Learning and Individual Differences 42 (2015) 70–76

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Effects of mathematics items' language demands for language minority students: Do they differ between grades?

Nicole Haag [a,*], Alexander Roppelt [a], Birgit Heppt [b]

[a] Institute for Educational Quality Improvement (IQB), Humboldt-Universität zu Berlin, Germany
[b] Humboldt-Universität zu Berlin, Germany

ARTICLE INFO

Article history:
Received 1 December 2014
Received in revised form 19 June 2015
Accepted 15 August 2015

Keywords:
Second language learners
Academic language
Test fairness
Educational measurement
Mathematics

ABSTRACT

This study investigates the connection between the academic language demands of mathematics test items and the test performance of monolingual students and language minority students for 2796 third and 26,016 fourth grade students in Germany. We analyzed students' responses to the 126 mathematics test items used in both grades using a generalized linear mixed model and included students' socioeconomic status as a covariate. The results indicate that language minority students are not significantly disadvantaged by items with high academic language demands compared to German monolingual students in the same grade. However, language minority students as well as German monolingual students are more strongly affected by items with high academic language demands in third grade than in fourth grade. Thus, the impact of academic language demands seems to depend on grade level rather than on language minority student status.

In many countries, language minority students have consistently been shown to reach on average lower proficiency levels in mathematics than their native peers (Mullis, Martin, Foy, & Arora, 2012). In Germany, the performance disadvantage of students who do not exclusively speak German at home in fourth grade is about half a standard deviation (Haag, Böhme, & Stanat, 2012), which corresponds to about two thirds of the growth that can be expected to occur between third and fourth grade (Bloom, Hill, Black, & Lipsey, 2008; Reiss & Winkelmann, 2009). This substantial achievement gap has led to concerns regarding the validity of large-scale assessment items for language minority students. As language minority students on average have lower language skills in the test language (e.g., Genesee & Lindholm-Leary, 2012; Haag et al., 2012) even if they were born and raised in the receiving country, they might be particularly disadvantaged by items with high language demands. Such items may thus be less appropriate for assessing the mathematics competence of language minority students than of monolingual students.

Language demands at the item level can be operationalized using the concept of academic language. A considerable body of research has accumulated on language demands at the item level and their effects on performance of second language learners in content assessments (e.g., Haag, Heppt, Stanat, Kuhl, & Pant, 2013; Martiniello, 2009; Prediger & Özdil, 2011; Shaftel, Belton-Kocher, Glasnapp, & Poggio, 2006; Townsend, Filippini, Collins, & Biancarosa, 2012; Ufer, Reiss, & Mehringer, 2013; Wolf & Leon, 2009). The present study adds to the current state of research by comparing the relationship of academic language and item difficulty for third and fourth grade language minority students in Germany on a standardized national assessment instrument in mathematics.

1.1. Academic language

Academic language is the language that is spoken in the classroom to impart and acquire knowledge (Chamot & O'Malley, 1994). This language is different from everyday language in a number of ways, most notably in terms of lexical (e.g., general and specialized academic vocabulary, long and infrequent words, clause connectors) and syntactical features (e.g., long and syntactically complex sentences, passive voice constructions, large numbers of long noun phrases and prepositional phrases) (Heppt, Haag, Böhme, & Stanat, 2015; Bailey, 2000/2005; Bailey, Butler, Stevens, & Lord, 2007; Butler, Bailey, Stevens, Huang, & Lord, 2004). The terms general and specialized academic vocabulary refer to vocabulary that is typically used in the context of schooling rather than in everyday conversations and that is necessary to acquire knowledge. Whereas general academic vocabulary (e.g., *report*, *synthesize*) is used across disciplines, specialized academic vocabulary (e.g., *denominator*, *multiplication*) is associated with a specific school subject.

The mastery of academic language is assumed to play a pivotal role for educational success (Bailey, 2007; Bailey & Butler, 2003; Butler, Lord, Stevens, Borrego, & Bailey, 2004). Gaining proficiency in academic language is expected to be especially challenging for students from families with low socioeconomic status (SES) and language minority

* Corresponding author at: Institute for Educational Quality Improvement (IQB), Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany.
E-mail address: nicole.haag@iqb.hu-berlin.de (N. Haag).

http://dx.doi.org/10.1016/j.lindif.2015.08.010
1041-6080/

students (see Heppt et al., 2015 for a more detailed description of the expected disadvantages). It is assumed that this challenge persists even if students are reasonably fluent speakers in everyday interactions (Cummins, 1979, 2000). Syntactical aspects of academic language and general academic vocabulary are likely to be only rarely addressed explicitly in classroom discourse in Germany (Komor & Reich, 2008) and in the United States (Nagy & Townsend, 2012), as teachers' awareness of the particular linguistic difficulties of the language of schooling is often limited. Unless students' academic language development is more explicitly promoted in school, students have to rely on their possibilities to acquire academic language outside school. However, students differ in their opportunities to acquire these language skills in their home environments. Especially students from families with low SES and language minority students are at risk of entering school with only limited proficiency in academic language (e.g., Scheele, Leseman, Mayo, & Elbers, 2012; Schleppegrell, 2012).

1.2. Disparities of language minority students in mathematics associated with academic language demands

Various studies have shown a connection between the difficulty of mathematics test items and linguistic demands of these items. In addition, a number of studies investigated potentially differential effects of linguistically challenging items for second language learners and for monolingual students, albeit with somewhat inconsistent results concerning the significance of these effects (for a recent review, see Lane & Leventhal, 2015). Most of these investigations were conducted in English-speaking countries and used methods of *differential item functioning* (DIF). Several studies found effects of item length (Abedi & Leon, 1999; Abedi, Leon, Wolf, & Farnsworth, 2008; for Germany: Haag et al., 2013), thus suggesting that the amount of text is a crucial determinant of disparities related to students' language background. Whereas one study identified only lexical features to be predictive of second language learners' performance in mathematics (Abedi et al., 2008), other studies found effects of lexical as well as syntactical features (Haag et al., 2013; Martiniello, 2008).

Thus, previous research suggests that academic language in test items affects the performance of language minority students, but no conclusive results were found regarding the contribution of different features (Lane & Leventhal, 2015). It may be the case that the academic language demands of test items are more adequately represented by global, aggregated indicators of academic language rather than by single features. For example, Haag et al. (2013) showed in a commonality analysis that although descriptive, lexical, and syntactical features all had unique contributions in predicting DIF against second language learners, the largest part of DIF variance was simultaneously explained by several features together. Aggregation methods, such as sum scores or factor scores, have already been successfully used to describe academic language demands in both test items (Cawthon, Kaye, Lockhart, & Beretvas, 2012) and person utterances of children and their caregivers in various conversational settings (Leseman, Scheele, Mayo, & Messer, 2007; Scheele et al., 2012).

An aggregated measure of linguistic complexity for mathematics test items should only include those academic language features that are not considered core elements of mathematics proficiency. In particular, such a measure should not include math-specific academic vocabulary, which is regarded as part of the construct the mathematics test intends to measure, whereas other academic language features could be considered as posing primarily linguistic difficulties (cf. Abedi, 2009). From a theoretical point of view, disadvantages associated with math-specific vocabulary should not compromise the validity of a math test because this type of vocabulary is explicitly taught in classrooms as part of the curriculum (KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland), 2005). Also, math-specific academic vocabulary does not predict item DIF against second language learners in third grade (Haag et al., 2013).

1.3. Grade level differences in the effects of academic language on mathematics performance

There is a well-established connection of grade level and linguistic demands of texts (Deane, Sheehan, Sabatini, Futagi, & Kostin, 2006; Graesser & McNamara, 2011). However, the impact of these changing linguistic demands on second language learners' performance in content assessments such as mathematics is not entirely clear. So far, data comparing the linguistic demands of mathematics tests and the impact of these demands on students' performance over grades are sparse.

On the one hand, there is some evidence that even though academic language features of mathematics test items may be more frequent in higher grades, their influence on item difficulty is stronger in lower grades. For example, Shaftel et al. (2006) found that nonmathematical linguistic complexities only predicted mathematics item difficulty in grade 4, but not in grades 7 and 10. They could not identify any differential effects for English Language Learners (ELLs), indicating that the disparities in mathematics achievement cannot be explained by the test items' academic language demands in any of the grades. They therefore concluded that, regardless of students' home language, academic language impacts students in lower grades more than in higher grades.

On the other hand, language minority students may remain impaired by academic language over time. For example, Slama (2012) analyzed the trajectories of ELLs' academic language achievement from grades 7 to 12 and found that it takes the average ELL until grade 11 to learn English well enough to be schooled in mainstream classrooms. This indicates that a lot of language minority students are only slowly mastering academic language and are thus experiencing difficulties in their schooling career (Slama, 2014). Regrettably, models of typical academic language development are still lacking (Snow & Uccelli, 2009), making it difficult to describe the level of academic language students are expected to display in different grade levels.

In summary, previous research identified substantial performance gaps between monolingual students and language minority students. These disadvantages seem to be related to the academic language demands of test items. However, research on the effects of academic language demands over grade levels is scarce and inconclusive. There is some evidence that the effect of academic language demands is stronger in lower grades. However, the results of previous studies are difficult to interpret because both the test items' demands and the students' proficiencies differ over the grades.

2. Research question

The present study provides a fine-grained analysis on the effects of academic language in elementary school by comparing these effects for language minority students and monolingual students in third and fourth grade. Investigating two adjacent grades allows us to work with a constant set of items, thus holding the academic language demands and the mathematical content constant and focus on the effects of the differences in students' proficiencies. We expected that test items with high academic language demands are disproportionally more difficult for younger students than for older students (Hypothesis 1a) and more difficult for language minority students than for German monolingual students (Hypothesis 1b). Moreover, we expected the disadvantage of language minority students due to academic language in mathematics items to be more prominent in third grade than in fourth grade (Hypothesis 2).

3. Method

3.1. Sample

This study is based on two nationally representative German large-scale assessment tests in elementary school with an overlapping set of