.org/10.1515/ijsl-2023-0022

Morek, M., & Heller, V. (2012). Bildungssprache—Kommunikative, epistemische, soziale und interaktive Aspekte ihres Gebrauchs [Academic language—Communicative, epistemic, social, and interactive aspects of its use]. *Zeitschrift Für Angewandte Linguistik*, *2012*(57), 67–101. https://doi.org/10.1515/zfal-2012-0011

Peng, P., Lin, X., Ünal, Z. E., Lee, K., Namkung, J., Chow, J., & Sales, A. (2020). Examining the mutual relations between language and mathematics: A meta-analysis. *Psychological Bulletin, 146*(7), 595–634. https://doi.org/10.1037/bul0000231

Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E., & Benholz, C. (2015). Sprachkompetenz und Mathematikleistung – Empirische Untersuchung sprachlich bedingter Hürden in den Zentralen Prüfungen 10 [Language proficiency and mathematics performance – Empirical investigation of language-related barriers in the central examinations at grade 10]. *Journal für Mathematik-Didaktik*, *36*(1), 77–104. https://doi.org/10.1007/s13138-015-0074-0

Probst, P., Boulesteix, A. L., & Bischl, B. (2019). Tunability: Importance of hyperparameters of machine learning algorithms. *The Journal of Machine Learning Research*, *20*(1), 1934–1965.

Purpura, D. J., & Ganley, C. M. (2014). Working memory and language: skill-specific or domain-general relations to mathematics? *Journal of Experimental Child Psychology*, *122*, 104–121. https://doi.org/10.1016/j.jecp.2013.12.009

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Rakesh, D., Lee, P. A., Gaikwad, A., & McLaughlin, K. A. (2024). Associations of socioeconomic status with cognitive function, language ability, and academic achievement in youth: A systematic review of mechanisms and protective factors. *Journal of Child Psychology and Psychiatry*, 66, 417-439. https://doi.org/10.1111/jcpp.14082

Rank, A., Deml, I., Lenske, G., Merkert, A., Binder, K., Schulte, M., Schilcher, A., Wildemann, A., Bien-Miller, L., & Krauss, S. (2021). Eva-Prim: Evaluation von Sprachförderkompetenz und (bildungs)sprachlichen Leistungen von Schülerinnen und Schülern in Mathematik. [Eva-Prim: Evaluation of language support competence and (academic) language performance of students in mathematics]. In S. Gentrup, S. Henschel, K. Schotte, L. Beck & P. Stanat (Eds.), *Sprach- und Schriftsprachförderung wirksam gestalten* (pp. 105–124). W. Kohlhammer.

Rank, A., Schilcher, A., & Schulte, M. (2020). Sprache und mathematikunterricht. Wie hängt das zusammen? [Language and mathematics instruction: How are they connected?] *Die Grundschulzeitschrift*, *323*, 29–31.

Rank, A., Wildemann, A., Hartinger, A., & Tietze, S. (2019). Early steps into science and literacy – EASI Science-L. Wirkungen sprachlicher Anregungsqualität in naturwissenschaftlichen Bildungsangeboten auf die sprachlichen Fähigkeiten von Vorschulkindern. [Early Steps into Science and Literacy – EASI Science-L: Effects of linguistic stimulation quality in science education programs on the language skills of preschool children] In Stiftung Haus der kleinen Forscher (Eds.), *Wirkungen naturwissenschaftlicher Bildungsangebote auf pädagogische Fachkräfte und Kinder* (pp. 140–193). Verlag Barbara Budrich. https://doi.org/10.2307/j.ctvmx3jn8

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Gender differences in reading and writing achievement: Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*, *74*(4), 445–458. https://doi.org/10.1037/amp0000356

Schleppegrell, M. (2004). *The language of schooling: A functional linguistics perspective*. Lawrence Erlbaum Associates.

Schleppegrell, M. J. (2012). Academic language in teaching and learning: An introduction to the special issue. *The Elementary School Journal*, *112*(3), 409–418. DOI:10.1086/663297

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–22.

Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science, 328*(5977), 450–452. https://doi.org/10.1126/science.1182597

Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge Handbook of Literacy* (pp. 112–133). Cambridge University Press.

Statista Research Departement (2024). *Armutsgefährdungsquote von Kindern in Deutschland nach Migrationshintergrund von 2009 bis 2017* [Child poverty risk rate in Germany by migration background from 2009 to 2017]. https://de.statista.com/statistik/daten/studie/786132/umfrage/armutsgefaehrdungsquote-von-kindern-nach-migrationsstatus-in-deutschland/

Steinhoff, T. (2019). Konzeptualisierung bildungssprachlicher Kompetenzen. Anregungen aus der pragmatischen und funktionalen Linguistik und Sprachdidaktik: Rethinking Academic Language. [How Pragmatics, Functional Linguistics und Language Pedagogy can Contribute to the Discourse]. *Zeitschrift Für Angewandte Linguistik*, 2019(71), 327–352. https://doi.org/10.1515/zfal-2019-2019

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics*, *28*(1), 112 –118.

Stekhoven, D. J., & Stekhoven, M. D. J. (2013). Package 'missForest'. *R package version*, *1*, 21.

Ufer, S., Reiss, Kristina, & Mehringer, V. (2013). Sprachstand, soziale Herkunft und Bilingualität: Effekte auf Facetten mathematischer Kompetenz [Language proficiency, social background, and bilingualism: Effects on facets of mathematical competence]. In M. Becker-Mrotzek, K. Schramm, E. Thürmann, & H. J. Vollmer (Eds.), *Sprache im Fach: Sprachlichkeit und fachliches Lernen* (pp. 185–201). Waxmann.

Viesel-Nordmeyer, N., Ritterfeld, U., & Bos, W. (2020a). Die rolle von sprache und arbeitsgedächtnis für die entwicklung mathematischen lernens vom vorschul- bis ins grundschulalter [The role of language and working memory in the development of mathematical learning from preschool to primary school age]. *Lernen und Lernstörungen*, *9*(2), 97–110. https://doi.org/10.1024/2235-0977/a000291

Viesel-Nordmeyer, N., Ritterfeld, U., & Bos, W. (2020b). Welche Entwicklungszusammenhänge zwischen Sprache, Mathematik und Arbeitsgedächtnis modulieren den Einfluss sprachlicher Kompetenzen auf mathematisches Lernen im (Vor-)Schulalter? [Which developmental interrelations between language, mathematics, and working memory modulate the influence of language skills on mathematical learning in (pre)school age?] *Journal für Mathematik-Didaktik*, *41*(1), 125–155. https://doi.org/10.1007/s13138-020-00165-0

Volodina, A., Weinert, S., & Mursin, K. (2020). Development of academic vocabulary across primary school age: Differential growth and influential factors for German monolinguals and language minority learners. *Developmental Psychology*, *56*(5), 922–936. https://doi.org/10.1037/dev0000910

Wei, P., Lu, Z., & Song, J. (2015). Variable importance analysis: A comprehensive review. *Reliability Engineering & System Safety*, *142*, 399–432. 10.1016/j.ress.2015.05.018

Weiß, R. H., & Osterland, J. (2013). *CFT 1-R Grundintelligenztest Skala 1 – Revision* [CFT 1-R: Basic intelligence test, scale 1 – Revised version]. Hogrefe.

Wimmer, H., & Mayringer, H. (2014). *SLS 2-9: Salzburger Lese-Screening für die Schulstufen 2-9* [SLS 2-9: Salzburg reading screening for grades 2 to 9.]. Hans Huber.

Received: 31.1.2025, Revised: 28.5.2025, Approved: 26.6.2025