Article

The International Journal of Bilingualism

International Journal of Bilingualism
2018, Vol. 22(1) 123–137

DOI: 10.1177/1367006916658711
journals.sagepub.com/home/ijb
SAGE

Does more exposure to the language of instruction lead to higher academic achievement? A cross-national examination

Orhan Agirdag
Laboratory for Education and Society, University of Leuven, Belgium; Department of Educational Sciences, University of Amsterdam, Netherlands

Gudrun Vanlaar
Centre for Educational Effectiveness and Evaluation, University of Leuven, Belgium

Abstract

Aims and objectives: As some Programme for International Student Assessment (PISA) studies claimed that native speaking (NS) students outperform language minority (LMi) students, far-reaching inferences have been drawn by policymakers. However, previous PISA assessments were not appropriate because they only included a dichotomous home language variable. The main objective of this study is to gain a better understanding of how students' language background and use are related to academic achievement.

Design: The PISA data from 2012 provides a unique opportunity to fill this research lacuna as it includes a more elaborated questionnaire on language background and use.

Data and analysis: Multivariate three-level analyses are conducted on PISA 2012 data from 18 countries, covering about 5,000 schools and 120,000 students.

Findings: The results show that there is an achievement gap between LMi and NS students for both reading and math. After controlling for students and school characteristics, the LMi–NS achievement gap narrows, but remains significant. This holds true for most countries. However, language use per se is not the cause of underachievement: LMi students who more often speak a minority language with their parents do not achieve less. In some countries, speaking a minority language more often with parents is actually positively related to math and reading achievement. Nevertheless, speaking the instruction language in the school context is positively associated with math and reading achievement.

Originality and significance: This study revealed that the relation between language use and academic achievement is more complex than it was conceptualized in most previous PISA studies. Scholars need to go beyond the dichotomous approach to achieve a better understanding of

Corresponding author:
Orhan Agirdag, Laboratory for Education and Society, University of Leuven, Andreas Vesaliusstraat 2 – bus 3761, 3000 Leuven, Belgium.
Email: orhan.agirdag@gmail.com

language use. Our results show that linguistic diversity could function as an asset for academic performance, at least if a good balance between focus on minority languages at home and instruction language at school can be found.

Keywords
academic achievement, PISA, language use, language background, ethnic minorities

Introduction

Over the past two decades, a lot of attention has been given to international studies of student achievement such as the Programme for International Student Assessment (PISA), the Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS). In particular, the PISA program, organized every three years by the Organisation for Economic Co-operation and Development (OECD), has a significant influence on the development and implementation of national policies (Breakspear, 2012; Grek, 2009). PISA is a worldwide study of 15-year-old students‘ academic performance in OECD member and non-member nations. Receiving intense media attention, the PISA results trigger discussions about the need for educational reforms. Many countries have reformed their educational systems with reference to PISA results and PISA has become an important element of the educational policy-making process (Breakspear, 2012; Ertl, 2006; Gür, Çelik, & Özoğlu, 2012).

Regarding educational language policies, a substantial influence is exerted by the PISA studies (Cummins, 2008). These language policies are often informed by studies that use the PISA data to compare the academic achievement of native speaking (NS) students who speak another language than the instruction language at home, i.e. language minority (LMi) students (e.g. Dronkers & Kornder, 2014; Esser, 2006; Marks, 2005; Schnepf, 2007; Stanat & Christensen, 2006). Some of those studies find that LMi students perform lower on PISA achievement tests than NS students (Schnepf, 2007; Stanat & Christensen, 2006). Based on this correlation, inferences have been drawn about the potential consequences of speaking a minority language (ML) and exposure to it within the family or school context (see Cummins, 2008). Even the effectiveness of bilingual education is questioned, with policymakers and some scholars arguing for immersion in the instruction language (IL) at an early age. For instance, Esser (2006) states that speaking a ML is the most important hindrance for the structural integration of immigrants. In Denmark, the PISA test results are interpreted as if most ‘bilingual students’ (that is the official Danish term) are ‘functionally illiterate’ or that bilingual students lack functional reading skills (Laursen, 2013).

Some authors have used the *language spoken at home* variable in the PISA study as if it was a proxy for linguistic ability. For instance, a frequently cited article by Schnepf (2007, p. 533) states that: “Pupils’ language skills are estimated with one variable: whether pupils speak a foreign language at home.” However, language proficiency and language use should not be confused with each other, and the correlation between both is modest at best (Van Tubergen & Kalmijn, 2009). The problematic inference from exposure to one’s ML to linguistic deficiency is also present in official OECD reports. For instance, the official PISA 2012 report states:

> Not understanding the language of the country of residence upon arrival is a disadvantage; but so is insufficient exposure to that language outside school. Policies aimed at supporting immigrant students who do not speak the language of assessment at home should focus on both school and home (OECD, 2013a, p.80).

Hence, many policymakers responded to these PISA reports with a nativist reflex, and most visibly by increasing sociopolitical pressure on minorities to abandon their ML. Fostered by the spurious inferences made in the PISA reports, assimilation policies assumed that language maintenance in family and school contexts is a hindrance for the social and economic integration of minorities in broader society (Extra & Yagmur, 2005; Jaspers, 2008; Pulinx, Van Avermaet, & Agirdag, 2016). Thus, there is not only a call for increased use of the IL, but also a call for decreased use of the ML. For instance, Agirdag (2010) reports that in Flanders (Belgium), Turkish children are forbidden to speak their mother tongue, and punished for speaking Turkish with their peers; this is a common practice in many Flemish schools. Pulinx et al. (2016) show that the vast majority of teachers support the monolingual policies in Flemish education: 77 percent of the teachers agree that students should not be allowed to speak a foreign language at school; almost a third of the teachers believe that students should be punished for their own benefit for speaking their mother tongue. While this push towards monolingualism is mostly driven by noble motives, namely to stimulate intergroup contact and to protect students against underachievement, from a scholarly point of view, it is not clear whether the restriction of students' ML is beneficial.

Most importantly, previous PISA assessments (of 2000, 2003, 2006, and 2009) are by no means appropriate to examining the impact of students' language use on their academic achievement. These PISA assessments only included a dichotomous home language variable, that is, whether or not a student speaks a language other than the IL at home. First of all, this home language variable might simply be a proxy for unmeasured characteristics such as ethnicity, linguistic discrimination, or length of residence in a country. As such, it is unclear which aspects specifically trigger the reported underachievement of LMi students. Moreover, in these PISA studies the LMi group is often treated as a homogenous group, although many students do not fall into either of the two categories (LMi vs. NS); it is equally possible that they use the dominant language as much as they use their ML. Another important distinction, which is not captured by the dichotomous conceptualization of the home language variable, is that the language spoken by a student can vary with respect to the *interlocutor*, that is the conversation partner with whom a language is used (e.g. parents, siblings or friends), and context (e.g. at home, at school). In addition, the impact of exposure to a language on academic performance might depend on both interlocutor and context (see also Duursma et al., 2007). These are important distinctions if we want to understand how linguistic background is related to academic outcomes.

The main purpose of this study is to gain a better understanding of how students' language background and use are related to their academic achievement. In order to do so, we use the most recent PISA data from 2012. In contrast with previous PISA assessments, the PISA 2012 includes a more elaborated questionnaire on language background and use. While still far from perfect, the PISA 2012 data includes more detailed information about language use across different interlocutors and activities. As such, the PISA 2012 provides a unique opportunity to gain a better understanding about the relationship between language use and student achievement. In what follows, we briefly discuss the theoretical rationale. We then provide an overview of previous empirical studies. Next, we discuss the methods used in our empirical study, we summarize the results, and we conclude with a discussion of the research findings.

Theory

The impact of using and exposure to an ML in an educational context with a different IL came to be theorized from two opposing points of view in the literature. A deficit and an additive view of ML use and exposure can be distinguished. The deficit view builds mainly on the notions of competition between languages and time on task (Leseman, 2000). In this time on task perspective, it is argued that there is a *competitive relation* between ML and IL concerning learning time and children's cognitive resources in the acquisition process of ML and IL skills (Mancilla-Martinez & Lesaux, 2011; Pearson & Fernández, 1994). This implies that when two or more languages are acquired, the language input and the available learning time should be divided between these different languages; the net result is less input and learning time per language (see Scheele, Leseman, & Mayo, 2010). This suggests a zero-sum game: gains in one language come at a cost to the other. In this respect, ML development is a barrier to academic IL learning and school achievement. This argument is often invoked to endorse monolingual instruction programs that promote total immersion in the IL for disadvantaged LMi students as a precondition for school success (Rossell & Baker, 1996). As IL proficiency is necessary for school success, LMi students exposed more consistently to IL are expected to have higher academic achievement than those who use their ML more.

The additive view of ML use and exposure, on the other hand, builds on the notions of cognitive executive control (Bialystok, 1999; Bialystok, Martin, & Viswanathan, 2005), linguistic interdependence (Cummins, 1979), and multicultural capital (Agirdag, 2014). Bialystok (1999) argues that bilingual children develop cognitive mechanisms for controlling the use of their language systems. That is, they must control which of the languages they use in a given context and do so without intrusions. This process of *executive control* leads to more developed neurological mechanisms and this boosted cognitive functioning might result in improved executing of cognitive tasks (Best, Miller, & Naglieri, 2011). However, as this type of research is relatively new and evolving rapidly, we do not know yet how the effects of bilingualism on executive control affect school learning. A second perspective is the so-called *linguistic interdependence hypothesis* of Cummins (1979). It suggests that the level of ability in one language is partly a function of ability in a second language. Concepts and skills that have been developed in the ML can be transferred to the IL through a 'common underlying proficiency', composed of both conceptual and procedural knowledge and skills. A recent study by Edele and Stanat (2016) shows that the transfer between ML and IL is language-independent and more pronounced for students with higher levels of ML proficiency. A third mechanism, explained by Agirdag (2014), states that access to ML can function as *multicultural capital*, which is defined as "a distinct type of cultural capital that results from the retention of ethnic and linguistic cultural forms" (p. 461). Like cultural capital, the possession of multicultural capital is expected to have a positive impact on students' academic performance, mainly because it provides access to family and community-based cultural resources that can improve student achievement (see also Portes and Hao, 2002).

Previous studies

Previous studies on the consequences of ML use and exposure on academic performance roughly consist of three types. First, there are studies that use international data, mainly the PISA data, to examine how language background is related to math, reading or science achievement (e.g. Dronkers & Van der Velden, 2013; Levels, Dronkers, & Kraaykamp, 2008; OECD, 2013a; Schnepf, 2007; Stanat & Christensen, 2006). A second type of study makes use of (larger) national samples and investigates how students' language use is related to math and reading achievement (e.g. Agirdag, Jordens, & Van Houtte, 2014; Brenneman, Morris, & Israelian, 2007; Hannover et al., 2013). Third, there are studies that use smaller (local) samples of young children and focus mostly on vocabulary proficiency, which is an outcome that is closely related to reading achievement (e.g. Duursma et al., 2007; Hammer, Davison, Lawrence, & Miccio, 2009; Mancilla-Martinez & Lesaux, 2011; Prevoo et al., 2014). National studies of academic achievement and local studies that focus on vocabulary proficiency typically use a more elaborated conceptualization and measurement of language use (including measures of language spoken with different family members and school peers) than the PISA studies that include a dichotomous home language variable.

Using data from the 2003 PISA survey, Schnepf (2007) finds that LMi students achieve less than NS students do across ten OECD countries. The achievement gap between LMi and NS students is reported to be about 10 to 50 PISA points. Expressed in standard deviations (SD), this is a difference between 0.10 SD and 0.50 SD. However, after controlling for socioeconomic background and school characteristics, this gap is reduced to insignificant levels in some countries and shrunk to half its size in other countries. Levels et al. (2008) also used the PISA 2003 data and found that immigrants who speak an ML at home achieve about 10 points less (i.e. 0.10 SD) on mathematics than their NS peers, this after controlling for socioeconomic background, gender and family structure. Dronkers and Van der Velden (2013) use data from the PISA 2006 survey and report that immigrant students who speak the IL at home score 23 points higher (i.e. 0.23 SD) on the reading test than LMi students. The official OECD report of the PISA 2012 results states that the OECD math achievement gap between LMi and NS students is 36 points for the group of native-born students, and 12 points for the group of immigrant students. However, after controlling for SES, the LMi–NS gap drops to 26 points for the group of native-born students and to zero (insignificant) for the group of immigrant students (OECD, 2013a, pp. 233-234).

Studies that use larger national samples provide a mixed picture. In the US, Brenneman et al. (2007) focused on the effects of language use and preference of English-Spanish speaking children on their reading achievement. Using factor analysis, they distinguished between a “family factor” (including language use with mother, father, grandparents and other relatives) and an “outside factor” (including language preference during activities such as watching TV, reading books, and talking with friends). Language spoken with siblings loaded on both factors. Only the outside factor was significantly related to English reading achievement: a decrease in the use of Spanish outside family context predicted higher English reading and comprehension skills. Hannover et al. (2013) used a sample of 813 German children to examine how language spoken at home (with mother, father, and other relatives) and language spoken with peers (schoolmates and friends) was related to reading performance. They found that speaking German more often with peers had a small positive impact on reading, while speaking German at home was not significantly related to reading performance. In Belgium, Agirdag et al. (2014) studied the impact of Turkish speaking children’s language use at home (factor score including language spoken with mother, father and siblings) and at school (factor score including language spoken inside and outside the classroom) on their math achievement. They showed that neither language use at home nor language use at school had a significant effect on math achievement.

In an example of a study on vocabulary achievement, Duursma et al. (2007) investigated whether paternal, maternal and sibling preferences for English have an impact on Latino children’s English vocabulary. They found that only paternal preference for English was a significant predictor of English vocabulary: students whose father spoke more English score higher. Maternal and sibling preferences for English did not predict the level of English vocabulary. Also Hammer et al. (2009) reveal that maternal use of Spanish does not have an impact on children’s English vocabulary achievement. However, paternal, siblings’ or friends’ language use is not examined in their study. Similarly, Prevoo et al. (2014) examined the relation between maternal language use and Dutch vocabulary among Turkish children in the Netherlands. After controlling for SES and parental reading input, they found that maternal language is not related to Dutch vocabulary achievement. In a longitudinal study in the US, Mancilla-Martinez and Lesaux (2011) combined measures of language use of the father, mother, and siblings into a single language use variable and they identified three groups: mostly Spanish, equal amounts, and mostly English. They used growth modeling to examine the association between early patterns of home language use and vocabulary growth. With respect to English vocabulary achievement, they found that initially (at age 4.5) the mostly English group outperformed the other two language groups, while by age 12, the equal amounts and mostly Spanish groups surpassed the mostly English group.

Research questions

The purpose of this study is to gain a better understanding of the relationship between students' language background and use, and their academic achievement. Two research questions are explored. The first research question (RQ1) is: what is the gross and net association between language background (i.e. the first language learned at home) and academic achievement? The results of RQ1 will contribute to the literature by examining whether there is an achievement gap between LMi students and NS. From a time on task perspective (see above), we expect that NS outperform LMi students. The second research question (RQ2) is: what is the gross and net relationship between language use (with different persons and in different contexts) and academic achievement for LMi students? From an additive perspective on multilingualism, we expect that exposure to a ML is related to higher academic achievement outcomes. The reverse can be expected from a time on task perspective on multilingualism.

Methodology

Data

We used the data collected for PISA 2012. Around 510,000 students participated across 65 countries in PISA 2012. The participating students took a test that lasted two hours, which consisted of a mixture of open-ended and multiple-choice questions based on real-life situations (see OECD 2013b). Additional information was gathered from the school authorities, a parent questionnaire and a student background questionnaire. Importantly, a so-called "educational career" questionnaire with detailed information about LMi students' language use was offered as a national option; this was administered in 19 countries. The analyses in the present study use only the data from these countries: Austria, Wallonia (Belgium), Flanders (Belgium), Canada, Denmark, Finland, Germany, Hong Kong (China), Hungary, Ireland, Italy, Latvia, Luxembourg, Portugal, Serbia, Singapore, Slovak Republic, and Slovenia. South Korea was excluded from the analyses because of the low proportion of LMi students.

Statistical analyses

We conducted three-level multivariate regression analyses as the data set consists of a clustered sample of students nested within schools and schools nested within countries using Mplus 7, (Muthen & Muthen, 2012). To provide an answer for RQ1, we estimated the association between language background and achievement (see Table 2). Then, to address RQ2, we examined the influence of language use with different persons and context on academic achievement with the sample of LMi students (see Table 3).

Because the PISA data is cross-sectional, it is not possible to determine causal relationships. In order to reduce spurious statistical relationships, we controlled for variables other than language background and language use that are shown to have an impact on academic performance in previous studies with the PISA data (see Agirdag, Yazici, & Sierens, 2015; Dronkers & Kornder, 2014; Dronkers & Van der Velden, 2013). We controlled for students' background characteristics, students' academic profile and school characteristics as these variables might have an impact on students' academic performance and are unequally distributed among different language groups. However, to prevent over-adjustment or over-controlling, we examined both research questions in two ways: first without control variables (i.e. gross effects) and then with control variables (i.e. net effects).

Reading and math were analyzed in multivariate regression models. Missing data were dealt with using the full information maximum likelihood (FIML) method. FIML uses all available data to estimate parameters based on the available complete data as well as the implied values of the missing data, given the observed data (see Enders & Bandalos, 2001). Metric predictors were grand mean centered in the three-level regression analysis. Unstandardized effects are reported in the tables. The sizes of effects were interpreted in comparison to the average progress that a student makes during one school year, which was estimated by calculating the differences in academic achievement between two different grades. Student weights were used, normalized for cluster size at school and country level. To examine the robustness of our findings and significant effects, we did additional analyses for each country separately in a two-level model. We reported the number of countries where we could replicate the effects found in the general three-level model. The full regression models for each country are available upon request.

Exploratory variables

Language background is a variable that is based on the following question: *What is the first language you learned at home?* Students had three answer options: (1) "The first language I learned at home was IL or dialect(s)"; (2) "I learned IL or dialect(s) and another language at the same time at home"; (3) "The first language I learned at home was a language other than IL or dialect(s)". Students who chose option 1 were categorized as NS and those who chose option 2 and 3 as LMi. Only LMi students were asked to answer further questions about their language use.

Language use was measured using five indicators. LMi students were asked: *Which language do you usually speak with the following people*: (1) mother, (2) father, (3) siblings, (4) best friend, and (5) my schoolmates. For each interlocutor a student could choose between four options: (a) mostly my heritage language, (b) about equally often my heritage language and the IL, (c) mostly the IL, (d) not applicable. Option (c) was used as the reference category in the analyses and option (d) was set to missing.

Control variables

Gender is a dichotomous variable derived from PISA variable ST04Q01 (Female = 1; Male = 2).

Immigrant background is based on the information about students' country of birth and those of their parents. Students who are native born and have both parents born in the test country are referred to as 'natives'. Within the immigrant group, a distinction is made between first generation immigrants (students born abroad), and second generation immigrants (students born in the test country, but at least one parent born abroad).

SES is a continuous variable for which the PISA variable ESCS is used. The index measures the economic, social and cultural status of the parents. It is a composite measure derived from the occupational status of the parents, parental education level, and home resources.

Grade is a continuous variable derived from PISA variable GRADE. This indicates the relative grade, i.e. whether students are at the modal grade in a country (value of 0), or whether they are below (negative value) or above the modal grade level (positive value).

Grade repetition is a dichotomous variable derived from PISA variable REPEAT. It indicates whether a student has repeated a grade in the past (not repeated = 0; repeated = 1).

Table 1. Descriptive statistics.

Exploratory variables			Control variables			Outcomes		
	Mean	SD		Mean	SD		Mean	SD
Language background			Gender			Math	508.87	92.05
Native speaker (NS)	0.81		Girl	0.49		Reading	507.62	92.03
Language minority (LMi)	0.19		Boy	0.51				
L with mother			Immigrant background					
Mostly ML	0.54		Native born	0.87				
Equally ML & IL	0.15		Second generation	0.07				
Mostly IL	0.31		First generation	0.06				
L with father			SES	0.06	0.96			
Mostly ML	0.52		Grade	–0.13	0.55			
Equally ML & IL	0.14		Grade repetition					
Mostly IL	0.34		No	0.88				
L with siblings			Yes	0.12				
Mostly ML	0.39		Track = general					
Equally ML & IL	0.16		Other (vocational)	0.40				
Mostly IL	0.45		General	0.60				
L with best friend			School SES	0.05	0.55			
Mostly ML	0.33		School LMi	0.18	0.24			
Equally ML & IL	0.13		Age started learning IL	1.55	0.90			
Mostly IL	0.54							
L with school friends								
Mostly ML	0.29							
Equally ML & IL	0.11							
Mostly IL	0.60							

Note: Not weighted and not centered values.

Track is a dichotomous variable distinguishing students enrolled in a general track with students enrolled in other tracks. Programs with a general curricular content (PISA variable ISCEDO = 1) and programs designed to give access to the next program level (PISA variable ISCEDD = 1) are categorized as 'general track' (Track = 1). Students enrolled in other programs (e.g. vocational, prevocational) are categorized as 'other track' (Track = 0).

School SES is created by aggregating student-level SES within each school.

School % LMi refers to the proportion of LMi students within each school, and is used as a measure of between-school segregation (see Table 1 for descriptive statistics).

Age of IL learning is an interval variable and is based on PISA variable EC06Q01. Students were asked: '*How old were you when you started learning <instruction language>?*'. There were five options: (1) 0–3 years old; (2) 4–6 years old (3) 7–9 years old; (4) 10–12 years old; (5) 13 years or older.

Outcomes

The outcomes in this study are academic achievement in mathematics and reading. The PISA achievement tests are paper-and-pencil tests consisting of multiple-choice items and short essay questions. The emphasis of the items is on the application of acquired knowledge in real-life

Table 2. Multivariate three-level regression analysis. Associations between language background and academic achievement.

	Model 1 (gross effects)		Model 2 (net effects)	
	Math	Reading	Math	Reading
Intercept	485.11 (8.49)***	479.25 (7.66)***	456.78 (15.62)***	519.47 (13.40)***
Language background (ref = LMi)				
Native speaker (NS)	22.84 (3.56)***	24.477 (3.00)***	13.33 (13.33)***	15.34 (2.08)***
Gender = boy	—	—	20.67 (2.73)***	–27.78 (2.67)***
Immigrant (ref= 1st generation)				
Native born	—	—	–2.28 (4.49)	2.97 (5.22)
Second generation	—	—	–8.49 (2.15)***	–1.17 (2.14)
Grade	—	—	30.52 (2.75)***	27.06 (2.47)***
Grade repetition = yes	—	—	–30.31 (4.40)***	–29.31 (3.90)***
Track = general	—	—	26.62 (7.66)***	29.61 (6.34)***
SES	—	—	12.46 (2.84)***	11.29 (2.77)***
School SES	—	—	51.44 (6.33)***	50.36 (6.25)***
School % LMi	—	—	1.95 (14.20)	–15.15 (12.64)
Residual variance student-level	5025.83 (315.13)***	5035.084 (396.79)***	4364.81 (258.59)***	4337.44 (306.58)***
Residual variance school-level	2600.27 (462.21)***	2842.951 (542.46)***	1199.31 (202.72)***	1230.52 (178.87)***
Variance country-level	974.13 (369.72)***	798.065 (222.32)***	1441.83 (637.46)*	1015.59 (431.95)*
Number of observations	students = 124705; schools = 5257; countries = 18		students = 120998; schools = 5239; countries = 18	

Note: Unstandardized coefficients (b), standard errors (in parentheses), p-values (p) and variance components.
***p< .001; **p< .01; *p < .05.

contexts. The PISA data contain five plausible values (PV) for each achievement domain. The analyses were separately conducted on each PV, and the results were combined in order to obtain unbiased estimates (Rubin, 1987).

Findings

Research question 1

Estimates for the relationship between language background and academic achievement are shown in Table 2. Model 1 indicates that NS students' achievement is significantly higher than that of LMi students for both math ($b = 22.84$, $p < .001$) and reading ($b = 24.48$, $p < .001$). This finding holds true for all countries except for Serbia, Singapore (for math and reading), and Ireland (only math). In these countries, NS students achieve about the same as LMi students (see also Appendix A).

In Model 2, net effects are estimated by entering control variables. The achievement gap between LMi and NS students narrows considerably after control variables are included, but remains significant for both math ($b = 13.334$, $p < .001$) and reading ($b = 15.341$, $p < .001$). The size of the LMi–NS achievement gap corresponds approximately to the achievement progress that an average student makes in half a school year (see Table 2). If we look separately at the 18 countries, this finding holds true for 13 countries for math and 14 countries for reading.

Table 3. Multivariate three–level regression analysis. Associations between language use and academic achievement.

	Model 3 (gross effects)		Model 4 (net effects)	
	Math	Reading	Math	Reading
Intercept	486.54 (8.05)***	485.76 (7.61)***	457.54 (18.65)***	523.09 (12.88)***
L with mother (ref= mostly IL)				
Mostly ML	0.20 (5.44)	–3.70 (5.22)	5.70 (3.47)	2.25 (3.29)
Equally ML & IL	3.99 (4.17)	6.17 (4.03)	3.95 (3.01)	3.54 (3.55)
L with father (ref= mostly IL)				
Mostly ML	5.60 (6.32)	6.77 (6.56)	8.23 (3.82)*	10.63 (3.74)**
Equally ML & IL	–1.92 (5.48)	0.41 (5.60)	0.37 (2.94)	2.50 (2.99)
L with siblings (ref= mostly IL)				
Mostly ML	2.32 (5.45)	–7.31 (6.00)	1.77 (4.72)	–3.20 (5.44)
Equally ML & IL	–5.38 (3.08)	–8.48 (3.19)**	–3.06 (2.75)	–4.89 (3.08)
L with best friend (ref= mostly IL)				
Mostly ML	–8.78 (2.08)***	–9.73 (2.73)***	–4.90 (1.65)**	–8.44 (1.91)***
Equally ML & IL	–13.17 (2.21)***	–12.04 (2.78)***	–6.09 (1.84)***	–7.97 (2.04)***
L with school friends (ref= mostly IL)				
Mostly ML	–10.08 (7.66)	–7.04 (8.13)	–17.23 (4.39)***	–13.75 (4.43)**
Equally ML & IL	–12.81 (4.34)**	–15.41 (5.42)**	–9.71 (4.40)*	–10.34 (4.73)**
Age started learning IL	—	—	–4.55 (1.72)**	–9.16 (1.72)***
Gender = boy	—	—	18.60 (3.89)***	–28.34 (1.69)***
Immigrant (ref= 1st generation)				
Native born	—	—	–11.84 (5.83)*	–10.43 (5.61)
Second generation	—	—	–8.45 (2.75)**	–4.89 (2.53)
Grade	—	—	27.80 (3.13)***	27.03 (2.96)***
Grade repetition = yes	—	—	–26.33 (3.22)***	–24.58 (3.53)***
Track = general	—	—	20.45 (7.44)**	25.91 (5.28)***
SES	—	—	20.45 (7.44)***	12.79 (1.67)***
School SES	—	—	53.25 (4.84)***	52.00 (5.02)***
School % LMi	—	—	14.18 (9.94)	4.96 (9.65)
Residual variance student–level	5780.74 (221.86)***	5894.03 (304.35)***	4570.83 (329.05)***	4590.99 (277.92)***
Residual variance school–level	1191.10 (242.04)***	1323.53 (281.07)***	1060.95 (98.49)***	1113.00 (103.21)***
Variance country–level	842.89 (465.14)	666.89 (263.22)*	1761.41 (756.76)**	1301.01 (529.69)*
Number of observations	students = 31131; schools = 4632; countries = 17		students = 25773; schools = 4559; countries = 17	

Note: Unstandardized coefficients (b), standard errors (between parentheses), p–values (p) and variance components.
***p< .001; **p< .01; *p < .05.

Research question 2

Table 3 gives the estimates for the associations between language use and academic achievement. Model 3 suggests that language use is not unambiguously related to academic performance. Language spoken with parents is not significantly related to achievement, either for math or for reading. However, the language spoken with a best friend is significantly related to academic achievement: speaking mostly IL is related to higher levels of math and reading achievement than speaking mostly a ML or speaking ML and IL equally with a best friend.

In Model 4 (Table 3), control variables for students' background characteristics, students' academic profile, school characteristics and age of IL learning are included. It is now clear that speaking mostly ML with both parents is related to higher academic performance than speaking mostly IL, although the association is only statistically significant for the language spoken with the father (for math: $b = 8.329$, $p = 0.031$; for reading: $b = 10.635$, $p = 0.005$). However, we were not able to replicate this finding in most countries: when countries were analyzed separately, we found that speaking mostly a ML with the father was significantly related to higher academic performance than speaking mostly IL (for math and reading) only in Canada, Finland, and Singapore. In 14 other countries, there was no significant effect.

Compared to speaking mostly the IL, speaking mostly a ML with school friends is related to lower levels of math (b = -17.24, $p < 0.001$) and reading achievement (b = -13.76, $p = 0.002$). The effect size is equivalent to the achievement progress that a student makes in one half of a grade. However, the effect is only statistically significant in 6 countries (for both math and reading), while no significant differences were found in 11 countries. It should be noted that this effect is independent of the language composition of the student body in a school: the share of LMi students in a school is generally not significantly related to academic performance.

Regarding control variables, it is clear that *age of IL learning* is negatively related to academic performance (for math: b = -4.555, $p = 0.008$; for reading: b = -9.165, $p < 0.001$). This means that learning the IL at a younger age is related to higher levels of math and reading achievement (see Table 3, Model 4). Nevertheless, this effect is only statistically significant in five countries for math achievement and nine countries for reading achievement. Other control variables in the model are not the primary concern of this study.

Conclusion and discussion

In the past fifteen years, many countries have reformed their educational systems with reference to the PISA studies. With respect to language policies, a substantial influence is exerted by PISA reports. Some of the PISA reports claimed that NS students outperform LMi students. Based on this finding, inferences have been drawn about the consequences of speaking and exposure to a ML. Many policymakers responded to these PISA reports with a monolingual reflex, most visibly by increasing sociopolitical pressure on linguistic minorities to abandon their mother tongues (Cummins, 2008; Pulinx et al., 2016). However, it was not possible to examine the impact of students' language use because very limited information was available about the language use of students. That is, only a dichotomous language background variable was assessed. The most recent PISA data provide a unique opportunity to fill this research lacuna as the PISA 2012 included a more elaborated questionnaire on linguistic issues. As such, first research question (RQ1) of this study was: what is the association between students' language background (i.e. the first language learned at home) and their academic achievement? The second research question (RQ2) was: what is the relationship between language use (with different persons and within different context) and academic achievement for language minority students?

With respect to RQ1, the results of the multivariate three-level regression analyses revealed that there is indeed an achievement gap between LMi and NS students, and this holds true for both reading and math. After taking account of students' background characteristics, students' academic profile, and school characteristics, the LMi–NS achievement gap narrows but remains significant. Separate analyses made clear that this is true for almost all countries in our analysis.

With RQ2, we investigated whether exposure to IL or ML with different interlocutors and within different contexts is related to academic performance. We found that LMi students who speak a ML more often with their parents do not achieve less. In most countries, language use is unrelated to academic performance. In a few countries, such as Canada, Finland, and Singapore, speaking a ML with parents has a *positive* effect. Future research should examine how the differences between countries can be explained. These findings fail to confirm the time on task hypothesis and provide partial support for the additive perspective on multilingualism.

It is important to bear in mind a few weaknesses of this study. A potential limitation is related to the cross-sectional design of our data: we could only indirectly rule out selection effects.

Consequently, we had a limited measure of students' previous academic performance (i.e. grade, grade repetition, and track included as control variables). Future research with longitudinal data could partly overcome this problem. Second, we also controlled for immigrant background and made a distinction between first and second generation immigrants. However, we were unable to distinguish further between different countries of origin because many this information is not available for many countries that participate in PISA. And third, it is not possible to examine the relationship between first and second language skills with the PISA data as it only measures academic achievement in the national languages.

With regard to educational policy, our study clarifies that the relation between language use and academic achievement is a complex one that cannot be resolved with quick fixes. First, it is important to recognize that there *is* an achievement gap between LMi and NS students. As such, policymakers should address this inequity. Second, as language use *per se* does not trigger underachievement, sociopolitical pressure towards monolingualism should be avoided (see also Piller, 2001). More exposure to a ML at home does not decrease students' academic performance. This contradicts the point of view of official OECD reports (see OECD, 2013a).

Finally, exposure to the IL remains important, in particular in the school context. Indeed, we found that in several countries, speaking IL with schoolmates is positively associated with math and reading achievement. This might reflect the negative effect of within-school segregation. Exposure and speaking IL is equally important, as our results show that speaking IL with school friends is beneficial. To this end, within-school segregation (such as tracking and ability grouping) should be dismantled. Even language learning programs that aim to improve IL proficiency should minimize the separation of LMi and NS students. Moreover, bilingual education programs that do not segregate LMi and NS students are preferred above those that require such segregation. Recent literature shows that it is possible to provide multilingual education without segregating students in different groups (see Van Laere, Agirdag, & Van Braak, 2016).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported by a Veni-grant by the Netherlands Organisation for Scientific Research (NWO), Project no. 451-14-010.

References

Agirdag, O. (2010). Exploring bilingualism in a monolingual school system: insights from Turkish and native students from Belgian schools. *British Journal of Sociology of Education*, *31*, 307–321.

Agirdag, O. (2014). The long-term effects of bilingualism on children of immigration: student bilingualism and future earnings. *International Journal of Bilingual Education and Bilingualism*, *17*, 449–464.

Agirdag, O., Jordens, K., & Van Houtte, M. (2014). Speaking Turkish in Belgian primary schools: Teacher beliefs versus effective consequences. *Bilig: Journal of Social Sciences of the Turkish World*, *70*, 7–28.

Agirdag, O., Yazici, Z., & Sierens, S. (2015). Trends in pre-school enrolment in Turkey: unequal access and differential consequences. *Comparative Education*, *51*, 537–554.

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*, 327–336.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development*, *70*, 636–644.

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism*, *9*(1), 103–119.

Breakspear, S. (2012). *The Policy Impact of PISA: An Exploration of Normative Effects of International Benchmarking in School System Performance*. OECD Education Working Papers, No. 71, OECD Publishing.

Brenneman, M. H., Morris, R. D., & Israelian, M. (2007). Language preference and its relationship with reading skills in English and Spanish. *Psychology in the Schools*, *44*(2), 171–181.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Bilingual Education Paper Series*, *3*, 1–69.

Cummins, J. (2008). Essay review [Review of the book *Where immigrant students succeed: A comparative review of performance and engagement in PISA 2003*, by P. Stanat & G. Christensen]. *Curriculum Inquiry*, *38*(4), 493–497

Dronkers, J., & Kornder, N. (2014). Do migrant girls perform better than migrant boys? Deviant gender differences between the reading scores of 15-year-old children of migrants compared to native pupils. *Educational Research and Evaluation*, *20*(1), 44–66.

Dronkers, J., & Van der Velden, R. (2013). Positive but also negative effects of ethnic diversity in schools on educational achievement? An empirical test with cross-national PISA data. In M. Windzio (Ed.). *Integration and Inequality in Educational Institutions* (pp. 71–98). New York, NY: Springer

Duursma, E., Romero-Contreras, S., Szuber, A., Proctor, P., Snow, C., August, D., & Calderón, M. (2007). The role of home literacy and language environment on bilinguals‘ English and Spanish vocabulary development. *Applied Psycholinguistics*, *28*(1), 171–190.

Edele, A., & Stanat, P. (2016) The role of first-language listening comprehension in second-language reading comprehension. *Journal of Educational Psychology*, *108*(2), 163–180.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*, 430-457.

Ertl, H. (2006). Educational standards and the changing discourse on education: The reception and consequences of the PISA study in Germany. *Oxford Review of Education*, *32*, 619–634.

Esser, H. (2006). *Sprache und Integration: die sozialen Bedingungen und Folgen des Spracherwerbs von Migranten*. Frankfurt, Germany: Campus.

Extra, G., & Yagmur, K. (2005). Sociolinguistic perspectives on emerging multilingualism in urban Europe. *International Journal of the Sociology of Language*, *175*, 17–40.

Grek, S. (2009). Governing by numbers: The PISA ‘effect’ in Europe. *Journal of Education Policy*, *24*(1), 23–37.

Gür, B. S., Çelik, Z., & Özoğlu, M. (2012). Policy options for Turkey: A critique of the interpretation and utilization of PISA results in Turkey. *Journal of Education Policy*, *27*(1), 1–21.

Hammer, C. S., Davison, M. D., Lawrence, F. R., & Miccio, A. W. (2009). The effect of maternal language on bilingual children’s vocabulary and emergent literacy development during Head Start and kindergarten. *Scientific Studies of Reading*, *13*(2), 99–121.

Hannover, B., Morf, C. C., Neuhaus, J., Rau, M., Wolfgramm, C., & Zander-Musić, L. (2013). How immigrant adolescents’ self-views in school and family context relate to academic success in Germany. *Journal of Applied Social Psychology*, *43*(1), 175–189.

Jaspers, J. (2008). Problematizing ethnolects: Naming linguistic practices in an Antwerp secondary school. *International Journal of Bilingualism*, *12*(1–2), 85–103.

Laursen, H. P. (2013). Umbrellas and angels standing straight–a social semiotic perspective on multilingual children’s literacy. *International Journal of Bilingual Education and Bilingualism*, *16*, 690-706.

Leseman, P. P. (2000). Bilingual vocabulary development of Turkish preschoolers in the Netherlands. *Journal of Multilingual and Multicultural Development*, *21*(2), 93–112.

Levels, M., Dronkers, J., & Kraaykamp, G. (2008). Immigrant children’s educational achievement in western countries: Origin, destination, and community effects on mathematical performance. *American Sociological Review*, *73*, 835–853.

Mancilla-Martinez, J., & Lesaux, N. K. (2011). Early home language use and later vocabulary development. *Journal of Educational Psychology*, *103*, 535.

Marks, G. N. (2005). Accounting for immigrant non-immigrant differences in reading and mathematics in twenty countries. *Ethnic and Racial Studies*, *28*, 925–946.

Muthen, L. K., & Muthen, L. (2012). *Mplus version 7* [computer software]. Los Angeles, CA: Muthen & Muthen.

OECD (2013a). *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed*. Paris, France: OECD Publishing.

OECD (2013b). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. Paris, France: OECD Publishing.

Pearson, B. Z., & Fernández, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning*, *44*(4), 617–653.

Piller, I. (2001). Naturalization language testing and its basis in ideologies of national identity and citizenship. *International Journal of Bilingualism*, *5*(3), 259–277.

Portes, A., & Hao, L. (2002). The price of uniformity: Language, family and personality adjustment in the immigrant second generation. *Ethnic and Racial Studies*, *25*, 889–912.

Prevoo, M. J., Malda, M., Mesman, J., Emmen, R. A., Yeniad, N., Van Ijzendoorn, M. H., & Linting, M. (2014). Predicting ethnic minority children's vocabulary from socioeconomic status, maternal language and home reading input: Different pathways for host and ethnic language. *Journal of Child Language*, *41*, 963–984.

Pulinx, R., Van Avermaet, P., & Agirdag, O. (2016) Silencing linguistic diversity: The extent, the determinants and consequences of the monolingual beliefs of Flemish teachers. *International Journal of Bilingual Education and Bilingualism*. Epub ahead of print 27 October 2015. doi:10.1080/13670050.2015.1102860.

Rossell, C., & Baker, K. (1996). The educational effectiveness of bilingual education. *Research in the Teaching of English*, *30*(1), 7–74.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Scheele, A. F., Leseman, P. P., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*(1), 117–140.

Schnepf, S. V. (2007). Immigrants' educational disadvantage: An examination across ten countries and three surveys. *Journal of Population Economics*, *20*, 527–545.

Stanat, P., & Christensen, G. (2006). *Where immigrant students succeed: A comparative review of performance and engagement in PISA 2003*. Paris, France: OECD.

Van Laere, E., Agirdag, O., & Van Braak, J. (2016) Supporting science learning in linguistically diverse classrooms: Factors related to the use of bilingual content in a computer-based learning environment. *Computers in Human Behavior*. *57*, 428–441.

Van Tubergen, F., & Kalmijn, M. (2009). Language proficiency and usage among immigrants in the Netherlands: Incentives or opportunities? *European Sociological Review*, *25*(2), 169–182.

Author biographies

Orhan Agirdag, PhD, is an assistant professor at the Laboratory for Education and Society at KU Leuven, an adjunct professor at University of Amsterdam, and a former Fulbright Fellow at the University of California, Los Angeles. His main research interests include inequality in education, educational policy, teacher education, student identities, multilingualism, and religiosity.

Gudrun Vanlaar, PhD, is a post-doctoral researcher investigating which educational practices are the most effective for high-risk students. She is interested in multilevel modelling and causal inference techniques. She is affiliated with the Centre for Educational Effectiveness and Evaluation at KU Leuven, Belgium.

Appendix A. Academic achievement differences between native speakers (NS) and language minority (LMi) students across countries.

		N	Reading		Math	
			Mean	SD	Mean	SD
Austria	LMi	792	462.72	89.10	470.21	87.75
	NS	3605	505.80	84.78	523.33	85.71
Belgium (French)	LMi	732	487.30	97.36	490.69	91.95
	NS	2604	520.08	91.93	515.85	88.45
Belgium (Flemish)	LMi	721	482.24	95.69	484.00	97.25
	NS	3787	539.92	87.28	555.53	95.11
Canada	LMi	3309	517.12	93.01	516.20	89.26
	NS	16324	519.67	87.27	516.66	83.59
Denmark	LMi	1769	454.84	83.85	448.57	80.61
	NS	4926	506.09	80.79	509.12	79.63
Finland	LMi	1496	465.59	105.40	460.75	92.56
	NS	6856	529.18	89.57	524.21	80.85
Germany	LMi	832	488.32	86.09	486.63	89.99
	NS	3271	529.67	83.00	535.48	91.10
Hong Kong	LMi	458	516.90	87.58	524.57	91.31
	NS	3908	552.51	83.52	570.30	94.00
Hungary	LMi	190	456.27	113.68	455.80	106.66
	NS	4441	501.88	84.84	489.69	89.08
Ireland	LMi	382	508.04	87.81	497.90	81.79
	NS	4223	531.49	82.36	506.82	82.41
Italy	LMi	4277	466.65	97.30	467.90	89.01
	NS	23022	512.09	87.62	505.74	86.72
Latvia	LMi	822	479.94	86.03	483.02	80.92
	NS	3324	502.81	78.59	501.16	79.90
Luxembourg	LMi	1778	478.07	93.93	477.91	86.46
	NS	2683	518.58	96.18	518.38	89.89
Portugal	LMi	532	469.05	96.83	468.70	93.19
	NS	4851	493.04	89.16	492.06	91.23
Serbia	LMi	438	454.27	89.48	453.32	87.78
	NS	3807	453.29	87.72	454.38	87.73
Singapore	LMi	4068	539.28	98.64	571.26	103.47
	NS	1220	545.07	100.67	571.32	105.94
Slovak Republic	LMi	499	428.06	107.24	450.03	99.74
	NS	3726	484.40	95.84	499.11	98.56
Slovenia	LMi	776	420.77	93.73	446.81	84.42
	NS	4256	481.31	87.29	500.84	86.03

Means, standard deviations (SD) and number of observations (N).