

# Bilingual advantage in early EFL pronunciation accuracy of German 4th-graders

International Journal of Bilingualism  
2021, Vol. 25(3) 708–726  
© The Author(s) 2021



Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1367006921996811  
journals.sagepub.com/home/ijb



Teresa Kieseier 

University of Mannheim, Germany

## Abstract

**Aims and Objectives:** We compared speech accuracy and pronunciation patterns between early learners of English as a foreign language (EFL) with different language backgrounds. We asked (1) whether linguistic background predicts pronunciation outcomes, and (2) if error sources and substitution patterns differ between monolinguals and heterogeneous bilinguals.

**Methodology:** Monolingual and bilingual 4th-graders ( $N = 183$ ) at German public primary schools participated in an English picture-naming task. We further collected linguistic, cognitive and social background measures to control for individual differences.

**Data and Analysis:** Productions were transcribed and rated for accuracy and error types by three independent raters. We compared monolingual and bilingual pronunciation accuracy in a linear mixed-effects regression analysis controlling for background factors at the individual and institutional level. We further categorized all error types and compared their relative frequency as well as substitution patterns between different language groups.

**Findings:** After background factors were controlled for, bilinguals (irrespective of specific L1) significantly outperformed their monolingual peers on overall pronunciation accuracy. Irrespective of language background, the most frequent error sources overlapped, affecting English sounds which are considered marked, are absent from the German phoneme inventory, or differ phonetically from a German equivalent.

**Originality:** This study extends previous work on bilingual advantages in other domains of EFL to less researched phonological skills. It focuses on overall productive skills in young FL learners with limited proficiency and provides an overview over the most common error sources and substitution patterns in connection to language background.

**Significance/Implications:** The study highlights that bilingual learners may deploy additional resources in the acquisition of target language phonology that should be addressed in the foreign language classroom.

## Keywords

Bilingual advantage, early foreign language learning, speech production, error sources in early EFL phonology, substitution patterns

---

### Corresponding author:

Teresa Kieseier, University of Mannheim, L15 1-6, 68161 Mannheim, 68161, Germany.  
Email: kieseier@phil.uni-mannheim.de

## Introduction

The consequences of prior bilingualism for foreign language (FL) learning are inconclusive. While some studies found advantages for bilingual learners (e.g., Husfeldt & Bader-Lehmann, 2009), others have found no significant differences for either group (e.g., Wilden & Porsch, 2015) or even suggested a bilingual disadvantage (e.g., May, 2006). Evidence is also mixed with respect to the type of advantage bilinguals may have over monolinguals, ranging from general cognitive to language-specific accounts (see Antoniou et al., 2015 for overview). In addition, previous research suggests that bilingual advantages only manifest once individual and social background factors are taken into account (Hopp et al., 2019). As yet, most prior studies have focused on grammar and on listening or reading skills, while phonology has often been neglected (Core & Scarpelli, 2015). Particularly, studies on young learners with limited proficiency levels are scarce (Kopečková, 2016) and often focus on perception abilities (e.g., Gallardo del Puerto, 2007).

## Background

### *Non-native phonological acquisition*

Many studies have found parallels between first (L1) and second language (L2) phonology, for instance in acquisition order and learner strategies (Kiparsky & Menn, 1987). However, in L2 speech acquisition, learners have already established a phonetic inventory and possess knowledge of phonological processes (Gut, 2010). During the acquisition of an additional language, this previously acquired knowledge may activate cross-linguistic transfer processes, i.e., sound substitutions, cluster simplifications, sound insertion or omissions (Goldstein & Bunta, 2012; Major, 2001). Depending on (perceived) similarity between L1 and L2 sounds, this influence can lead to positive or negative transfer (Core & Scarpelli, 2015, see also Best & Tyler, 2007). Sounds with similar features between L1 and L2 may often cause more interference and be more difficult to acquire than sounds that differ more strongly (e.g., Wode, 1983).

One focus in FLL has been the question whether certain tendencies in speech production are tied to language-specific characteristics or hold universally. Eckmann (1985) stresses the importance of universal tendencies in his Markedness Differential Hypothesis, which suggests that unmarked sounds (frequent occurrence across languages) are acquired before marked ones. Overall, the interaction of both universal and language-transfer processes whose relative influence can change during the acquisition process is widely supported. For instance, Major's (2001) Ontogeny Phylogeny Model (OPM) claims that L2 phonological acquisition includes three processes: transfer of native language phonology, acquisition of target phonology, and universals. He shows that with increasing proficiency, transfer processes eventually become less frequent, while developmental processes gain importance.

If more than one language has been acquired previously, the picture becomes increasingly complex as learners have multiple possible candidates for cross-linguistic transfer. Findings in third (L3) language acquisition range from support for transfer from the L1 or L2, to combined transfer depending on factors such as typological closeness and proficiency level (see Cabrelli Amaro, 2012 for overview).

Previous research showed that phonological acquisition in subsequent language learning cannot be viewed in isolation but is influenced by a multitude of factors. The role of age in non-native language acquisition has been extensively researched and it has been suggested that an early age of onset is particularly crucial in the area of phonology in order to reach native-like competence (for overview see Ioup, 2008). Other factors include individual differences in terms of linguistic history

such as age of acquisition or length of exposure, general proficiency and literacy status in each language as well as cognitive and social variation between speakers (see Gut, 2010; Hirosh-Degani, 2018). The type of exposure to the FL also plays a role (Hirosh-Degani, 2018). In formal language-learning contexts, FL pronunciation is usually not the focus and seldomly explicitly taught or discussed (Eckmann et al., 2003).

Overall, outcomes in non-native pronunciation may be influenced by a variety of linguistic and personal factors and depend on the type of contrast and language combination investigated.

### *Phonological outcomes in monolinguals and bilinguals*

As in other linguistic domains, studies investigating differential effects between monolingual and bilingual learners with respect to phonological outcomes have rendered mixed results. Findings with respect to perceptive and productive abilities range from advantages (Goldstein & Bunta, 2012; Kehoe et al., 2001) or disadvantages for bilingual learners (Fabiano-Smith & Goldstein, 2010a; Gildersleeve-Neumann et al., 2008) to no group differences (Goldstein et al., 2005; MacLeod et al., 2011). On top of that, researchers disagree on what type of advantage bilinguals may have over monolingual peers. Some support a general bilingual advantage (e.g., Bartolotti & Marian, 2012; Tremblay & Sabourin, 2012), though there is no consensus on whether this advantage may be attributed to enhanced cognitive abilities (heightened executive function or conflict resolution skills) or more diverse linguistic experience. However, the notion of a general cognitive bilingual advantage has been challenged in recent meta studies (e.g., Lehtonen et al., 2018). Linguistic effects seem more likely as bilinguals can use previous language learning experience and draw from a larger phonetic repertoire (e.g., Marx & Mehlhorn, 2010). Others propose a specific bilingual advantage as learners use cross-linguistic similarities for positive transfer (e.g., Patihis et al., 2015).

Monolingual and bilingual phonological skills have further been compared for additional language learning. In a study on the perception of non-native Japanese stop contrasts, Enomoto (1994) found that multilingual EFL learners outperformed their monolingual English peers. A more recent study by Patihis et al. (2015) also suggested that bilingualism can be a positive resource in perception abilities as heterogeneous multilingual speakers performed significantly better than monolinguals when discriminating unknown Korean phonemes. Findings showed that bilingual advantages only manifested for languages with contrasts similar to the Korean ones. Antoniou et al. (2015) compared the ability of English monolinguals, Korean-English, and Mandarin-English bilinguals to distinguish phonetic contrasts in two different artificial languages after a short training phase. Both bilingual groups outperformed their monolingual peers. Importantly, learners' success was influenced by the similarity between artificial and native language with respect to the contrasts.

For productive abilities, González-Ardeo (2001) compared pronunciation skills between Spanish monolingual and Spanish-Basque bilingual learners of English. Ratings on mispronunciations and overall intelligibility found no significant group differences. Similarly, Lloyd-Smith et al. (2017) found no differences between German-Turkish bilingual heritage speakers and monolingual controls with respect to global accent in FL English. Although less common, differences between monolingual and bilingual achievements in FL phonology have also been investigated for child learners with limited overall proficiency. Interestingly, some studies with younger FL learners are more in favor of a bilingual advantage, which may provide evidence for the above-mentioned role of age of acquisition in FL pronunciation. Morales Reyes et al. (2017) showed that 4–8-year-old Korean-English bilinguals performed significantly better than English monolinguals in the production of Spanish rhotics. Similarly, Kopečková (2016) compared monolingual 11–12-year-old German FL users to bilinguals growing up with German and an additional L1. Results showed that

bilinguals outperformed monolinguals in their pronunciation of English and Spanish rhotic sounds. Kopečková further stressed the importance of controlling for intra-bilingual differences, since her findings suggest interactions between type of bilingual experience and FL outcomes.

In sum, findings on the effect of bilingualism on FL phonology have been heterogeneous, focused on perception studies, older learners or specific language-contrasts. Therefore, it is crucial to include background factors in order to tease apart findings from different age groups as well as learners with varying linguistic history and experience. Advantages may be more pronounced for younger than older learners, and for more experienced bilinguals (e.g., Cenoz, 2013; Gallardo del Puerto, 2007; Kopečková, 2016).

### *English phonology features and learnability*

This section outlines the main characteristics of English segmental phonology and summarizes English sounds identified as challenging in the FLL context. In both parts, German phonology is used as the main basis of comparison due to the population of FL learners examined in the present study. The English (Received Pronunciation) sound system consists of 12 full vowels (/i:, ɪ, u:, ʊ, ə, ɜ:, ɔ:, ɛ, ʌ, ɒ, æ, ɑ:/), 8 diphthongs (/aɪ, aʊ, ɔɪ, eɪ, oʊ, ɪə, eə, ʊə/), and 24 consonants (/p, b, k, g, t, d, f, v, θ, ð, s, z, ʃ, ʒ, tʃ, dʒ, m, n, ŋ, w, ɹ, j, l, h/). Despite many shared consonants, German lacks the interdental English fricatives /θ/ and /ð/ and the bilingual approximant /w/. Moreover, /r/ is realized as a uvular fricative [ʁ] or trill [R] in Standard German, while English uses the post-alveolar approximant [ɹ] (Knight et al., 2007). Both English and German distinguish between long and short vowels but they often differ slightly in terms of their exact place of articulation. Importantly, the German front rounded sounds /y/, /Y/, /œ/ and /ø/ lack English equivalents, while the English front open-low sound /æ/ does not exist in standard German. In addition, German only has three diphthongs (/aɪ, aʊ, ɔɪ/) (König & Gast, 2012).

Aside from phonemic differences, both languages also have different allophonic variants. For instance, in English, /l/ is realized as the so-called ‘velarized’ or ‘dark’ [ɫ] in syllable-final position (e.g., ‘milk’ [mɪɫk]) (Carr, 1993). Characteristically, plosives are devoiced in word-final position in German but not in English (König & Gast, 2012). English and German also differ in their phonotactics. While both languages allow complex consonant clusters, German only permits postalveolars before nasals (‘[ʃ]nee’) in initial clusters, whereas in English only the combination s + nasal (‘[s]now’) is possible (König & Gast, 2012).

Several studies on the acquisition of FL English pronunciation have identified some of the most challenging sounds, sometimes universally, sometimes depending on language combination. For example, the production of interdental fricatives has been intensively studied for learners of different ages and language backgrounds (Hanulíková & Weber, 2010; Major, 2001). Interestingly, though /θ/ and /ð/ are considered marked phonemes in general and acquired late even by natives (Fabiano-Smith & Goldstein, 2010), substitution patterns seem to differ depending on the learners’ language background. While European French learners often replace the voiceless interdental fricative with /t/, German speakers prefer /s/ (Hanulíková & Weber, 2010). Other error sources for German learners of English include pronunciation of word-final voiced obstruents, and the distinction between /v/ and /w/ (Major, 2001).

The rich English vowel system also causes problems for learners with more restricted inventories. For instance, Turkish learners of English often struggle with the production of diphthongs (Demircioglu, 2013). In contrast, the elaborate and often fine-grained differences between the German vs. English vowel inventories result in inaccurate transfer or assimilation (e.g., Bohn & Flege, 1992; Kautzsch, 2010; Wode, 1983).

Altogether, while some English sounds seem to be universally problematic (e.g., interdental fricatives), others pose language-specific challenges due to small differences between typologically close languages (i.e., English and German).

### *The current study*

Research comparing FLL pronunciation of learners with monolingual and bilingual backgrounds is scarce, and most studies focused on domains other than phonology, older learners, or perception. The few available studies examined specific language pairings or single features. Since early FLL predominantly focuses on oral skills, overall pronunciation ability is crucial in communication development (e.g., Gilakjani, 2012). Further, especially at limited proficiency levels, L1 influences FL pronunciation (e.g., Major, 2001), which can shed light on multilingual cross-linguistic processes.

Against this backdrop, we investigate differences in overall pronunciation accuracy as well as different error and transfer patterns for early EFL learners of diverse language backgrounds in German primary school. We ask the following research questions:

- 1) Do monolingual and bilingual learners of English as a foreign language in primary school exhibit different accuracy values in phonological production?
  - 1a) Does L1 significantly predict accuracy scores in the bilingual group?

We hypothesize primary school bilingual learners of EFL to outperform their monolingual peers in phonological production accuracy due to their richer phonetic repertoire and increased language learning experience (e.g., Marx & Mehlhorn, 2010) – at least once controlled for important individual, linguistic, and cognitive covariates which have been found to modulate a possible bilingual effect in FL learning (see Antoniou et al., 2015; Gut, 2010).

- 2) Do monolinguals and bilinguals show different error and substitution patterns in EFL speech production?
  - 2a) Does L1 influence the occurrence of error patterns within the bilingual group?

We predict that all learners will have problems with universally difficult or marked English sounds (e.g., Major, 2001) and, thus, show similar error patterns irrespective of language background. In addition, we expect German to be the main source of transfer due to actual and perceived typological closeness between German and English (Marx & Mehlhorn, 2010) and language-dominance (Gut, 2010). For certain error types, substitution choices may vary according to background language (Hanulíková & Weber, 2010).

## **Method**

### *Participants*

The sample consisted of 183 4th-grade students from six German public primary schools. Participants had a mean age of 10.26 years and bilinguals averaged 9.34 years of contact with German at the beginning of data collection. Most were dominant in German (see also Table A1 in Appendix). At the point of testing, they had received 4 years of two 45-minute English lessons per week. English input varied strongly on a class level since teachers had very heterogeneous backgrounds with respect to educational paths, proficiency level, experience in FLL teaching, etc.

The sample consisted of 80 monolingual (44%) and 103 bilingual (56%) students, and we defined all speakers who had acquired a language in addition to German productively and/or perceptively before entering school as bilingual. The bilingual group was heterogeneous with 22 different L1s: Afghan, Albanian, Arabic, Bosnian Bulgarian, Chinese, French, Greek, Italian, Croatian, Kurdish, Persian, Polish, Roma, Romanian, Russian, Serbian, Spanish, Tamil, Turkish, Hungarian, Vietnamese. The largest L1 subgroups were Turkish ( $n = 40$ ), Kurdish ( $n = 11$ ), Albanian ( $n = 10$ ), and Italian ( $n = 8$ ). The proportion of bilingual students varied between schools ranging from 22% to 87% ( $M = 56\%$ ;  $SD = 27\%$ ).

## Materials

To test overall pronunciation ability in English, we developed a picture-naming task including 23 single words. By using a more controlled task we wanted to ensure comparable productions across participants and were further taking into consideration practical time constraints in the schools. We selected items on the basis of school materials and included only nouns and adjectives to match the actual proficiency level of the learners. The objective was to find items that were considered easy enough for learners with very limited productive proficiency, but at the same time contained a wide range of segmental features. Although the present paper does not focus on specific contrasts but overall pronunciation accuracy, we made sure that features, particularly the ones reported as problematic for EFL learners, were at least doubled across stimuli. After all items had been selected, matching colored pictures were chosen and put into a Power Point presentation (see Table A2 in Appendix for full item list).

In addition to age and gender, a number of linguistic, cognitive and social variables were collected. Receptive vocabulary knowledge was assessed with the British Picture Vocabulary Scale (BPVS3, Dunn et al., 2009), and receptive grammatical knowledge in English using the Test for Reception of Grammar (TROG-2, Bishop, 2003). We also collected productive vocabulary knowledge via a category fluency task (adapted from Delis et al., 2001) in English, German and the respective L1s of the bilingual students.

For cognitive abilities, phonological awareness was tested with a phoneme manipulation task (following Weber et al., 2007). In addition, we tested executive control with the Simon task (Simon, 1969), general cognitive skills using the CFT 20-R (Weiß, 2006), and working memory with a forward digit span (adapted from HAWIK-IV; Petermann & Petermann, 2008).

Lastly, information on language history and social background was assessed with a detailed parental questionnaire. Social variables included were parent education (years in school), net household income, SES (ISEI, computed according to occupational status following Ganzeboom et al., 1992) and cultural capital (number of German books in household).

## Procedure

Participants were tested individually by trained research assistants in separate rooms at their respective schools during regular school hours. Consent was given by parents beforehand and participation was voluntary. The individual testing sessions lasted about 30–45 minutes and were audio recorded. All tests were paper-pencil based with the exception of the computer-based Simon task and the picture-naming task. The BPVS, TROG and CFT-20R were conducted in group test sessions in the classroom during regular English lessons. The Simon task, BPVS, phoneme manipulation task, Digit Span and the CFT had already been collected at the end of grade 3 due to practical time constraints in the schools. All other tests were administered at the end of grade 4. Parent questionnaires were distributed through the teachers. For item non-responses in the crucial social

variables education, net household income, SES, and cultural capital, the values were imputed via a maximum likelihood-estimation with an EM-algorithm using the other factors as predictors.

## Analysis

Audio recordings were transcribed and coded, and for the standardized tests, we computed scores according to the respective manuals. Production data from non-German L1s were translated and coded by native speakers. Codes were performed and double-checked by trained research assistants.

The picture-naming task was rated for pronunciation accuracy by three independent raters with a background in linguistics. Two raters were proficient L2 English speakers, the third was a German-English bilingual speaker. Inter-rater analyses demonstrated a high level of overlap ( $\alpha = .981$ ). All non-agreement cases were resolved during group discussions. Ratings were done at item level initially. If the target item was produced, it was rated for overall pronunciation accuracy, which resulted in a percentage score of correctly produced items relative to overall productions. Participants produced a total of 2549 target items of which 2237 were pronounced correctly, which resulted in an accuracy score of 87% (Table A1). In a second step, if the item was pronounced incorrectly, the raters documented the error(s) within that item on the feature level. Afterwards, all errors as well as the types of processes occurring during these mispronunciation instances were counted.

For RQ1, we performed *t*-tests comparing overall monolingual and bilingual pronunciation accuracy on the group level. Afterwards, to account for the hierarchical data structure of our sample, we ran a linear mixed-effects regression analysis using the *lme4* package in *R* (Bates et al., 2015). We included all linguistic, cognitive and social background variables at the individual level. Since we further wanted to control for the heterogeneous compositions as well as sizes of schools and classes, we included proportion of bilingual students as well as mean SES at the school as factors at the institutional level. As our participants were nested in L1 groups, classes and schools, we included them as random effects school\*class\*L1 (Baltes-Götz, 2013). First, we conducted the Null model, which only included these random effects. Model 1 then initially included all individual, linguistic and cognitive background factors as well as institutional factors. Following Cheng et al. (2010) we then used a stepwise-backwards method, excluding the weakest predictor until we arrived at the most parsimonious model (see also Hox, 2010). In Model 2, we included social background factors available for a sub-group of participants and fitted it to Model 1 without any further model optimization.

For RQ2, we compared the types of errors and their frequency relative to overall errors between monolinguals and bilinguals performing *t*-tests. In addition, for the most frequent error sources, we traced the respective processes (substitutions) and compared those between the two groups.

These analyses for RQ1 and RQ2 were repeated for the group of bilingual learners only, comparing overall performance and error patterns of different L1 groups.

## Results

### *Bilingual advantage in overall production accuracy?*

An initial independent samples *t*-test revealed no significant differences between monolingual and bilingual learners in terms of overall pronunciation accuracy ( $t(181) = -0.39, p = .697$ ) and both groups scored rather highly in terms of correct productions. Monolinguals performed significantly better on measures for receptive ( $t(182) = 3.51, p = .001$ ) and productive vocabulary ( $t(182) =$

3.08,  $p = .002$ ) in English, receptive grammar in English ( $t(182) = 3.90, p = .000$ ) and productive vocabulary in German ( $t(182) = 2.59, p = .010$ ). Moreover, monolinguals outperformed their bilingual peers with respect to phonological awareness ( $t(181) = 3.42, p = .001$ ) and general cognitive abilities ( $t(178) = 3.96, p = .000$ ). In addition, the monolingual group reached significantly higher scores on all social variables (for overview of descriptive statistics see Table A1 in Appendix).

In a next step, to account for the nested data structure and to control for linguistic, personal and social differences between learner groups, we performed a linear mixed-effects regression analysis. The resulting models are summarized in Table 1.

In the best-fitting model with all linguistic and cognitive background variables (Model 1), bilingualism was a significant predictor of pronunciation accuracy. In addition, receptive vocabulary in English, working memory and phonological awareness reached significance. In Model 2 including social variables, bilingualism stayed a robust predictor of production accuracy, while receptive vocabulary in English and working memory also significantly contributed.

In addition, we performed a second regression model following the same procedure for the bilingual learners only to test whether L1 group significantly impacted pronunciation accuracy (see Table A3 in Appendix). We also added productive L1 vocabulary as predictor to this model. Neither L1 variable significantly predicted pronunciation accuracy in the models with or without social variables. The only significant predictor was receptive vocabulary in English.

### *Error and substitutions patterns*

For RQ2, we focused on all incorrect productions, and how these were distributed overall and for monolinguals and bilinguals separately. Table 2 summarizes the most common error sources across the full sample.

All errors with occurrences above five are listed and all remaining errors were summarized as “other.” Errors affected vowels, consonants as well as clusters. In a next step, we focused on the five most common error sources: /ɪ/, /θ/, /aʊ/, /s/-clusters and /w/. We decided to include all other error occurrences involving sibilants (/s/ and /ʃ/) and merged those with the /s/-clusters to form the category “sibilants.” We then compared these five error categories and their relative occurrence between monolinguals and bilinguals (see Table 3). Four of the most common errors overlapped for monolinguals and bilinguals: /ɪ/, /θ/, /aʊ/ and sibilants, though the order of frequency was different. The most frequent error source for monolingual German learners was /ɪ/ (25%), followed by sibilants (16%), /θ/ (14%), /w/ (9%) and /aʊ/ (8%). For bilinguals, the predominant error source was /θ/ (22%), followed by /ɪ/ (18%), /aʊ/ (14%) and sibilants (11%), while /w/ was not among the top-five error sources.

The overall five most common errors accounted for 72% of total errors for monolinguals and 69% for bilinguals. Independent sample *t*-tests revealed that monolinguals and bilinguals did not differ significantly in the overall percentage of errors made. In addition, out of the five most common error types, only /w/ differed significantly between groups but was generally rare. The bilingual group was very heterogeneous, yet an ANOVA comparing occurrence of the top-five errors between bilingual L1 groups found no significant differences.

For each of the top-five error categories (/ɪ/, /θ/, /w/, /aʊ/, and /s/-substitution single or in clusters), we also compared frequencies of substitution patterns relative to overall errors in each group. The English approximant /ɪ/ was either substituted with /w/ ([ɪŋ] > [wɪŋ] (‘ring’)) or with the German fricative /ɣ/ ([ɪŋ] > [ɣɪŋ] (‘ring’)). In both speaker groups, the first option was most frequent. Bilinguals substituted with /w/ 81%, and with /ɣ/ 19% of the time. Monolinguals opted for the /w/ substitution in 84% and for the German transfer in 16% of the cases.



**Table 1.** Comparison between monolingual and bilingual learners: predictors of pronunciation accuracy (picture naming) in grade 4 at institutional and individual levels (hierarchical mixed linear regression); controlled for effects of school, class affiliation and L1.

Parameters	Null model	Model 1	Model 2
<b>Fixed effects</b>			
Intercept	0.87*** (0.01)	0.87*** (0.01)	0.87*** (0.01)
<b>Institutional level</b>			
Proportion of bilinguals/school	–	n.s.	n.s.
Mean SES/school	–	n.s.	n.s.
<b>Individual level</b>			
<b>Linguistic factors</b>			
Productive vocabulary German	–	n.s.	n.s.
Bilingualism (0=monolingual, 1=bilingual)	–	0.03* (0.01)	0.03* (0.01)
Productive vocabulary English	–	n.s.	n.s.
Receptive vocabulary English	–	0.04** (0.01)	0.04** (0.01)
Receptive grammar English	–	n.s.	n.s.
<b>Personal &amp; cognitive factors</b>			
Gender (0=male, 1=female)	–	0.02 <sup>ns</sup> (0.01)	0.02 <sup>ns</sup> (0.01)
Age (months)	–	n.s.	n.s.
Nonverbal IQ	–	n.s.	n.s.
Working memory	–	0.03** (0.01)	0.03*** (0.01)
Executive control	–	n.s.	n.s.
Phonological awareness	–	0.02* (0.01)	0.02 <sup>ns</sup> (0.01)
<b>Social factors</b>			
Parent's nonresponse	–	n.s.	–
SES parents	–	–	0.01 <sup>ns</sup> (0.02)
Education parents	–	–	–0.02 <sup>ns</sup> (0.01)
Net-income family	–	–	–0.0009 <sup>ns</sup> (0.02)
Cultural capital	–	–	0.01 <sup>ns</sup> (0.02)
<b>Random parameters</b>			
<b>Variance</b>			
Within L1cs group	0.02	0.01	0.01
Between L1cs group	0.001	0.0001	0.001
N	174	174	138 <sup>a</sup>
–2 Restricted Log Likelihood	–212.1	–249.0	–195.4
Number of parameters	3	8	12

Note. Dependent variable: Pronunciation accuracy (grade 4).

Notation: unstandardized estimates ( $\beta$ ), standard error (SE).

Predictor variables z-standardized.

ns = not significant, \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

Null Model without fixed predictors  $\beta$  (SE).

Model 1: Null model + institutional, linguistic, personal and cognitive predictors  $\beta$  (SE).

Model 2: Model 1 with social predictors  $\beta$  (SE).

Random effects: L1 groups within schools and classes (L1cs groups).

Models optimized via X<sup>2</sup>-comparison. Predictors not improving model fit removed (n.s.).

<sup>a</sup>Only cases with full information from parents; missing values caused by item non-response ML-estimated (EM-algorithm).

For the interdental fricative /θ/, a number of different substitution options were found: it was either substituted with plosives (/t/, /d/, e.g., [mauθ] > [maut], ‘mouth’), other fricatives (/s/, /f/, e.g., [mauθ] > [mauf]), or omitted completely ([mauθ] > [mau]). Figure 1 shows that for both

**Table 2.** Descriptives of mispronunciations: Overall occurrences of error sources (number of participants), examples error types.

Type	N	Example
Errors overall	325 (135)	–
/ɹ/	69 (50)	[ɹɪŋ] > [wɪŋ] ('ring')
/θ/	60 (57)	[maʊθ] > [maʊt] ('mouth')
/aʊ/	37 (31)	[maʊs] > [moʊs] ('mouse')
(initial) /s/-clusters	22 (20)	[ˈstrʌ:bəɹi:] > [ˈstrɑ:(ɹ)bəɹi:] ('strawberry')
/w/	20 (19)	[ˈwɪndəʊ] > [ˈvɪndəʊ] ('window')
/s/	16 (16)	[maʊs] > [maʊf] ('mouse')
/g/ (word-final)	14 (14)	[dɒg] > [dɒk] ('dog')
/ɔ:/	14 (13)	[bɔ:t] > [bɛt] ('ball')
/e/	11 (11)	[ˈɛlɪfənt] > [ˈi:lɪfənt] ('elephant')
/ə/	10 (9)	[ˈɛlɪfənt] > [ˈɛlɪfɛnt] ('elephant')
/ɪ/	9 (9)	[fɪ] > [fi:] ('fish')
dark ɪ	8 (8)	[æpt] > [æpl] ('apple')
/p/	8 (8)	[æpt] > [æbt] ('apple')
/ʃ/	5 (4)	[ʃi:p] > [si:p] ('sheep')
/ɑ:/	5 (5)	[stɑ:(ɹ)] > [steə(ɹ)] ('star')
/æ/	5 (5)	[kæt] > [kɛt] ('cat')
other	12 (10)	e.g., [waɪt] > [wi:t] ('white')

**Table 3.** Most common error source compared between monolingual and bilingual learners (occurrences and % out of overall errors, *t*-tests, effect size).

Error Category	monolingual	bilingual	<i>t</i> -test	<i>d</i> <sub>Cohen</sub>
/ɹ/	37 (0.25)	32 (0.18)	<i>t</i> (323) = 1.06 <sup>ns</sup> , <i>p</i> = .289	0.13
/θ/	21 (0.14)	39 (0.22)	<i>t</i> (318.8) = -1.85 <sup>ns</sup> , <i>p</i> = .065	0.20
/aʊ/	12 (0.08)	25 (0.14)	<i>t</i> (310.5) = -1.56 <sup>ns</sup> , <i>p</i> = .119	0.16
sibilants	24 (0.16)	19 (0.11)	<i>t</i> (263.5) = 1.13 <sup>ns</sup> , <i>p</i> = .258	0.13
/w/	14 (0.09)	6 (0.03)	<i>t</i> (231.2) = 2.03*, <i>p</i> = .044	n.a.

groups the predominant substitution option was /f/, yet this tendency was stronger for monolinguals (57%) as bilingual substitutions were more diverse.

The diphthong /aʊ/ was either changed to /oʊ/ or substituted with either of the long monophthongs /o:/ or /u:/. In both groups, /oʊ/ was the most frequent substitution (bilinguals 76%, monolinguals 58%). For bilinguals the second most frequent choice was /o:/ (12%), while monolinguals preferred /u:/ (25%).

The approximant /w/ was either substituted with /ɹ/ ([ˈwɪndəʊ] > [ˈɹɪndəʊ], 'window') or the voiced fricative /v/ ([ˈwɪndəʊ] > [ˈvɪndəʊ]). For bilinguals, /v/ was the more frequent substitution in 67% of the cases, while monolinguals preferred /w/ for /ɹ/ (93%).

The fricative /s/ was substituted with another fricative (/ʃ/, /f/, or /θ/, e.g., [ˈspɑɪdə] > [ˈʃpɑɪdə], 'spider'), with the plosive /t/ ([maʊs] > [maʊt], 'mouse') or omitted (e.g., [ˈspɑɪdə] > [ˈpaɪdə]). Sometimes, the entire cluster was substituted with /d/ (e.g., [steə(ɹ)z] > [deə(ɹ)z], 'stairs'). The distributions between these options are compared for bilinguals and monolinguals in Figure 2. Bilinguals preferred substitution with /s/ (50%), while for monolinguals substitution choices were more evenly distributed, the most common substitution being /f/ (29%).

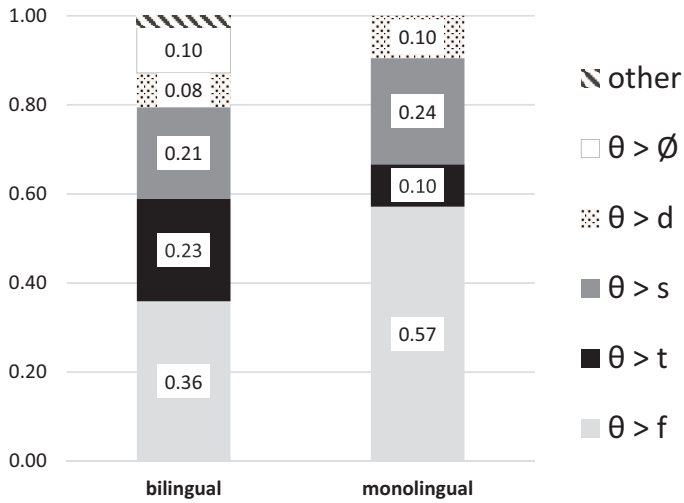


Figure 1. Interdentalsubstitutions bilingual versus monolingual language background.

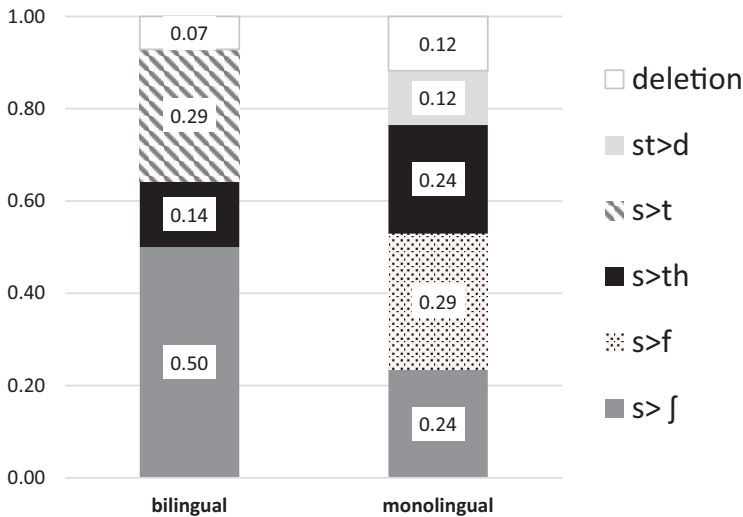


Figure 2. /s/-substitutions bilingual versus monolingual language background.

A Pearson’s Chi-Square test revealed that the substitution patterns between monolingual and bilinguals reached significance for the substitution patterns of the categories /w/,  $X^2(1, N = 20) = 87.9, p = .005$ , and /s/,  $X^2(5, N = 31) = 12.6, p = .027$ .

### Discussion

The objective of the present contribution was to compare overall phonological production in 183 primary school learners of English with monolingual or bilingual language background. We asked (1) whether language background significantly predicts pronunciation accuracy of single-word

items during a picture-naming task in 4th-grade EFL learners, and (2) whether error patterns and processes are influenced by language background.

With regard to RQ1, the bilinguals' EFL pronunciation was significantly more target-like, compared to their monolingual peers – once controlled for social, linguistic and cognitive background factors. The present findings thereby support previous research reporting bilingual advantages in other areas of FLL (e.g., Husfeldt & Bader-Lehmann, 2009) and in terms of phonological production skills for child learners (e.g., Kopečková, 2016). Interestingly, the advantage seems to be more pronounced in the phonological domain. At group level, monolinguals performed significantly better on all linguistic measures, except English pronunciation accuracy, where the two groups performed equally well (Table A1, see also Hopp et al., 2019). A possible explanation for this discrepancy could be that bilinguals were particularly receptive to pronunciation cues. The majority of bilinguals in our sample were not literalized in their non-German L1s and had not received any formal training in their L1 either. Hence, they are used to differentiating between German and their additional L1 based on pronunciation cues. Since FL teaching in English was also mostly oral, bilinguals may have an advantage which is heightened in the phonological area due to their experience and daily practice in paying attention to language contrasts in the speech input.

More specifically, we asked whether English pronunciation benefits from a general bilingual skill or other than German L1 transfer (RQ1a). Within-bilingual comparisons showed, however, that L1 did not predict pronunciation accuracy. This is in favor of a general bilingual advantage in FLL pronunciation (e.g., Tremblay & Sabourin, 2012) and contradicts some previous research finding language-specific patterns, which, unlike our contribution, often focused on single phoneme contrasts (Patihis et al., 2015). It is possible that, while typology and language combination affect outcomes more strongly for specific contrasts, additional linguistic knowledge manifests as a resource for bilinguals for overall pronunciation accuracy when a broader variety of sounds is considered. Moreover, in light of the present results, we may have to adjust our understanding of the nature of a general bilingual advantage (see, e.g., Antoniou et al., 2015). Since cognitive factors (executive function) and phonological awareness (phoneme manipulation) were controlled for, a general bilingual advantage does not seem to be based on cognitive factors alone. Rather, their general multilingual experience may have equipped learners with enhanced knowledge and strategies functioning as a resource in FL production. A longitudinal study on learners' development of FL phonology (see, e.g., Kopečková, 2016) could shed further light on the type of advantage and if bilinguals already had an edge over monolinguals at the start of FL exposure or whether they were equipped with better language learning strategies which expedited their acquisition process.

Turning to RQ2, we took a closer look at the occurrences of specific error types. First, our results show that the most common error sources concern /ɪ/, /θ/, /w/, /aʊ/ and /s/-clusters, which is consistent with the set acquired late during L1 English acquisition (Fabiano-Smith & Goldstein, 2010b) and typical challenges for FLL of English found previously (Major, 2001). As predicted from the German-English phoneme contrasts, the most common error sources were sounds not present (/θ/, /w/) or with different phonetic realization (/ɪ/) in German. Hence, as predicted based on Major's OPM (2001), errors can be explained by a combination of universal difficulty (markedness) and transfer processes.

Comparing monolingual and bilingual error patterns revealed that error sources for both groups largely overlapped. This suggests that German phonology functioned as the base for cross-linguistic interference for all learners, which corresponds to findings by Lloyd-Smith et al. (2017). This can be attributed to effects of language dominance (Table A1) or (perceived) typological closeness between the two languages (Marx & Mehlhorn, 2010). The bilabial approximant /w/ was the only error which was significantly more frequent for monolinguals than bilinguals, although it occurred rarely overall. This tendency could be explained by the fact that, unlike German, some other

background languages, e.g., Spanish (Fabiano-Smith & Goldstein, 2010) or Kurdish (Rahimpour & Saedi Doveise, 2011), have /w/ in their inventory. Moreover, monolinguals showed more mispronunciations of (/ɪ/), which again could mean that monolinguals, at least tentatively, have more problems with sounds that have close but non-identical equivalents in German compared to their bilingual peers. However, the lack of a significant effect of heritage language with respect to error distribution within the bilingual group (RQ2a) weakens the role of bilinguals' L1s as source for positive transfer.

We also compared substitution patterns for the most frequent error types between monolingual and bilingual learners and found significant differences for /s/ and /w/. For /s/, substitution patterns are diverse for monolinguals, while bilinguals preferably palatize to /ʃ/, which is surprising since particularly in clusters (/ʃ/pinne versus /s/pider), we would expect a strong interference for monolinguals. For /w/, monolinguals predominantly opted for a substitution with /ɪ/, although this sound is problematic in other circumstances, while /v/ is present in the German inventory. Similar tendencies can be found for the substitution of /au/, a diphthong present in German, with /ou/, which is absent from the German inventory. These findings could be explained as an attempt by the learners to sound as English as possible and therefore trying to substitute with a less "German-sounding" variant. Additional remarks collected during the task further strengthen the claim that parallels between languages are not yet obvious to the learners. Although the difference in substitution types for /θ/ based on language group did not reach significance, bilingual substitution patterns were more diverse. Interestingly, /f/ substitution was the most frequent variant in both groups, but particularly dominant in monolingual Germans, while previous studies often report /s/ as the preferred option (e.g., Hanulíková & Weber, 2010). Altogether, it seems that while errors are similarly distributed, learners will make different substitution choices for some sounds.

As the present study provides an overall picture of pronunciation skills and error sources in early learners of EFL with various language backgrounds, it has several limitations. The very low oral proficiency of the learners made the item selection process difficult, resulting in a high number of non-productions and relatively small group sizes. To make definite predictions about specific error and substitution patterns, future research should study the most common error types more closely and specifically include a larger number of items to increase overall occurrences. In addition, the cognate status of many "easy" English words (e.g., 'mouse' versus 'Maus') may have biased productions. Future research should include a more diverse output and include connected speech samples instead of limiting itself to single word items in isolation. While our sample authentically reflected the linguistic diversity in German schools, future studies should focus on larger numbers of specific L1 groups to trace transfer patterns more thoroughly and possibly also include different types of bilinguals (age/order of acquisition, etc.). Lastly, we could only approximate controlling the students' spoken English input by including class affiliation in our analyses, since some teachers refused to provide speech samples.

To conclude, the present study shows that a bilingual advantage in FL pronunciation is already visible in young learners at the primary school level but only when background factors are considered. In addition, though the same error sources are predominant across language backgrounds, substitution patterns may vary between monolingual and bilingual learners. These results indicate that bilingualism may lead to an interaction of both general and specific advantages depending on type of contrast investigated. This contribution underlines the importance of taking into account the complex interplay of background factors in order to accurately reflect individual differences between learners and identify the type of resources bilingual learners are equipped with.

The present findings further suggest that individual learner resources need to be considered more strongly in the FLL classroom, especially in terms of pronunciation, which is a crucial component in developing communication skills (Gilakjani, 2012). Using phonology could be a starting

point, where learners can benefit from cross-linguistic comparisons and bilinguals' additional linguistic knowledge, which can eventually be transferred to other linguistic domains.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the German Federal Ministry for Education and Research (BMBF) as part of the research initiative "Sprachliche Bildung und Mehrsprachigkeit" (FKZ 01JM1401; 2014–2017).

## ORCID iD

Teresa Kieseier  <https://orcid.org/0000-0002-8112-0682>

## References

- Antoniou, M., Liang, E., Ettliger, M., & Wong, P. C. M. (2015). The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition*, 18(4), 683–695.
- Baltes-Götz, B. (2013). *Analyse von hierarchischen linearen Modellen mit der SPSS Prozedur MIXED*. Zentrum für Informations-, Medien- und Kommunikationstechnologie (ZIMK): Universität Trier.
- Bartolotti, J., & Marian, V. (2012). Language learning and control in monolinguals and bilinguals. *Cognitive Science*, 36(6), 1129–1147.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. In O.-S. Bohn & M. J. Munro (Eds.), *Language learning & language teaching. Language experience in second language speech learning (Vol. 17)*, pp. 13–34. John Benjamins Publishing Company.
- Bishop, D. V. M. (2003). *Test for Reception of Grammar: TROG-2: Manual (version 2)*. Pearson.
- Bohn, O. -S, & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131–158.
- Cabrelli Amaro, J. (2012). L3 phonology: An understudied domain. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third language acquisition in adulthood* (pp. 33–60). John Benjamins.
- Carr, P. (1993). *Phonology. Modern linguistics series*. Macmillan.
- Cenoz, J. (2013). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching*, 46(1), 71–86.
- Cheng, J., Edwards, L. J., Maldonado-Molina, M. M., Komro, K. A., & Muller, K. E. (2010). Real longitudinal data analysis for real people: Building a good enough mixed model. *Statistics in Medicine*, 29(4), 504–520.
- Core, C., & Scarpelli, C. (2015). Phonological development in young bilinguals: Clinical implications. *Seminars in Speech and Language*, 36(2), 100–108.
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D-KEFS)*. The Psychological Corporation.
- Demircioglu, M. D. (2013). The pronunciation problems for Turkish learners in articulating of the diphthongs in English learning. *Procedia-Social and Behavioral Sciences*, 106, 2983–2992.
- Dunn, L. M., Dunn, D. M., Styles, B., & Sewell, J. (2009). *British picture vocabulary scale: BPVS-III*. Lloyd. GL Assessment.
- Eckman, F. R. (1985). Some theoretical and pedagogical implications of the markedness differential hypothesis. *Studies in Second Language Acquisition*, 7, 289–307.
- Eckman, F. R., Elreyes, A., & Iverson, G. K. (2003). Some principles of second language phonology. *Second Language Research*, 19(3), 169–208.
- Enomoto, K. (1994). L2 perceptual acquisition: The effect of multilingual linguistic experience on the perception of a "less novel" contrast. *Edinburgh Working Papers in Applied Linguistics*, 5, 15–29.

- Fabiano-Smith, L., & Goldstein, B. A. (2010a). Phonological acquisition in bilingual Spanish-English speaking children. *Journal of Speech, Language, and Hearing Research*, 53(1), 160–178.
- Fabiano-Smith, L., & Goldstein, B. A. (2010b). Early-, middle-, and late-developing sounds in monolingual and bilingual children: An exploratory investigation. *American Journal of Speech-Language Pathology*, 19(1), 66–77.
- Gallardo del Puerto, F. (2007). Is L3 phonological competence affected by the learner's level of bilingualism? *International Journal of Multilingualism*, 4(1), 1–16.
- Ganzeboom, H. B., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science*, 2(3), 119–128.
- Gildersleeve-Neumann, C. E., Kester, E. S., Davis, B. L., & Peña, E. D. (2008). English speech sound development in preschool-aged children from bilingual English-Spanish environments. *Language Speech and Hearing Services in Schools*, 39(3), 314–328.
- Goldstein, B. A., & Bunta, F. (2012). Positive and negative transfer in the phonological systems of bilingual speakers. *International Journal of Bilingualism*, 16, 388–401.
- Goldstein, B. A., Fabiano, L., & Washington, P. S. (2005). Phonological skills in predominantly English-speaking, predominantly Spanish-speaking, and Spanish-English bilingual children. *Language, Speech, & Hearing Services in Schools*, 36(3), 201–218.
- González-Ardeo, J. M. (2001). Engineering students and ESP in the Basque Country: SLA versus TLA. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Looking beyond second language acquisition: Studies in tri- and multilingualism* (pp. 75–95). Stauffenburg.
- Gut, U. (2010). Cross-linguistic influence in L3 phonological acquisition. *International Journal of Multilingualism*, 7(1), 19–38.
- Hanulíková, A., & Weber, A. (2010). Production of English interdental fricatives by Dutch, German, and English speakers. In *New Sounds 2010. Sixth international symposium on the acquisition of second language speech*. Adam Mickiewicz University (pp. 173–178).
- Hirosh, Z., & Degani, T. (2018). Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin & Review*, 25(3), 892–916.
- Hopp, H., Vogelbacher, M., Kieseler, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of proficiency in the minority and the majority language. *Learning and Instruction*, 61, 99–110.
- Hox, J. J. (2010). *Multilevel analysis. Techniques and applications* (2nd ed.). Routledge.
- Husfeldt, V., & Bader-Lehmann, U. (2009). Englisch an der Primarschule: Erfahrungen aus der Schweiz. In G. Engel, B. Groot-Wilken, & E. Thürmann (Eds.), *Englisch in der Primarstufe - Chancen und Herausforderungen: Evaluation und Erfahrungen aus der Praxis* (pp. 111–1239). Cornelsen.
- Ioup, G. (2008). Exploring the role of age in the acquisition of a second language phonology. In J. Hansen-Edwards, & M. Zampini. (Eds), *Phonology and second language acquisition* (pp. 176–203). Johns Benjamin Publishing Company.
- Kautzsch, A. (2010). Exploring L1 transfer in German learners of English: High front vowels, high back vowels and the BED/BAD distinction. *Research in Language*, 8(1), 63–84.
- Kehoe, M., Trujillo, C., & Lleó, C. (2001). Bilingual phonological acquisition: an analysis of syllable structure and VOT. In K. F. Cantone & M. O. Hinzelin (Eds). *Proceedings of the colloquium on structure, acquisition and change of grammars: Phonological and syntactic aspects* (pp. 38–54). Universität Hamburg, Arbeiten zur Mehrsprachigkeit.
- Kiparsky, P., & Menn, L. (1987). On the acquisition of phonology. In G. Loup & S. H. Weinberger (Eds.). *Interlanguage phonology. The acquisition of a second language sound system: Series on issues in second language research* (pp. 23–52). Newbury House Publishers.
- Knight, R.-A., Villafana Dalcher, C., & Jones, M. J. (2007, August). A real-time case study of rhotic acquisition in Southern British English. In *International congress of phonetic sciences XVI* (pp. 1581–1584). Saarland University, Saarbrücken, Germany.
- König, E., & Gast, V. (2012). *Understanding English-German contrasts* (Vol. 29). Schmidt.

- Kopečková, R. (2016). The bilingual advantage in L3 learning: A developmental study of rhotic sounds. *International Journal of Multilingualism*, 13(4), 410–425.
- Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, 144(4), 394–425.
- Lloyd-Smith, A., Gyllstad, H., & Kupisch, T. (2017). Transfer into L3 English. *Linguistic Approaches to Bilingualism*, 7(2), 131–162.
- MacLeod, A. A., Laukys, K., & Rvachew, S. (2011). The impact of bilingual language learning on whole-word complexity and segmental accuracy among children aged 18 and 36 months. *International Journal of Speech-Language Pathology*, 13(6), 490–499.
- Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second-language phonology*. *Second language acquisition research series*. Erlbaum.
- Marx, N., & Mehlhorn, G. (2010). Pushing the positive: Encouraging phonological transfer from L2 to L3. *International Journal of Multilingualism*, 7(1), 4–18.
- May, P. (2006). Englisch-Hörverstehen am Ende der Grundschulzeit. In W. Bos & M. Pietsch (Eds.), *KESS 4-Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen*. (Vol. 1, pp. 203–224). Waxmann.
- Morales Reyes, A., Arechabaleta-Regulez, B., & Montrul, S. A. (2017). The acquisition of rhotics by child L2 and L3 learners. *Journal of Second Language Pronunciation*, 3(2), 242–266.
- Patihis, L., Oh, J. S., & Mogilner, T. (2015). Phoneme discrimination of an unrelated language: Evidence for a narrow transfer but not a broad-based bilingual advantage. *International Journal of Bilingualism*, 19(1), 3–16.
- Petermann, F., & Petermann, U. (2008). HAWIK-IV. *Kindheit und Entwicklung*, 17(1), 71–75.
- Rahimpour, M., & Saedi Doveise, M. (2011). A phonological contrastive analysis of Kurdish and English. *International Journal of English Linguistics*, 1(2), 73–82.
- Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, 81(1), 174–176.
- Tremblay, M.-C., & Sabourin, L. (2012). Comparing behavioral discrimination and learning abilities in monolinguals, bilinguals and multilinguals. *The Journal of the Acoustical Society of America*, 132(5), 3465–3474.
- Weber, J., Marx, P., & Schneider, W. (2007). Die vorschulische Förderung der phonologischen Bewusstheit. In H. Schöler & A. (H.) Welling (Eds.), *Sonderpädagogik der Sprache* (pp. 746–761). Hogrefe.
- Weiß, R. (2006). *CFT 20-R.: Grundintelligenztest Skala 2*. Manual. Hogrefe.
- Wilden, E., & Porsch, R. (2015). Die Hör- und Leseverstehensleistungen im Fach Englisch von Kindern am Ende der Grundschulzeit unter besonderer Berücksichtigung von lebensweltlicher Ein- und Mehrsprachigkeit. In M. Kötter & J. Rymarczyk (Eds.), *Englischunterricht auf der Primarstufe* (pp. 59–80). Peter Lang.
- Wode, H. (1983). Phonology in L2 acquisition. In H. Wode (Ed.), *Papers on language acquisition, language learning and language teaching* (pp. 175–187). Groos.

### Author biography

Teresa Kieseier is an academic staff member and doctoral student at the department of English at the University of Mannheim (Germany). Her research interests include L1 and child L2/L3 acquisition, multilingualism and phonology.



## Appendix

Table A1. Descriptive statistics of linguistic, personal, cognitive and social factors.

	Full sample	monolinguals	bilinguals	t-tests	$d_{Cohen}$
<i>Linguistic factors</i>					
Correct pronunciation E	0.87 (0.13)	0.86 (0.14)	0.87 (0.13)	$t(181) = -0.39^{ns}$	.07
Receptive vocabulary E	39.68 (6.31)	41.47 (5.69)	38.27 (6.45)	$t(182) = 3.51^{***}$	.52
Productive vocabulary E	11.25 (5.69)	12.36 (5.22)	9.91 (5.45)	$t(182) = 3.08^{**}$	.46
Receptive grammar E	45.98 (10.16)	48.17 (8.16)	43.0 (9.48)	$t(182) = 3.90^{***}$	.58
Productive vocabulary G	21.77 (7.31)	23.17 (7.10)	20.41 (7.27)	$t(182) = 2.59^*$	.38
Productive vocabulary LI	11.05 (7.17)	—	11.05 (7.17)	—	—
<i>Individual factors</i>					
Sex (0 = male; 1 = female)	0.45 (0.50)	0.41 (0.49)	0.48 (0.50)	$t(182) = -0.92^{ns}$	.14
Age (months)	111.12 (5.76)	110.35 (5.50)	111.72 (5.92)	$t(180) = -1.59^{ns}$	.24
Phonological awareness	0.57 (0.36)	.67 (0.33)	0.49 (.36)	$t(181) = 3.42^{***}$	.52
Basic cognitive skills	101.98 (14.97)	106.78 (14.78)	98.22 (14.08)	$t(178) = 3.96^{***}$	.60
Working memory	7.79 (1.70)	8.06 (1.86)	7.57 (1.54)	$t(182) = 1.95^{ns}$	.29
Executive control	66.33 (46.88)	60.68 (41.64)	70.78 (50.37)	$t(182) = -1.46^{ns}$	.22
<i>Social factors</i>					
Education parents (years)	11.52 (1.63)	12.04 (1.45)	11.06 (1.65)	$t(144.9) = 3.84^{***}$	.63
SES parents	49.46 (10.05)	53.92 (8.24)	45.62 (9.91)	$t(144.8) = 5.55^{***}$	.91
Net income	3764.25 (2050.84)	4468.68 (2063.90)	3157.91 (1846.28)	$t(145) = 4.06^{***}$	.67
Cultural capital	237.56 (384.56)	392.86 (506.79)	111.83 (161.11)	$t(78.0) = 4.40^{***}$	.78

Note. Descriptive statistics full sample, monolingual versus bilingual group: Mean (SD), t-tests, effect size.

For standardized tests: Higher mean = better test result.

ns = not significant, \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

Social values only available for subgroup with parent information, missing values caused by item non-response ML-estimated (EM-algorithm).

**Table A2.** Item list picture naming task.

---

cat  
 dog  
 apple  
 ball  
 red  
 white  
 mouse  
 ring  
 spider  
 spoon  
 star  
 toothbrush  
 window  
 bathroom  
 mouth  
 feather  
 street  
 stairs  
 sheep  
 fire  
 fish  
 strawberry  
 elephant

---

**Table A3.** Bilingual learners: predictors of pronunciation accuracy (picture naming) in grade 4 at institutional and individual levels (hierarchical mixed linear regression); controlled for effects of school, class affiliation and LI.

Parameters	Null Model	Model 1	Model 2
<b>Fixed effects</b>			
Intercept	0.89*** (0.01)	0.92*** (0.02)	0.89*** (0.02)
<b>Institutional level</b>			
Proportion of bilingual/school	–	n.s.	n.s.
Mean SES/school	–	n.s.	n.s.
<b>Individual level</b>			
<i>Linguistic factors</i>			
Productive vocabulary German	–	n.s.	n.s.
LI group	–	–0.002 <sup>ns</sup> (0.002)	–0.0003 <sup>ns</sup> (0.002)
Productive vocabulary LI	–	–0.01 <sup>ns</sup> (0.01)	–0.02 <sup>ns</sup> (0.01)
Productive vocabulary English	–		n.s.
Receptive vocabulary English		0.06*** (.01)	0.06*** (0.02)
Receptive grammar English		n.s.	n.s.
<i>Personal &amp; cognitive factors</i>			
Gender (0=male, 1=female)	–	n.s.	n.s.
Age (months)	–	n.s.	n.s.
Nonverbal IQ	–	n.s.	n.s.
Working memory	–	n.s.	n.s.

(Continued)

**Table A3.** (Continued)

Parameters	Null Model	Model 1	Model 2
Executive control	–	n.s.	n.s.
Phonological awareness		n.s.	n.s.
<i>Social factors</i>			
<i>Parent's nonresponse</i>	–	n.s.	–
SES parents	–	–	–0.01 <sup>ns</sup> (0.02)
Education parents	–	–	–0.03 <sup>ns</sup> (0.02)
Net-income family	–	–	0.003 <sup>ns</sup> (0.02)
Cultural capital	–	–	0.003 <sup>ns</sup> (0.03)
Random parameters			
<i>Variance</i>			
Within LIcs group	0.01	0.01	0.01
Between LIcs group	0.0003	0.0000	0.0002
N	95	95	69 <sup>a</sup>
–2 Restricted Log Likelihood	–140.2	–157.1	–120.0
Number of parameters	3	6	10

Note. Dependent variable: Pronunciation accuracy (grade 4).

Notation: unstandardized estimates ( $\beta$ ), standard error (SE).

Predictor variables z-standardized.

ns = not significant, \* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ .

Null Model without fixed predictors  $\beta$  (SE).

Model 1: Null Model + institutional, linguistic, personal and cognitive predictors  $\beta$  (SE).

Model 2: Model 1 with social predictors  $\beta$  (SE).

Random effects: LI groups within schools and classes (LIcs groups).

Models optimized via  $X^2$ -Comparison. Predictors not improving model fit removed (n.s.).

<sup>a</sup>Only cases with full information from parents; missing values caused by item nonresponse ML-estimated (EM-algorithm).