Does bilingualism correlate with or predict higher proficiency in L3 English? A contrastive study of monolingual and bilingual learners

Eliane Lorenz,a Sharareh Rahbari,b Ulrike Schackowc and Peter Siemundd

Abstract

This study investigates whether English in instructed settings is more successfully acquired by learners who are already bilingual in comparison to those with a monolingual background. There remains substantial controversy regarding potential advantages of bilingual speakers in their acquisition of additional languages, especially in heritage speaker contexts. We here contribute to this discussion by analysing the English C-test results of 1,718 bilingual and monolingual students of grades 7 and 9, sampled in schools across Germany. The bilingual students speak either Russian or Turkish (heritage language) and German (majority language). The monolingual control group was raised in German only. The main predictor variables are reading fluency and comprehension in German and the heritage languages. Additional predictor variables include school type, school year, socioeconomic status, cognitive ability, amongst others. Using correlation and regression analyses we test if reading fluency and comprehension impact proficiency in English and if bilingual students enjoy advantages over their monolingual German peers. The results reveal no systematic advantage of bilingual students, although we find significant correlations between reading

Affiliations

a Norwegian University of Science and Technology (NTNU), Norway
email: eliane.lorenz@ntnu.no (corresponding author)
b University of Hamburg, Germany
email: sharareh.rahbari@uni-hamburg.de
c University of Hamburg, Germany
email: ulrike.schackow@gmail.com
d University of Hamburg, Germany
email: peter.siemund@uni-hamburg.de
fluency and comprehension and C-test results. School type, cognitive skills, among others, are predictors for English performance, whereas socioeconomic status returns no significant effect.

**Keywords:** advantages; bilingualism; heritage speakers; interdependence; multilingualism.

1 Introduction and background

We here probe into the widely held assumption that previous language experience qua bilingual or multilingual upbringing fosters the acquisition of additional languages. Such a ‘multilingual boost’ (Berthele and Udry, 2019) could be evidenced for balanced bilinguals acquiring English, especially in the Basque Country and Catalunya (Agustín-Llach, 2019; Cenoz, 2003, 2013; Cenoz and Valencia, 1994; Lasagabaster, 1998, 2001; Safont Jordà, 2003; Sagasta Errasti, 2003; Sanz, 2000), and more recently also for bilingual heritage speakers in their acquisition of English, most prominently in the context of Germany but partially also other countries (Edele, Kempert and Schotte, 2018; Fleckenstein, Möller and Baumert, 2018; Göbel, Rauch and Vieluf, 2011; Hirosh and Degani, 2018; Hopp et al., 2019; Klieme et al., 2006; Maluch et al., 2015; Maluch, Neumann and Kempert, 2016; Sanders and Meijers, 1995; Schwartz et al., 2007; Siemund and Mueller, 2020; van Gelderen et al., 2003). It is especially this latter context of bilingual heritage speakers that we are interested in here.

The discussion of bilingual and multilingual advantages remains controversial for at least the following reasons. First, the type of bilingualism influences the results. Cenoz (2003: 82) argues that multilingual advantages for additional language acquisition surface most prominently with balanced bilinguals but are less obvious for subtractive bilinguals (see also Cummins, 1979; Lambert, 1974). Second, the acquisition of general proficiencies, such as oral and written skills, appears more susceptible to multilingual boosting than specific skills related to grammatical development (Cenoz, 2003:80; Lorenz, 2019; Siemund and Lechner, 2015; Siemund and Mueller, 2020). Third, there is the problem of whether observable multilingual advantages are causally related to language development or to more general cognitive skills (Berthele and Udry, 2019; Berthele and Vanhove, 2020). Fourth, there are differences in relation to language background. Maluch et al. (2015: 82) report impressive English achievement gains for the Chinese-German and Polish-German bilingual learners in their sample, though less augmentation for the Turkish-German and Arabic-German groups. Bilingual group differences are also discussed in Fleckenstein et al. (2018). Fifth, multilingual advantages (operationalised with measures assessing lexical development, general language achievement
and grammatical performance) appear to be more easily detectable in younger cohorts, though they fade as students grow older (Hopp, Vogelbacher et al., 2019; Maluch et al., 2016; Siemund and Lechner, 2015). Sixth, the influence of the majority language generally seems to be stronger. In regard to linguistic interdependence, the majority language exerts a positive influence (Edele et al., 2018). Positive influence from the heritage languages appears more difficult to observe. With respect to metalinguistic awareness, Angelovska (2018:142, 147) reports that it is most frequently the dominant language that is activated in L3 written production. Seventh, the context of language use matters, with educational contexts producing results different from a home or family context (Hirosh and Degani, 2018:899). Eighth, the socioeconomic status, here especially parental education and to a lesser degree household income, appears to be a strong predictor for an advantage (or respectively disadvantage) in L3 acquisition in a school context (de Angelis, 2015). And ninth, bilinguals who are fully literate in both languages outperform their partially literate bilingual peers in the L3 (Rauch, Naumann and Jude, 2012).

An important problem for the unambiguous validation of bilingual and multilingual advantages lies in the methodologies used. No study known to us is based on randomised samples, and we here cannot overcome this problem either for ethical or logistic reasons. Thus, the cohorts typically differ in size and internal composition. Moreover, the language combinations vary and so do the proficiencies of the students in their background languages. Most studies use different test designs and focus on disparate aspects of language. For example, Göbel et al. (2011) and Klieme et al. (2006) rely on data collected within the DESI-study (Deutsch-Englisch Schülerleistungen-International) that sorts students according to their first language, i.e., into ‘German,’ ‘non-German’ and ‘multilingual’ (more than one first language). Maluch et al. (2015) worked with a 6th-grade elementary school sample, measuring English language achievement with a Cloze test and German proficiency with a reading test. The language background of the students \( n = 939 \) bilinguals versus \( n = 1,896 \) monolinguals was identified by parents’ and students’ reports. Maluch et al. (2016), in contrast, is based on a sample of 1,032 students, discriminating between German monolingual students \( n = 820 \), mixed dominant German students \( n = 119 \), mixed dominant non-German students \( n = 63 \) and non-German only students \( n = 30 \), as determined by parental reports (heritage languages include Turkish, Arabic, Chinese, Polish, French, Italian, Russian, Vietnamese and others). Language achievement in English was assessed with a Cloze test in grades 6 and 8. In Maluch and Kempert (2019), language background is captured using students’ reports (bilingual: \( n = 456 \), monolingual: \( n = 839 \)), while the research focus is on listening and reading comprehension in English as a foreign language. Edele et al. (2018) established language
background on the basis of student questionnaires and measured proficiency in English and German in terms of reading comprehension, though using listening comprehension tests for Russian and Turkish. They analysed 8,752 German 10th-grade students (\(n = 7,964\) monolinguals, \(n = 436\) Turkish-German students, \(n = 352\) Russian-German students). Fleckenstein et al. (2018) analysed \(n = 541\) bilingual students in grade 9, comparing them with \(n = 3,289\) monolingual German students. The bilingual students were recruited from ten schools and, besides German, have French, Greek, Italian, Polish, Portuguese, Russian, Spanish or Turkish as their background language. They used reading comprehension tests to assess students’ proficiency in English and the background languages. The study by Hopp, Vogelbacher et al. (2019) is primarily concerned with lexical development and uses a comparatively small sample: 200 students (\(n = 88\) monolingual German, and \(n = 112\) bilingual) at the end of grade 3.

2 Research questions

Bilingualism and bilingual education have been argued to impact positively on cognitive development and educational attainment. In this research strand, bilingualism tends to be understood in a narrower sense, namely in terms of ‘cognitive academic language proficiency’ (CALP; Cummins, 2001: 58), which contrasts with ‘basic interpersonal communicative skills’ (BICS). Hulstijn (2015:19–20) proposes another classification scheme distinguishing between ‘basic language cognition’ (BLC) and ‘higher (extended) language cognition’ (HLC). BLC refers to the knowledge that all (adult) native speaker share (Hulstijn, 2015:32, 55; see also Hulstijn, 2019:158). HLC, which could tentatively be described as academic language, needs to be taught explicitly (i.e., depends on the educational history) and more importantly, native speakers differ along this dimension (Hulstijn, 2015:22; Hulstijn, 2019:164–65). Others suggest that it would perhaps be more adequate to regard this opposition as a continuum instead of a dichotomy (Baker, 2001:170–72; Dąbrowska 2012:332; Kecskes and Paap, 2000:xviii–xx; but for a critical discussion see Hulstijn, 2015:54–56).

Cummins (1979) formulated two hypotheses, namely the ‘threshold’ and the ‘interdependence’ hypotheses, regarding academic language competence in the bilingual speaker. Cummins (1979:230) proposes that positive cognitive effects and hence higher educational attainment can be expected in situations of additive bilingualism, though not necessarily in cases of dominant bilingualism. The threshold hypothesis requires a certain level of academic language competence to be exceeded before positive cognitive effects of bilingualism become visible (Cenoz, 2003:82). As for the interdependence hypothesis and the bilingual mind, Cummins (1979:233) explains that academic language
competence in the one language can be expected to impact positively on such
capability in the other language. This entails important consequences for
the education of heritage speakers, as language capability in the majority
language depends, inter alia, on heritage language competence. In other words,
it may be advantageous to develop the heritage language of heritage speakers
academically. However, Berthele and Vanhove (2020:553–554) argue that the
interdependence between languages remains difficult to operationalise and that
alternative principles (language giftedness or more general cognitive skills)
could be evoked to account for observable effects. Berthele and Udry (2019)
fail to substantiate such bilingual advantages for heritage speakers.

Bilingualism and multilingualism are widely believed to foster metalinguis-
tic awareness (Bialystok, 2001; de Angelis, 2007; Herdina and Jessner, 2002;
Jessner, 2008). We believe that metalinguistic awareness is the main driving
force behind bilingual advantages for language acquisition (Jessner’s ‘M-factor’
[2008:275]; Hirosh and Degani, 2018:892, 907), yet again, it seems to matter
what type of bilingual speaker is considered. Rauch et al. (2012) argue that
biliteracy is the key aspect for heightened metalinguistic awareness. Spellerberg
(2016) even reports a negative correlation between bilingualism and metalin-
guistic awareness in the context of heritage speakers. El Euch (2010) did not
find differences regarding metalinguistic awareness between bilinguals and
trilinguals.

Berthele and Vanhove (2020:551) point out that what is at issue in the
research area described above is crosslinguistic influence of a specific type.
While it is completely uncontroversial that the languages in the bilingual or
multilingual mind influence one another, traditionally conceptualised in terms
of facilitation and interference (see Odlin, 2003), the debate about multilin-
gual advantages suggests facilitation effects of a more general cognitive type.
These are assumed to impact different kinds of language proficiency (Hirosh
and Degani, 2018:893).

The present study addresses two questions, namely (i) whether there are cor-
relations between reading fluency and comprehension in the background lan-
guages (heritage and majority language) and proficiency in English as a third
language measured in terms of a C-test, and (ii) whether any such correlations
persist if additional confounding variables are controlled for. Following many
other studies cited above, we thus search for interdependence in multilingual
contexts, asking if heritage bilingualism impacts the acquisition of English as
a third language positively or not.

3 Methodology

The study reported here is part of a larger research project on multilingual
development (Multilingual development: a longitudinal perspective [MEZ,
This longitudinal project studies the linguistic development of German secondary school students from school year 7 (age 12 and 13) and year 9 (age 14 and 15). The study involves four waves of data collection between 2016 and 2019. It focuses on bilingual cohorts of Russian-German and Turkish-German heritage speakers, contrasting these with monolingual German control groups. The students of English can be classified as beginners/intermediate learners. 8

3.1 Participants
We here focus exclusively on the data of the first wave of data collection (carried out between January and March 2016) in which a total of 1,875 students were approached. Of these, 1,818 eventually participated in the sampling procedures. All bilingual students had to go through a so-called locator test (Hellrung et al., 2017:18) to ensure that they had at least some proficiency in the heritage language. For this test, the students were played audio recordings of a native speaker of Russian, respectively Turkish, reading out 8 sentences in the target language (Russian or Turkish). For each sentence the students were shown several pictures out of which only one illustrated the heard content. In addition, one hundred of the students had to be excluded due to contradictory statements regarding their background languages. Table 1 shows the final distribution of the students.

Table 1. Participants per language group

| Background languages                     | Number of students |
|------------------------------------------|--------------------|
| German monolingual (GER)                 | 914                |
| Russian-German bilingual (RUS-GER)       | 319                |
| Turkish-German bilingual (TUR-GER)       | 485                |
| Total                                    | 1,718              |

GER = German, RUS = Russian, TUR = Turkish

For the majority of the bilingual participants, the heritage language is largely confined to the home context, in particular when conversing with their parents. Nearly 40% reported using the heritage language mainly or exclusively with their parents. Less than 10% said they use the heritage language with their siblings. Among their friends and peers in school, only approximately 5% indicated using the heritage language mainly or exclusively. Moreover, all participants grow up in Germany and attend German schools, which means that they are surrounded by German on a daily basis (and most likely during the larger part of the day). Hence, the bilinguals are classified as dominant in German. As we will show later, the scores in the reading comprehension
and fluency tests in German and the heritage languages confirm this assessment: the means of the two scores obtained in the German version are significantly higher (based on t-test comparisons) than the scores in the Russian and Turkish versions.

The students were interviewed on two days that followed within the period of about one week. They were tested in regard to their linguistic and cognitive abilities. In addition, questionnaires were distributed collecting various background variables like socioeconomic status, gender, age, school type, the students’ self-concept, their motivation to learn English and their self-assessment regarding the learning of languages.

3.2 Instruments used
The two main instruments of the current study were a reading fluency and comprehension test (LGVT – Lesegeschwindigkeits- und Verständnistest) as well as a C-test. The reading fluency and comprehension test measured the receptive skills of students (Schneider, Schlagmüller and Ennemoser, 2016). This test was conducted in German as well as in the respective heritage language (Russian or Turkish). The students received gapped texts containing options to fill the gaps, out of which only one option made sense in the given context. The text needed to be fully understood to identify the correct option. This procedure measured text comprehension. The test contained 47 gaps and had to be completed within a time span of six minutes. For each gap, the students received two points for choosing the correct option, zero if the gap was skipped or left empty, and a negative point (−1) if the wrong option was chosen. This established a total range of −47 to +94 points for each student. The second measure provided by this test concerned reading fluency, which was measured by how much text (number of words) the students were able to process out of a total of 2,105 words. As the original version of the test was only available for German, it was translated and adapted for the heritage languages Russian and Turkish. The stories used in the heritage language versions were not identical to the German story, but they came from the same genre and addressed a similar topic.

We here use only one test measuring reading comprehension and fluency in German and the heritage languages to approximate proficiency in the respective languages. This is not uncontroversial and there are numerous studies, especially Trapman et al. (2017:252), which have shown that reading comprehension and fluency are interconnected with vocabulary and grammatical knowledge and may even be linked via a causal or reciprocal relationship. It would therefore be desirable to include not just one component, as was done in the current study, but to employ several other measures as well (see Trapman et al., 2017:243–45). This, however, was not feasible due to the large battery of
test instruments the participants had to complete on two successive days at each measuring point.

We measured lexical and grammatical competence in English by means of a C-test. Here, students had to complete end-clipped words in a text and also needed to adjust the words grammatically, for instance decide between have, has or had for the first gap in: Some schools ha__ contacts all ov__ Europe. We used four short texts on different topics that the students were asked to complete within 20 minutes. The four texts contained a total of 90 end-clipped words and a combined score for all four texts was computed and finally used. For each gap, the students received one point for correct and zero for incorrect answers. Misspellings or incompletely filled gaps were considered incorrect. Therefore, the C-test scale ranged from 0 to 90.

C-test measures are not uncontroversial, as Drackert and Timukova (2020) have demonstrated in a recent study. They used a Russian C-test with two different learner populations, namely foreign language learners of Russian and Russian heritage speakers. They showed that the heritage language learners had difficulties in producing the correct spelling for the end-clipped words, even though they excelled in retrieving the right meaning of the text (Drackert and Timukova, 2020:123–24). As they further explain, ‘to complete a C-test gap correctly, a learner needs to have the item in their vocabulary, to identify the item correctly based on the context, and to produce its correct grammatical and orthographical form’ (2020:123), at least if only completely correct items are counted as correct (instead of allowing partially correct answers). We here assess both vocabulary as well as grammatical knowledge. Since the participants in the current study are all foreign language learners, unlike Drackert and Timukova’s (2020) participants, we decided to use a dichotomous scoring throughout.

For cognitive ability we used the subtest N2 of the standardised German cognitive ability test KFT 4–12+ R (Heller and Perleth, 2000) designed for students from grade 4 to 12 to measure visual-spatial ability. Each student received 25 test items in total. For correct answers they gained one point each and for incorrect answers zero, which established a scale ranging from 0 to 25 points.

We operationalised the socioeconomic status with the Highest International Socio-Economic Index (HISEI) of the families. This index was calculated using the students’ father’s and mother’s occupations. Their professions were established according to the International Standard Classification of Occupations code, the ISCO-08-categorisation (International Labor Office, 2012). In a subsequent step, these codes were transferred to the International Socio-Economic Index (ISEI) (Ganzeboom, 2010). For the present analysis, we used the HISEI values that may range from 16 up to 90.
The students’ self-concept in relation to language learning was measured by their degree of agreement to the following statements: (i) ‘it is easy for me to learn languages,’ (ii) ‘my ability to acquire additional languages is high,’ (iii) ‘I’m not talented in acquiring additional languages’ and (iv) ‘I’m frightened to make mistakes when using a foreign language.’ Their motivation to learn English is inferred from their degree of agreement with the statements: (i) ‘learning English is fun,’ (ii) ‘the school subject English is boring,’ (iii) ‘it is important for me to have good grades in English’ and (iv) ‘it is important for me to be good in English.’ The rating scales ranged from one (strongly disagree) to four (strongly agree). For the students’ self-assessment regarding the learning of English, they had to rate their abilities in pronunciation, writing, grammar and punctuation in general, grammar of speaking, grammar of writing, and their vocabulary and lexical knowledge, using a six-point rating scale ranging from ‘excellent’ to ‘very bad.’

4 Results

The following sections draw on three types of qualitatively different results. First, the dataset will be briefly dissected along various dimensions such as school type, cognitive abilities, socioeconomic background and results from the reading fluency and comprehension test as well as the English C-test. We here contrast the three different language groups (GER, RUS-GER and TUR-GER) and inspect the data accordingly. Second, we investigate the correlational strength between the scores of the reading comprehension and fluency tests (in German and the heritage languages) and the scores of the English C-test. And third, we will run two regression analyses to assess the impact of all predictor variables on the model of variance, with the English C-test results as the response variable.

4.1 Response and predictor variables in isolation

Table 2 offers a closer look at the distribution of the individual language groups according to the type of school and the socioeconomic background. The division between Gymnasium and other is nearly equal for the monolinguals as well as the bilinguals. We only differentiate between the university-bound school track (i.e., Gymnasium or GY) and the vocational school tracks (i.e., other), following Lechner and Siemund (2014) as well as Maluch and Kempert (2019). There is an observable (proficiency and performance) gap between students attending Gymnasium and those attending any of the other types of schools, whereas we typically find little variation among the students attending the latter types (Lechner and Siemund, 2014:334).
When considering the socioeconomic status (HISEI), it becomes apparent that the monolingual German students on average have higher scores than either bilingual group. This reaches statistical significance \((F(2, 1512) = 88.62, p < .001 \eta^2 = 0.10)\). Common statistics report that young adolescents from immigrant families have a lower socioeconomic status than their monolingual German peers (see for example Stanat et al., 2016:463).

Table 2. Distribution of language groups according to type of school and socioeconomic status

| Type of school | Socio-economic status (HISEI)* |
|---------------|-------------------------------|
|               | GY          | other        | Mean  | Median | SD     |
| GER absolute  | 476         | 438          | 55.19 | 56.00  | 20.20  |
| in %          | 52.08       | 47.92        |       |        |        |
| RUS-GER absolute | 152        | 167          | 43.68 | 36.35  | 19.66  |
| in %          | 47.65       | 52.35        |       |        |        |
| TUR-GER absolute | 232        | 253          | 40.22 | 31.72  | 18.83  |
| in %          | 47.84       | 52.16        |       |        |        |
| Total absolute | 860         | 858          | 49.37 | 47.83  | 20.89  |
| in %          | 50.06       | 49.94        |       |        |        |

GY = Gymnasium (university-bound school track), SD = standard deviation, GER = German, RUS = Russian, TUR = Turkish

* Mean/median/standard deviation (SD) are based on 1,515 students, information of \(n = 203\) students is missing. This includes \(n = 53\) German monolinguals, \(n = 36\) Russian-German bilinguals and \(n = 114\) Turkish-German bilinguals.

The results of the cognitive ability test, the reading fluency and comprehension test as well as the C-test are summarised in Table 3. The German monolingual group achieved the highest scores in the German reading fluency and comprehension test (Reading comprehension: \(F(2, 1668) = 47.31, p < .001 \eta^2 = 0.05\); Reading fluency: \(F(2, 1637) = 11.7, p < .001 \eta^2 = 0.01\)) and also in the English C-test (\(F(2, 1669) = 10.89, p < .001 \eta^2 = 0.01\)). The Russian-German bilingual group obtained slightly better results in the German reading test when considering reading comprehension in comparison to the Turkish-German group, as well as the C-test in English. The former is significant \((p < .001)\), however, the difference for the English C-test does not reach statistical significance \((p = .100)\). The same holds true for reading fluency. Even though the Turkish-German group on average has a slightly higher value than the Russian-German group, this difference is not statistically significant \((p = .780)\).
Table 3. Test scores: mean, median, standard deviation (SD)

| Cognitive ability test* | Reading comprehension German* | Reading fluency German* | Reading comprehension heritage language* | Reading fluency heritage language* | C-Test English* |
|-------------------------|------------------------------|-------------------------|------------------------------------------|----------------------------------|----------------|
| GER                     | Mean 17.43                   | 31.21                   | 861.33                                   | –                                | 53.18          |
|                         | Median 19.00                 | 30.00                   | 821.00                                   | –                                | 55.00          |
|                         | SD 5.61                      | 10.14                   | 223.14                                   | –                                | 18.74          |
| RUS-GER                 | Mean 17.44                   | 28.52                   | 805.79                                   | 5.58                             | 405.29         |
|                         | Median 19.00                 | 29.00                   | 779.00                                   | 3.00                             | 313.00         |
|                         | SD 5.30                      | 9.82                    | 217.89                                   | 10.06                            | 300.58         |
| TUR-GER                 | Mean 14.05                   | 25.63                   | 807.18                                   | 8.07                             | 560.63         |
|                         | Median 15.00                 | 26.00                   | 780.00                                   | 7.00                             | 475.00         |
|                         | SD 6.51                      | 10.23                   | 243.97                                   | 9.10                             | 296.76         |
| Total                   | Mean 16.49                   | 29.16                   | 836.34                                   | 7.12                             | 499.70         |
|                         | Median 18.00                 | 29.00                   | 815.00                                   | 5.00                             | 437.00         |
|                         | SD 6.02                      | 10.39                   | 229.63                                   | 9.55                             | 307.76         |

GER = German, RUS = Russian, TUR = Turkish

* Cognitive ability is based on 1,675 students (NA = 43); Reading comprehension in German is based on 1,671 students (NA = 47); Reading fluency in German is based on 1,640 students (NA = 78 students); Reading comprehension in the heritage language is based on 714 students (NA = 93); Reading fluency in the heritage language is based on 699 students (NA = 108); C-Test in English is based on 1,672 students (NA = 46).

Regarding their cognitive abilities, there is a statistically significant difference across the three groups \(F(2, 1672) = 56.69, p < .001, \eta^2 = 0.06\). The German monolinguals and the Russian-German group rank each significantly above the Turkish-German group (GER vs TUR-GER: \(p < .001\); RUS-GER vs TUR-GER: \(p < .001\)), while the difference between the German monolinguals and the Russian-German bilinguals is not significant (\(p = .640\)). Finally, we note that the values for reading comprehension in the heritage languages are drastically below those measured for German. This time, the Turkish-German cohort outperforms the Russian-German cohort, both \(t\)-test comparisons return statistically significant differences (Reading comprehension in the heritage language: \(t(553.55) = -3.33, p < .001, d_{\text{cohen}} = 0.26\); Reading fluency in the heritage language: \(t(574.88) = -6.68, p < .001, d_{\text{cohen}} = 0.52\)).

The results for the variables self-concept, motivation and self-assessment are shown in Table 4. What the data basically reveal is that we do not find especially strong variation across the three groups distinguished here. Generally speaking, the students score highest when self-assessing their skills in English,
and the motivational scores are marginally above those for self-concept. The small difference observed for self-assessment is not significant, though it is for self-concept and motivation (Self-assessment: $F(2, 1666) = 2.04, p = .131, \eta^2 = 0.01$; Self-concept: $F(2, 1667) = 8.25, p < .001, \eta^2 = 0.01$; Motivation: $F(2, 1667) = 6.33, p = .002, \eta^2 = 0.01$). The Turkish-German bilinguals see themselves as more capable of learning languages, which is significant, when compared to the monolingual German students ($p < .001$), but not when compared to the Russian-German bilinguals ($p = .061$). Similarly, they perceive themselves as more highly motivated. This difference is significant both when compared to the monolingual cohort ($p = .001$) and when compared to their Russian-German peers ($p = .015$). There is no statistically significant difference between the German monolinguals and the Russian-German bilinguals (Self-concept: $p = .250$; Motivation: $p = .930$).

**Table 4.** Self-concept, motivation and self-assessment

|        | Self-concept* | Motivation* | Self-assessment* |
|--------|---------------|-------------|------------------|
| GER    | Me 2.64       | 4.14        | 4.14             |
|        | Mdn 2.75      | 4.33        | 4.33             |
|        | SD 0.66       | 0.91        | 0.91             |
| RUS-GER| Me 2.70       | 4.04        | 4.04             |
|        | Mdn 2.75      | 4.17        | 4.17             |
|        | SD 0.63       | 0.93        | 0.93             |
| TUR-GER| Me 2.79       | 4.17        | 4.17             |
|        | Mdn 2.75      | 4.33        | 4.33             |
|        | SD 0.62       | 0.86        | 0.86             |
| Total  | Me 2.69       | 4.13        | 4.13             |
|        | Mdn 2.75      | 4.33        | 4.33             |
|        | SD 0.64       | 0.90        | 0.90             |

GER = German, RUS = Russian, TUR = Turkish, SD = standard deviation

* Self-concept is based on 1,670 students (NA = 48); Motivation is based on 1,670 students (NA = 48); Self-assessment is based on 1,669 students (NA = 49).

**4.2 Dependencies between languages – Correlation analyses**

In order to address research question (i), i.e., whether there are dependencies between the scores of the background languages German and the heritage language Russian or Turkish and the scores obtained in the English C-test, we ran several correlations. Before running the correlations, we visually inspected the data to assure linearity and tested for homoscedasticity and autocorrelation by using the car package in R (Fox and Weisberg, 2019). None of the assumptions
required for correlation analyses were violated. In addition, the sample is large enough, i.e., we can assume normality.

We consider the entire dataset and also each language group independently, including separate correlations for the younger and the older cohorts (see Table 6 in the Appendix). For almost all combinations between the two scores of the German proficiency tests (reading comprehension and reading fluency in German) and the English test results we observe moderately positive correlations. This means that there is indeed a dependency between proficiency in German, here operationalised through reading fluency and reading comprehension, and the performance in English, based on four C-tests. The only exception is the older cohort of the Turkish-German bilinguals. Here, we only find a weak, but still positive correlation between reading fluency in German and the English C-test. The remaining correlations centre around an effect size of .5 (Pearson’s $r$). Furthermore, in all three groups, we notice marginally stronger correlations for the younger cohort compared to their older peers, with the exception of the Russian-German bilinguals for reading fluency. Next, we analysed the effect size between the proficiency in the heritage language (reading comprehension and reading fluency in the heritage language) and English. We monitor mainly weak but significant correlations between reading comprehension in Russian or Turkish and the English C-test scores, and no significant dependency between reading fluency in the heritage language and the English C-test. Again, the older cohort of the Turkish-German students performs differently than the rest. The dependency between reading comprehension in Turkish and English is comparably higher for this group. In addition, there is a significantly negative correlation between reading fluency and the English C-test scores. This comparison reveals that the effect between English and German is equally strong for both monolingual and bilingual participants and that the dependency between the heritage languages and English, if any, is weak.

4.3 Dependencies between languages and extralinguistic variables – Regression analyses

We also looked for dependencies between the language test scores by including the effects of the other predictor variables discussed above on the basis of multiple linear regression analyses. We visually inspected the dataset with scatterplots and histograms, screened for linearity and homoscedasticity, and tested for normality (Shapiro-Wilk normality test), autocorrelation (Durbin-Watson test) and multicollinearity (Variance Inflation Factors or VIF) using the car and dplyr packages in R (Fox and Weisberg, 2019; Wickham et al., 2020). The Durbin-Watson statistic was close to 2, which means that we can
reject autocorrelation, and the Shapiro-Wilk normality test did not return any violations of normally distributed residuals. Furthermore, none of the VIF-scores was greater than 3, thus, we did not detect multicollinearity in our dataset.

For the first regression analysis, we disregarded the heritage language proficiency test (as this would only consider the bilingual but not the monolingual students) and built a model that includes the remaining predictor variables. In order to adhere to the principle of parsimony, which says that only those variables should be part of the model that substantially add to its explanatory power (Levshina, 2015:149), we used the best subset method from the package leaps (Lumley, 2017) in R to select the relevant predictor variables. The final model 1 contains only eight predictor variables, because reading fluency in German, socioeconomic status15 and self-concept did not turn out to contribute significantly to explaining the variance in the English C-test scores.

Table 5. Multiple linear regression analysis

| Parameter                                      | Regression model 1 | Regression model 2 |
|------------------------------------------------|--------------------|--------------------|
| Intercept                                      | −45.22*** (2.65)   | −32.134*** (4.06)  |
| Reading comprehension German                   | 0.35*** (0.03)     | 0.37*** (0.05)     |
| Reading fluency German                         | −                  |                   |
| Reading comprehension heritage language        | −0.10* (0.05)      |                   |
| Reading fluency heritage language              | −0.01** (0.00)     |                   |
| Group: Russian-German                          | −0.22ns (0.68)     |                   |
| Group: Turkish-German                          | −1.60* (0.63)      |                   |
| Cognitive ability test                         | 0.53*** (0.05)     | 0.40*** (0.08)     |
| Gender: male                                   | −1.59** (0.52)     |                   |
| School: other                                  | −11.89*** (0.58)   | −12.76*** (0.95)   |
| School year                                    | 6.47*** (0.26)     | 5.68*** (0.43)     |
| Socio-economic status                          |                    |                   |
| Self-concept                                   | 1.13* (0.55)       |                   |
| Self-assessment                                | 7.09*** (0.36)     | 6.81*** (0.58)     |
| No. of observations*                           | 1,661              | 648                |
| $R^2$                                          | 0.71               | 0.68               |
| $p$-value                                      | < .001             | < .001             |

Estimates with standard error in parenthesis

Significance codes: ns = not significant; ns $p > .05$, * $p > .01$, ** $p > .001$, *** $p > .001$

Model 1 includes 57 missings and Model 2 includes 156 missings.
The contribution of the variables in the model confirms our expectations (Table 5). Attending a vocational-track secondary school has a negative effect on the C-test results, and attending school year 9 instead of year 7 has a positive effect. Also, female students outperform their male peers. In addition, having a higher motivation to learn English and self-assessing one’s own performance in English positively affects the C-test score. With increasing reading comprehension in German, we also observe a significantly higher score for English. This confirms that there is a correlation between reading comprehension in German and proficiency in English ($r = 0.57$, $p < .001$). The results for the group variable are disparate. While there is a significantly negative effect for the Turkish-German bilinguals when compared to the monolingual German students, the same effect cannot be observed for the Russian-German bilinguals.\textsuperscript{16} Moreover, when comparing the estimates of type of school and school year with those belonging to the group of Turkish-German bilinguals, the predicted change in the C-test scores is much larger for the two former variables than for the latter.

In a second regression model, we now exclusively consider the bilingual students ($n = 804$), as we also want to assess the impact of heritage language proficiency on proficiency in English. The data were again visually inspected and screened to test for the same assumptions as mentioned above. The identical results remain, i.e., the residuals are normally distributed, we can reject autocorrelation and we did not detect multicollinearity. We again used the best subset method to select the relevant predictor variables. The final model includes seven predictors: reading comprehension (German), reading comprehension (heritage language), reading fluency (heritage language), cognitive ability, school type, school year and English self-assessment (Table 5). As before, the socioeconomic status is not included in the final model, as it does not significantly contribute to explaining the variance in the English C-test scores, i.e., it does not make the model better. Apart from that, the group variable (Russian-German versus Turkish-German), reading fluency in German, gender, self-concept and motivation do not add to explaining the variance in the English C-test scores either and were excluded from the regression model. The output of the second model is comparable to the first model. The only new information is the influence of the heritage language. Here, both reading comprehension and reading fluency have a significant impact. The former has a positive effect, just as for reading comprehension in German, but the estimate is smaller; while the latter even shows a small negative effect on the C-test results in English.
5 Discussion

Besides interpreting the above findings, we also try to evaluate them in relation to previous studies in the field. This is not easy, as most studies focus on different groups of participants, have different sample sizes and use different measures and tests. In view of these differences, every study in a sense is unique and difficult to compare with other studies. Of course, they all share the interest of identifying the impact of previously acquired languages on subsequent language acquisition.

5.1 Dependencies between the background languages and English

Following Cummins (1979), proficiency in one language, more precisely academic language competence, is said to have a positive influence on the same register in another language. Even though this interdependence was based on the two languages of a bilingual person, the concept has been extended to third or additional language acquisition as well (see Berthele and Udry, 2019; Lasagabaster, 2001). We argue that our findings are compatible with Cummins’ hypothesis, as we report a moderately positive correlation between German and English for all groups. However, we cannot say whether there is a causal relationship between German reading scores and proficiency in English, as other factors may intervene (see also Berthele and Vanhove, 2020:553). Nevertheless, we consider this positive interdependence plausible, as German is the language of instruction – at least in lower grades – in English lessons of German schools, so that students without sufficient competencies in German can be expected to find it harder to follow English lessons. The findings of the current analysis are also in line with several previous studies (Edele et al., 2018; Hopp, Vogelbacher et al., 2019; Maluch et al., 2015) that establish a positive relationship between proficiency in the background and instructional language German and in the foreign language English.

Nevertheless, reading proficiency in the background languages Russian and Turkish did not emerge as equally strong predictors for English proficiency in our sample, even though we found positive effects for reading comprehension in Russian and Turkish on English. Conversely, reading fluency in the heritage language in correlation with English returned no significant relationship (except for a weak negative correlation among the Turkish-German bilinguals in school year 9). We consider this finding important, as it underlies much of Cummins’ (1979) reasoning, and may even be compatible with his idea of thresholds. According to this theory, a speaker needs to have at least a certain level of proficiency in a language before beneficial effects on another language can be expected (Cummins, 1979:229). This could explain why we
see a moderately strong correlation between German and English, but why the correlation is not nearly as strong between the heritage languages and English. As we reported in Section 4.1, the results of the two reading comprehension and fluency tests differ crucially in that the individual scores obtained for the heritage language tests are drastically lower than those for German. This observation illustrates the unbalanced status of the two languages of the bilingual heritage speakers. The low performance in the heritage language version of the reading comprehension and fluency test could mean that the academic language proficiency in Russian and Turkish is too weak to have a positive effect on English. In other words, it may be advantageous or perhaps even necessary to develop the heritage language academically (see Section 5.3). It may also be likely that one would perhaps find stronger positive effects once general language proficiency is taken into consideration, and not just reading fluency and comprehension, as in the present study.

Furthermore, in all three groups, we notice stronger correlations for the younger cohorts compared to their older peers when looking at the dependencies between German and English. This holds true for the correlations between reading comprehension as well as reading fluency in German and the C-test scores in English, except for the latter comparison in the Russian-German group. This seems to tap into how language dependencies develop over time. One could perhaps tentatively argue that at a younger age, the languages of both monolinguals and bilinguals are less separated and cognitively treated as one to a greater extent. Studies investigating younger as well as older participants would be needed to confirm this initial indication.

It remains puzzling why the three groups show different correlation strengths contingent on age group and the tests used. This may indicate that the assumption of simple bilingual or multilingual effects fails to capture the complex reality adequately. The effect of different language backgrounds is also documented in Maluch et al. (2015:82) in which Chinese-German and Polish-German bilinguals outperform Arabic-German and Turkish-German bilinguals in English language achievement. There must be additional, perhaps cultural factors, which we did not control for in the present study, that explain the differences between the Russian-German and the Turkish-German subgroups.

We can summarise that the proficiencies of the students in their respective languages are interrelated, but that the resulting picture is complex. The instructional language German has the strongest effect on proficiency in English. It must remain open whether this interdependence is due to linguistic factors or general cognitive skills (Berthele and Udry, 2019:14; Berthele and Vanhove, 2020:562; Hirosh and Degani, 2018:915).
5.2 Interdependence when controlling for additional variables

Once additional predictor variables are included to explain the variance in the response variable English, we can make some striking observations. Intriguingly, socioeconomic status did not turn out to contribute significantly to explaining the performance in English. This is surprising, as in other studies this variable had a significant influence on language acquisition (see, for example, Cenoz, 2013; Lechner and Siemund, 2014), and because we observed a significant difference between the monolingual students on the one hand, and the bilingual students on the other (overall, the latter had a lower socioeconomic status than the former). However, this social dimension may be partially contained in the variable school-type. Moreover, according to de Angelis (2015), educational success and, hence, performance in the school subject English as a foreign language, may ultimately be related to the educational status of the parents, instead of socioeconomic status. She argues that income and occupation may be less reliable in assessing socioeconomic status, and that the educational level of the parents may be more objective (de Angelis, 2015:441). The context is different to the current study (de Angelis’ study is located in South Tyrol in Italy), yet it may be worthwhile to single out educational status of the parents in future studies.

Much in line with Berthele and Vanhove (2020) as well as Berthele and Udry (2019), variables pertaining to cognition clearly show a strong effect on the C-test results in English. The results of the cognitive ability test have a positive impact on the C-test scores, and the type of school also affects the C-test results crucially. Students who attend university-bound secondary schools, and who are therefore associated with overall better performance in school (likely related to cognitive advantages, higher motivation, perhaps also higher social status), score significantly better in English than those who attend vocational-track schools.

In a similar vein, motivation has a positive effect on the results in English, provided we consider the entire population and only include the test scores of the German reading comprehension task (regression model 1). It does not play a significant role once we only consider the bilinguals (regression model 2). Also, higher ratings of one’s own self-assessment in English have a positive effect on the C-test scores. This nicely shows that the students are more or less realistic regarding their own skills in the foreign language English.

Once all variables are controlled for, we only see marginal differences across the three groups. We observe a negative effect among the Turkish-German bilinguals, when compared to their monolingual peers. This is perhaps surprising, as it challenges the idea of a bilingual advantage. No such relationship can be observed in the group of Russian-German bilinguals (even though there is a small negative effect). However, once we include heritage language proficiency,
we observe no group differences anymore. Of all the variables, this one seems to have only little explanatory power and is superseded by other factors.

Moreover, reading comprehension in German has a positive effect on the C-test scores in English, while reading fluency does not. We interpret this to mean that comprehension is more closely tied to cognition, whereas students can report fluency without having understood the text. The two scores measure different skills, demonstrating that quality is more important than quantity.

5.3 Advantages of previous multilingual experience
As already indicated, in this study we fail to identify an advantage of the bilingual participants over their monolingual German peers. This may be due to several factors. Van Gelderen et al. (2003:22) argue that their equally negative results may be related to social background, which they did not control for. The current study shows that although the scores for the socioeconomic status of the monolinguals are significantly higher than of their bilingual peers, the effect on the C-test results in English is not significant.

Moreover, several recent studies argue that the (positive) influence of background languages on subsequent language development is higher at a lower age. For example, Hopp, Vogelbacher et al. (2019) report that the bilingual students in their sample attest higher proficiencies in vocabulary and grammar tests in comparison to a monolingual control group, though only in the lower age cohorts. Bilingual advantages, apparently, disappear with increasing age. In a similar vein, Lorenz et al. (2019), in a study of pronominal object placement in Russian-German and Turkish-German heritage bilinguals, find stronger traces of crosslinguistic influence and more potential for bilingual advantages in the younger cohorts (nicely tying in with earlier observations regarding subject-verb agreement made in Siemund and Lechner, 2015). In our view, such findings are plausible since schooling can be expected to exert levelling forces of some magnitude, with students converging on a common school standard over time and initially positive effects disappearing (see also Hopp, Jakisch et al., 2020). Since we here examined students at the ages of 12/13 and 14/15, whereas Hopp, Vogelbacher et al. (2019) studied primary school children, the bilinguals may have already moved out of this propitious phase.

Another point worth mentioning is the type of language skill and the language constellation used to predict a relationship between mono- or bilingualism and additional language acquisition. Schwartz et al. (2007:45–46) attest that bilingualism alone did not contribute to predicting higher proficiency in English, but that biliteracy had a positive effect in their study on L3 English literacy with L1 Russian and L2 Hebrew. Rauch et al. (2012) come to the same conclusion in their study on Turkish-German learners of English.
Furthermore, the bilinguals in the present study are dominant in German and weaker in their heritage language, comparable to heritage bilinguals more generally (see Montrul, 2016), which sets them apart from balanced bilinguals. There is recurring evidence that ascribes balanced bilinguals an advantage in foreign language acquisition over monolinguals (see again Section 1). Hence, one could infer that if bilinguals’ proficiencies in the heritage language are strengthened, for instance via heritage language tuition, they should then also demonstrate this so-called multilingual boost. Berthele and Vanhove (2020:563), however, are quite critical of heritage language instruction and resulting positive effects on transfer in third language acquisition. Even though they do not argue against heritage language instruction per se, they see no evidence for the claim that heritage language instruction necessarily results in advantages for additional language acquisition. We concur. It would really be too simplistic to assume that with higher proficiency in the heritage language, heritage bilinguals necessarily enjoy a language learning advantage over their monolingual peers. Too many other factors, in addition to the ones addressed above, play a role here, including intricate identity issues.

Nevertheless, we are still convinced that there is a connection between heritage language proficiency and the acquisition of additional languages. However, it would seem that for heritage bilingualism to be beneficial, there needs to be additional support in the foreign language classroom (see Bonnet and Siemund, 2018, for an overview). Hopp, Vogelbacher et al. (2019:108) remark that the course material that is used for foreign language teaching in German schools is largely designed for monolingual German students, but that the needs of the bilingual students hardly get addressed. Cummins (2013:298) also sees the importance of including multilingual practices by stating that it is ‘reasonable to argue that learning efficiencies can be achieved if teachers explicitly draw students’ attention to similarities and differences between their languages and reinforce effective learning strategies in a coordinated way across languages.’ Students, especially those who have a comparably low proficiency in their heritage language, may in fact not be aware of their potential resources. The activation of prior linguistic knowledge, and hence the use of the entire linguistic repertoire, could significantly improve the language skills of the bilinguals (Cenoz and Gorter, 2017:9). This can be expected to boost metalinguistic awareness, which, as we argued in Section 2, is the main driving force behind multilingual advantages. Furthermore, Melo-Pfeifer (2018:207–08) also places emphasis on a pluralistic approach in the foreign language classroom in order to overcome the monolingual bias and to pay justice to the understanding that languages are not separate entities. 18 This does not mean that the value of the English language classroom is to be de-emphasised, but the learning context could be enriched with additional resources. While all
these are theoretical considerations, Hopp, Jakisch et al. (2020) go one step further. Based on preliminary findings of an intervention study that introduces multilingual elements in the English language classroom, they indeed demonstrate advantages for students (interestingly, both monolinguals as well as multilinguals profited equally strongly). Moreover, even though less class-time was devoted to English proper, these students did not lag behind their peers who received regular English tuition (Hopp, Jakisch et al., 2020).

6 Summary and conclusions

The present study failed to detect a general bilingual or multilingual advantage, which is in line with, amongst others, Berthele and Udry (2019) and van Gelderen et al. (2003). Even though the dataset investigated here attests (some) positive correlations between heritage language proficiency and proficiency in English, as well as clearly more robust correlations between proficiency in the majority language and proficiency in English, we did not identify generally stronger correlations or higher C-test scores for the bilingual cohorts. If at all, we noticed a small negative effect for the Turkish-German cohorts, when compared to their German monolingual peers (see also Maluch et al., 2015).

More conclusive than bona fide linguistic differences between mono- and bilinguals or dependencies between the background languages, however, are a number of interesting extralinguistic variables that turned out to contribute significantly to explaining the performance in the foreign language English. Intriguingly, these are equally important for all participants. Generally, attendance of the university-bound school type (Gymnasium) and the school year (grade 9) had the clearest positive effects. Cognitive skills, self-assessment in English and partly motivation and gender are also substantial predictors for the performance in English. Contrary to other studies (for example, Cenoz, 2013; Lechner and Siemund, 2014), the socioeconomic status of the participants does neither positively nor negatively affect the results in the foreign language (see also de Angelis, 2015, for evidence that parental education is a determining factor). However, our results correspond to those presented in Hopp (2019).

These findings are based on reading comprehension and fluency in German and the heritage languages Russian and Turkish, as well as a general C-test in English, comparing monolingual German students with two bilingual groups, i.e., Russian-German and Turkish-German heritage speakers. We here see a methodological problem, as – contingent on the linguistic skills as well as the bilingual groups tested – one could perhaps arrive at differing results (Trapman et al., 2017). Given the political and educational implications of the kind of research pursued here, this needs to be a warning to all serious scholarly work.
Work on additional and different bilingual groups as well as the inclusion of other linguistic phenomena and language tests is urgently required.

**About the authors**

Eliane Lorenz is a postdoctoral researcher in English linguistics and multilingualism at the Norwegian University of Science and Technology (NTNU), Trondheim. She is currently part of the project ‘Acquisition of English in the Multilingual Classroom’ and the sociolinguistics project ‘Language, Attitudes and Repertoires in the Emirates.’ She completed her PhD in English Linguistics at the University of Hamburg in 2019. Her research interests include L2 and L3 acquisition, multilingualism, contrastive linguistics, varieties of English and corpus linguistics.

Sharareh Rahbari is a PhD student in English Linguistics at the University of Hamburg. She completed her undergraduate studies in ‘Language, Culture & Translation’ at Johannes Gutenberg University Mainz and attained a master’s degree in English and American Studies at the University of Hamburg. Her research interests lie in the area of second and third language acquisition, multilingualism, English as a foreign language, language contact and crosslinguistic influence.

Ulrike Schackow completed a Master of Education in English and Russian at the University of Hamburg in 2018. During her studies, she worked on reading comprehension, English as a third language, as well as the influence of social background variables on foreign language acquisition. In 2019, she started to work as a high school teacher in Germany.

Peter Siemund is Professor and Chair of English Linguistics at the University of Hamburg. His areas of interest include World Englishes, the use of English as an additional language and multilingual development. Further information can be found at www.eng-ling.uni-hamburg.de.

**Acknowledgements**

The study reported here is part of a larger research project on multilingual development (Multilingual development: a longitudinal perspective, University of Hamburg; www.mez.uni-hamburg.de). The financial support of the German Federal Ministry of Education and Research is hereby gratefully acknowledged.
### Appendix

**Table 6.** Correlation matrix with Pearson’s $r$ and significance

|   | 1 C-test all | 2 C-test GER | 3 C-test GER-7 | 4 C-test GER-9 | 5 C-test RUS-GER | 6 C-test RUS-GER 7 | 7 C-test RUS-GER 9 | 8 C-test TUR-GER | 9 C-test TUR-GER 7 | 10 C-test TUR-GER 9 |
|---|-------------|--------------|---------------|--------------|-----------------|-------------------|------------------|----------------|-----------------|-----------------|
| 1 | LGVT GER 1 all | .57***       |               |              |                 |                   |                  |                |                 |                 |
| 2 | LGVT GER 1 GER | .56***       |               |              |                 |                   |                  |                |                 |                 |
| 3 | LGVT GER 1 GER-7 | .55***      |               |              |                 |                   |                  |                |                 |                 |
| 4 | LGVT GER 1 GER-9 |               | .52***        |              |                 |                   |                  |                |                 |                 |
| 5 | LGVT GER 1 RUS-GER |               |               | .61***       |                 |                   |                  |                |                 |                 |
| 6 | LGVT GER 1 RUS-GER 7 |               |               |             | .58***          |                   |                  |                |                 |                 |
| 7 | LGVT GER 1 RUS-GER 9 |               |               |             |                 | .54***            |                  |                |                 |                 |
| 8 | LGVT GER 1 TUR-GER |               |               |             |                 |                   |                  | .54**          |                 |                 |
|   | 1 C-test all | 2 C-test GER | 3 C-test GER-7 | 4 C-test GER-9 | 5 C-test RUS-GER | 6 C-test RUS-GER 7 | 7 C-test RUS-GER 9 | 8 C-test TURGER | 9 C-test TURGER 7 | 10 C-test TURGER 9 |
|---|-------------|-------------|--------------|-------------|----------------|-------------------|-------------------|----------------|-------------------|-------------------|
| 9 | LGVT GER 1  |              |              |             |                |                   |                   |                |                   | 0.56***           |
|   | TURGER-7    |              |              |             |                |                   |                   |                |                   |                   |
| 10| LGVT GER 1  |              |              |             |                |                   |                   |                |                   | 0.41***           |
|   | TURGER-9    |              |              |             |                |                   |                   |                |                   |                   |
| 11| LGVT GER 2  | 0.47***      |              |             |                |                   |                   |                |                   |                   |
|   | all         |              |              |             |                |                   |                   |                |                   |                   |
| 12| LGVT GER 2  |              |              |             |                |                   |                   |                |                   | 0.50***           |
|   | GER         |              |              |             |                |                   |                   |                |                   |                   |
| 13| LGVT GER 2  |              |              |             |                |                   |                   |                |                   | 0.47***           |
|   | GER-7       |              |              |             |                |                   |                   |                |                   |                   |
| 14| LGVT GER 2  |              |              |             |                |                   |                   |                |                   | 0.43***           |
|   | GER-9       |              |              |             |                |                   |                   |                |                   |                   |
| 15| LGVT GER 2  |              |              |             |                |                   |                   |                |                   | 0.54***           |
|   | RUS-GER     |              |              |             |                |                   |                   |                |                   |                   |
| 16| LGVT GER 2  |              |              |             |                |                   |                   |                |                   | 0.48***           |
|   | RUS-GER-7   |              |              |             |                |                   |                   |                |                   |                   |
| 17| LGVT GER 2  |              |              |             |                |                   |                   |                |                   | 0.54***           |
|   | RUS-GER-9   |              |              |             |                |                   |                   |                |                   |                   |
|   | 1 C-test all | 2 C-test GER | 3 C-test GER-7 | 4 C-test GER-9 | 5 C-test RUS-GER | 6 C-test RUS-GER 7 | 7 C-test RUS-GER 9 | 8 C-test TUR-GER | 9 C-test TUR-GER 7 | 10 C-test TUR-GER 9 |
|---|-------------|--------------|---------------|---------------|----------------|------------------|------------------|----------------|------------------|------------------|
| 18 | LGVT GER 2 TUR-GER | | | | | | | 38*** | | |
| 19 | LGVT GER 2 TUR-GER-7 | | | | | | | 43*** | | |
| 20 | LGVT GER 2 TUR-GER-9 | | | | | | | | .15* |
| 21 | LGVT HL 1 all | | | | | | | | |
| 22 | LGVT HL 1 GER | | | | | | | | |
| 23 | LGVT HL 1 GER-7 | | | | | | | | |
| 24 | LGVT HL 1 GER-9 | | | | | | | | |
| 25 | LGVT HL 1 RUS-GER | | | | | | 24*** | | |
| 27 | LGVT HL 1 RUS-GER-9 | | | | | | | .22* | |
|   | 1 C-test all | 2 C-test GER | 3 C-test GER-7 | 4 C-test GER-9 | 5 C-test RUS-GER | 6 C-test RUS-GER 7 | 7 C-test RUS-GER 9 | 8 C-test TUR-GER | 9 C-test TUR-GER 7 | 10 C-test TUR-GER 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 28 | LGVT HL 1 TUR-GER |  |  |  |  |  |  |  |  | .37*** |
| 29 | LGVT HL 1 TUR-GER-7 |  |  |  |  |  |  |  |  | .22** |
| 30 | LGVT HL 1 TUR-GER-9 |  |  |  |  |  |  |  |  | .35*** |
| 31 | LGVT HL 2 all |  |  |  |  |  |  |  |  |  |
| 32 | LGVT HL 2 GER |  |  |  |  |  |  |  |  |  |
| 33 | LGVT HL 2 GER-7 |  |  |  |  |  |  |  |  |  |
| 34 | LGVT HL 2 GER-9 |  |  |  |  |  |  |  |  |  |
| 35 | LGVT HL 2 RUS-GER |  |  |  |  |  |  |  |  | -.04ns |
| 36 | LGVT HL 2 RUS-GER-7 |  |  |  |  |  |  |  |  | -.10ns |
|   | 1 C-test all | 2 C-test GER | 3 C-test GER-7 | 4 C-test GER-9 | 5 C-test RUS-GER | 6 C-test RUS-GER 7 | 7 C-test RUS-GER 9 | 8 C-test TURGER | 9 C-test TURGER 7 | 10 C-test TURGER 9 |
|---|-------------|-------------|---------------|--------------|----------------|------------------|-----------------|----------------|-----------------|-----------------|
| 37| LGVT HL 2 RUS-GER-9 |              |               |              |                 |                   | .09**          |                |                 |                 |
| 38| LGVT HL 2 TURGER   |              |               |              |                 |                   | -.04**         |                |                 |                 |
| 39| LGVT HL 2 TURGER-7 |              |               |              |                 |                   | -.04**         |                |                 |                 |
| 40| LGVT HL 2 TURGER-9 |              |               |              |                 |                   | -.17*          |                |                 |                 |

Abbreviations: LGVT = reading fluency and comprehension test, LGVT 1 = reading comprehension, LGVT 2 = reading fluency, GER = German, RUS = Russian, TUR = Turkish, HL = heritage language, C-Test = C-Test in English

Significance codes: ns = not significant; **p > .05, * .05 >= p > .01, ** .01 >= p > .001, *** .001 >= p > 0

**Notes**

1. ‘Multilingual boost’ refers to the assumed advantage in the acquisition of additional languages by speakers who have access to two or more languages, due to heightened metalinguistic awareness, enhanced cognitive skills, as well as linguistic transfer and the ability to access language skills such as reading or writing from the previously acquired languages (Berthele and Udry, 2019:1; see also de Angelis, 2007).

2. Sanders and Meijers (1995) as well as van Gelderen et al. (2003) fail to identify positive effects of previous multilingual experience for the acquisition of English in different groups of heritage speakers in the Netherlands.

3. For an assessment of multilingual advantages for cognitive (executive) control, see Duñabeitia and Carreiras (2015), Lehtonen et al. (2018), as well as Paap, Johnson and Sawi (2016).

4. We here make reference to ‘general language proficiency’ in the sense of Cenoz (2003:74), who refers to a rather broad or holistic category, comprising oral, written or both language skills. This needs to be distinguished from specific language skills, such as the discrimination of sounds.
(Cenoz, 2003:78). Similarly, Hulstijn (2015:21) defines language proficiency as ‘knowledge of language and the ability to access, retrieve and use that knowledge in listening, speaking, reading or writing.’

5 The two age cohorts in Hopp, Vogelbacher et al. (2019) represent two neighbouring school grades (years 3 and 4) and show that even after only a short time, the initial bilingual advantages in lexical development decrease. Moreover, this insight comes from a longitudinal dataset, in contrast to the cross-sectional designs used in Maluch et al. (2016) and Siemund and Lechner (2015).

6 We should keep in mind that heritage speakers, even though they are strictly speaking native speakers of that language, differ from other native speakers in terms of ‘higher language cognition’ (HLC) (Hulstijn, 2019:163).

7 We are well aware that we can equate neither reading comprehension and fluency nor the performance in a C-test with ‘proficiency’, since both represent specific aspects of the concept of proficiency. Nevertheless, there is sufficient overlap between the two proficiency measures to justify this research question. Whereas the former relies exclusively on reading competences (i.e., reading comprehension and reading fluency), the latter is an aggregate measure that captures reading skills along with production skills (i.e., knowledge about grammar, lexis and spelling) (see also Drackert and Timukova, 2020:123).

8 In Germany, English as a school subject starts either in school year 1 or 3 (primary school), or in year 5 (secondary school), as a mandatory subject, depending on the federal state. During primary education, however, English classes typically focus more on simple vocabulary and only later, during secondary schooling, are language skills such as reading, writing, listening and oral production explicitly taught. Hence, the younger cohort is at least in the third year of English tuition in an instructed setting and the older cohort in the fifth. Therefore, they are classified as beginners or intermediate learners of English.

9 The reliability of this test, as reported by Schneider et al. (2016:27), based on correlations of a retest measure, is .88 for reading comprehension and .80 for fluency, at a significance level of \( p < .001 \).

10 Within the current project, it was only feasible to use a subset of this cognitive test, instead of the entire test or other scores, such as the more comprehensive IQ score. The participants were already presented with an extensive battery of different tests, distributed over two test days (approximately one week apart) for each measuring point. It was impossible to expand this without the possibility of test weariness. Each participant was presented with 25 test items. These differed, depending on the age cohort (i.e., the year 7 students were presented with items 11 to 35, and the year 9 students with items 21 to 45).

11 Cronbach’s \( \alpha \) was used to measure the reliability of this test. The obtained values are .90 for the younger cohort (grade 7) and .85 for the older cohort (grade 9) (MEZ, 2020).

12 A strong agreement to statements iii and iv (self-concept) and statement ii (motivation) indicates a low self-concept or motivation, whereas a strong agreement to the other statements indicates the opposite. This was accounted for in the final score by inverting the corresponding scales before calculating these two variables.

13 For the calculations, we use R (R Core Team, 2019). A note of caution: the dataset is not complete, there are missing values causing fewer data points to be used for some analyses (indicated in the relevant sections). We decided against data imputation as the missing values are most likely not missing at random. In addition, we understand the scores obtained in the English C-test as the independent variable and all other variables, i.e., the test scores for German and the heritage language, the score of the cognitive test, as well as the remaining extralinguistic variables as the explaining variables.

14 Normally, the school type Gymnasium attracts the high performers, while the other secondary-school types take in students at the average or low performance levels.
15 We here add that the estimate of the socioeconomic status (HISEI), when included in the regression model, turned out to be extremely small (.004) at a significance level of \( p = .756 \).

16 Note that the estimate for the Russian-German bilinguals is also negative, though not statistically significant.

17 These authors argue that this is the main predictor, supplanting linguistic interdependence.

18 However, current didactic approaches may work towards separating them. This could also relate back to what we have said earlier (see Section 5.1) about age differences and cognitive separation of languages.

References

Agustín-Llach, M. D. P. (2019) The impact of bilingualism on the acquisition of an additional language: Evidence from lexical knowledge, lexical fluency, and (lexical) cross-linguistic influence. *International Journal of Bilingualism* 23.5: 888–900. https://doi.org/10.1177/1367006917728818

Angelovska, T. (2018) Cross-linguistic awareness of adult L3 learners of English: A focus on metalinguistic reflections and proficiency. *Language Awareness* 27: 136–52. https://doi.org/10.1080/09658416.2018.1431243

Baker, C. (2001) *Foundations of Bilingual Education and Bilingualism*. Clevedon: Multilingual Matters.

Berthele, R. and Udry, I. (2019) Multilingual boost vs. cognitive abilities: Testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism*. Advance online publication. https://doi.org/10.1080/14790718.2019.1632315

Berthele, R. and Vanhove, J. (2020) What would disprove interdependence? Lessons learned from a study on biliteracy in Portuguese heritage language speakers in Switzerland, *International Journal of Bilingual Education and Bilingualism* 23.5: 550–66. https://doi.org/10.1080/14790050.2017.1385590

Bialystok, E. (2001) *Bilingualism in Development: Language, Literacy and Cognition*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511605963

Bonnet, A. and Siemund, P. (eds) (2018) *Foreign Languages in Multilingual Classrooms*. Amsterdam: John Benjamins. https://doi.org/10.1075/hslld.7

Cenoz, J. (2003) The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism* 7.1: 71–87. https://doi.org/10.1177/13670069030070010501

Cenoz, J. (2013) Defining multilingualism. *Annual Review of Applied Linguistics* 33: 3–18. https://doi.org/10.1017/S026719051300007X

Cenoz, J. and Gorter, D. (2017) Minority languages and sustainable translanguaging: Threat or opportunity? *Journal of Multilingual and Multicultural Development* 38.10: 901–12. https://doi.org/10.1080/01434632.2017.1284855

Cenoz, J. and Valencia, J. (1994) Additive trilingualism: Evidence from the Basque Country. *Applied Psycholinguistics* 15.2: 195–207. https://doi.org/10.1017/S0142716400005324

Cummins, J. (1979) Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research* 49.2: 222–51. https://doi.org/10.3102/00346554049002222

Cummins, J. (2001) *Language, Power and Pedagogy: Bilingual Children in the Crossfire*. Clevedon: Multilingual Matters. https://doi.org/10.21832/9781853596773
Hirosh, Z. and Degani, T. (2018) Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin and Review* 23.3: 892–916. https://doi.org/10.3758/s13423-017-1315-7

Hopp, H. (2019) Cross-linguistic influence in the child third language acquisition of grammar: Sentence comprehension and production among Turkish-German and German learners of English. *International Journal of Bilingualism* 23.2: 567–83. https://doi.org/10.1080/1367006917752523

Hopp, H., Jakisch, J., Sturm, S., Becker, C. and Thoma, D. (2020) Integrating multilingualism into the early foreign language classroom: Empirical and teaching perspectives. *International Multilingual Research Journal* 14.2: 146–62. https://doi.org/10.1080/13670050.2017.1385590

Hopp, H., Vogelbacher, M., Kieseier, T. and Thoma, D. (2019) Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction* 61: 99–110. https://doi.org/10.1016/j.learninstruc.2019.02.001

Hulstijn, J. (2015) *Language Proficiency in Native and Non-native Speakers: Theory and Research*. Amsterdam: John Benjamins. https://doi.org/10.1075/litt.41

Hulstijn, J. (2019) An individual-differences framework for comparing nonnative with native speakers: perspectives from BLC theory. *Language Learning* 69: 157–83. https://doi.org/10.1111/lang.12317

International Labor Office (2012) *International Standard Classification of Occupations. ISCO-08. Volume 1: Structure, Group Definitions and Correspondence Tables*. Geneva: ILO.

Jessner, U. (2008) A DST model of multilingualism and the role of metalinguistic awareness. *The Modern Language Journal* 2: 270–83. https://doi.org/10.1111/j.1540-4781.2008.00718.x

Kecskes, I. and Paap, T. (2000) *Foreign Language and Mother Tongue*. Mahwah, NJ: Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410606464

Klieme, E., Eichler, W., Helmke, A., Lehmann, R. H., Nold, G., Rolff, H.-G., … and Willenberg, H. (2006) *Unterricht und Kompetenzerwerb in Deutsch und Englisch: Zentrale Befunde der Studie Deutsch-Englisch-Schülerleistungen-International (DESI)*. Frankfurt am Main: DIPF.

Lambert, W. E. (1974) Culture and language as factors in learning and education. In F. E. Aboud and R. D. Mead (eds) *Cultural Factors in Learning and Education*. Bellingham, WA: Fifth Western Washington Symposium on Learning.

Lasagabaster, D. (1998) Learning English as an L3. *ITL Review of Applied Linguistics* 121–22: 51–84. https://doi.org/10.1075/itl.121-122.05las

Lasagabaster, D. (2001) The effect of knowledge about the L1 on foreign language skills and grammar. *International Journal of Bilingual Education and Bilingualism* 4.5: 310–31. https://doi.org/10.1080/13670050108667735

Lechner, S. and Siemund, P. (2014) The role of language external factors in the acquisition of English as an additional language by bilingual children in Germany. In H. Paulasto, L. Meriläinen, H. Rionheimo and M. Kok (eds) *Language Contacts at the Crossroads of Disciplines* 319–45. Cambridge: Cambridge Scholars Publishing.

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A. and Antfolk, J. (2018) Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin* 144.4: 394–425. https://doi.org/10.1037/bul0000142

Levshina, N. (2015) *How to Do Linguistics with R Data Exploration and Statistical Analysis*. Amsterdam: John Benjamins. https://doi.org/10.1075/z.195
Lorenz, E. (2019) *The use of tense and aspect in the additional language English by monolingual speakers and bilingual heritage speakers*. Doctoral Dissertation, University of Hamburg. urn:nbn:de:gbv:18-106368

Lorenz, E., Bonnie, R. J., Feindt, K., Rahbari, S. and Siemund, P. (2019) Cross-linguistic influence in unbalanced bilingual heritage speakers on subsequent language acquisition: Evidence from pronominal object placement in ditransitive clauses. *International Journal of Bilingualism* 23.6: 1410–30. https://doi.org/10.1177/13670069187991296

Lumley, T. (2017) Leaps: Regression subset selection. R package version 3.0. Retrieved from https://CRAN.R-project.org/package=leaps

Maluch, J. T. and Kempert, S. (2019) Bilingual profiles and third language learning: The effects of the manner of learning, sequence of bilingual acquisition, and language use practices. *International Journal of Bilingual Education and Bilingualism* 22.7: 820–22. https://doi.org/10.1080/13670069.2017.1322036

Maluch, J. T., Kempert, S., Neumann, M. and Stanat, P. (2015) The effect of speaking a minority language at home on foreign language learning. *Learning and Instruction* 36: 76–85. https://doi.org/10.1016/j.learninstruc.2014.12.001

Maluch, J. T., Neumann, M. and Kempert, S. (2016) Bilingualism as a resource for foreign language learning of language minority students? Empirical evidence from a longitudinal study during primary and secondary school in Germany. *Learning and Individual Differences* 51: 111–18. https://doi.org/10.1016/j.lindif.2016.09.001

Melo-Pfeifer, S. (2018) The multilingual turn in foreign language education. Facts and fallacies. In A. Bonnet and P. Siemund (eds) *Foreign Languages in Multilingual Classrooms* 191–212. Amsterdam: John Benjamins. https://doi.org/10.1075/hlsld.7.09mel

MEZ – Mehrsprachigkeitsentwicklung im Zeitverlauf (2014–2019). Projektkoordination: Prof. Dr. Dr. h. c. Ingrid Gogolin, Universität Hamburg; © MEZ 2017.

MEZ – Mehrsprachigkeitsentwicklung im Zeitverlauf (2020). Testhandbuch Querschnitt. Hamburg (Universität Hamburg, MEZ-Projekt), Mimeo.

Montrul, S. (2016) *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139030502

Odlin, T. (2003) Cross-linguistic influence. In C. Doughty and M. H. Long (eds) *The Handbook of Second Language Acquisition* 436–86. Malden, MA: Blackwell. https://doi.org/10.1002/9780470756492.ch15

Paap, K. R., Johnson, H. A. and Sawi, O. (2016) Should the search for bilingual advantages in executive functioning continue? *Cortex* 74: 305–14. https://doi.org/10.1016/j.cortex.2015.09.010

R Core Team (2019) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Rauch, D. P., Naumann, J. and Jude, N. (2012) Metalinguistic awareness mediates effects of full biliteracy on third-language reading proficiency in Turkish-German bilinguals. *International Journal of Bilingualism* 16.4: 402–18. https://doi.org/10.1177/1367006911425819

Safont Jordà, M. P. (2003) Instructional effects on the use of request acts modification devices by EFL learners. In A. Martínez Flor, E. Usó Juan and A. Fernández Guerra (eds) *Pragmatic Competence and Foreign Language Teaching* 211–32. Castelló de la Plana: Publicacions de la Universitat Jaume I.

Sagasta Errasti, M. P. (2003) Acquiring writing skills in a third language: The positive effects of bilingualism. *International Journal of Bilingualism* 7.1: 27–42. https://doi.org/10.1177/13670069030070010301
Sanders, M., and Meijers, G. (1995) English as L3 in elementary school. *ITL Review of Applied Linguistics* 107.8: 59–78. https://doi.org/10.1075/itl.107-108.04san
Sanz, C. (2000) Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied Psycholinguistics* 21.1: 23–44. https://doi.org/10.1017/S0142716400001028
Schneider, W., Schlagmüller, M. and Ennemoser, M. (2016) *Lesegeschwindigkeits- und -verständnistest für die Klassenstufen 5–12*. 2nd edition. Göttingen: Hogrefe.
Schwartz, M., Geva, E., Share, D. L. and Leikin, M. (2007) Learning to read in English as a third language: The cross-linguistic transfer of phonological processing skills. *Written Language and Literacy* 10.1: 25–52. https://doi.org/10.1075/wll.10.1.03sch
Siemund, P. and Lechner, S. (2015) Transfer effects in the acquisition of English as an additional language by bilingual children in Germany. In H. Peukert (ed.) *Transfer Effects in Multilingual Language Development* 147–60. Amsterdam: John Benjamins. https://doi.org/10.1075/hsl.4.07sie
Siemund, P. and Mueller, J. T. (2020) Are multilinguals the better academic ELF users? Evidence from a questionnaire study measuring self-assessed proficiencies. In A. Mauranen and S. Vetchinnikova (eds) *Language Change: The Impact of English as a Lingua Franca* 234–66. Cambridge: Cambridge University Press.
Spellerberg, S. M. (2016) Metalinguistic awareness and academic achievement in a linguistically diverse school setting: a study of lower secondary pupils in Denmark. *International Journal of Multilingualism* 13.1: 19–39. https://doi.org/10.1080/14790718.2015.1053891
Stanat, P., Böhme, K., Schipolowski, S. and Haag, N. (eds) (2016) *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ender der 9. Jahrgangsstufe im zweiten Ländervergleich*. Münster, New York: Waxmann.
Trapman, M., van Gelderen, A., van Schooten, E. and Hulstijn, J. (2017) Reading comprehension level and development in native and language minority adolescent low achievers: Roles of linguistic and metacognitive knowledge and fluency. *Reading and Writing Quarterly* 33: 239–57. https://doi.org/10.1080/10573569.2016.1183541
Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Snellings, P., Simis, A. and Stevenson, M. (2003) Roles of linguistic knowledge, metacognitive knowledge and processing speed in L3, L2 and L1 reading comprehension: A structural equation modelling approach. *International Journal of Bilingualism* 7.1: 7–25. https://doi.org/10.1177/13670069030070010201
Wickham, H., François, R., Henry, L. and Müller, K. (2020) *dplyr: A Grammar of Data Manipulation*. R package version 1.0.1. https://CRAN.R-project.org/package=dplyr