The Role of General and Subject-specific Language Skills when Learning Mathematics in Elementary School

Stefan Ufer · Katrin Bochnik

Abstract The role of skills in the language of instruction for mathematics learning is well established by longitudinal studies at the primary school level. Explanations for these relations lead to the question: are they mostly due to general, domain-over-arching language skills, or does the command of subject-specific language registers play an important role? Integrating prior research threads, we propose two instruments to measure subject-specific language skills in mathematics: One measuring mathematical vocabulary, and one measuring mathematical text comprehension. We report on a longitudinal study with $N = 237$ German grade 3 students, which investigated the predictive value of these instruments beyond prior arithmetic skills, general language skills, and control variables such as general cognitive skills and socio-economic status on students' later arithmetic skills. We applied a multidimensional assessment model to measure arithmetic skills. Apart from replicating the prominent role of general language skills found in earlier studies, our results indicate a substantial, additional role of subject-specific language skills for the development of mathematics skills. These relations could be identified for knowledge of mathematical vocabulary, as well as for mathematical text comprehension. The results indicate that fostering subject-specific language skills already at the primary school level is not only one of many goals of mathematics instruction, but is a core prerequisite to supporting mathematical skill acquisition.

Keywords Development of mathematics skills · Language skills · Subject-specific language skills · Non-German family language · Mathematical vocabulary

S. Ufer (✉) · K. Bochnik
Mathematisches Institut, Ludwig-Maximilians-Universität, Theresienstraße 39, 80333 München, Germany
E-Mail: ufer@math.lmu.de

Die Rolle allgemein- und fachsprachlicher Kompetenzen für das Mathematiklernen in der Grundschule

Zusammenfassung Die Bedeutung von sprachlichen Fähigkeiten in der Unterrichtssprache für das Mathematiklernen ist durch längsschnittliche Studien im Primarbereich umfangreich belegt. Erklärungsansätze für diesen Zusammenhang führen zu der Frage, ob hier im Wesentlichen allgemeine, domänenübergreifende sprachliche Fähigkeiten oder vielmehr das Beherrschen fachspezifischer Sprachregister eine zentrale Rolle spielen. Aufbauend auf bestehenden Forschungsansätzen schlagen wir zwei Instrumente zur Messung fachspezifischer Sprachfähigkeiten vor: Eines zur Messung des mathematischen Fachwortschatzes, und eines zur Messung des mathematischen Textverständnisses. Wir berichten Ergebnisse einer Längsschnittstudie mit $N = 237$ deutschen Drittklässlern, die den Vorhersagewert dieser Instrumente für die spätere Arithmetikleistung unter Kontrolle der arithmetischen Vorleistung, allgemeiner sprachlicher Fähigkeiten und Kontrollvariablen wie allgemeinen kognitiven Fähigkeiten und dem sozioökonomischen Status untersucht. Zugrunde gelegt wurde eine mehrdimensionale Konzeptualisierung arithmetischer Fähigkeiten. Neben der Replikation früherer Studien zur Bedeutung allgemeiner Sprachfähigkeiten weisen unsere Ergebnisse auf einen substantiellen, zusätzlichen Einfluss fachspezifischer Sprachfähigkeiten für die Entwicklung mathematischer Fähigkeiten hin. Diese Zusammenhänge konnten sowohl für Wissen zum mathematischen Fachwortschatz, als auch für das mathematische Textverständnis nachgewiesen werden. Die Ergebnisse deuten an, dass eine Förderung fachspezifischer Sprachfähigkeiten bereits in der Grundschule nicht nur eines unter vielen Zielen des Mathematikunterrichts, sondern auch eine wesentliche Voraussetzung für erfolgreiche mathematische Lernprozesse ist.

Schlüsselwörter Entwicklung mathematischer Kompetenzen · Sprachkompetenz · Fachsprachliche Kompetenzen · Nicht-deutsche Familiensprache · Mathematischer Fachwortschaft

1 Introduction

National and international studies have repeatedly demonstrated that mathematics skills are strongly correlated with students' skills in the language of instruction. In particular regarding students with a migrant background (Tarelli et al. 2012) and students whose family language differs from the language of instruction (Heinze et al. 2007), language skills have been identified as an important explanatory variable for their disadvantages in mathematics learning, above and beyond measures of socio-economic status (Tarelli et al. 2012). Current research from different perspectives strives to understand this relation between language skills and mathematics learning. General education research has adopted a differentiated view on language skills to identify those facets of language skills which are important for learning processes, including mathematics learning (e.g. grammar, vocabulary, and listening comprehension: Paetsch et al. 2015). From a subject-specific perspective, the role

of subject-specific language registers has been discussed (e.g. Prediger and Wessel 2013), which must be mastered to construct meaning about mathematical concepts during instruction. This raises the question on the role of subject-specific language skills that go beyond mastering everyday language registers or more sophisticated demands of general academic language registers (Wessel and Erath 2018). While first promising attempts to conceptualise and operationalise mathematical language skills exist (Bochnik and Ufer 2016a; Schindler et al. 2019), our knowledge about the relation of such subject-specific language skills to the learning of mathematics skills remains limited.

Furthermore, in prior research, mathematics skills have often been conceptualised as a one-dimensional construct (Prediger et al. 2018). This contrasts with claims that language skills should be of particular importance for certain, more conceptual facets of mathematics skills (Heinze et al. 2007). First attempts to differentiate these more conceptual facets from other, more schematic skills, however, were mostly explorative and based on instruments that were not designed for such a differentiation (e.g. Ufer et al. 2013). On the other hand, models to conceptualise multidimensional mathematics assessments have been proposed in the past (Suurtamm et al. 2016). Thus, studying the role of general and subject-specific language skills when learning different dimensions of mathematics skills is a desideratum that can extend our understanding of the complex interaction between language skills and mathematics learning.

Accordingly, the goal of this contribution is to investigate the relation between general and subject-specific language skills and interindividual differences in the development of mathematics skills, applying a one-dimensional as well as a multi-dimensional conceptualisation of mathematics skills.

Our research interest in the role of language is from the perspective of migration-related disparities in mathematics skills. Thus, we will introduce control variables that proved relevant in this field to frame our study. Based on this, we present the theoretical mechanisms from the literature that explain the well-established finding that the development of mathematics skills is linked to students' language skills (Sect. 1.1). Subsequently, we present theoretical approaches on the role of subject-specific language skills, and introduce our understanding of this construct (Sect. 1.2). Prior findings on varying relations of language skills to different subdimensions of mathematics skills are presented in Sect. 1.3. Here, we connect these findings to the SPUR framework, which we used to conceptualise subdimensions of mathematics skills in arithmetic.

1.1 Language Skills in the Language of Instruction, and the Development of Mathematics Skills in Elementary School

Much interest in the role of language for mathematics learning in the literature stems from results on the comparably lower mathematics skills of students who come from families with a migrant background (Tarelli et al. 2012), or speak different languages at home than the language of instruction (Prediger et al. 2018; Ufer et al. 2013; Heinze et al. 2007). Two explanatory mechanisms for these disparities are discussed in the literature.

First, the families' socio-economic status, in particular the economic and cultural resources that a family can draw upon to support the students' learning, are discussed as a cause of migration-related disparities. Apart from the *Highest International Socio-Economic Index of Occupational Status* (HISEI; e.g. Gebhardt et al. 2013), the *Books-at-home Index* (Paulus 2009) has frequently been used as a proxy-measure of socio-economic status, which takes the approximate number of books in the students' home as a measure of the cultural capital of the family. From this perspective, it was shown that, in Germany, socio-economic status is particularly strongly related to mathematics skills (Baumert and Schümer 2001; Ehmke and Jude 2010) and that socio-economic status partially explains migration-related disparities in cross-sectional studies (Haag et al. 2012). Some longitudinal studies have demonstrated the relations of similar strength socio-economic status and language skills to mathematics learning gain (Paetsch et al. 2016) in school age. Morgan et al. (2011) found that only socio-economic status, and not reading skills, predicted mathematical learning gain in Kindergarten children. However, other longitudinal studies have not identified a relation of socio-economic status to mathematical learning gain in secondary school (Ehmke et al. 2006) and primary school (Ufer et al. 2013) when controlling for language skills. To summarise, existing evidence indicates that language skills do predict mathematical learning gain in school age beyond parents' socio-economic status, but it also warrants controlling measures of socio-economic status when studying the relation between language skills and mathematics learning.

As a second mechanism, students' language skills have been studied in the past. Indeed, migration-related differences in reading skills have been documented by several large-scale studies (Tarelli et al. 2012; Haag et al. 2012). A positive relation between mathematics skills and language skills has been demonstrated for reading comprehension at the end of grade 4 in Germany (Pietsch and Krauthausen 2006), but also as early as pre-school age (Grüßing and Schmitman gen. Pothmann 2007). Several cross-sectional studies have revealed that language skills have explanatory power regarding mathematics skills beyond measures of general cognitive ability (Heinze et al. 2007) or socio-economic status (e.g., Prediger et al. 2018; Ufer et al. 2013). In longitudinal analyses, not only overarching measures of language skills (Ufer et al. 2013), but also specific measures of reading comprehension (Paetsch et al. 2016) and vocabulary (Paetsch et al. 2015) could be identified as significant predictors of mathematics skills in elementary school. In contrast, grammatical skills could not explain a significant share of variance beyond these other two language measures (Paetsch et al. 2015). Different explanations have been proposed in the past for these effects of language skills on mathematics learning (Bochnik 2017).

1.1.1 Explanation by Text-based Test Instruments

A first explanation starts from the text-based test instruments used in most of the studies reported above. It might be that a certain share of the relation between language skills and measured mathematical learning gain is actually caused within the test situation itself, since students with lower language skills might have problems reading the test items. Along this line, prior research has identified features of mathematical test items, which specifically disadvantage learners with lower language

skills (e.g. Haag et al. 2013; Shaftel et al. 2006). These studies indicate that it is not technical mathematical vocabulary which poses problems for learners with low language skills, but complex grammatical constructions such as noun constructions, unknown or ambiguous terms, or the length of the text in the tasks (Haag et al. 2013; Martiniello 2008). Research on language accommodations of mathematics tests has studied the effects of eliminating or reducing language-related difficulties in tests by splitting long sentences, paraphrasing uncommon phrases or words, providing additional time, or implementing other kinds of support for learners with a home language different from the language of instruction (Abedi et al. 2004). In their meta-analysis of studies on these language accommodations, Kieffer et al. (2009) found only weak positive effects. Providing English-Language-Learners with dictionaries had the relatively largest, but still small effect. These results indicate that difficulties in test comprehension cannot plausibly be seen as the only explanation for the observed language-related disparities in mathematics skills, and other explanations must be taken into consideration.

1.1.2 Learning-related Explanations

These explanations of language-based disparities have, for example, focused on students' possibility of using *language-based opportunities to learn* during mathematics instruction. This comprises (a) comprehension of explanations by the teacher or fellow students (Erath et al. 2018), participation in the classroom discourse (Bochnik and Ufer 2016b), and access to written materials in mathematics instruction, such as textbooks (Österholm and Bergqvist 2013). Research on multilingual classrooms has illustrated that students are less involved in the classroom discourse and contribute their ideas less frequently, if their home language differs from the language of instruction (Civil 2008; Gorgorió and Planas 2001). This is particularly important since school-based mathematics learning is conceptualised as a discursive practice in research on language in mathematics learning (e.g. Moschkovich and Zahner 2018). For example, Steenpaß and Steinbring (2014) point out that even though concrete manipulatives are used to support mathematical communication in classrooms, the structure of these materials and the way mathematical concepts are encoded in their use is mediated verbally during instruction. Moreover, (b) research from very different perspectives has pointed out that weak language skills might exclude students from a meaningful participation in this discourse (Civil and Planas 2004; Abedi and Herman 2010; Heinze et al. 2007). Finally (c), other authors have described the role of language during mathematics learning and construction of mathematical meaning itself. Such approaches conceptualize mathematical thinking and knowing as a completely (Sfard 2008) or at least partially language-based process. This includes not only using language as one among several possible representations of mathematical concepts (e.g. Lesh et al. 2003), but also private or mental speech, such as during mathematical calculation (e.g. Ostad 2013), or self-explanations and verbalisations of mathematical operations based on symbols or concrete manipulatives (Sarama and Clements 2016).

   To describe students' difficulties and instructional support regarding these three uses of language in mathematics learning, the concept of language register has been

used successfully in the past (Prediger and Wessel 2018). Halliday (1978) defines a register as a 'configuration of semantic resources that the member of the culture associates with a situation type and is the meaning potential that is accessible in a given social context'. Registers are usually associated with specific linguistic features such as lexical features (e.g. vocabulary used), grammatical constructions, and text types, but also certain styles and forms of discourse. In this vein, past research has made the distinction between everyday registers (relating to Basic Interpersonal Communications Skills, BICS; Cummins 2008) and academic language registers (relating to Cognitive Academic Language Proficiency, CALP; Cummins 2008). The latter register is mostly associated with lexical features such as specific vocabulary or complex word constructions, and grammatical features such as complex sentences, passive voice, or noun phrases (e.g. Cummins 2008). The results described above on the explanatory power of skills in the language of instruction for the development of mathematics skills illustrate that mastering this academic register is an important prerequisite for benefitting from mathematical learning opportunities in school (Schleppegrell 2001). However, apart from these mostly subject-independent conceptualisations of language skills, the role of subject-specific language registers in mathematics learning has also been discussed.

1.2 Conceptualising and Measuring Subject-specific Language Skills

From a subject-related perspective, language serves as one of multiple, interlinked representations that constitute the meaning and content of mathematical concepts. Among symbolic, graphical, and concrete representations (e.g. in the form of manipulatives), verbal language is considered one central form of representation (Lesh et al. 2003). Prediger and Wessel (2013) characterise the use of language to describe mathematical concepts during classroom discourse by a specific register. This *verbal school register* is considered a part of the academic language register described by Schleppegrell (2001). Prediger and Wessel (2013) situate this register between an everyday register, which relates to Cummins BICS concept (Cummins 2008), and a *technical register*, which describes the disciplinary use of language in mathematics as a school subject. They hypothesise that the verbal school register is at least partially subject-specific, that it is crucial for an epistemic use of language during mathematics learning, but that usual classroom instruction offers few learning opportunities in school to acquire this register. Outside school, learning opportunities regarding this register are considered to vary with the degree of support offered by the students' families, resulting in restricted access to this register for learners, for example, from families who speak languages other than the language of instruction (e.g. Schleppegrell et al. 2004).

Related but not identical to the differentiation between school and technical register, some authors differentiate between formal and meaning-related registers (Prediger and Pöhler 2015; Prediger and Wessel 2018). In this perspective, the formal register is restricted to elaborating symbolic representations and techniques (Prediger and Wessel 2018), including technical terms which should be mastered after learning a mathematical concept (Prediger and Pöhler 2015). The meaning-related register, in contrast, is connected to the description of phenomena underlying math-

ematical concepts (e.g. bundling quantities as a basis of the decimal system). This register overlaps with the everyday register, the school register, but also technical terms for concepts learned earlier (Prediger and Wessel 2018). Thus, what was part of the formal register while learning one concept seems to be seen as part of the meaning-related register that is necessary to acquire more advanced concepts. Based on qualitative analyses, the meaning-related register is assumed to form a central basis for the acquisition of conceptual understanding (Götze 2018). In this sense, formal registers relate to outcomes of mathematics learning, while meaning-related registers relate to prerequisites as well as outcomes of mathematics learning.

1.2.1 Subject-specific Language Skills

In this contribution, we conceptualise *subject-specific language skills* as a construct describing how individual students can master formal and meaning-related aspects of the mathematical verbal school register. Considering the role of meaning-related language as a prerequisite of mathematics learning, we include technical language here to the extent that it is used to describe representations, examples, or phenomena related to previously learned mathematical concepts. Such language comprises the command of vocabulary specific to the school subject, but also more complex skills such as extracting mathematical structures described verbally, for example, in word problems (cf. Gabler and Ufer n.d.). Due to its close connection to knowledge of other representations of mathematical concepts, this construct overlaps naturally with mathematical knowledge and skills (Härtig et al. 2015), as well as subject-overarching (general) language skills (Dyrvold et al. 2015). Past research, however, has shown that it is possible to conceptualise and measure such subject-specific language skills sufficiently, independently from mathematics skills as well as general language skills (Bochnik and Ufer 2017). Different approaches have been developed in mathematics education but also for other subjects, to operationalise these subject-specific language skills.

1.2.2 Subject-specific Vocabulary

The specific vocabulary of mathematics, in particular technical terms, is a central tool for mathematical precision and accuracy (Maier and Schweiger 1999). The terms used in school to discuss mathematical concepts are part of the lexical level of the formal register. A particular difficulty is assumed in technical terms which have a different or more specific meaning in the context of (school) mathematics than in everyday language (e.g. 'function' or 'bundling'; Heinze et al. 2011). Based on interaction analyses of classroom transcripts, Schütte (2009) hypothesises that even though subject-specific vocabulary is often explicitly introduced in elementary school mathematics instruction, the terms are often automatized with little embedding into concrete discursive practices. When conceptualising students' subject-specific skills in terms of vocabulary, a consensus seems to be that 'knowing' a mathematical term cannot be completely independent of understanding the related mathematical concept to some extent (Dyrvold et al. 2015). However, while understanding that a concept relates to (potentially) deep connections between different

representations as well as real-world phenomena (Lesh et al. 2003), knowing a specific term may describe the ability to match the term to one or more examples or (alternative) representations of the concept, in contrast to examples or representations of other concepts. In our understanding, knowledge of subject-specific vocabulary goes beyond merely knowing technical terms. Subject-specific language skills in our sense include being able to connect vocabulary from the formal register to corresponding (acquired) mathematical concepts (e.g. 'increasing' or 'decreasing' with additive-subtractive structures). In this sense, knowledge of mathematical vocabulary requires some (perhaps shallow) understanding of the corresponding concept, but this should of course not be misunderstood as an indicator of deep understanding of the concept.

Schindler et al. (2019) report that over 60% of elementary school students in their samples from Switzerland and Germany could correctly identify examples for technical terms for arithmetic concepts such as 'addition' or 'difference'. Lower solution rates were found for geometric terms. Bochnik and Ufer (2017) report comparable solution rates (59% on average) for items similar to those used by Schindler et al. (2019), but higher solution rates (83% on average) when third-graders were asked to name mathematical concepts such as 'remainder'. Moreover, they found a significant correlation between knowledge of mathematical vocabulary and mathematics skills measured by a separate test. Bochnik and Ufer (2017) report that knowledge of subject-specific vocabulary cross-sectionally explained differences in mathematics skills beyond socio-economic status, general cognitive skills, and general language skills. In a study on science classrooms, Taboada (2012) demonstrates that knowledge of subject-specific vocabulary predicts science reading comprehension beyond knowledge of more general vocabulary. However, even in our broad conceptualisation, a sole focus on subject-specific vocabulary can be regarded as a reductionist perspective on subject-specific language skills, leading to the demand for alternative measures that go beyond knowledge of subject-specific vocabulary.

1.2.3 Subject-specific Text Comprehension

Current works on language in mathematics learning argue for the role of a meaning-related register of subject-specific language skills (Götze 2018). As indicated above, knowledge of mathematical vocabulary might be shallow in the sense that it does not connect to a deeper understanding of the underlying concepts. Götze (2018) reports that a substantial amount of vocabulary-related teaching is not connected to the conceptual meaning of the vocabulary. Thus, more comprehensive measures of subject-specific language skills are required, which cover meaning-related facets. However, such measures have rarely been proposed in the literature, so far. Most approaches start from the concept of C-Tests (Eckes and Grotjahn 2006). These consist of a set of short texts, in which parts of certain words have been erased. To complete these words, it is necessary to reconstruct the meaning of the whole text from the available parts using local and global coherences in the original text. As Schütte (2009) illustrates in his case studies, much of the contents of classroom discourse remain implicit. Inferring missing information in a C-Test from the available information may be seen as parallel to reconstructing implicit or missed informa-

tion when following a verbal classroom discourse. For this reconstruction, a merely shallow understanding of subject-specific language, in the sense of a formal register, is most likely not sufficient, but the meaning of the underlying concepts and their relations are required to infer the missing information. C-Tests are considered effective, valid, and reliable measures covering a broad range of components of language skills (Eckes and Grotjahn 2006). Özcan (2013) constructed an adapted version of C-Tests to measure subject-related language skills in chemistry, and found moderate positive correlations to subject-matter knowledge. Höttecke et al. (2017) studied C-Tests to measure domain-independent (general) language skills as well as subject-specific language skills in physics and sports. Based on moderate latent correlations between the three constructs, they could separate the two subject-specific language skill measures from the general language skill measure empirically. While the subject-related language skills in physics correlated substantially with the physics grade, the respective correlation was not significantly different from zero for the sports test. To explain the latter finding, the authors argue that subject-related language skills might play a minor role when assigning grades in sports classes. Bochnik and Ufer (2016a) developed an instrument for mathematical text comprehension based on C-Tests, and report that it explains differences in mathematics skills beyond socio-economic status, general cognitive skills, general language skills, and knowledge of subject-specific vocabulary cross-sectionally.

A major challenge in conceptualising and measuring subject-specific language skills is to separate the construct from general language skills and measures of mathematics skills. For example, Schindler et al. (2019) tried to reduce the linguistic and mathematical demands of their mathematical vocabulary items as much as possible by using a multiple-choice format with very brief and focused instructions. In contrast, a study by Bae et al. (2015) operationalised knowledge of mathematical vocabulary by asking the students to provide definitions for mathematical terms. It can be assumed that such measures placed much broader demands on students' mathematical knowledge than selecting from pre-defined alternatives relating to the corresponding concepts. Solving the items of Bae et al. (2015) requires much deeper knowledge of the concepts behind the technical terms to construct the required definitions. Results on the separability of the subject-specific and general language skills are scarce. However, using confirmatory factor analyses, Bochnik (2017) could demonstrate that subject-specific language skills as measured by her instruments can be separated psychometrically from general language skills and mathematics skills.

To summarise, first approaches to measure subject-related language skills have been developed in the past, including approaches that go beyond subject-specific vocabulary. First studies have provided evidence that subject-specific language skills can be separated empirically from subject-specific knowledge with respect to skills as well as general language skills (Özcan 2013; Höttecke et al. 2017). For example, in a latent variable analysis of data from a cross-sectional study with third-graders, Bochnik and Ufer (2016a) provide evidence that subject-specific language skills partially mediate the relation between general language skills and mathematics skills. Moreover, training studies from secondary school indicate a possible causal connection between subject-specific language skills and mathematics skills (e.g. Prediger

and Wessel 2018). However, it remains an open question to what extent subject-specific language skills are of particular importance for mathematics learning beyond other factors. For example, it might be argued that knowledge of mathematical vocabulary is of little relevance to explain mathematical learning progress over time, because this vocabulary is not connected to a deep conceptual understanding in the sense of a meaning-related register (Moschkovich 2002), or because it is explicitly treated in mathematics instruction anyway (Schütte 2009; Schindler et al. 2019). On the other hand, if a considerable amount of content, apart from technical vocabulary, remains implicit in classroom discourse (Schütte 2009), this might speak for the necessity to reconstruct subject-specific coherences when attending to classroom discourse, as it is measured by instruments based on C-Tests.

1.3 A Multi-dimensional Perspective on Mathematics Skills

Much research on the role of language skills for mathematics learning has conceptualised mathematics skills as a one-dimensional construct, but there are calls for a more differentiated perspective (Suurtamm et al. 2016). Based on results from the SOKKE study, for example, Heinze et al. (2007) and Ufer et al. (2013) argue that language skills might be more important to gain deep conceptual insights into mathematics, and of less importance to acquire more or less schematic calculation skills. Paetsch and Felbrich (2015) proposed that language skills should be particularly important when learning to solve mathematical word problems. Furthermore, educational standards worldwide proposed a multi-facetted view of mathematics skills, focusing on a variety of conceptual fields, mathematical practices, and levels of conceptual demand (NCTM 2000; CCSSI 2010; KMK 2004). A prominent differentiation has been between knowledge of mathematical concepts versus the ability to perform mathematical operations efficiently and correctly. Extending such works, Bleiler and Thompson (2013) describe a so-called SPUR approach towards mathematics assessment that differentiates the four dimensions *performing (schematic) operations (S), using properties to solve mathematical problems (P), using mathematical concepts as a functional tool outside mathematics (U)*, and *dealing with representations of mathematical concepts (R)*. While the framework originates from a curriculum development approach, these task types pose different demands on students' mathematical skills and they can be assumed to involve language skills in different ways to decode the task, to solve the task, and to acquire the knowledge necessary to solve the task.

Tasks of dimension S require that operations be performed, e.g. symbolic arithmetic procedures. They usually only contain a few words in the task text (e.g. 'Calculate!'). The necessary operations can—in principle—be solved using explicitly learned solution strategies, which can be applied on a purely symbolic level without necessarily making cognitive use of (meaning-related) language. Moreover, it has been argued that symbolic schemata might be easier to grasp in classroom instruction than a deeper understanding of the properties and relations of mathematical concepts (Heinze et al. 2007; Ufer et al. 2013). All this speaks against a strong role of language skills in performing such operations and learning the corresponding skills. Indeed, only weak language-related differences were found in grades 1 and 2

for performance on tasks of this type (Heinze et al. 2007; Ufer et al. 2013). However, results on the relation between procedural and conceptual knowledge indicate that knowledge of procedures is also acquired in interaction with conceptual knowledge (Schneider 2006; Rittle-Johnson et al. 2015). Learning to cope with more conceptual demands has, in turn, been connected to language skills in past studies (Heinze et al. 2007; Ufer et al. 2013). Thus, it remains an open question as to whether relations between language and mathematics skills might not be more pronounced in later grades of primary school.

For the other three types of assessment tasks, students' performance can be assumed to rely strongly on their conceptual knowledge: Rich knowledge of the concepts underlying a problem has been found to be a major predictor of mathematical problem-solving skills (Ufer et al. 2008). Knowledge about the meanings of mathematical concepts in real-world situations is considered an important prerequisite to apply mathematics to the problems in the real world (Stern 1994). Finally, it has been argued that the transfer between semantic representations such as visualisations and manipulatives of mathematical concepts and their symbolic representation is central to conceptual understanding in mathematics (Lesh et al. 2003). As noted above, in prior studies, the acquisition of conceptual knowledge has been considered to be specifically sensitive to students' language skills (Heinze et al. 2007; Ufer et al. 2013), so that it can be expected that these three dimensions, P, U, and R, are particularly sensitive to students' general and subject-specific language skills.

Among these three dimensions, tasks to use concept properties to solve mathematical problems (P) represent conceptual understanding of mathematical concepts most broadly. Even if language demands within the text can be limited for problem-solving items, the strong connection of problem-solving skills to conceptual understanding speaks for a close relation between performance on these tasks and subject-specific language skills. Using mathematical concepts in real-world situations (U) may be assumed to demonstrate the strongest relation to mathematical text comprehension skills, since reconstructing verbal descriptions of mathematical structures from the problem text is seen as a central barrier in word-problems (Paetsch and Felbrich 2015). This holds for solving the test items, but also during classroom instruction, because world-problems can be assumed to be mainly presented in text form also during instruction. In contrast, mathematical representations are usually explained verbally during classroom discourse (Steenpaß and Steinbring 2014). Connecting these representations to their mathematical meaning during instruction will most likely rely strongly on the availability of subject-specific vocabulary. To summarise, if there is any inter-individual relation between subject-specific language skills and mathematics skills beyond general language skills, performance in and development of the three conceptual dimensions (P, U, R) can be expected to correlate with subject-specific language skills. Among these, a pronounced connection can be expected between subject-specific text comprehension with skills and *using mathematical concepts to solve real-world problems* (U) as well as between subject-specific vocabulary with skills and *dealing with mathematical representations* (R).

Although these first hypotheses can be proposed in relation to the specific importance of general language skills to acquire different facets of mathematics skills, their differential dependency on general and subject-specific language skills has not

yet been approached empirically in the literature. The most common hypothesis is that the acquisition of conceptual understanding (mirrored specifically in the P, U, and R facets) will demonstrate a stronger link to language skills than skills relating to the execution of more or less schematic procedures.

2 A Longitudinal Study on the Role of Subject-specific Language Skills

2.1 Goals and Research Questions

Based on prior findings, the role of general skills in the language of instruction when learning mathematics can be seen as mostly established for the elementary school age (Heinze et al. 2007; Ufer et al. 2013) as well as for secondary school (Prediger et al. 2018). However, open questions remain. Firstly, several authors claim that language skills are particularly important for more conceptual demands such as problem-solving based on concept properties, functional use of mathematics, and representation of mathematical concepts, as compared to more schematic, technical demands (Heinze et al. 2007). However, the most evidence for this hypothesis stems from research in the first year of school. Its generalisability to later years of (elementary) schooling is not straightforward, since technical aspects of mathematics, such as calculation strategies or algorithms, also build on sophisticated mathematical ideas such as place value, or properties of arithmetic operations. Thus, one goal of the current study was to investigate the role of language skills for mathematics learning with grade 3 learners, and to analyse the effects of language skills on a variety of dimensions of mathematics skills. Our specific focus was on arithmetic, which forms a central part of the contents of elementary school mathematics instruction.

Q1   To what extent do general language skills predict mathematics skills at the end of grade 3, beyond prior mathematics skills, cultural capital, and general cognitive skills?

Theoretical and prior research indicates that general language skills in the language of instruction are related to the development of mathematics skills. Thus, we expected a significant contribution of general language skills above prior mathematics skills and other control variables when predicting overall mathematics skills at the end of grade 3. However, based on prior results from studies in earlier grades (Heinze et al. 2007), we expected that this would hold for the three subdimensions that reflect conceptual understanding (P, U, R), but not for schematic skills (S).

Secondly, prior research has highlighted the potential role of subject-specific language skills for mathematics learning. Although there are promising results of first intervention studies establishing learning opportunities for mathematics-related language skills, research explicitly measuring such skills and investigating their role for learning mathematics is scarce. This holds in particular for the elementary school age, longitudinal studies, and a nuanced view on the broad concept of subject-related language skills, ranging from having a command of mathematical vocabulary to the comprehension of mathematical structures presented in verbal form. To contribute to

filling this research gap, we included different measures of subject-related language skills.

Q2 To what extent can prior subject-specific language skills predict mathematics skills at the end of grade 3, above and beyond the general language skills, prior mathematics skills, cultural capital, and general cognitive skills?

Based on prior statements on the role of subject-specific language skills, we expected that they would predict overall mathematics skills at the end of grade 3 beyond the variables analysed for question Q1. Again, we expected a significant relation for the three dimensions that involve conceptual understanding, but not for schematic skills.

Q3 Which of the two facets of subject-specific language skills (mathematical vocabulary and mathematical text comprehension) contributes more strongly when predicting mathematics skills at the end of grade 3?

Little prior research has studied different facets of subject-specific language skills. Based on prior research, both facets can be expected to relate to learning gain regarding understanding of mathematical concepts. The first hypotheses described in 1.3 lead to the expectation that subject-specific text comprehension will be linked to learning to use concepts to solve real-world problems (U), while knowledge of subject-specific vocabulary will relate to learning to interpret and use mathematical representations (R).

2.2 Method: Sample, Instruments, Analysis Strategy

To answer the questions raised above, we conducted a longitudinal study with grade 3 students from eight elementary schools in southern Germany. Out of an initial sample of 383 students, $N = 237$ (111 female, 126 male) remained in the longitudinal sample after 8 of 24 classes did not participate in the second measurement. The number of students who did not speak German at home differed only marginally between the whole sample (39.8%) and the selected subsample (39.7%). This indicates that dropout from the study was not connected to students' language background. The students participated voluntarily and with informed parental consent after the research team had approached their teachers. The study consisted of two measurement times, one in autumn 2013 (T1, see also Bochnik and Ufer 2016a) and one in summer 2014 (T2), each consisting of two sessions on different days within two weeks.

2.2.1 Instruments—Mathematics Skills (T1 and T2)

For each of the two measurements, four scales of mathematics tasks were developed focusing on different dimensions of mathematics skills. These dimensions were based on an adaptation of the SPUR framework to the regional elementary school curriculum. Items for the dimension *schematic skills* (T1: 9 items; T2: 10 items) focused on symbolic two-digit (T1) and three-digit (T2) addition and subtraction respectively as well as single-digit multiplication and division. Items for the dimension *problem-solving based on properties* (T1: 15 items; T2: 16 items) of mathematical

concepts involved non-standard tasks that were posed in an intra-mathematical context (e.g. write down all multiplications that have 18 as a result). The dimension *using mathematical concepts* (T1: 6 items; T2: 8 items) involved single- and multistep word problems. Finally, *representing mathematical concepts* (T1: 18 items; T2: 18 items) focused on representing numbers and operations using mathematical visualisations such as the place-value table, base-ten-blocks, the number line, or a 10×10 point grid for multiplications. Figs. 1, 2, 3 and 4 in the appendix provide example items for each scale. All items were piloted and adapted repeatedly to include only those language demands which were necessary to communicate the task requirements. Separate mean scores were calculated for each of the four dimensions at each measurement. Moreover, we generated an overall score of mathematics skills for each of the two measurements by averaging the z-standardised scores from the four dimensions. The test scores were not linked between the two measurements, so that the scores cannot be compared directly between measurements of the same scale at different times.

2.2.2 Instruments—Subject-specific Language Skills (T1)

Three scales of items to measure subject-specific language skills at T1 were developed to study their relation to the development of mathematics skills. These scales mirror the differentiation of knowledge of subject-specific vocabulary and text comprehension.

Regarding *mathematical vocabulary*, two scales were developed. One scale (7 items) presented the students with tasks or pictures, in which a mathematical symbol or a part of the picture are marked with an arrow. Children were asked to describe the meaning of this symbol with an appropriate term. In some cases, a few written words provided the students with the start of a short sentence that had to be completed with the respective term (cf. Fig. 5 for an example item). Students' answers were coded as correct if the correct word could be identified. Thus, correct spelling was not necessary for a correct answer. Since students were asked to actively produce technical terms in these items, we refer to this scale as *knowledge of active subject-specific vocabulary*. The second subject-specific vocabulary scale (7 items) consisted of items that presented students with one single subject-specific term each. These terms referred to a mathematical concept (e.g. 'Verdopplungsaufgabe'—doubling task) or concrete operations that can be described arithmetically (e.g. 'dazukommen'—increasing). Below each term, five different possible descriptions of this operation with mathematical symbols were provided (cf. Fig. 6). The students were asked to select the alternative that best fitted the given term. Since students were provided with subject-specific vocabulary and had to select an appropriate example, we refer to this scale as *knowledge of passive subject-specific vocabulary*. Five alternative options were provided to reduce the chance of guessing the correct solution, while not overstraining students with too many alternatives. The terms used in both scales comprised typical technical terms from elementary school instruction (e.g. smaller number, neighbour number, remaining the same). Both scales did not presuppose deep conceptual knowledge of the related concepts, but required students to provide (active scale) a mathematical

term for examples or representations of a mathematical concept, or to identify (passive scale) the correct example or representation for a given mathematical term. For both scales of mathematical vocabulary, students had 10 min to complete all items. The ratio of correctly solved items per scale was calculated for each scale. As a measure of subject-specific vocabulary, we calculated the mean of the scores for the two scales.

To measure *mathematical text comprehension*, three short texts were developed that referred to mathematical concepts. One text provided a story from everyday life (cf. Fig. 7), two texts resembled written adaptations of texts that could arise in typical interactions in the mathematics classroom (e.g. describing a strategy to calculate $34 + 7$). The texts consisted of 43–67 words. In each text, we removed 7–10 words, leaving only the first one to four letters. To create these gaps, we deleted parts of words that carry the central mathematical meaning of the text, but which can be reconstructed from the mathematical meaning of the rest of the text. In particular, it was not possible to reconstruct the words using only the surface structure of the text. Students had 15 min to complete all three texts. We coded each completed word independently as correct or incorrect. Again, correct spelling was not necessary as long as a word could be identified that fitted the meaning of the text. For each text, the ratio of correctly filled-in gaps was calculated and the mean of these scores from the three texts was calculated.

To generate an overall indicator for subject-specific mathematics skills, we calculated the mean of the z-standardised scores for active mathematical vocabulary, passive mathematical vocabulary, and mathematical text comprehension.

2.2.3 Instruments—Control Variables (T1)

As a control variable, general language skills were measured using the SFD 3–4 (Hobusch et al. 2002), an approved instrument which has been used in prior research. It covers general vocabulary, prepositions, articles, and listening comprehension. General cognitive skills were surveyed with the Culture Fair Test, Scale 1 (CFT 1; Cattell et al. 1997). As a rough proxy measure of cultural capital, we asked students to estimate the number of books in their homes on a scale from 'very few (0–10)', 'one book shelf (11–25)', 'one bookcase (26–100)', and 'two bookcases (101–200)' up to 'more than 200 books'. Each answer choice was accompanied by a picture showing a number of books in the respective range arranged as in the answer option. This is an approved instrument to measure cultural capital using student reports at the elementary school level (Paulus 2009). To survey the language spoken at home, we asked each student to report the languages she or he can speak, which languages she or he most often uses to communicate with mother, father, and siblings (if available) and which language their parents use to communicate with each other. For every question, a range of typical languages was provided for selection. Children could select multiple languages, and add other languages which were not on the list, and were asked to underline the language they speak most. Based on these data, a language spoken at home different from the language of instruction was coded if the underlined language spoken with both parents was not German.

Table 1 Mean solution rates, standard deviations, and internal consistencies for the instruments used in the analyses

Instrument	First measurement (T1)			Second measurement (T2)		
	M	SD	α (N$_{it}$)	M	SD	α (N$_{it}$)
Mathematics skills (over all four dimensions)						
Overall score	0.62	0.19	0.81 (4)	0.54	0.20	0.83 (4)
Schematic mathematical skills	0.81	0.22	0.73 (8)	0.71	0.27	0.74 (8)
Problem-solving using properties of mathematical concepts	0.57	0.25	0.81 (14)	0.36	0.24	0.80 (14)
Use of mathematical concepts	0.60	0.26	0.64 (6)	0.54	0.27	0.71 (8)
Representing mathematical concepts	0.48	0.22	0.74 (15)	0.54	0.21	0.74 (15)
Subject-specific language skills (over all three dimensions)						
Active mathematical vocabulary	0.87	0.17	0.62 (7)	–		
Passive mathematical vocabulary	0.64	0.22	0.54 (7)	–		
Mathematical text comprehension	0.52	0.22	0.75 (3)	–		
Control measures						
General cognitive skills (IQ, CFT 1)	107.2	13.7	0.77 (36)	–		
General language skills (SFD 3–4)	0.81	0.16	0.91 (47)	–		
Cultural capital	3.50	1.13	–	–		

Note that scores from measurements of the same scale at different times cannot be compared directly, because the different scales were not linked between measurements

Cultural capital from 1: 0–10 books, 2: 11–25 books, 3: 26–100 books, 4: 101–200 books to 5: more than 200 books

M mean value, *SD* standard deviation

Table 1 provides an overview of descriptive statistics and internal consistencies for the instruments used in the analyses. For some scales, the number of items differs from the number of developed items, as some items did not fit quality criteria in the evaluation process (Bochnik 2017)[1]. Table 11 in the appendix contains information about the language-related features of the test items in this study and, for comparison, in the studies by Haag et al. (2013) and Paetsch et al. (2015).

2.2.4 Data Analysis

Multiple linear regression analyses were used to analyse the relations between the different constructs. The overall mathematics skill score at T2 was used as a dependent variable, and also the scores at T2 relating to each subdimension of the SPUR model. To describe inter-individual differences in students' development of mathematics skills and to model learning gain, mathematics skills in the first measurement were included as a covariate in all regression analyses. In a first baseline model (model 0), language spoken at home (German vs. non-German), cultural capital, and general cognitive skills were included as predictors. Subsequently, general

---

[1] Results of confirmatory factor analyses for the mathematics skills scales at T1 and T2 and for the scales for subject-specific language skills can be found in Table 7 in the appendix. Tables 8, 9 and 10 in the appendix contain the manifest correlations between the scales.

language skills (model 1) and the overall score for subject-specific language skills (model 2) were added to this initial model. In models 3a and 3b, the overall score for subject-specific language skills was replaced by the separate scores for mathematical vocabulary and mathematical text comprehension. Dependencies between students from the same class were controlled using the TYPE=COMPLEX option in the MPLUS software (Muthén and Muthén 2017).

3 Results

Tables 2, 3, 4, 5 and 6 show the results of planned stepwise regression analyses to investigate questions Q1–Q3.

3.1 Preliminary Analyses: Control Variables

In a first step, regression analyses were conducted for the overall mathematics skill score and for each of the four subdimension scores at T2, using the respective T1 measures as well as language spoken at home, cultural capital, and general cognitive skills as independent variables (Tables 2, 3, 4, 5 and 6, models 0, 0.S, 0.P, 0.U, and 0.R). In all models, the respective measure of mathematics skills from T1 was a significant predictor (e.g. Table 2, model 0, $\beta = 0.74$, $p < 0.001$). Students with better skills at T1 demonstrated better skills at T2, as expected. As prior research has found (Ufer et al. 2013), the language spoken at home was not a significant predictor of learning gain, when prior skills, cultural capital, and general cognitive skills were controlled (e.g. Table 2, model 0, $\beta = 0.02$, $p = 0.59$). For the overall score, cultural capital was the only additional significant predictor (Table 2, model 0, $\beta = 0.10$, $p = 0.01$), and this relation was significant for the dimensions *problem-solving with conceptual properties* (Table 4, model 0.P, $\beta = 0.16$, $p < 0.01$), *using concepts* (Table 5, model 0.U, $\beta = 0.14$, $p < 0.001$), and *representing concepts* (Table 6, model 0.R, $\beta = 0.11$, $p = 0.01$), while it was slightly weaker for *schematic skills* (Table 3, model 0.S, $\beta = 0.08$, $p = 0.07$). As for general cognitive skills, its relation to overall learning gain was only a tendency (Table 2, model 0, $\beta = 0.07$, $p = 0.07$), although it demonstrated a significant relation to learning gain for each of the four separate dimensions.

3.2 Prediction by General Language Skills (Question Q1)

When including general German language skills in the models, they significantly predicted mathematics skills at the end of grade 3 beyond prior mathematics skills (Table 2, model 1, $\beta = 0.15$, $p < 0.001$). The same pattern could be observed for *schematic skills* (Table 3, model 1.S, $\beta = 0.17$, $p < 0.01$), and was substantially stronger for *using concepts* (Table 5, model 1.U, $\beta = 0.30$, $p < 0.001$) and *representing concepts* (Table 6, model 1.R, $\beta = 0.29$, $p < 0.001$). For *problem-solving* (Table 4, model 1.P, $\beta = 0.06$, $p = 0.48$), general language skills showed no relation to learning gain.

To summarise, the results regarding question Q1 for the overall score are in line with prior findings in the literature: General language skills relate to mathematical

Table 2 Regression models for the dependent variable *overall mathematics score (T2)*

Independent variable	Model 0 Baseline model			Model 1 General language skills			Model 2 Subject-specific language skills			Model 3a Mathematical vocabulary			Model 3b Mathematical text comprehension		
	β	SE	p	β	SE	p	β	SE	p	β	SE	p	β	SE	p
Overall mathematics score (T1)	0.74	0.03	<0.001***	0.68	0.03	<0.001***	0.64	0.03	<0.001***	0.64	0.03	<0.001***	0.62	0.04	<0.001***
Language spoken at home (GE)	0.02	0.04	0.59	–0.01	0.04	0.75	0.00	0.04	0.93	0.00	0.04	0.93	–0.01	0.04	0.83
Cultural capital	0.10	0.04	0.01*	0.07	0.04	0.09ᵗ	0.08	0.04	0.08ᵗ	0.07	0.04	0.08ᵗ	0.08	0.04	0.05ᵗ
General cognitive skills	0.07	0.04	0.07ᵗ	0.06	0.04	0.16	0.05	0.05	0.26	0.05	0.05	0.25	0.05	0.05	0.30
General language skills	–	–	–	0.15	0.04	<0.001***	0.13	0.04	<0.01**	0.14	0.04	<0.01**	0.13	0.04	<0.01**
Subject-specific language skills															
Overall score	–	–	–	–	–	–	0.10	0.05	0.046*	–	–	–	–	–	–
Mathematical vocabulary	–	–	–	–	–	–	–	–	–	0.09	0.05	0.05ᵗ	–	–	–
Math. text comprehension	–	–	–	–	–	–	–	–	–	–	–	–	0.11	0.05	0.03*
R²	0.70	0.03	–	0.71	0.03	–	0.72	0.03	–	0.72	0.03	–	0.72	0.03	–

β standardized regression coefficient, *SE* standard error of β

ᵗ *p* < 0.10, **p* < 0.05, ***p* < 0.01, ****p* < 0.001

Table 3 Regression models for the dependent variable *schematic skills (T2)*

Independent variable	Model 0.S Baseline model			Model 1.S General language skills			Model 2.S Subject-specific language skills			Model 3a.S Mathematical vocabulary			Model 3b.S Mathematical text comprehension		
	β	SE	p	β	SE	p	β	SE	p	β	SE	p	β	SE	p
Schematic skills (T1)	0.51	0.08	<0.001***	0.46	0.09	<0.001***	0.45	0.08	<0.001***	0.45	0.08	<0.001***	0.44	0.09	<0.001***
Language spoken at home (GE)	−0.02	0.04	0.74	−0.06	0.05	0.20	−0.04	0.05	0.40	−0.04	0.05	0.40	−0.05	0.05	0.28
Cultural capital	0.08	0.05	0.07[t]	0.05	0.04	0.24	0.04	0.04	0.27	0.04	0.04	0.28	0.05	0.04	0.18
General cognitive skills	0.16	0.06	<0.01**	0.12	0.05	0.01*	0.08	0.05	0.13	0.08	0.05	0.11	0.08	0.05	0.13
General language skills	–	–	–	0.17	0.06	<0.01**	0.10	0.07	0.15	0.11	0.07	0.13	0.09	0.07	0.17
Subject-specific language skills															
Overall score	–	–	–	–	–	–	0.19	0.07	<0.01**	–	–	–	–	–	–
Mathematical vocabulary	–	–	–	–	–	–	–	–	–	0.18	0.07	<0.01**	–	–	–
Math. text comprehension	–	–	–	–	–	–	–	–	–	–	–	–	0.19	0.06	<0.01**
R^2	0.38	0.07	–	0.40	0.07	–	0.42	0.07	–	0.42	0.07	–	0.42	0.07	–

β standardized regression coefficient, *SE* standard error of β

[t] $p<0.10$, *$p<0.05$, **$p<0.01$, ***$p<0.001$

Table 4 Regression models for the dependent variable *problem solving with concept properties (T2)*

Independent variable	Model 0.P Baseline model			Model 1.P General language skills			Model 2.P Subject-specific language skills			Model 3a.P Mathematical vocabulary			Model 3b.P Mathematical text comprehension		
	β	SE	p	β	SE	p	β	SE	p	β	SE	p	β	SE	p
Problem solving (T1)	0.52	0.07	<0.001***	0.50	0.08	<0.001***	0.50	0.07	<0.001***	0.49	0.07	<0.001***	0.48	0.08	<0.001***
Language spoken at home (GE)	0.03	0.05	0.56	0.01	0.05	0.81	0.02	0.05	0.75	0.02	0.05	0.76	0.01	0.05	0.77
Cultural capital	0.16	0.05	<0.01**	0.15	0.06	<0.01**	0.15	0.06	<0.01**	0.15	0.06	<0.01**	0.15	0.05	<0.01**
General cognitive skills	0.21	0.06	<0.001***	0.20	0.06	<0.01**	0.19	0.07	<0.01**	0.19	0.07	<0.01**	0.19	0.07	<0.01**
General language skills	–	–	–	0.06	0.09	0.48	0.06	0.10	0.56	0.06	0.10	0.55	0.05	0.09	0.61
Subject-specific language skills															
Overall score	–	–	–	–	–	–	0.03	0.06	0.58	–	–	–	–	–	–
Mathematical vocabulary	–	–	–	–	–	–	–	–	–	0.04	0.06	0.54	–	–	–
Math. text comprehension	–	–	–	–	–	–	–	–	–	–	–	–	0.05	0.08	0.50
R²	0.53	0.05	–	0.53	0.05	–	0.53	0.05	–	0.53	0.05	–	0.53	0.05	–

β standardized regression coefficient, *SE* standard error of *β*

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table 5 Regression models for the dependent variable *using concepts (T2)*

Independent variable	Model 0.U Baseline model			Model 1.U General language skills			Model 2.U Subject-specific language skills			Model 3a.U Mathematical vocabulary			Model 3b.U Mathematical text comprehension		
	β	SE	p	β	SE	p	β	SE	p	β	SE	p	β	SE	p
Using concepts (T1)	0.55	0.05	<0.001***	0.44	0.06	<0.001***	0.40	0.06	<0.001***	0.41	0.06	<0.001***	0.38	0.07	<0.001***
Language spoken at home (GE)	0.08	0.05	0.13	−0.01	0.05	0.89	0.01	0.05	0.90	0.01	0.05	0.90	−0.01	0.05	0.92
Cultural capital	0.14	0.03	<0.001***	0.09	0.03	<0.01**	0.09	0.03	<0.01**	0.09	0.03	<0.01**	0.10	0.03	<0.01**
General cognitive skills	0.13	0.04	0.001**	0.09	0.05	0.05^t	0.06	0.05	0.24	0.06	0.05	0.21	0.05	0.05	0.31
General language skills	–	–	–	0.30	0.06	<0.001***	0.25	0.06	<0.001***	0.26	0.06	<0.001***	0.23	0.06	<0.001***
Subject-specific language skills															
Overall score	–	–	–	–	–	–	0.17	0.06	<0.01**	–	–	–	–	–	–
Mathematical vocabulary	–	–	–	–	–	–	–	–	–	0.16	0.06	<0.01**	–	–	–
Math. text comprehension	–	–	–	–	–	–	–	–	–	–	–	–	0.19	0.07	<0.01**
R^2	0.50	0.05	–	0.55	0.04	–	0.57	0.04	–	0.57	0.04	–	0.57	0.04	–

β standardized regression coefficient, SE standard error of β

^{t}p<0.10, *p<0.05, **p<0.01, ***p<0.001

Table 6 Regression models for the dependent variable *representing concepts (T2)*

Independent variable	Model 0.R Baseline model			Model 1.R General language skills			Model 2.R Subject-specific language skills			Model 3a.R Mathematical vocabulary			Model 3b.R Mathematical text comprehension		
	β	SE	p	β	SE	p	β	SE	p	β	SE	p	β	SE	p
Representing concepts (T1)	0.45	0.07	<0.001***	0.38	0.07	<0.001***	0.33	0.07	<0.001***	0.33	0.07	<0.001***	0.32	0.07	<0.001***
Language spoken at home (GE)	0.11	0.07	0.12	0.03	0.07	0.71	0.04	0.07	0.56	0.04	0.07	0.57	0.03	0.07	0.66
Cultural capital	0.11	0.05	0.03*	0.04	0.06	0.49	0.04	0.06	0.50	0.04	0.06	0.50	0.05	0.07	0.43
General cognitive skills	0.13	0.05	0.01*	0.07	0.06	0.24	0.05	0.06	0.38	0.05	0.06	0.36	0.03	0.06	0.60
General language skills	–	–	–	0.29	0.06	<0.001***	0.25	0.06	<0.001***	0.26	0.06	<0.001***	0.22	0.05	<0.001***
Subject-specific language skills															
Overall score	–	–	–	–	–	–	0.14	0.05	<0.01**	–	–	–	–	–	–
Mathematical vocabulary	–	–	–	–	–	–	–	–	–	0.14	0.05	<0.01**	–	–	–
Math. text comprehension	–	–	–	–	–	–	–	–	–	–	–	–	0.21	0.05	<0.001***
R^2	0.36	0.06	–	0.41	0.06	–	0.42	0.06	–	0.42	0.06	–	0.43	0.06	–

β standardized regression coefficient, *SE* standard error of β

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

learning gain (i.e. predicted mathematics skills in grade 3, controlling for prior mathematics skills). That this relation is significant not only for *using concepts (U)* and *representing concepts (R)*, but with a slightly weaker relation also for *schematic skills (S)*, differs from prior cross-sectional findings (Heinze et al. 2007) as well as longitudinal findings for grade 1 (Ufer et al. 2013). Finally, also not in line with the hypotheses found in prior research, general language skills did not relate to learning gain regarding *problem solving based on concept properties (P)*.

3.3 Prediction by Subject-specific Language Skills (Questions Q2 and Q3)

Regarding subject-specific language skills (question Q2), the predictivity of an overall score combining all three indicators (active vocabulary, passive vocabulary, mathematical text comprehension) was first examined. For the overall score (Table 2, model 2, $\beta = 0.10$, $p = 0.046$), as well as for the single subdimensions *schematic skills* (Table 3, model 2.S, $\beta = 0.19$, $p < 0.01$), *using concepts* (Table 5, model 2.U, $\beta = 0.17$, $p < 0.01$), and *representing concepts* (Table 6, model 2.R, $\beta = 0.14$, $p < 0.01$), subject-specific language skills significantly predicted mathematics skills in grade 3 beyond general language skills and the respective prior mathematics skills. While the strengths of these relations were similar, subject-specific language skills did not significantly predict *problem solving based on concept properties* (Table 4, model 2.P, $\beta = 0.03$, $p = 0.58$).

In two final models, the overall score for subject-specific language skills was replaced by the scores for mathematical vocabulary (models 3a) with respect to mathematical text comprehension (models 3b) to analyse the relative predictivity of these two measures (question Q3). Due to the substantial correlation between the two measures ($r = 0.72$, $p < 0.001$), we did not analyse them in a joint regression model. For the overall score, mathematical text comprehension demonstrated a similar relation to learning gain as the overall measure of subject-specific language skills (Table 2, model 3b, $\beta = 0.11$, $p = 0.04$), but the respective relation did not reach significance for mathematical vocabulary (Table 2, model 3a, $\beta = 0.09$, $p = 0.05$). For the subdimensions *schematic skills* (Table 3, model 3a.S, 3b.S), *using concepts* (Table 4, model 3a.U, 3b.U), and *representing concepts* (Table 6, model 3a.R, 3b.R), both indicators showed similar predictivity as the overall score for subject-specific language skills. Descriptively, these relations seemed to be slightly stronger for mathematical text comprehension than for mathematical vocabulary, in particular when predicting learning gain in *representing concepts* (Table 6, models 3a.R, 3b.R). To test this difference statistically, we compared the fit of two models including both predictors, mathematical text comprehension, and mathematical vocabulary: One model with the two standardised regression coefficients constrained to be equal, and one with freely estimated standardised regression coefficients. A Chi-square test of model fit did not indicate a significantly worse fit of the constrained model compared to the model with free standardised regression coefficients ($\chi^2 (1) = 1.13$, $p = 0.29$). The same pattern was observed for the overall score ($\chi^2 (1) = 0.01$, $p = 0.92$), schematic skills ($\chi^2 (1) = 0.01$, $p = 0.93$), using problem solving based on concept properties ($\chi^2 (1) = 0.05$, $p = 0.82$), and using concepts ($\chi^2 (1) = 0.23$, $p = 0.63$).

4 Discussion

The main goal of the current contribution was to study the predictive value of mathematics-specific language skills beyond prior mathematics skills, general language skills, general cognitive abilities, and cultural capital on mathematics learning during grade 3. Furthermore, apart from a one-dimensional conceptualisation of mathematics skills, a differentiated perspective was adopted in accordance with the SPUR-approach (Bleiler and Thompson 2013). We will first discuss the results for the one-dimensional conceptualisation, and then for the four subdimensions of mathematics skills. We will derive final conclusions in light of the study's limitations and highlight central implications for research and development.

4.1 Results from a One-dimensional Perspective on Mathematics Skills

Regarding the variables which predict overall mathematical learning gain, our results point to a central role of language skills above related measures, such as general cognitive abilities, language spoke at home, and cultural capital. These results on general language skills replicate prior results from the primary (Heinze et al. 2007; Ufer et al. 2013) and secondary level (Prediger et al. 2018) and underpin the well-established relevance of general language skills for mathematics learning.

Our results on subject-specific language skills extend these results. In our study, subject-specific language skills demonstrated a significant relation to mathematical learning gain beyond general language skills and the other control variables (e.g. Table 2, model 2). This result indicates that it is not only important to master aspects of the academic language register in general. Having a certain command of a subject-specific school register as conceptualised, for example, by Prediger and Wessel (2013) has additional impact on elementary school students' mathematical learning progress. Intervention studies in secondary school have provided first evidence (Prediger and Wessel 2018) that a causal effect might stand behind this relation. This underpins the importance of subject-specific language support which is closely connected to the concepts at hand during mathematics instructions.

Beyond intervention studies, in which usually only a few, quite broad dimensions of language skills and support can be studied systematically, our longitudinal, correlational approach allows us to explore the importance of different facets of subject-specific language skills separately. Of course, the inclusion of relevant competing explanatory variables, such as cultural capital, general cognitive skills, and of course general language skills, is vital to arrive at valid conclusions. Controlling for these variables, our results indicate that in particular mathematical text comprehension relates to the acquisition of mathematics skills. Knowledge of mathematical vocabulary demonstrated a descriptively positive, but not significant relation to learning gain for the overall measure of mathematics skills. This underpins hypotheses from qualitative research and theoretical assumptions on the role of skills relating to meaning-related registers that go beyond isolated knowledge of technical terms (Moschkovich 2002; Götze 2018). Skills in mathematical text comprehension were measured using an instrument based on the idea of C-Tests. Similar instruments of subject-specific language skills have also been developed for other school subjects

in the past (Höttecke et al. 2017; Özcan 2013), but results on their longitudinal predictivity have been rare, up to now. The current study is the first to provide direct longitudinal evidence regarding the importance of the two facets of subject-specific language skills for the acquisition of mathematics skills.

4.2 Results from a Multi-dimensional Perspective on Mathematics Skills

Following hypotheses on diverging effects of language skills on different dimensions of mathematics skills, we applied an instrument to measure mathematics skills that was conceptualised to provide separate scores for each of the four SPUR dimensions (Bleiler and Thompson 2013). The results here largely reproduce the effects of language skills observed for the overall learning gain: Subject-specific language skills predict learning beyond general language skills and the other control variables. In these cases, the relations can be traced back to both indicators, mathematical vocabulary and mathematical text comprehension, with a slightly stronger relation for mathematical text comprehension.

Studying the effects of subject-specific text comprehension was a central innovation in our study. As expected, this measure demonstrated significant relations to learning gain in *representing mathematical concepts (R)* and in *using mathematical concepts (U)*. Both dimensions are typical content of elementary school mathematics instruction, and both dimensions have been hypothesised to be strongly related to language skills before (Steenpaß and Steinbring 2014 for representations; Paetsch and Felbrich 2015 for word problems), because the meaning of mathematical representations is negotiated mostly verbally in the classroom. Similarly, mastering the different verbal means to represent mathematical concepts in word problems (cf. Gabler and Ufer n.d.) seems not only related to mathematical vocabulary, but to mathematical text comprehension. Thus, these results illustrate the importance of subject-specific text comprehension beyond other measures such as general language skills for some dimensions of mathematics skills, and they extend prior research on approaches to measure such skills (Höttecke et al. 2017; Özcan 2013).

Moreover, the relation to mathematical vocabulary is interesting, since Haag et al. (2013) find that mathematical vocabulary is not a reason for second language learners' difficulties in mathematics assessments. Although it has been pointed out that mathematical vocabulary is usually explained explicitly in the classroom (Haag et al. 2013), it seems plausible that its predictive power in our study mirrors the role of subject-specific vocabulary when learning and teaching mathematics. The descriptively slightly stronger relation of mathematical text comprehension with learning gain is, again, in line with theoretical assumptions and hypotheses from qualitative research on the role of meaning-related registers (Moschkovich 2002; Götze 2018).

Due to the restricted sample size and the strong correlation between the two measures of subject-specific language skills, we could not find any indications that one of the two measures is a stronger predictor of mathematics skills than the other. This is particularly interesting since prior analysis of data from the first measurement has demonstrated that knowledge of mathematical vocabulary can be differentiated from mathematical text comprehension (manifest correlation r=0.73; cf. Table 8

in the appendix). It remains an important direction for further research to examine the relation between and the effect of these two measures in detail and empirically test assumptions on the specific role of skills regarding meaning-related registers. Our test of knowledge of mathematical vocabulary only requires a relatively shallow connection between mathematical vocabulary and conceptual knowledge. Of course, it is a desirable goal of mathematics instruction in general that students make deep connections with the meanings of these terms (Götze 2018). However, it remains an open question as to whether these differences in deep connections can further explain learning, or whether knowledge of vocabulary, as we operationalised it, might not be sufficient to support further learning. This would be of specific interest since such understanding might be easier to acquire in initial learning, and could form a basis for deeper elaboration in subsequent learning.

Beyond these commonalities to the one-dimensional analysis, there are interesting deviations from the picture observed there. Firstly, both general as well as subject-specific language skills did not explain learning gain regarding *problem solving based on concept properties* (dimension P) beyond the control variables. For this measure, learning gain was best explained by general cognitive abilities and the families' cultural capital. This result contradicts expectations from the literature that language skills are most central for more conceptual dimensions of mathematics skills. One reason for this result could be that we deliberately selected non-routine tasks to measure subdimension P. These tasks require students to develop their own solution approaches, based on their knowledge of the involved concepts. At least analyses from secondary school have demonstrated that such tasks are quite rarely implemented in classroom instruction (Jordan et al. 2008). Thus, being able to develop strategies to approach these tasks might be less a matter of school-based learning processes and perhaps more one of informal learning opportunities outside school. Individually, the use of such informal learning processes can plausibly be assumed to be related to general cognitive skills, and it is also plausible that they are more intensively initiated in families that have more cultural capital at their disposal. A similar explanation would not apply to representing and using mathematical concepts (which showed significant relations to language skills), since the representations and word problems used in our study are typical content of elementary school mathematics instruction. However, further research is necessary to investigate this result in more detail.

Secondly, regarding the hypothesis that general and subject-specific language skills are of lower importance for the acquisition of more *schematic aspects of mathematics skills* (dimension S) than of more conceptual dimensions, our study could not replicate the cross-sectional results (Heinze et al. 2007) and longitudinal results (Ufer et al. 2013) from grade 1. Although the relation of general language skills was descriptively weaker for schematic skills than for representing and using concepts, schematic skill related significantly to general and subject-specific language skills. One explanation might be the difference in mathematical content between grade 1 and grade 3: While numbers until 20 in grade 1 can, in principle, be represented and discussed as compositions of single objects, this is no longer possible for the numbers up to 1000 in grade 3. Here, handling these numbers, also in terms of calculation techniques and procedures, requires substantial conceptual

discourse regarding the place-value system (Fuson et al. 1997), as well as calculation strategies (Heinze et al. 2018). Thus, the increasingly complex conceptual background of seemingly schematic calculation procedures in grade 3 might lead to a stronger impact of language skills on the acquisition of calculation procedures when progressing through the elementary school grades. This interpretation would be in line with findings indicating bidirectional relations between conceptual and procedural knowledge when learning mathematics (Schneider 2006; Rittle-Johnson et al. 2015).

Finally, there is an interesting difference between our results on *using mathematical concepts* and Paetsch and Felbrich's (2015) analysis of learning gain in mathematical word problems during grade 3. While we find a relation between initial language skills and mathematics skills at T2, when controlling for mathematics skills at T1, Paetsch and Felbrich found that initial language skills do not predict learning gain in word problems. In their study, the same students who progressed in language skills during grade 3 also improved in solving word problems. On the one hand, this result might be due to methodological differences: While our study applied autoregressive models to investigate learning gain, Paetsch and Felbrich explicitly modelled learning gain as a latent IRT variable. A different explanation might lie in the different samples: While our study did not select students with a specific profile, Paetsch and Felbrich's analysis only included students who spoke German as a second language. In combination, both results might indicate that language skills are important to make progress in terms of word problems (our study), but that this relation is not observable if language skills are particularly low at the start of grade 3. In the latter case, e.g. for many students with German as a second language, it might be more important to reach a sufficient level of language skills during grade 3, even if their skills were comparably low initially.

4.3 Limitations

Although the presented study can add new insights on the role of general and subject-specific language skills in mathematics learning, the results should be taken with some qualifications. Naturally, the design of our study did not allow us to describe the development of mathematics skills in absolute terms, but only to relate interindividual differences in this development to other variables. Since the latter was the main purpose of our study, we assume that this is a minor limitation. Further, we decided, firstly, to focus on arithmetic which represents only a part of the contents of elementary mathematics instruction (KMK 2004). The results of Schindler et al. (2019) indicate that subject-specific vocabulary might involve more nuanced complexities in geometry, for example in terms of mathematical class inclusion (e.g. every square is a rectangle). Thus, the effects of subject-specific vocabulary might be even stronger in geometry if handling geometric terms indeed turns out to be more difficult. As regards mathematical text comprehension, studying this construct in other mathematical content areas will be an important step for future research. Although arithmetic skills form a central goal of elementary school instruction, an extension of this line of research to geometry and, possibly, chance and data might be of interest.

Secondly, the longitudinal, correlational design of our study poses some fundamental restrictions. For example, although a substantial number of relevant control variables were taken into account, strong causal conclusions cannot be drawn from this kind of study. Experimental intervention studies are necessary to provide such causal evidence. First studies which show an effect of interventions focusing on subject-specific language on mathematics skills exist for secondary school level (Prediger and Wessel 2018). The study by Schindler et al. (2019) at the elementary school level did not report effects on mathematical skills beyond mathematical vocabulary. Although such intervention studies are of vital interest to establish causal connections between general and subject-specific language skills and mathematics skills, they are restricted in the resolution of different dimensions of general or subject-specific language skills. Correlational studies, such as the one presented here, can provide more differentiated insights into the complex relation between different facets of language skills and mathematics skills.

Thirdly, we decided to use self-developed instruments in our study. This restricts comparability to other studies (which also use a variety of instruments). The 'schematics skills' scale, however, contains similar tasks as can be found in most standardised tests of elementary school mathematics (e.g. Roick et al. 2004). The final word on the validity of the SPUR framework to describe mathematics skills broadly has not yet been spoken. However, applying this framework connects our work to current assessment and teaching approaches, which emphasise the mathematical practices students conduct with a given concept, i.e. how students make use of a concept (e.g. CCSSI 2010; KMK 2004). Concerning reliability, in particular the instruments on subject-specific language skills warrant further optimisation. In particular, the reliability coefficient for active vocabulary is low. Unless this is resolved, the findings might be taken with a grain of salt and require further studies with improved instruments.

Finally, our study with $N = 237$ grade 3 students has, of course, a restricted power to identify relations between students' learning prerequisites and their learning gain. In particular, the large number of control variables makes it difficult to identify cumulative effects beyond their explanatory power. For example, comparing the predictive strength of mathematical vocabulary and mathematical text comprehension with sufficient power would require a larger sample size, due to the high correlation between the two measures. Our study could reliably detect some relations between subject-specific language skills and mathematics skills. When considering the size of the regression coefficients descriptively, it seems plausible that studies with larger samples will be able to establish more of the detailed relations. For the current study, this means that insignificant coefficients should not be taken as evidence for strong claims that certain learning prerequisites are unimportant. However, for educational practice, it is primarily vital to identify those prerequisites that have the strongest relation to learning gain. Although larger studies might provide an even more nuanced picture in the future, the results obtained with our sample still have important implications.

4.4 Conclusions and Implications

To our knowledge, the present study is the first to measure different dimensions of mathematics-specific language skills explicitly and to relate them to mathematical learning gain longitudinally. Our results underpin the importance of general *and* subject-specific language skills for mathematics learning already in elementary school. This indicates that approaches to support students' language acquisition within subject-related instruction (e.g. Schindler et al. 2019; Wessel and Erath 2018) may provide effective means to foster mathematical learning beyond the acquisition of subject-specific language skills. Accordingly, developing such approaches based on educational theories of the role of language in mathematics learning as well as empirical evidence on the mechanisms assumed in these theories remains a central desideratum (Gabler and Ufer n.d.). In this context, it will be of interest if migration- or language-related disparities in mathematics achievement can be reduced by a stronger focus on subject-specific language skills in classroom education.

Our study integrates two perspectives on subject-specific language skills: A perspective focusing on subject-specific vocabulary (Schindler et al. 2019; Taboada 2012) and a perspective focusing on measures of subject-specific text comprehension, often with instruments based on the idea of C-Tests (Höttecke et al. 2017; Özcan 2013). On the one hand, it was found that both types of measures can reliably be combined into a comprehensive measure of subject-specific language skills. Moreover, the effects of these two indicators seemed to be similar in our study. Future research is necessary to investigate if a focus on only one of the two facets, mathematical vocabulary or mathematical text comprehension, is suited to support students' learning.

Finally, our study aimed to investigate hypotheses regarding the varying role of general and subject-specific language skills on the acquisition of deep knowledge about mathematical concepts vs. of more schematic calculation procedures. This differentiation was made in an ad-hoc manner in prior studies, e.g. based on subscale analyses (Heinze et al. 2007) or factor analyses (Ufer et al. 2013) of existing tests. For this study, we designed an instrument to measure mathematics skills based on a multi-dimensional conceptualisation of mathematics skills (SPUR approach: Bleiler and Thompson 2013). The results point towards a more pronounced role of language skills also for the acquisition of calculation procedures in grade 3 than observed on the prior studies in grade 1. If this pattern of results can be sustained, it would indicate a danger that language-related problems in mathematics might be difficult to observe based on calculation skills in lower school grades, but accumulate and extend to all dimensions of mathematics skills with increasing complexity of school mathematics. This warrants an early diagnosis of schematic as well as conceptual mathematics skills, but also of general and subject-specific language skills already at early levels of elementary school to provide instructional support as early as possible. Instruments such as those developed for this study might be adapted to these earlier levels and serve as a basis for these diagnostics.

Appendix

Item Examples and Translations

Fig. 1 Example item for the dimension *schematic mathematical skills* (T1) (Translation: 'Calculate.')

Rechne aus.

38 + 43 = ☐   27 + 68 = ☐

Fig. 2 Example item for the dimension *problem-solving using properties of mathematical concepts* (T1) (Translation: 'Put in the correct digits')

Setze die richtigen Ziffern ein.

2☐ + 17 = ☐3

Fig. 3 Example item for the dimension *use of mathematical concepts* (T1) (Translation: 'Markus has 11 strawberries. He has 6 strawberries less than Anna. How many strawberries has Anna?')
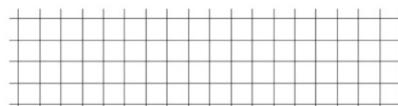
Markus hat 11 Erdbeeren. Er hat 6 Erdbeeren weniger als Anna. Wie viele Erdbeeren hat Anna?

Anna hat ☐ Erdbeeren.

Fig. 4 Example item for the dimension *representing mathematical concepts* (number line) (T1) (Translation: 'Which numbers could that be?')
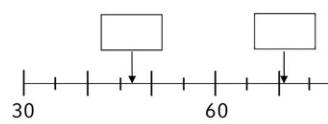
Welche Zahlen könnten das sein?

30   60

Fig. 5 Example item for active mathematical vocabulary (T1) (Translation: '7 is ...')

7 ist ...

7 < 13
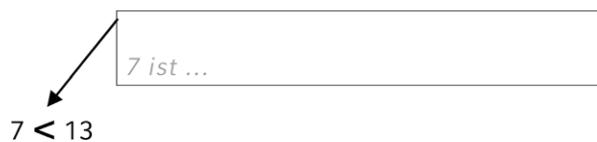
Fig. 6 Example item for passive mathematical vocabulary (T1) (Translation: 'Doubling 14')

Verdopplungsaufgabe zu 14				
☐	☐	☐	☐	☐
14-14	14+2	14+14	28:2	14:2

Fig. 7 Excerpt from an example text for subject-specific text comprehension. The English translation of the item does not contain gaps, since it was not possible to design gaps in the translation that would exactly correspond to the gaps in the original German item (Translation: 'Thomas and Timo get pocket money every week. They both get the same amount of pocket money. Timo buys a bigger bag of sweets than Thomas every week. Timo pays more.')

Taschengeld

Thomas und Timo bekommen jede Woche Taschengeld.

Die beiden bekommen gleich v——————— Taschengeld.

Timo kauft jede Woche eine gr——————— Tüte

Süßigkeiten als Thomas. Timo bezahlt m———————.

Er hat n——————— dem Einkauf w——————— Geld als

Thomas.

Jetzt hat T——————— mehr Geld als T———————.

Sie freuen sich auf das nächste Taschengeld.

Further Information on the Applied Instruments

Table 7 Results of confirmatory factor analyses for mathematics skills (T1 and T2) and subject-specific language skills (T1)

Model	N	Df	χ^2	χ^2/df	RMSEA	CFI
Mathematics skills T1 (4 factors)	383	854	1025.3	1.20	0.023	0.94
Mathematics skills T2 (4 factors)	237	939	1095.1	1.17	0.026	0.92
Subject-specific language skills T1 (3 factors)	383	116	204.2	1.76	0.045	0.91

The sample of $N = 383$ for T1 includes students from seven further classrooms which did not participate in the longitudinal study

The larger sample was used for the confirmatory factor analysis to obtain more reliable estimates

Relevant cutoff criteria:

χ^2/df $0 \leq \chi^2/df \leq 2$ for good fit (Schermelleh-Engel et al. 2003)

CFI CFI ≥ 0.90 for acceptable fit (Hu and Bentler 1999)

RMSE RMSEA ≤ 0.05 for good fit (Schermelleh-Engel et al. 2003)

Table 8 Bivariate manifest Pearson correlations between the measures of mathematics skills and subject-specific language skills at T1

	Measure (T1)				
Measure (T1)	P	U	R	VO	TC
Schematic mathematical skills (S)	0.65	0.47	0.34	0.31	0.41
Problem-solving using properties (P)	–	0.57	0.49	0.47	0.58
Use of mathematical concepts (U)	–	–	0.60	0.45	0.58
Representing mathematical concepts (R)	–	–	–	0.51	0.55
Mathematical vocabulary (VO)	–	–	–	–	0.73
Mathematical text comprehension (TC)	–	–	–	–	–

$N = 237$, all correlations are significantly different from zero
$p < 0.001$

Table 9 Bivariate manifest Pearson correlations of the measures of mathematics skills and subject-specific language skills at T1 with mathematics skills at T2

	Measure (T2)			
Measure (T1)	S	P	U	R
Schematic mathematical skills (S)	0.59	0.47	0.43	0.49
Problem-solving using properties (P)	0.50	0.68	0.62	0.62
Use of mathematical concepts (U)	0.52	0.51	0.68	0.53
Representing mathematical concepts (R)	0.45	0.48	0.54	0.56
Mathematical vocabulary (VO)	0.40	0.40	0.49	0.45
Mathematical text comprehension (TC)	0.45	0.49	0.59	0.53

$N = 237$, all correlations are significantly different from zero
$p < 0.001$

Table 10 Bivariate manifest Pearson correlations between the measures of mathematics skills at T2

	Measure (T2)		
Measure (T2)	P	U	R
Schematic mathematical skills (S)	0.57	0.51	0.47
Problem-solving using properties of mathematical concepts (P)	–	0.63	0.57
Use of mathematical concepts (U)	–	–	0.58
Representing mathematical concepts (R)	–	–	–

$N = 237$, all correlations are significantly different from zero
$p < 0.001$

Table 11 Language-related characteristics of the items in the current study, and in the studies by Haag et al. (2013) and Paetsch et al. (2015)

Dimension	This article					Haag et al. (2013)	Paetsch et al. (2015)				
	S	P	U	R	*All*	*All*	A	D/H	M	W	*All*
# Items on scale	16	26	10	27	*79*	*56*	8	6	8	8	*30*
Descriptive features (per item)											
# Words	2.0	6.0	18.6	9.2	*7.9*	*19.8*	1.0	3.0	2.5	20.0	*6.9*
# Number symbols	1.0	2.3	1.8	1.4	*1.7*	–	–	–	–	–	–
# Sentences	1.0	1.3	4.2	1.7	*1.7*	*2.7*	1.0	1.0	0.9	3.8	*1.7*
Lexical features (per item)											
# Academic words	1.1	1.2	0.0	1.0	*1.0*	*1.6*	1.0	0.0	0.1	0.3	*0.4*
# Subject-specific vocabulary words	0.1	1.0	0.0	1.1	*0.7*	*0.9*	0.0	1.0	0.0	0.0	*0.2*
Grammatical features (per item)											
# Connectors	0.0	0.0	0.5	0.4	*0.2*	–	0.0	0.0	0.1	0.4	*0.1*
# Nominalisations	0.0	0.1	0.0	0.0	*0.0*	–	0.0	1.0	0.0	0.4	*0.3*
# Nominal phrases	0.0	1.3	8.3	1.9	*2.1*	*5.0*	–	–	–	–	–
# Prepositional phrases	0.0	0.3	1.1	0.3	*0.3*	*1.4*	–	–	–	–	–
# Simple sentences	1.0	1.2	4.2	1.4	*1.6*	*2.2*	–	–	–	–	–
# Complex sentences	0.0	0.0	0.0	0.3	*0.1*	*0.5*	–	–	–	–	–

Common items for first and second measurement were counted only once

A items on arithmetic operations, *D/H* items on doubles/halves, *M* items on measurement, *W* word problems

References

Abedi, J., & Herman, J. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record*, *112*(3), 723–746.

Abedi, J., Hofstetter, C. H., & Lord, C. (2004). Assessment accommodations for English language learners: implications for policy-based empirical research. *Review of Educational Research*, *74*(1), 1–28.

Bae, Y. S., Hickson, L. & Chiang, H.-M. (2015). Mathematical word problem solving ability of children with autism spectrum disorder and their typically developing peers. *Journal of Autism and Developmental Disorders*, 45(7), 2200–2208.

Baumert, J., & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele & W. Schneider, et al. (Eds.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 323–410). Opladen: Leske + Budrich.

Bleiler, S. K., & Thompson, D. R. (2013). Multidimensional assessment of CCSSM. *Teaching Children's Mathematics*, *19*(5), 292–300.

Bochnik, K. (2017). *Sprachbezogene Merkmale als Erklärung für Disparitäten mathematischer Leistung: differenzierte Analysen im Rahmen einer Längsschnittstudie in der dritten Jahrgangsstufe*. Münster: Waxmann.

Bochnik, K., & Ufer, S. (2016a). Die Rolle (fach-)sprachlicher Kompetenzen zur Erklärung mathematischer Kompetenzunterschiede zwischen Kindern mit deutscher und nicht-deutscher Familiensprache. *Zeitschrift für Grundschulforschung*, *9*(1), 135–147.

Bochnik, K., & Ufer, S. (2016b). Measuring language-related opportunities to learn in primary mathematics classrooms. In C. Csíkos, A. Rausch & J. Szitányi (Eds.), *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 115–122). Szeged, Hungary: PME.

Bochnik, K., & Ufer, S. (2017). Fachsprachliche Kompetenzen im sprachsensiblen Mathematikunterricht der Grundschule. Implikationen einer Studie zur sprachbezogenen Analyse mathematischer Leistungsunterschiede zwischen Lernenden mit deutscher und nicht-deutscher Familiensprache. In D. Leiss, A. Neumann & K. Schwippert (Eds.), *Sprachförderung im Fach Mathematik – Forschungsstand und Herausforderungen im Verlauf der Schulzeit* (pp. 81–89). Münster: Waxmann.

Cattell, R. B., Weiß, R. H., & Osterland, J. (1997). *Grundintelligenztest Skala 1: CFT 1*. Göttingen: Hogrefe.

CCSSI (2010). *Common core state standards for mathematics*. Washington D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers.

Civil, M. (2008). Language and mathematics: Immigrant parents' participation in school. In O. Figueras, J.L. Cortina, S. Alatorre, T. Rojano & A. Sepúlveda (Eds.), *Proceedings of the joint meeting of the 32nd conference of the international group for the psychology of mathematics education and the XX North American chapter* (Vol. 2, pp. 329–336). Morelia: PME.

Civil, M., & Planas, N. (2004). Participation in the mathematics classroom: does every student have a voice? *For the Learning of Mathematics*, *24*(1), 7–12.

Cummins, J. (2008). BICS and CALP: empirical and theoretical status of the distinction. In *Encyclopedia of language and education* (pp. 487–499). Berlin Heidelberg: Springer.

Dyrvold, A., Bergqvist, E., & Österholm, M. (2015). Uncommon vocabulary in mathematical tasks in relation to demand of reading ability and solution frequency. *Nordic Studies in Mathematics Education*, *20*(1), 101–128.

Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, *23*(3), 290–325.

Ehmke, T., & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller & M. Prenzel, et al. (Eds.), *PISA 2009: Bilanz nach einem Jahrzehnt* (pp. 231–254). Münster: Waxmann.

Ehmke, T., Hohensee, F., Siegle, T., & Prenzel, M. (2006). Soziale Herkunft, elterliche Unterstützungsprozesse und Kompetenzentwicklung. In PISA-Konsortium Deutschland (Ed.), *PISA 2003* (pp. 225–248). Münster: Waxmann.

Erath, K., Prediger, S., Quasthoff, U., & Heller, V. (2018). Discourse competence as important part of academic language proficiency in mathematics classrooms: the case of explaining to learn and learning to explain. *Educational Studies in Mathematics*, *99*(2), 161–179.

Fuson, K. C., Wearne, D., Hiebert, J. C., Murray, H. G., Human, P. G., Olivier, A. I., et al. (1997). Children's conceptual structures for multidigit numbers and methods of multidigit addition and subtraction. *Journal for Research in Mathematics Education*, *28*(2), 130–162.

Gabler, L., & Ufer, S. (n.d.). Sprachliche Flexibilität von Grundvorstellungen zu Addition und Subtraktion – Eine Vorstudie zu einem Förderkonzept für die zweite Jahrgangsstufe. *Journal für Mathematikdidaktik*, under revision

Gebhardt, M., Rauch, D., Mang, J., Sälzer, C., & Stanat, P. (2013). Mathematische Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Ed.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (pp. 275–308). Münster: Waxmann.

Gorgorió, N., & Planas, N. (2001). Teaching mathematics in multilingual classrooms. *Educational Studies in Mathematics*, *47*(1), 7–33.

Götze, D. (2018). Fostering a conceptual understanding of division: A language-and mathematics-integrated project in primary school. In N. Planas & M. Schütte (Eds.), *Proceedings of the IV ERME Topic Conference 'Classroom-based research on mathematics and language'* (pp. 73–80). Dresden: ERME.

Grüßing, M., & Schmitman gen. Pothmann, A. (2007). Ohne Zahlen keine Welt und ohne Wörter guckt man sich nur an. *Grundschulunterricht*, *54*(7–8), 28–33.

Haag, N., Böhme, K., & Stanat, P. (2012). Zuwanderungsbezogene Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Eds.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs* (pp. 209–235). Münster: Waxmann.

Haag, N., Heppt, B., Stanat, P., Kuhl, P., & Pant, H. A. (2013). Second language learners' performance in mathematics: disentangling the effects of academic language features. *Learning and Instruction*, *28*, 24–34.

Halliday, M. A. K. (1978). *Language as social semiotic: the social interpretation of language and meaning*. London: Edward Arnold.

Härtig, H., Bernholt, S., Prechtl, H., & Retelsdorf, J. (2015). Unterrichtssprache im Fachunterricht – Stand der Forschung und Forschungsperspektiven am Beispiel des Textverständnisses. *Zeitschrift für Didaktik der Naturwissenschaften*, *21*(1), 55–67.

Heinze, A., Arend, J., Gruessing, M., & Lipowsky, F. (2018). Instructional approaches to foster third graders' adaptive use of strategies: an experimental study on the effects of two learning environments on multi-digit addition and subtraction. *Instructional Science*, *46*(6), 869–891.

Heinze, A., Herwartz-Emden, L., & Reiss, K. (2007). Mathematikkenntnisse und sprachliche Kompetenz bei Kindern mit Migrationshintergrund zu Beginn der Grundschulzeit. *Zeitschrift für Pädagogik*, *53*(4), 562–581.

Heinze, A., Herwartz-Emden, L., & Reiss, K. (2011). Die Rolle von Kenntnissen der Unterrichtssprache beim Mathematiklernen. Ergebnisse einer quantitativen Längsschnittstudie in der Grundschule. In S. Prediger & E. Özdil (Eds.), *Mathematiklernen unter Bedingungen der Mehrsprachigkeit – Stand und Perspektiven der Forschung und Entwicklung in Deutschland* (pp. 11–33). Münster: Waxmann.

Hobusch, A., Lutz, N., & Wiest, U. (2002). *Sprachstandsüberprüfung und Förderdiagnostik für Ausländer- und Aussiedlerkinder: SFD 3–4*. Horneburg: Persen.

Höttecke, D., Ehmke, T., Krieger, C., & Kulik, M. A. J. (2017). Vergleichende Messung fachsprachlicher Fähigkeiten in den Domänen Physik und Sport. *Zeitschrift für Didaktik der Naturwissenschaften*, *23*(1), 53–69.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Jordan, A., Krauss, S., Löwen, K., Blum, W., Neubrand, M., Brunner, M., et al. (2008). Aufgaben im COACTIV-Projekt: Zeugnisse des kognitiven Aktivierungpotentials im deutschen Mathematikunterricht. *Journal für Mathematik-Didaktik*, *29*(2), 83–107.

Kieffer, M. J., Lesaux, N. K., Rivera, M., & Francis, D. J. (2009). Accommodations for English language learners taking large-scale assessments: a meta-analysis on effectiveness and validity. *Review of Educational Research*, *79*(3), 1168–1201.

KMK (2004). *Bildungsstandards im Fach Mathematik für den Primarbereich: Jahrgangsstufe 4*. Neuwied: Luchterhand.

Lesh, R., Cramer, K., Doerr, H., Post, T., & Zawojewski, J. (2003). Using a translation model for curriculum development and classroom instruction. In R. Lesh & H. Doerr (Eds.), *Beyond constructivism. Models and modeling perspectives on mathematics problem solving, learning, and teaching*. Mahwah: Lawrence Erlbaum.

Maier, H., & Schweiger, F. (1999). *Mathematik und Sprache: Zum Verstehen und Verwenden von Fachsprache im Mathematikunterricht*. Wien: ÖBV & HPT.

Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, *78*(2), 333–368.

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: who falls increasingly behind? *Journal of Learning Disabilities*, *44*(5), 472–488.

Moschkovich, J. (2002). A situated and sociocultural perspective on bilingual mathematics learners. *Mathematical Thinking and Learning*, *4*(2–3), 189–212.

Moschkovich, J., & Zahner, W. (2018). Using the academic literacy in mathematics framework to uncover multiple aspects of activity during peer mathematical discussions. *ZDM Mathematics Education*, *50*(6), 999–1011.

Muthén, L., & Muthén, B. (2017). *Mplus (Version 8)*. Los Angeles: Muthén & Muthén.

NCTM (2000). *Principles standards for school mathematics*. Reston: NCTM.

Ostad, S. A. (2013). Private speech use in arithmetical calculation: contributory role of phonological awareness in children with and without mathematical difficulties. *Journal of Learning Disabilities*, *46*(4), 291–303.

Österholm, M., & Bergqvist, E. (2013). What is so special about mathematical texts? Analyses of common claims in research literature and of properties of textbooks. *ZDM—The International Journal on Mathematics Eduction*, *45*(5), 751–763.

Özcan, N. (2013). *Zum Einfluss der Fachsprache auf die Leistung im Fach Chemie. Eine Förderstudie zur Fachsprache im Chemieunterricht*. Berlin: Logos Verlag.

Paetsch, J., & Felbrich, A. (2015). Longitudinale Zusammenhänge zwischen sprachlichen Kompetenzen und elementaren mathematischen Modellierungskompetenzen bei Kindern mit Deutsch als Zweitsprache. *Psychologie in Erziehung und Unterricht*, *63*(1), 16–33.

Paetsch, J., Felbrich, A., & Stanat, P. (2015). Der Zusammenhang von sprachlichen und mathematischen Kompetenzen bei Kindern mit Deutsch als Zweitsprache. *Zeitschrift für Pädagogische Psychologie*, *29*, 19–29.

Paetsch, J., Radmann, S., Felbrich, A., Lehmann, R., & Stanat, P. (2016). Sprachkompetenz als Prädiktor mathematischer Kompetenzentwicklung von Kindern deutscher und nicht-deutscher Familiensprache. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *48*, 27–41.

Paulus, C. (2009). Die Bücheraufgabe zur Bestimmung des kulturellen Kapitals bei Grundschülern. http://psydok.psycharchives.de/jspui/handle/20.500.11780/3344. Accessed 26 Jan 2020.

Pietsch, M., & Krauthausen, G. (2006). Mathematisches Grundverständnis von Kindern am Ende der vierten Jahrgangsstufe. In W. Bos & M. Pietsch (Eds.), *KESS 4. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Schulen* (pp. 143–164). Münster: Waxmann.

Prediger, S., & Pöhler, B. (2015). The interplay of micro-and macro-scaffolding: an empirical reconstruction for the case of an intervention on percentages. *ZDM Mathematics Education*, *47*(7), 1179–1194.

Prediger, S., & Wessel, L. (2013). Fostering German-language learners' constructions of meanings for fractions: design and effects of a language-and mathematics-integrated intervention. *Mathematics Education Research Journal*, *25*(3), 435–456.

Prediger, S., & Wessel, L. (2018). Brauchen mehrsprachige Jugendliche eine andere fach- und sprachintegrierte Förderung als einsprachige? *Zeitschrift für Erziehungswissenschaft*, *21*(2), 361–382.

Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E., & Benholz, C. (2018). Language proficiency and mathematics achievement. *Journal für Mathematik-Didaktik*, *39*(Supplement 1), 1–26. Translation of an article in the same journal from 2015, see https://doi.org/10.1007/s13138-015-0074-0.

Rittle-Johnson, B., Schneider, M., & Star, J.R. (2015). Not a one-way street: bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review*, *27*(4), 587–597.

Roick, T., Gölitz, D., & Hasselhorn, M. (2004). *Deutscher Mathematiktest für dritte Klassen: DEMAT 3+*. Manual: Beltz.

Sarama, J., & Clements, D.H. (2016). Physical and virtual manipulatives: what is 'concrete'? In P.S. Moyer-Packenham (ed.), *International perspectives on teaching and learning mathematics with virtual manipulatives* (pp. 71–93). Heidelberg: Springer.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Schindler, V., Moser-Opitz, E., Cadonau-Bieler, M., & Ritterfeld, U. (2019). Überprüfung und Förderung des mathematischen Fachwortschatzes der Grundschulmathematik – eine empirische Studie. *Journal für Mathematik-Didaktik*, *40*(1), 1–35.

Schleppegrell, M.J. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, *12*(4), 431–459.

Schleppegrell, M.J., Achugar, M., & Oteíza, T. (2004). The grammar of history: enhancing content-based instruction through a functional focus on language. *Tesol Quarterly*, *38*(1), 67–93.

Schneider, M. (2006). Konzeptuelles und prozedurales Wissen als latente Variablen: Ihre Interaktion beim Lernen mit Dezimalbrüchen. https://depositonce.tu-berlin.de/handle/11303/1605. Accessed 26 Jan 2020.

Schütte, M. (2009). *Sprache und Interaktion im Mathematikunterricht der Grundschule: Zur Problematik einer Impliziten Pädagogik für schulisches Lernen im Kontext sprachlich-kultureller Pluralität*. Münster: Waxmann.

Sfard, A. (2008). Thinking as communicating: human development, the growth of discourses, and mathematizing. *TMME*, *5*, 429–436.

Shaftel, J., Belton-Kocher, E., Glasnapp, D., & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of English language learners and students with disabilities. *Educational Assessment*, *11*(2), 105–126.

Steenpaß, A., & Steinbring, H. (2014). Young students' subjective interpretations of mathematical diagrams: elements of the theoretical construct 'frame-based interpreting competence'. *ZDM—The International Journal on Mathematics Eduction*, *46*, 3–14.

Stern, E. (1994). Wie viele Kinder bekommen keinen Mohrenkopf? Zur Bedeutung der Kontexteinbettung beim Verstehen des quantitativen Vergleiches. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *26*(1), 79–93.

Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., et al. (2016). *Assessment in mathematics education: large-scale assessment and classroom assessment*. Heidelberg: Springer.

Taboada, A. (2012). Relationships of general vocabulary, science vocabulary, and student questioning with science comprehension in students with varying levels of English proficiency. *Instructional Science*, *40*(6), 901–923.

Tarelli, I., Schwippert, K., & Stubbe, T. C. (2012). Mathematische und naturwissenschaftliche Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, H. Wendt, O. Köller & C. Selter (Eds.), *TIMSS 2011: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 247–267). Münster: Waxmann.

Ufer, S., Heinze, A., & Reiss, K. (2008). Individual predictors of geometrical proof competence. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano & A. Sepulveda (Eds.), *Proceedings of the Joint Meeting of PME 32 and PME-NA XXX, Morelia, Mexico* (Vol. 4, pp. 361–368). Morelia: PME.

Ufer, S., Reiss, K., & Mehringer, V. (2013). Sprachstand, soziale Herkunft und Bilingualität: Effekte auf Facetten mathematischer Kompetenz. In M. Becker-Mrotzek, K. Schramm, E. Thürmann & H. J. Vollmer (Eds.), *Sprache im Fach* (pp. 185–202). Münster: Waxmann.

Wessel, L., & Erath, K. (2018). Theoretical frameworks for designing and analysing language-responsive mathematics teaching–learning arrangements. *ZDM Mathematics Education*, *50*(6), 1053–1064.