Hans-Georg Müller and Christoph Schroeder*

On the influence of the first language on orthographic competences in German as a second language: a comparative analysis

https://doi.org/10.1515/applirev-2020-0145
Received November 30, 2020; accepted June 5, 2022; published online June 24, 2022

Abstract: This article presents an empirical study on selected aspects of the orthographic performance of mono- and bilingual students in German. Particular attention is paid to the question whether a connection can be drawn between the orthographic performance of the bilingual students in German and their first language. For this purpose, a corpus is analyzed with data from approx. 800 test students. One group is monolingual in German, one is bilingual German–Turkish and one is bilingual German–Russian. A highly significant influence of Russian resp. Turkish on the performance to differentiate between tense and lax vowels is found. This corresponds to expectations insofar as in opposition to German, neither in Turkish nor in Russian, tense and lax vowels are distinguished phonologically. In other areas of orthography, this influence of the first language is not measurable or not to the same extent.

Keywords: German; orthography; Russian; second language writing; Turkish; vowels

1 Introduction

The aim of our contribution is to present empirically reliable data on the acquisition of orthography in German as a second language. We focus bilingual students aged 10–18 with Russian or Turkish as their first language and German as their early second language and use a larger corpus containing orthographic exercise data. Our aim is to examine whether significant differences can be found between the orthographic performance of these three groups and to what extent possible

*Corresponding author: Christoph Schroeder, Department of German Studies, University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany, E-mail: schroedc@uni-potsdam.de.
https://orcid.org/0000-0003-1188-7746
Hans-Georg Müller, Department of German Studies, University of Potsdam, Potsdam, Germany

© 2022 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.
differences can, in the case of the bilingual students, be traced back to phonological contrasts between their first language (Russian or Turkish) and their early second language German.

In the first part of our article, we shortly summarize the state of research in the areas that touch on our topic. In the second part, we present an empirical study in which data from a spelling trainer corpus is analyzed. In the third part, we discuss our results against the background of previous studies.

2 State of research

In languages with alphabetic orthographies, the development of orthographical skills is typically related to the degree of phonological awareness in the language at issue, including the degree to which phonological distinctions of this language are perceived (Mann and Wimmer 2002; Weinrich and Fay 2007). This accounts for German more than for more ‘deeper’ orthographies (Goswami et al. 2005). It has also been shown to be bear consequences for the acquisition of orthography by dialect-speaking children (Bühler et al. 2018). Broadly speaking, there is no doubt that also the acquisition of orthography in the second language is constrained by the degree of phonological awareness in the language at issue, and that first-language phonological knowledge plays a role in the development of phonological awareness in the second (Hulstijn 2011; Mehlem 2015; Munro and Bohn 2007).

Our paper, then, touches on three areas of research. Firstly, we have to concern ourselves with the contrastive phonology of German with Russian and Turkish, the first languages of the learners we are investigating, see Section 2.1. Following that, in Section 2.2, we are concerned with research on the acquisition of orthography in a second language. Thirdly, the way beginning writers spell vowels is very much dependent on how they perceive and process them. Accordingly, the perception and processing of vowels in the second language plays a role as the second research area that we have to consider (unlike, for example, the spelling of words, where grammatical and lexical knowledge is more important), see Section 2.3. Section 2.4 summarizes the state of research.

2.1 German, Turkish–German, Russian–German: relevant features

2.1.1 German

German has a vowel system of 16 vowel monophones and three diphthongs. The monophonemic vowels can be divided into two main classes, namely those which
can occur in stressed syllables and those which cannot. The latter class consists of a highly reduced vowel, in some environments centralized close-mid, in some centralized open-mid, and sometimes even zero (Maas 2006). The first class can again be divided into two classes, namely tense and lax vowels. The distribution between both is very much related to syllable structure. Stressed vowels occur in open syllables and in syllables with a simple coda. Lexical stems usually have either one or two syllables, and the word stress is usually on a root syllable of a lexical word. In monosyllabic words, syllables can have up to three consonants in the onset and up to four consonants in the coda.

In orthography, tense vowels in open syllables are not marked as such. However, in closed syllables with a tense vowel, there is a tendency to signal the tense vowel as such. In that case not the vowel itself is marked, but either the vowel grapheme is doubled, or <h> follows the vowel grapheme, or, in the case of <i>, the vowel grapheme is followed by <e>. Vowel grapheme doubling takes place exclusively with the vowels <a>, <o>, and <e>, and in specific words. The following <h> combines only with sonorants in the syllable coda, e.g. <Lehm> for [le:m] “clay”. On the other hand lax vowels are marked as such only where there can be an opposition to tense vowels, namely in syllables with a simple coda. Here, the marking is done by duplicating the grapheme representing the coda, as long as it is a monograph, or, in the case of <k>, by writing <ck>, e.g. <satt> for [zat] “full”, <Glück> for [glʏk] “luck”.

For reasons of comparison, which play a role in our statistical analysis in Chapter 3, some more features of German orthography need to be mentioned. First, the orthography has a tendency to follow the principle of morphological consistency, that is, to spell a particular morpheme always in the same form, as long as this does not contradict phonographic principles. Second, upper case letters mark the nucleus of noun phrases, and third, compounds, which have a high productivity in the language, are written in closed, i.e. compound form.

2.1.2 Turkish–German

The sound inventories of Turkish and German do not differ much in the consonant area. In the field of vowel phonemes, it is remarkable that diphthongs and central vowel qualities such as [ɪ], [ʊ], [ɐ] are present in German, but not in Turkish. In Turkish, on the other hand, for the eight vowels the oppositions front/back and spread/rounded vowels are exploited to a maximum, so that rather qualities of the cardinal vowels are used. However, the use of the central vowel [a] is also reported (cf. Coşkun 2003), but it is more tense than the German schwa. There are no diphthongs, at most, there are phoneme sequences of one vowel and the semi-vowel [j]. Vowel tenseness in Turkish is not distinctive; it occurs only lexically and
as compensation for the deletion of stem-final [ɣ] (cf. van der Hulst and van de Weijer 1991). The Turkish syllable has a comparably simple structure (V, VC, CV, or CVC). The syllables of an orthographical word show a strong tendency towards successive phonological assimilation of the characteristics ‘front’ and ‘back’, the so-called ‘vowel harmony’. Also, the suffixes are determined by the vowel harmony, with additional labial assimilation. Word accent in Turkish is to be described as ‘phonetically weak’; it lies on the final syllable of the (phonological) word (cf. Kabak and Vogel 2001). Turkish is a syllable-counting language in the sense of Auer (1993).

2.1.3 Russian–German

As in Turkish and unlike German, Russian does not have diphthongs either, neither is tenseness distinctive. Altogether, the Russian vowel system is more reduced than the German (and the Turkish) one; Russian disposes of six vowel phonemes. The two decisive differences concerning vowel quality between Russian and German are that Russian has no front rounded vowels, only back ones, and that Russian has less differentiations in the centralized vowel area than German. In the area of syllable structure there are more similarities between Russian and German than between Turkish and German. Russian is also an accent-counting language with complex syllable onsets and final devoicing. Vowels in unstressed syllables, however, are also strongly reduced. Word accent is more variable in Russian than in German (Bergmann and Mehlhorn 2010; Kümmel 2017; Yanushevskaya and Bunčić 2015).

2.2 Vowels in second language acquisition

With ‘second language acquisition’, we refer specifically to the acquisition of a second language in (pre-school) childhood and in a setting, where the (second) language is also the dominant language of the social environment and the school. This distinguishes early second language acquisition from both adult second language acquisition and foreign language acquisition (Paradis 2007).

For adult second language acquisition, it can be regarded as confirmed that sound contrasts between the first and the second language generally have an effect on the accessibility of differentiations. Kerschhofer-Puhalo (2014) uses an identification test in order to examine the sound perception of German vowels with of 173 adult learners of German as a foreign language with 10 different first languages. On the one hand, there are language-specific patterns in which systematic differences between the vowel systems of the first language and German pose the greatest
challenges. At the same time the differentiation between tense and lax vowels, but also the differentiation of vowels within the respective vowel class remains difficult in all language groups, as long as vowel identification is not supported by lexical information (here: only nonsense words). However, intensity of input seems to make a difference, as Nimz (2015) shows. She investigates the perception of vowel length differences in adolescents with L1 Turkish who were exposed to intensive instruction of German as a foreign language (GfL) and went to school in a German-language high school in Istanbul (Turkey). In the comparison with a monolingual German control group of the same age, perceptual difficulties were found in the GfL-learners only where the opposition between tense versus lax was accompanied by a qualitative difference at the same time, namely in [uː] versus [u] (e.g., <spuken> [ʃpuːkən] “haunt” vs. <spucken> [ʃpʊkən] “spit”).

As for the development of language-analytical skills in multilingual children, Mehlem (2015) makes it clear that sound contrasts between languages also have a fundamental effect on the accessibility of differentiations in child second language acquisition and bilingual first language acquisition. Thus, differentiations that are present in one language but not in the other are more difficult to access than differentiations that are present in both languages. Kerschhofer-Puhalo (2009) also transfers her results of the adult perception study to the acquisition of orthographic skills in German as a second language of children with Turkish and Berber as first languages. In the analysis of the error categories found here the author assumes that the phenomena found for adult foreign language are also more or less valid for the acquisition of second languages with children.

Several studies support the assumptions of Kerschhofer-Puhalo (2009) and Mehlem (2015), but also add further differentiations.

In an experimental design, Højen and Flege (2006) investigate vowel discrimination in English among three groups of adults, namely (i) Spanish monolinguals, (ii) English monolinguals and (iii) Spanish–English bilinguals who acquired English as a second language during childhood in an (American) English-speaking environment. The Spanish monolinguals achieved only random results, and both the English monolinguals and the English–Spanish second speakers achieved high scores. However, significant differences between the bilinguals and the English monolinguals were found in two out of three differentiations, whose degree of difficulty was rated as ‘high’ by the researchers due to the differences between the two languages. Højen and Flege (2006) conclude that the bilinguals had completed a largely successful perceptual learning process, but that the “underlying perceptual systems” (Højen and Flege 2006: 3072) of the monolingual and the multilinguals were not identical.
Another perception study is Darcy and Krüger (2012), again with an experimental design. They investigate vowel discrimination in Turkish–German bilingual children aged 10–11 who had acquired German as early second language (in kindergarten). The authors compare them to a group of German monolingual children of the same age. The bilingual children attended an elementary school with a balanced bilingual program. Significant differences were found where the vowel systems of Turkish and German showed the greatest differences, namely in the discrimination of [iː] ~ [ı] and in the discrimination of the tense front vowels [iː] ~ [eː]. The bilingual children also had greater difficulty than the monolingual children in discriminating [eː] and [ɛ], although the differences were not significant.

Walkenhorst (2020) carries out a pseudoword perception study with three groups of school children in their first six months after school enrolment. These are German monolinguals, German–Turkish bilinguals and German–Russian bilinguals. Some of the bilingual children had acquired German as early second language, and some had acquired both of their languages simultaneously. The study focuses on the differentiation between tense and lax vowels and has a design similar to that of Darcy and Krüger (2012). Concerning the German–Russian bilingual children, the results did not differ markedly from those of the monolinguals. Also, in the German–Turkish bilingual group, only minor and statistically non-significant differences were found across vowels in the overall view. Thus, this group only distinguished the [iː - ɪ]-contrast less frequently than the monolingual group – a similar, only less pronounced, result as in Darcy and Krüger (2012). In a next step, Walkenhorst (2020) subdivided the Turkish-speaking bilingual group based on parent-reported frequencies of using German and Turkish. She distinguishes a Turkish-dominant, a German-dominant, and a balanced group with respect to language use. Interestingly, the lowest vowel distinction rate was found in the balanced group, and the second lowest was in the Turkish-dominant group. The German-dominant group, on the other hand, showed “a higher discrimination rate across vowels and for front and back rounded vowels than the German-monolingual group” (Walkenhorst 2020: 105).

Walkenhorst (2020) is the only study that combines a perception study of vowel discrimination with a study of writing development. We will come to this in the next section.

---

1 It should be added, however, as Walkenhorst (2020) also points out, that with the group formation according to dominance, the groups were quite small, namely 10–12 children. See, however, Kern and Walkenhorst (2017), where exploratively the Russian-speaking group is also subdivided according to language dominance criteria.
2.3 Acquisition of orthography in the second language (with a focus on German)

Again, we need to differentiate between second language acquisition in pre-school child age, and both adult second language acquisition and foreign language acquisition. The latter are usually based on a previous acquisition of literacy and spelling in the first language and must accordingly take into consideration the orthographic knowledge of the students in their first language (Weth and Schroeder, to appear). For the students whose data we analyze, however, we expect the second language to be the first language in which they learn to read and write. It is therefore rather unlikely that they make use of orthographic knowledge from their first language (see, however, the discussion in Section 4).

Overall, the few research results that exist point in different directions. On the one hand, there is empirical evidence that students who are still in the process of building up lexical and grammatical knowledge in their second language have greater difficulties in orthographic domains which demand grammatical and lexical knowledge (Bahr et al. 2015 for English as a second language; Thomé 1987; Fix 2002 for German as a second language). Examples of such areas are the spelling of the orthographical word, employment of upper and lower case and morphologically based spellings (Siebert-Ott et al. 2011). Other studies suggest that children with German as their second language make slightly more spelling mistakes than children with German as their first language, without a specific error category being particularly highlighted (Grießhaber 2004; Jeuk 2009, 2012, all for primary school children). Bulut (2018) on the other hand presents a longitudinal survey of the acquisition of spelling in first and second grades (527 monolingual German children, 110 bilingual and multilingual children; no information given about the ‘other’ languages of the bilingual and monolingual children). The author shows that the bi- and the multilingual children perform significantly better than the monolinguals in all domains investigated, i.e. marking of the differentiation between tense and lax vowels, morphological consistency in final devoicing, spelling of schwa.

With regard to the areas of orthography that demand a phonological analysis the expectation is repeatedly expressed that differences between the first and the second language pose particular challenges for the second language writer, especially in the area of syllable structure and the systematic differentiation between tense and lax vowels (Bredel 2012, 2013; Bredel et al. 2011: Ch. 6; Grießhaber 2004; Röber-Siekmeyer 2003).

Walkenhorst (2020) however, argues against this. She based her perception study of the written language acquisition (see Section 2.2) on a longitudinal design
(first grade and first half of second grade) with 12 of the children. Her results show no or only a weak connection between vowel-specific discriminations and the children’s spellings.

Becker (2011), on the other hand, assumes that there is indeed a connection between the phonological systems of the two languages in bilingual children’s spellings. Her conclusion is based on a qualitative longitudinal study of the acquisition of literacy by nine children with Turkish as their first and German as their second language and 11 monolingual German children. The two groups were studied from the first to the fourth grade and underwent various tests. Particularly in the area of vowel spelling, there are major differences between the children with German as their first language and those with German as their second language: “The L2 children had considerable difficulties with the spelling of long and short vowels. This impression runs almost throughout primary school” (Becker 2011: 225, transl.: CS).2

The quantitative study of monolingual and multilingual fourth-graders by Steinig et al. (2009) supports Becker’s results. The marking of lax vowels appears here as the category with the greatest differences between monolingual and multilingual children.3

It must remain open how the differences between the monolingual and multilingual students are to be interpreted. Şahiner (2012, 2018), in a case study of the acquisition of vowel spelling in a small group of Turkish–German bilingual and German monolingual children in the initial phase of acquiring the written language sees different dynamics at work between the groups, concerning stages of development. Here, too, as with the results of Becker (2011), differences in syllable structure between the two languages seem to play a central role. At the end of the first class, however, there are no more differences between the students concerning their orthographic performance. In Becker’s study (2011), too, the bilingual fourth-graders seem to be taking greater steps forward than the monolingual students in terms of the spelling of tense and lax vowels, but they do not seem to catch up with

2 Becker (2011) does not give a statistical overall view of her results, but draws the above cited result from the case studies of the individual children. In the fourth grade, however, a pseudo-word dictation is carried out, the aim of which is to collect data on the spelling of tense and lax vowels. This was carried out twice, with an interval of six month. In the first round, the monolingual German children performed appropriately on average 3.2 times out of six possible differentiations between the spelling of tense and lax vowels, while the multilingual children only performed once. In the second implementation, the monolingual German children performed the differentiation between tense and lax vowels appropriately on average 3.6 times, while the multilingual children performed it appropriately three times (on average) (cf. Becker 2011: 173–179). Thus, the multilingual people seem to have taken greater steps than the monolingual ones.

the monolinguals (see note 2). Richter (2008), on the other hand, finds the same stages of development in a longitudinal study of German–Italian students in the fourth, fifth, sixth and ninth grades of an Italian–German bilingual school, albeit with a time lag. The representation of tense vowels (esp. the employment of <h> in order to represent tense vowels, as well as double vowel spelling) belongs to the areas in which the bilingual students are 2–3 years behind the others.4

2.4 Summary of the state of research

We summarize the state of research, concentrating on the points that are crucial for us: With respect to the vowel system, clear contrasts between Turkish and Russian on the one hand and German on the other can be identified. Turkish and Russian do not systematically use the tenseness opposition in their phoneme systems, whereas German does. There is evidence that these contrasts have an influence on vowel perception with respect to the acquisition of German as a foreign language by adults (Kirschhofer-Puhalo 2014; Nimz 2015). Regarding the acquisition of German in early childhood, results are not so clear. Research results indicate that possibly dominance relations in the use of the two languages of the bilinguals play a role (Darcy and Krüger 2012; Walkenhorst 2020). Even more unclear are the relationships with regard to the acquisition of orthography. On the one hand, Bulut (2018) shows a ‘bilingual advantage’; on the other hand, Becker (2011), Steinig et al. (2009), and Richter (2008) point to particular difficulties of bilinguals especially in the area of vowel writing, possibly pointing to connections with their ‘other’ languages. Thirdly, Walkenhorst (2020) concludes that while bilingual children may make different vowel differentiations than monolinguals; these are, however, not reflected in spelling strategies.

3 Empirical investigation

In this part, we present a corpus-based examination of differences between the orthographic performance of monolingual German students and those with a Turkish or Russian first language background. In particular, we aim to understand to what extent these differences can be attributed to the phonological contrasts of their respective first languages. Four performance parameters are estimated for each test

---

4 See especially Table 7.14 in Richter (2008: 220) on the spelling of <ch> on syllable onset and with tense vowels, Table 7.15, p. on double vowel spelling *<dofe> versus <doofe> (‘silly’) on p. 227, Table 7.39 “Periods of acquisition of the individual spelling error ranges”, pp. 344–345.
person: One parameter represents the performance in spelling tasks with, and another one without consideration of (the representation of) vowel tenseness/laxness. The other two parameters represent performances in upper and lower case spelling as well as in separate and compound spelling. Those latter variables serve as control values in order to better estimate the causes of possible group differences.

3.1 The data corpus

The data we investigate are part of the Orthografietrainer corpus and contain personal details of 809 test subjects and their performance in various orthographic tasks. The website Orthografietrainer.net offers a varied range of thematically sorted orthographic exercises for school classes and individuals on the orthographic sections “spelling”, “upper and lower case”, “separate and compound spelling” and “comma placement”. In addition, grammatical exercises are offered. Each exercise consists of 30 test sentences and represents a specific orthographic sub-problem (e.g. double consonants, employment of <h>, capital letter with nominalization, etc.). The tasks are supposed to be solved online by processing each exercise sentence (e.g. insertion of letters, spaces, or commas, changes from lowercase to uppercase etc.). The trainees receive feedback on the correctness of their decision immediately after processing each exercise sentence. Wrongly edited sentences are presented again during the program run.

For our analysis, we selected exercises that both represent typical problems of German orthography and permit a broad analysis of student’s orthographic performance. According to the hypotheses of the study, we focussed on spelling tasks that require the consideration of vowel tenseness (such as inserting single or double consonants, differentiating between tense and lax /i/ by selecting <i> resp. <ie>, inserting double vowels or adding <h>, see above), and on the other hand exercises for which the tenseness opposition in vowels is irrelevant (morphologically based spelling, writing of consonants, foreign words, etc.). In addition, exercises from the core areas of upper and lower case (concrete and abstract nouns, nominalizations, number words) as well as separate and compound spellings (verb prefixes, compounds of adjectives verbs and nouns etc.) were used for comparison. For the performance scaling, only the first approach to each exercise sentence was included to the corpus.

For privacy reasons, the declaration of personal data on Orthografietrainer.net is voluntary. People who have been registered for more than three months are asked to provide details of their personal background on the following topics:

– geographical location (state, federal state / canton),
– age, gender, school type, grade level,
- family background (first language, language(s) used in communication with parents, parents’ educational background, number of books in the household),
- school and extracurricular preferences (favorite subjects, hobbies),
- attitude to typical tasks of the German classes (reading, writing, rating of own performance in reading and writing tasks, school grade).

Test subjects have the option of objecting to data collection in its entirety or to all individual questions. The corpus presented in this study contains only data from subjects who had completed a sufficient number of exercises and provided voluntary information about themselves, in particular about their first language. These criteria applied to 809 subjects of which 182 named Turkish and 157 Russian as their first language. The rest of the subjects reported German to be their first language.

3.2 Methods of scaling and analysis

In order to ensure an adequate scaling of the orthographic performance, the correct and incorrect answers to each orthographic task were scaled according to the Item Response Theory (IRT). We processed the data using the Rasch model for dichotomous item responses, assuming that the test results of the subjects are representing one single performance parameter in each test section and that the probability of solving each item correctly increases monotonically with increasing individual competence. The Rasch model provides both, one difficulty parameters for each test item and one performance-parameter for each subject.

In our study, four Rasch parameters were estimated for each subject in the tasks “spelling tasks with [resp. without] the need to consider vowel tenseness opposition”, “upper and lower case” and “separate and compound spelling”. The scaling was executed with the program ConQuest 2.0 with a maximum of 10,000 iterations. The estimations consistently provided an acceptable model quality.5

5 The validity of the Rasch model is assumed in the estimation of the model parameters and can only be checked indirectly within the IRT. Rough violations of the model assumptions can lead to a termination of the estimation process, which did not occur during any of our program runs. In addition, problematic items usually show bad fit values, indicating violations of the model assumptions, which was also unproblematic in the present study: The fit statistics (weighted fit) showed a good to at least acceptable model fit for more than 95% of the items. Further validity criteria allow the comparison of the fit between different models, which was omitted in this study. In similar studies, orthographic performance data (especially from the Orthograffetrainer corpus) had proven to be sufficiently stable to maintain the validity of the Rasch model (Müller 2016: 258–267; Müller et al. 2018: 77 f.).
In order to examine how the first language influences the spelling performance in the second, we used a Generalized Linear Mixed Model (GLMM). In contrast to simple linear models, the GLMM provides the opportunity to consider random variables whose influence on the dependent variable is likely but not in focus of the investigation—typically group variables in cohort surveys. Since the test subjects of the study are students registered by their German teacher, each subject could be assigned to both a specific teacher and class. The model estimation showed that the variable “teacher” had a weak but significant influence, slightly increasing the model quality. The consideration of the variable “class” in contrast did not lead to any further model improvement.

In addition to the variable “first language”, all available personal variables were included into the estimation in order to be able to assess and, if necessary, control their possible influence on the orthographic performance parameters. The following variables proved to be significant and were therefore included into the final statistic model:

- first language
- grade level
- gender
- federal state/canton (automatically grouped to optimize model quality)
- school type (automatically grouped, see note 10)
- the two self-assessment variables “Reading at school is easy for me” resp. “Writing at school is easy for me” (both ordinal scaled)
- the two self-assessment variables “I like reading at school” resp. “I like writing at school” (again, ordinal scaled)

### 3.3 Data evaluation

The majority of the significant predictor values within the models occurred in the expected way and coincide with the findings from similar studies (such as the influence of grade level and gender, cf. Müller 2016: 319). Also, the influence of the
geographical region and the type of school were expectable in the way they occurred (Müller 2016: 331). The details will be discussed in the following section.

In contrast to the variable “first language”, which showed a significant influence on the orthographic performance parameters (see below), the variables “Which language do you speak presently with your mother/with your father?” did not lead to a significant result. This is remarkable, for it indicates that differences in the orthographic performance related to the first language seem to be caused more by an early language imprinting than by the recent linguistic behaviour.

Noteworthy is also the significant influence of the variable “Reading at school is easy for me”, while the variable “Writing at school is easy for me” reached the significance level only in two of the models (see below). A reverse result would have been rather expectable. Only in two models the variables on reading and writing preferences (“reading [resp. writing] at school is a task I like to do”) showed a significant impact on the orthographic performance parameters. Neither the variables of the school grade in German nor the school and extracurricular preferences (favorite subjects, hobbies) were able to further elucidate the personal parameters.

In the following sections, we present Generalized Linear Mixed Models for all four examined orthographic sub-areas and discuss their relation to the first language of the subjects. An overall interpretation of the joint analysis results is completing the section.

3.3.1 Personal performance in spelling tasks with the need to consider vowel tenseness/laxness

Figure 1 shows the significant predictive variable on the performance parameter “spelling tasks with the need to consider vowel tenseness/laxness”. The strength of the lines symbolizes the importance of the variable to elucidate the variance of the students’ performance parameter. According to the model the orthographic performance parameter is significantly influenced by the variables “grade level” and “school type”, which from a didactic point of view corresponds fully to the expectations. Also, the reading variable has a high influence, which represents individual preferences in dealing with written language.

The variable “first language”, has a stronger influence than the geographical variables “state” and “federal state/canton”. In fact, it is even stronger than the
variable “gender” which is usually very strong in orthographic performance tests (Adams and Simmons 2019; Müller 2016: 326).

The coefficients of the linear mixed model and the statistical characteristics are shown in Table 1. The data indicate that subjects with German as their first language have a performance advantage over their classmates with Russian or Turkish as first language. This advantage extends to approximately three grade levels and is more than twice as high as the effect of the gender variable. The comparatively high standard error (std.err = 0.107) indicates a high variation within the groups, which was expectable, since orthographic performances usually spread considerably among students (Müller 2016: 206; Müller et al. 2018: 87). Nevertheless, the $p$-level ($p = 0.003$) confirms the variable to be highly significant and the coefficient ($b = 0.319$) indicates a strong influence on the students’ performances.

All other personal variables show an expected influence on the orthographic performance. The automatic grouping arranged the types of school as “group 0”...
Table 1: GLMM – dependent variable: personal performance in spelling tasks with the need to consider vowel tenseness/laxness.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std.err</th>
<th>90% Confidence interval</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.700</td>
<td>0.266</td>
<td>1.178 - 2.223</td>
<td>6.382</td>
</tr>
<tr>
<td>Language = German</td>
<td>0.319</td>
<td>0.107</td>
<td>0.109 - 0.530</td>
<td>2.975</td>
</tr>
<tr>
<td>Class level</td>
<td>0.116</td>
<td>0.016</td>
<td>0.085 - 0.146</td>
<td>7.442</td>
</tr>
<tr>
<td>Gender = m</td>
<td>-0.140</td>
<td>0.050</td>
<td>-0.239 - 0.042</td>
<td>-2.790</td>
</tr>
<tr>
<td>Federal state (group 1)</td>
<td>-0.196</td>
<td>0.086</td>
<td>-0.366 - 0.042</td>
<td>-2.277</td>
</tr>
<tr>
<td>Federal state (group 3)</td>
<td>-0.333</td>
<td>0.141</td>
<td>-0.609 - 0.057</td>
<td>-2.366</td>
</tr>
<tr>
<td>School type (group 1)</td>
<td>-0.545</td>
<td>0.100</td>
<td>-0.742 - 0.347</td>
<td>-5.419</td>
</tr>
<tr>
<td>School type (group 2)</td>
<td>-0.261</td>
<td>0.075</td>
<td>-0.408 - 0.114</td>
<td>-3.475</td>
</tr>
<tr>
<td>School type (group 3)</td>
<td>-0.206</td>
<td>0.100</td>
<td>-0.402 - 0.011</td>
<td>-2.070</td>
</tr>
<tr>
<td>School type (group 4)</td>
<td>-0.867</td>
<td>0.182</td>
<td>-1.225 - 0.510</td>
<td>-4.762</td>
</tr>
<tr>
<td>reading_easy = 2</td>
<td>-0.362</td>
<td>0.099</td>
<td>-0.556 - 0.168</td>
<td>-3.185</td>
</tr>
<tr>
<td>reading_easy = 3</td>
<td>-0.179</td>
<td>0.056</td>
<td>-0.289 - 0.069</td>
<td>-3.185</td>
</tr>
<tr>
<td>writing_easy = 1</td>
<td>-0.361</td>
<td>0.183</td>
<td>-0.719 - 0.002</td>
<td>-1.973</td>
</tr>
<tr>
<td>writing_easy = 2</td>
<td>-0.234</td>
<td>0.087</td>
<td>-0.404 - 0.064</td>
<td>-2.702</td>
</tr>
</tbody>
</table>

which are typically associated with higher learning achievements (high school, pedagogical and technical college, etc.), while the other school types indicate a gradational reduction of the performance parameter compared to this fixed point.¹⁰ For the reading_easy and writing_easy variables, higher coefficients indicate a greater affinity for reading and writing, so that low performance parameters correspond with low values of that variables and thus, also meet the expectations.

Figure 2 graphically illustrates the group differences of the first language variable. It shows that the performance parameter differs sparsely between Russian and Turkish as first languages, while the difference to the group with German as the first language is high. Scattering among all groups is broad.

¹⁰ A detailed presentation of this analysis is omitted here for reasons of space. To state the most important facts: School type 1 included the secondary schools of the German, Austrian, and Swiss school systems (ISS, Austrian HS, NMS). Type 2 included technical secondary schools Austrian High Schools, and Technical High School. Type 3 included vocational schools, Vocational Schools and High Schools (BMS, BFS), as well as Austrian General High School (AHS). Finally, type 4 comprised elementary schools, special schools and special education schools. As expected, the significance of school type was 0.352, which was clearly higher than that of all other variables, followed by grade level (0.182), attitude toward reading (0.131), and geographic region (0.131). The influence of first language, with an importance of 0.078, showed only slightly more than one-fifth the importance of school type, but occurred consistently across all school types. An interaction term between language and school types showed no significant result.
The reasons for the highly significant performance advantage of the monolingual German students compared to their classmates with Russian and Turkish language background can only be determined by considering further orthographic performance data. At first glance, the differences may support the assumption that a non-German language background is a general disadvantage in learning development, as it has repeatedly been demonstrated by the major international student assessments like PISA or TIMSS (see OECD 2010: 67, 2016: 248; Wendt et al. 2016: 320 f.). The examination of the other performance parameters will show that this view is not fully sufficient and requires differentiation.

### 3.3.2 Personal performance in spelling tasks without the need to consider vowel tenseness/laxness

Table 2 presents the coefficients of the Generalized Linear Mixed Model for performance in the spelling tasks without the need to consider (the orthographic representation of) vowel tenseness/laxness. The characteristics of the variables show clear parallels to the first model, especially in the influence of the grade level, the school type, and the variables for reading and writing preferences.
However, in contrast to the first model, there is no longer any influence of the variable “first language”. Furthermore, both geographic variables as well as the gender variable miss the significance level. The performance to solve spelling tasks in consonant writing thus turns out to be supra-individually more stable and less dependent on personal factors than the orthographic representation of vowel tenseness vs. laxness.

The most important result of this model is that the “first language” variable does not play a significant role anymore, and is included in the table only to illustrate the statistical characteristics. The coefficients for each first language subpopulation are very similar to each other, have a different order than in the first model, and are far from significance. Hence, the assumption that only a more difficult social background for the bilingual speakers causes their performance differences falls short. Rather, we assume that the performance differences of the first model are caused by the specific requirements to consider the vowel tenseness opposition. Since this opposition has no phonological value in Russian and Turkish, students with this language background show measurably larger problems in differentiating between tense and lax vowels, but not in the marking of other spelling regularities. Thus, the highly significant performance differences in the labelling of vowel qualities indicate that phonological characteristics of the first languages seem to have a strong effect on the spelling behaviour of the test subjects.

Table 2: GLMM – dependent variable: personal performance in spelling tasks without the need to consider vowel tenseness/laxness.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std.err</th>
<th>95% Confidence interval</th>
<th>T</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.056</td>
<td>0.234</td>
<td>0.597</td>
<td>1.515</td>
</tr>
<tr>
<td>Class level</td>
<td>0.047</td>
<td>0.014</td>
<td>0.020</td>
<td>0.075</td>
</tr>
<tr>
<td>School type (group 1)</td>
<td>−0.236</td>
<td>0.094</td>
<td>−0.420</td>
<td>−0.051</td>
</tr>
<tr>
<td>School type (group 2)</td>
<td>−0.149</td>
<td>0.066</td>
<td>−0.279</td>
<td>−0.018</td>
</tr>
<tr>
<td>reading_easy = 3</td>
<td>−0.126</td>
<td>0.061</td>
<td>−0.245</td>
<td>−0.006</td>
</tr>
<tr>
<td>like_writing = 3</td>
<td>0.212</td>
<td>0.090</td>
<td>0.034</td>
<td>0.389</td>
</tr>
<tr>
<td>like_writing = 2</td>
<td>0.171</td>
<td>0.075</td>
<td>0.024</td>
<td>0.317</td>
</tr>
<tr>
<td>like_writing = 1</td>
<td>0.158</td>
<td>0.068</td>
<td>0.024</td>
<td>0.293</td>
</tr>
<tr>
<td>Language = German</td>
<td>0.035</td>
<td>0.094</td>
<td>−0.149</td>
<td>0.219</td>
</tr>
<tr>
<td>Language = Russian*</td>
<td>−0.057</td>
<td>0.132</td>
<td>−0.316</td>
<td>0.203</td>
</tr>
</tbody>
</table>

*aThe coefficient of the German–Turkish bilingual speakers was chosen by the program as a reference point with no parameter estimations.
3.3.3 Performance in separate and compound spelling and capitalisation

Tables 3 and 4 show excerpts of the coefficient matrices of the GLMM for the performance variables “separate and compound spelling” and “upper and lower case”. The presentation of the other social variables is omitted here, since these correspond, apart from minor differences, to the coefficients of the last two models.

The influence of the first language proves to be significant in both models, but lags behind the differences in the marking of vowel tenseness/laxness. In fact, the effects are about two thirds of the size of the first model. In addition, the personal parameter “separate and compound spelling” varies broadly within the language groups, which is reflected in a large standard error and an only moderate significance level ($p = 0.015$).

The performance differences in both performances require a renewed interpretation of the overall situation. At first, the data seem to reinforce the hypothesis

### Table 3: GLMM – dependent variable: personal performance in separate and compound spelling.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std.err</th>
<th>95% Confidence interval</th>
<th>$T$</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.232</td>
<td>0.289</td>
<td>0.666</td>
<td>1.798</td>
</tr>
<tr>
<td>Language = German</td>
<td>0.275</td>
<td>0.112</td>
<td>0.054</td>
<td>0.495</td>
</tr>
<tr>
<td>Language = Russian</td>
<td>0.233</td>
<td>0.153</td>
<td>-0.068</td>
<td>0.534</td>
</tr>
<tr>
<td>Gender = m</td>
<td>-0.107</td>
<td>0.051</td>
<td>-0.206</td>
<td>-0.008</td>
</tr>
</tbody>
</table>

[omitted variables: federal state, school type, reading_easy, like_writing]

### Table 4: GLMM – dependent variable: personal performance in upper and lower case.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Std.err</th>
<th>95% Confidence interval</th>
<th>$T$</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.708</td>
<td>0.200</td>
<td>1.316</td>
<td>2.100</td>
</tr>
<tr>
<td>Language = German</td>
<td>0.258</td>
<td>0.076</td>
<td>0.108</td>
<td>0.408</td>
</tr>
<tr>
<td>Language = Russian</td>
<td>-0.023</td>
<td>0.107</td>
<td>-0.233</td>
<td>0.186</td>
</tr>
<tr>
<td>Class level</td>
<td>0.098</td>
<td>0.011</td>
<td>0.076</td>
<td>0.119</td>
</tr>
<tr>
<td>Gender = m</td>
<td>-0.077</td>
<td>0.034</td>
<td>-0.144</td>
<td>-0.009</td>
</tr>
</tbody>
</table>

[omitted variables: federal state, school type, reading_easy, writing_easy]
of a general performance advantage of students with German as their first lan-
guage over those with German as their second language. On the other hand, the
effects of the language variable do not reach the magnitude of the first model, but
remain clearly behind it.

The interpretation is also difficult because we know too little about the written
language socialization of the bilingual study population. Thus, we cannot fully
exclude that the bilingual students also use orthographic knowledge from their
first language in order to solve the spelling task in the second. German, in oppo-
sition to Russian and Turkish uses capital letters in order to mark syntactic units,
i.e. the nucleus of noun phrases. Consequently, the performance of monolingual
German speakers could reflect their monolingual written language experience,
while the bilinguals have to cope with two differently structured writing systems.

This contrastive hypothesis is supported by peculiarities concerning word
spellings, where the students with Turkish as their first language, in contrast to
both other groups, show a lower average personal parameter. Even though the
differences between German–Turkish and the German–Russian bilinguals miss
the significance level ($p = 0.128$), the performance parameter of German–Russian
bilingual students is 0.233 points higher than the average performance of the
German–Turkish subpopulation, while the parameter of monolingual German
speakers shows only a slight additional increase of 0.042 points.\footnote{The fact that
nevertheless the coefficient for the German monolinguals, but not for the German–
Russian bilinguals exceeds the significance level, is due to the higher group size of the monolingual
German subpopulation, whose numerical superiority can compensate for the high variance within
the group and cause a significant result.} Again, it is
tempting to relate this either to the phonetically weak word accent of Turkish,
which, moreover, is not lexically determined (see Section 2.1.2), or to differences in
the spelling of the orthographic word in German and Turkish (see Fuhrhop 2008 for
German and Menz and Schroeder 2015 for Turkish). However, as mentioned above,
this interpretation cannot be more than a first approach to the phenomena and
needs further empirical support. At this point, the findings are only reported in
order to serve as a starting point for possible follow-up studies. In particular, the
writing experience in the first language may have to be taken into account.

3.3.4 Interpretation of the overall situation

Several typical predictors of performance were found to be stable in all models,
especially the variables class level and gender, school type, geographic charac-
teristics, and reading and writing preference. In contrast to these very stable fac-
tors, the influence of the first language variable requires further examination.
Three of four performance models showed a significant performance advantage of the students with German as their first language, which was particularly high in tasks requiring to consider vowel tenseness vs. laxness, but absent in other spelling tasks. This lack of differences in one basic orthographic subtask suggests that explanations assuming just a general, non-linguistic disadvantage of students with non-German first language are not fully sufficient. Rather, the particular linguistic characteristics of Turkish and Russian, both not requiring the consideration of vowel tenseness/laxness, provide an attractive explanation for the very domain-specific performance differences within the spelling tasks. Since this opposition is sound based, it establishes a direct link between experiences in spoken and written language. So it is quite likely to influence the students writing behaviour stronger than other orthographical tasks and to cause the described performance differences. As mentioned above it would be too speculative to interpret the performance differences of tasks with upper and lower case and separate and compound spelling as caused by the grammatical and/or orthographic features of Russian and Turkish since we know too little about the literacy experiences of the students in their respective first language.

However, the performance differences in upper and lower case and separate and compound spelling are not nearly as high as in the marking of vowel tenseness/laxness. Thus the data most likely suggest a superposition of two effects, one language-specific and one language-non-specific. In this perspective, the language-non-specific effect would indicate a more general performance disadvantage of students with German as their second language, as has been observed in other empirical studies (see citing above). The language-specific effect, on the other hand, seems to be caused by the characteristics of the particular first language.

4 Conclusions

The examination of the data presented here partly confirms and partly contradicts assumptions often expressed in the research literature, but so far were rarely empirically investigated. First of all, it can be generally confirmed that students with German as their second language make slightly more spelling mistakes than their peers with German as their first language (cf. Grießhaber 2004; Jeuk 2009, 2012). Striking about our study, however, is that these differences were particularly pronounced in spelling tasks with the need to consider vowel tenseness and laxness, but not apparent in other spelling tasks. The differences in upper and lower case and in separate and compound spelling were significant, but less pronounced.
We interpret these quantity differences as a result of the superposition of different effects: The acquisition of German as an early second language may have a slightly decelerating effect on the development of orthographic competences (Becker 2011; Richter 2008). In particular, however, the tense opposition in German vowels seems to be a challenge, if the first language does not make a systematic distinction here.

In that respect, our results correspond to the findings of Becker (2011: 225) and Steinig et al. (2009: 279). Our data confirm these findings in terms of inferential statistics and exclude a number of alternative hypotheses by taking into account personal variables such as age, gender and school effects. Furthermore, since we measured the group differences not on the basis of error frequencies, but rather on the basis of a probabilistic performance model, the influence of possible random properties of the test items can be largely controlled.

At the same time, our findings contradict Bulut’s (2018) findings which indicate a bilingual advantage; more detailed reconciliations are still pending here.

As mentioned before, Becker (2011) and Şahiner (2018) found that bilinguals catch up with their performance deficits in vowel writing more quickly than monolingual German students. In all our models, however, the class level proved to be a constantly occurring variable. Although group differences in tasks with upper and lower case and separate and compound were much smaller among older students, it cannot be assumed that orthographic problems of students with German as their second language simply disappear on their own.

Finally, our findings contradict Walkenhorst’s (2020) assumptions: Even though we did not conduct perception studies but inferred from spelling performance, we believe we can conclude that a first language without systematic differentiation between tense and lax vowels leads to corresponding discrimination difficulties in German, which in turn may cause corresponding problems in spelling. In our view, this finding is also an important argument for the importance of phonological awareness in the acquisition of orthography.

Questions also remain open in our investigation. We have not been able to explain why significant group differences could also be measured in capitalization and hyphenation, but not in consonant spelling or morphological spelling. Why should there be such a ‘bilingual disadvantage’? We have brought into play a possible written language socialization in the non-German first language; however, we cannot pursue this further here. Also, the performance parameters in all groups show very high scatter. This indicates that in future research, further personal parameters need to be taken in account when predicting the orthographic performance of the test subjects. This includes, for the bilingual students, a more detailed account of their bilingual repertoires and of the age of onset in the acquisition of the second language (Paradis 2007).
Nevertheless, we hope that this paper will provide a clearer empirical basis for discussions about the interrelation between the first and the second language with regard to the development of orthographic competencies in the second language.

References


