Abstract As a consequence of globalization and migration, the number of children receiving literacy instruction in their second language (L2) is high and still increasing. Therefore, teachers need instruction methods that are effective in both L1 and L2 learners. Here, we investigate the effectiveness of a computerized training program combining phoneme perception, phonological awareness, and systematic phonics, in a sample of second-graders ($N = 26$) instructed in German as L2. Based on prior evidence concerning (1) literacy acquisition in L2 and (2) effects of literacy development on oral language abilities, we expected significant training effects on children’s literacy skills and vocabulary knowledge. The children of the training group worked through the program during school lessons, 20 min per day, for a period of 8 weeks. The controls continued to receive standard classroom instruction. German tests of phonological awareness, reading, spelling, and vocabulary were performed at three time points (pretest, immediate posttest, and follow-up after 9 weeks). Analyses confirmed that improvements in phonological awareness, spelling, and vocabulary between pretest and posttest were stronger in the training group when compared to the controls. For spelling and vocabulary, these effects were still significant at follow-up. Effect sizes were medium to high. For the reading measures, no group differences were found. In sum, the results yield further evidence for the effectiveness of phonics-based literacy instruction in L2 learners, and for the beneficial effects of basic literacy skills on novel word learning.

Keywords Second language learners · Phonics instruction · Phonological representations · Orthographic facilitation · Computer-based intervention · Orthographic representations

Introduction

As a result of globalization and migration, many children receive literacy instruction in their second language (L2). In Germany, the proportion of primary
school children with a migration background is 37% on average, and more than 60% in metropolitan areas (Autorengruppe Bildungsberichterstattung 2018; Klatte et al. 2017a). About 20% of these children are learning German as L2 (Koschollek et al. 2019). Main languages spoken at home differ from the language of instruction in 22% and 15% of the students in the US and EU, respectively (Camarota and Zeigler 2015; Wendt and Schwippert 2017). Studies consistently show that L2 learners perform worse in literacy tests when compared to their monolingual peers (Farnia and Geva 2013; Klatte et al. 2017a; Melby-Lervåg and Lervåg 2014; Weis et al. 2019; Wendt and Schwippert 2017). These group differences are at least partly mediated by differences in proficiency in the language of instruction (Wendt and Schwippert 2017). However, lower socioeconomic status and parental education, and less favorable home literacy environments also contribute to L2 learners’ difficulties in literacy acquisition (Autorengruppe Bildungsberichterstattung 2018; Klatte et al. 2017a; Niklas and Schneider 2013; Wendt and Schwippert 2017).

Due to long-lasting and intensive research, a lot is known concerning early literacy instruction and intervention for children instructed in their native language (L1; e.g. Galuschka et al. 2014; Teale et al. 2018). Concerning literacy instruction in L2, research has mainly focused on English language learners. Findings indicate that the development of basic reading skills is highly comparable for L2 and L1 learners (for a recent review, see Schaars et al. 2019), and that evidence-based instruction programs designed for native children are also effective in students instructed in L2 (Goldenberg 2011; Ludwig et al. 2019; Richards-Tutor et al. 2016). However, less evidence is available for more transparent orthographies such as German. In the current study, we investigated whether and to what extent a computerized grapho-phonological training program that has been successfully evaluated in monolingual German children is also effective in immigrant children learning German as L2. Based on the empirical evidence described below, we expected beneficial training effects (1) on early literacy skills, and (2) transfer effects on German vocabulary.

Training components and their role in L1 and L2 literacy acquisition

In the current study, training was administered via the computerized training program Lautarium (Klatte et al. 2017b), which has proved effective in children with dyslexia (Klatte et al. 2014, 2017b), first-grade at-risk children (Klatte et al. 2016), and typically developing readers (Klatte et al. 2018). Lautarium fosters basic literacy skills through a combination of systematic phonics instruction and auditory-phonological training. Phonics instruction means that children are taught to read and spell syllables and words via phonics-based strategies, i.e., by acquisition and application of grapheme-phoneme and phoneme-grapheme correspondences. Several systematic reviews and meta-analyses consistently proved the effectiveness of phonics-based instruction in fostering literacy acquisition in struggling and typically developing readers (Ehri et al. 2001; Galuschka et al. 2014; McArthur et al. 2012, 2018). Phonics-based interventions help the children to grasp the alphabetic principle, i.e., to understand how letters in written words map onto phonemes in spoken words (Hatcher et al. 2004; Snowling and Hulme, 2012). According to the “self-teaching hypothesis” proposed by Share (1995), phonological decoding, i.e., the ability to decode unfamiliar words through application of grapheme-phoneme mappings, enables acquisition of word-specific orthographic representations, which in turn provide the foundation of rapid word recognition. The proposed mechanism has been confirmed in simulation studies (Ziegler et al. 2014). Fostering decoding abilities thus set the ball rolling for acquisition of orthographic knowledge and fluent reading.

Several studies confirmed that, for L2 learners, phonics instruction is as effective, or even more effective when compared to children instructed in L1 (Ginns et al. 2019; Goldenberg 2011; Ludwig et al. 2019). Importantly, this holds true regardless of oral L2 proficiency, i.e., even students with poor oral language scores profit from phonics-based instruction (Goldenberg 2011; Gunn et al. 2005; Manis et al. 2004).

In Lautarium, phonics instruction is combined with intensive auditory-phonological training, i.e., phoneme perception and phonological awareness. Phonological awareness, i.e., the ability to consciously access and manipulate the sound units of language,
has proved fundamental for reading and spelling acquisition in numerous studies since the 1980s (e.g. Bradley and Bryant 1983; Wagner and Torgesen 1987). Phonological awareness in kindergarten predicts later literacy skills (for reviews, see Melby-Lervåg et al. 2012; Pfost 2015), and children with literacy disorders exhibit severe difficulties in phonological awareness tasks when compared to typically developing children matched for chronological age, or reading age (Melby-Lervåg et al. 2012). Furthermore, training phonological awareness in preschoolers fosters reading and spelling acquisition in the early grades (Ehri et al. 2001; Fischer and Pfost 2015). Concerning phonological awareness in L2 learners, mixed results have been reported. Jongejan et al. (2007) found comparable performance levels and developmental trajectories for phonological awareness from Grade 1 to 4 in children with English as a first vs. second language. In a meta-analysis of 43 studies using phonemic awareness tests, a small but significant effect was found in favor of the first-language learners (Melby-Lervåg and Lervåg 2014). For L2 learners of German, some studies reported phonological awareness performance comparable to native German children (Czapka et al. 2019; Duzy et al. 2013a, b; Limbird et al. 2014; Limbird and Stanat 2006), whereas others found group differences in favor of the native children (Blatter et al. 2013; Pröscholdt et al. 2013; Schöppe et al. 2013; Weber et al. 2007). Like in native-speaking children, phonological awareness predicts reading and spelling skills in L2 learners of English (Harrison et al. 2016; Jongejan et al. 2007; Quiroga et al. 2002) and German (Czapka et al. 2019; Duzy et al. 2013a; Limbird et al. 2014). Furthermore, phonological awareness training fosters early literacy skills in L2 (Blatter et al. 2013; Schöppe et al. 2013; Weber et al. 2007; Yeung et al. 2013).

In contrast to other phonics-based programs, Lautarium includes intensive training of phoneme perception. Phoneme categories of the native language are established within the first year of life (Kuhl 2004; Werker and Tees 1984), but continue to be refined up until adolescence (Hazan and Barrett 2000). Phoneme perception predicts reading acquisition in typically developing children (Boets et al. 2011), and is deficient in children with reading disorders (Klatte et al. 2013; Manis et al. 1997; McBride-Chang 1995). Several authors argue that speech perception deficits are causally involved in disordered reading acquisition (Hornickel et al. 2009; Noordenbos and Serniclaes 2015; Ziegler et al. 2009). Poor phoneme perception may result in less specified (Vandermosten et al. 2020) or noisy phonological representations in the mental lexicon which, in turn, affect processing, storage, and access to phonological information, and hamper the building-up of phoneme-grapheme-correspondences (Elbro and Jensen 2005; Swan and Goswami 1997). Concerning second-language learners, studies revealed deficits in L2 phoneme perception in migrant preschool children learning Dutch (Janssen et al. 2017) and English (McCarthy et al. 2014) as L2, and in primary school children learning German as L2 (Darcy and Krüger 2012). Deficits in L2 speech perception (indicated by speech recognition in noise) were even found in bilingual adults speaking fluent, accent-free English, indicating less robust L2 phoneme representations despite high L2 proficiency (Rogers et al. 2006; Tabri et al. 2011).

Phoneme perception tasks in Lautarium require discrimination and identification of consonants and vowel lengths. Vowel length perception is included since, in German language, vowel length is phonemic and orthographically marked. For example, the spoken German words kann (/kan/, [can]) and Kahn (/ka:n/, [barge]) differ only in vowel length. In orthography, vowel length is marked by the letters following the vowel according to specific rules. Thus, in order to understand and use the respective rules, a child needs to be able to classify a vowel as short or long. Impaired vowel length perception has been reported for German children with literacy disorders (Klatte et al. 2013; Landerl 2003; Steinbrink et al. 2014), and for children learning German as L2 (Brunner 2012).

 Altogether, the available evidence indicates that the training components of Lautarium may be efficient in fostering early literacy skills in L2 learners. In the following section, we will summarize the evidence underlying our second hypothesis, stating that Lautarium training fosters vocabulary knowledge in L2 learners.

**Effects of literacy acquisition on spoken language in L1 and L2 learners**

Despite huge individual differences, as a group, second-language learners with an immigrant background have lower proficiency in the language of
instruction when compared to their monolingual peers (Melby-Lervåg and Lervåg 2014; Schaars et al. 2019; Wendt and Schwippert 2017). This disadvantage is especially strong for children exposed to another language at home (Hoff 2013; Wendt and Schwippert 2017). Proficiency in the language of instruction is closely related to school achievement, especially with respect to literacy skills (Prevoo et al. 2016). Concerning the latter, studies consistently showed that L2 vocabulary knowledge reliably predicts reading (text) comprehension, but has a minor or no effect on word decoding and phonological awareness (Babayigit, 2015; Knell 2018; Melby-Lervåg and Lervåg 2014; Prevoo et al. 2016). These findings led to the general conclusion that instruction in early literacy skills in L2 should not be postponed until a certain vocabulary size is acquired (Goldenberg 2011; Knell 2018). Here, we go a step further, arguing that L2 vocabulary acquisition may be enhanced by training early reading skills in L2. This assumption is based on recent evidence concerning the effects of reading acquisition on phonological long-term representations, and their role in learning new words. Especially, it has been argued that learning to read in an alphabetic system fosters the building-up of fine-grained, precise phonological representations on the lexical (i.e., words) and sub-lexical levels (syllables and phonemes), which in turn support precise encoding and maintenance of novel phonological forms. Different lines of evidence confirm this view. First, studies on categorical speech perception revealed that literacy acquisition improves precision of phoneme boundaries in adults (Serniclaes et al. 2005) and children (Burnham 2003; Hoonhorst et al. 2011). Second, it has been shown that phonological awareness predicts vocabulary acquisition in L1 and L2 learners (Kalia et al. 2018; Marecka et al. 2018), and that training of phonological awareness (i.e., blending, deletion, and substitution of phonemes in words) improves later recall of the target words much more than rhyme training or teaching definitions (Melby-Lervåg and Hulme 2010). In addition, training of phonological awareness and letter-sound mappings in preschoolers fosters long-term learning of new, unfamiliar names associated with cuddle toys (de Jong et al. 2000). Third, studies revealed beneficial effects of reading instruction on verbal short-term memory (for review, see Demoulin and Kolinsky 2016). Especially, it has been shown that word decoding skills in 6-year-olds predict subsequent growth in nonword repetition (Nation and Hulme 2011), which in turn is a valid predictor of vocabulary acquisition in L1 (Baddeley et al. 1998) and L2 (Hummel and French 2016; Service 1992). Fourth, studies with school-age children revealed that exposure to printed words enhances learning of unfamiliar pseudowords (Jubenville et al. 2014; Ricketts et al. 2009; Rosenthal and Ehri 2008) and of novel L2 words (Chambre` et al. 2020; Hu 2008). For example, in the Jubenville et al. (2014) study, children had to learn novel labels for unfamiliar objects presented pictorially. In one learning condition, printed words were presented in addition to the pictures and spoken labels. Print reliably facilitated the acquisition and recall of the words in both monolingual and bilingual children. This “orthographic facilitation effect” is still more pronounced when students have to decode the words during the learning phase (Chambre` et al. 2020), and is attributed to the interaction of orthographic and phonological representations in the mental lexicon. Especially, it has been argued that, during reading acquisition, increasing orthographic knowledge permanently refines phonological word representations (Muneaux and Ziegler 2004), and that lexical representations integrating semantic, phonological and orthographic knowledge are especially stable and easily retrieved (Perfetti and Hart 2002).

Based on this evidence, we expect the intensive phonics and phonological training provided by Lau-tarium to foster both literacy skills and vocabulary knowledge in immigrant children with German as a second language.

Method

Participants

Participants were 26 s-grade children (14 male, mean age 7;11 SD 5 months) with an immigrant background learning German as a second language. The children were recruited from four classes of a primary school in Bavaria. Children from two classes (n = 12; 5 male, mean age 7;9, SD 4.2 months) were assigned to the training group. The remaining children (n = 14, 9 male, mean age 8;0, SD 5.4 months) served as controls. One child from the control group did not participate in the follow-up test, so the number of controls was reduced to 13.
Written informed consent was provided by the parents. The study was approved by the responsible Bavarian school authority and data protection officer. However, applicable data protection rules did not allow inquiry of the children’s heritage languages.

Design

Short- and long-term effects of Lautarium-training were assessed by means of a pretest–posttest-follow-up design (see Fig. 1). The study started with a pretest in the middle of the school year. Thereafter, the training group performed Lautarium during school lessons, 5 times per week, for 20–30 min, over a period of 8 weeks. The control group continued to receive standard classroom instruction. The training phase was followed by posttest. The follow-up test was performed after an interval of 9 weeks, in which both groups attended regular classroom lessons.

Children’s reading, spelling, phonological awareness, and vocabulary scores were assessed at each measurement point (pretest, posttest, follow-up). In addition, a nonverbal reasoning task was included in the pretest in order to rule out group differences in general cognitive abilities. Testing was done in the morning in a separate classroom in the school. Tests of reading comprehension, spelling, phonological awareness and nonverbal intelligence were administered in groups. Decoding skills (reading aloud) and vocabulary (picture naming) were assessed individually.

Diagnostic tests

Word and pseudoword decoding were assessed by means of the “One Minute Test of Reading Fluency”, a subtest from the “Salzburg Reading and Spelling Test” (SLRT-II; Moll and Landerl 2014). Two lists are presented to each child, one containing words and one containing pseudowords constructed according to German phonotactic rules. The task is to read the words and pseudowords out aloud correctly as fast as possible. Raw scores represent the number of items correctly read within 1 min.

Reading comprehension on the level of words was assessed by means of a subtest of the “Test of Reading comprehension for Grade 1 to 6” (ELFE 1-6; Lenhard and Schneider 2006). The children have to select, out of four alternatives, the word that matches a target picture. Raw scores represent the number of items completed correctly within 3 min.

For spelling, a standardized German spelling test for second-graders was used, in which the children have to write down 15 words and 3 sentences according to dictation (Hamburg Spelling Test, HSP 2; May 2012b, c). Five raw scores are derived from this test: number of correct words, number of correct graphemes, and success in application of letter-sound mappings (“alphabetic strategy”), spelling rules, and morphemic knowledge. As the norms provided in this test are applicable only at the end of the second school year, a further version of the “Hamburg Spelling Test” (HSP 1+; May 2012a) with norms for the middle of the school year was additionally included in the pretest.
Phonological awareness was assessed by means of three subtests from the “Kaiserslauterer Gruppentest zur Lautbewusstheit” (Klatte et al. manuscript in preparation). The subtests required speeded identification, deletion, and substitution of sounds in words presented pictorially. Concerning identification, the children have to decide whether or not a specific speech sound is present in a number of target words. In case of “yes”-answers, the children also have to indicate the position of the sound in the word (beginning, middle, end). For sound deletion and substitution, the children have to indicate, for each target word, which word emerges when the second phoneme is eliminated or substituted, respectively. For each subtest, raw scores represent the number of items correctly solved within 3 min.

For vocabulary assessment, a subtest from a German test battery for language proficiency in 5- to 10-year-old children was included (SET 5-10; Petermann et al. 2010). The test consists of 40 pictures representing objects and actions. For each picture, the child has to pronounce the appropriate German word. The raw score represents the number of items correctly solved.

For assessment of nonverbal abilities, 16 matrices increasing in difficulty were selected from a subtest of nonverbal intelligence in children. The test is part of a widely used developmental test battery for primary school children in Germany (Esser et al. 2008). For each matrix, the missing element has to be chosen out of 5 alternatives. The raw score represents the number of items correctly solved.

Training

Training was administered via the computer program Lautarium (Klatte et al. 2017a, b). The program combines auditory-phonological training with training of grapheme-phoneme-mappings and reading and spelling of transparent words. In addition, rapid access from written words to meaning is included by means of a word-to-picture matching task with fast presentation times.

Structure and contents of the training program

Lautarium uses two types of building blocks representing phonemes and the corresponding graphemes, respectively. Phonemes are represented by blocks with pictures of easy-to-name objects. The phoneme represented by a specific block is the initial sound of the picture’s verbal label (e.g., the picture of a ball represents the phoneme/b/). Graphemes are represented by blocks with a single letter or a letter combination that is usually used for a specific phoneme in German orthography (Thomé 2000), e.g., the phoneme/a/is usually represented by the grapheme ⟨ei⟩.

The speech material comprises about 1400 German nouns and 2500 pseudowords differing in phonological structure (CV, VCV, CVC, CCV, VCCV). Each speech item is implemented through high-resolution recordings produced by two professional speakers (a male and a female). For about 40% of the nouns, pictorial presentation is also available.

Tasks belonging to different training domains are intermixed, and within each domain, children start with simple tasks and proceed to more difficult ones. Due to the combinations of tasks and materials, Lautarium consists of 58 different exercises. Each exercise comprises 10–30 trials, depending on task complexity. Responses are followed by immediate feedback (correct, incorrect, time-out). In case of errors or time-outs, the respective trial is repeated until the correct answer is provided in time. Depending on the percentage of trials correctly solved in the first attempt, the child either has to repeat the respective exercise, or proceeds to the following one.

Before starting a new task, children have to work through an interactive instruction that provides explanation, practice with examples, and informative feedback. In addition, Lautarium is equipped with a token system fostering focused, intensive training.

Phoneme perception

Phoneme perception tasks require discrimination (same/different response to pairs of pseudowords) and identification of plosives and vowel lengths. Plosive perception tasks proceed from simple syllable structures to more complex structures in which the target plosives are embedded in consonant clusters. For the latter, phoneme perception deficits in children with literacy disorders are especially pronounced (Klatte et al. 2013). In vowel lengths identification tasks, target words are presented auditory or pictorially, and children have to classify the lengths of the vowels as short or long.
Phonological awareness

Phonological awareness tasks require sound-to-word matching, segmentation, blending, and matching of initial or final sounds in words and pseudowords. For sound-to-word matching, a phoneme block is presented, followed by two words presented auditorily (each synchronized with a response button) or pictorially. The children have to decide whether or not the target phoneme is present in one of the words. For segmentation, words are presented auditorily or pictorially. The corresponding phoneme blocks have to be selected and arranged in the correct order. In blending tasks, a sequence of phoneme blocks constituting a word is presented, and the children have to select the correct word by clicking on the corresponding picture. Matching of initial or final sounds is trained by means of odd-one-out tasks originally used by Bradley and Bryant (1983). Sequences of three words or pseudowords are presented, each synchronized with a response button. The children have to decide which of the three items differs from the others with respect to the initial or final sound.

Grapheme-phoneme correspondences

In Lautarium, training of grapheme-phoneme-mappings and phonological training are interlinked. For example, when the correct phonemes have been selected in phoneme identification or segmentation tasks, the corresponding graphemes have to be assigned.

Reading and spelling

In Lautarium, newly acquired phonological skills and letter knowledge are immediately applied in reading and spelling tasks. Reading tasks comprise matching printed words to spoken words (and vice versa), and matching pictures to printed words. The latter task focuses on reading speed, aiming to foster direct access from print to meaning and thus enlarge the children’s sight word vocabulary. Spelling tasks require segmentation of target words presented auditorily or pictorially into their constituent graphemes.

Statistical analyses

In the spelling test, raw scores reflecting application of spelling rules and morpheme knowledge were highly correlated (pretest: $r = 0.77$, posttest: $r = 0.71$, follow-up: $r = 0.85$) and thus summed up to a single score termed “orthographic strategy” for the analyses.

Potential group differences in learning gains between pretest and posttest, and between pretest and follow-up were assessed by means of analysis of covariance (ANCOVA), with the pretest score from the respective test included as covariate (O’Connell et al. 2017; Rausch et al. 2003). In case of violation of ANCOVA assumptions, repeated measurement analyses of variance (ANOVAs) were performed including measurement point as within-subject and treatment group as between-subject factor. For significant results ($p < 0.05$), effect sizes corrected for pretest scores ($d_{corr}$) were calculated (Klauer 1989). As performing multiple analyses bears the danger of cumulation of type 1 error, we verified our results by performing repeated measurement multivariate analyses of variance (MANOVAs) for each of the outcome domains (reading, spelling, phonological awareness). MANOVA statistics are provided in the appendix.

Results

For the sample as a whole, one-sample t-tests revealed that standard scores (T-scores) for reading comprehension, spelling, and vocabulary were significantly below the means of the norm samples (one sample t-tests, all $p < 0.05$), with effect sizes ranging from $d = 1.86$ for vocabulary to $d = 0.46$ for reading comprehension (see Table 1). Concerning decoding, mean scores for words and pseudowords corresponded to the percentile rank bands 47–49 and 50–53, respectively (no T-scores were available for this test). Thus, as expected from prior research, deficits in reading comprehension, spelling, and—most pronounced—vocabulary were confirmed in the current sample of L2 learners, with decoding skills being in the average range.

Training effects

Performance raw scores and statistical results concerning treatment effects at posttest and at follow-up
are provided in Tables 2 and 3, respectively. For all variables, pretest scores did not differ significantly between groups (t-tests, all ps > 0.05).

**Reading** No significant treatment effects were found, neither for word comprehension nor for decoding of words and pseudowords (all ps > 0.21).

**Spelling** Due to inhomogeneity of regression slopes for number of correct words, number of correct graphemes, and application of letter-sound-mappings ("alphabetic strategy"), training effects at posttest were analyzed by repeated measurement ANOVAs, with time (pretest vs. posttest) as within-subject factor and treatment group as between-subject factor. These analyses revealed non-significant main effects of group (F(1,24) < 1 in all cases), and significant (correct words: F(1,24) = 26.23, p < 0.001; correct

### Table 1: Mean standard scores (T-Scores) at pretest and deviations from the norm sample means (T = 50) for vocabulary, spelling, and reading (N = 26)

| Variable          | Score | M     | SD    | t(1, 25)b | p   | d     |
|-------------------|-------|-------|-------|-----------|-----|-------|
| Vocabulary        | Tb    | 34.85 | 8.16  | -9.47     | < 0.001 | 1.86 |
| Spelling          |       |       |       |           |      |       |
| Correct words     | T     | 45.65 | 6.97  | -3.18     | 0.004 | 0.62 |
| Correct graphemes | T     | 46.23 | 7.14  | -2.69     | 0.012 | 0.53 |
| Alphabetic strategy| T    | 47.08 | 8.08  | -1.84     | 0.077 |       |
| Orthographic strategy| T  | 45.65 | 7.34  | -3.02     | 0.006 | 0.59 |
| Reading           |       |       |       |           |      |       |
| Comprehension, words | T   | 47.18 | 6.08  | -2.37     | 0.026 | 0.46 |
| Decoding, words   | PRc   | 47–49 | Not applicable |       |      |       |
| Decoding, pseudowords | PRc | 50–53 | Not applicable |       |      |       |

a t test for one sample to M = 50
b T-Scores
c Percentile range

### Table 2: Mean performance in training and control group and treatment effects at posttest (n = 26)

| Variable          | Pretest | Posttest | Results and effect sizes |
|-------------------|---------|----------|--------------------------|
|                   | Training group, n = 12 | Control group, n = 14 | Training group, n = 12 | Control group, n = 14 | Posttest |
|                   | M      | SD | M | SD | M | SD | F | p | dcorr |
| Nonverbal intelligencea | 13.08 | 0.79 | 11.86 | 2.91 |       |       |       |         |
| Phoneme awareness  |       |     |   |    |   |    |   |     |       |
| Identification     | 22.92 | 5.68 | 22.93 | 5.59 | 26.25 | 8.32 | 24.29 | 6.12 | F(1, 23) = 0.50 | 0.488 | – |
| Deletion           | 5.17  | 3.27 | 6.71  | 2.87 | 10.75 | 3.22 | 8.57  | 3.69 | F(1, 23) = 4.90 | 0.037 | 1.13 |
| Substitution       | 3.00  | 2.13 | 5.07  | 2.97 | 8.08  | 3.48 | 7.29  | 3.91 | F(1, 23) = 4.17 | 0.053 | 1.00 |
| Reading            |       |     |   |    |   |    |   |     |       |
| Comprehension      | 21.25 | 4.71 | 23.21 | 7.20 | 28.08 | 6.11 | 29.07 | 7.70 | F(1, 23) = 0.15 | 0.704 | – |
| Decoding, words    | 39.75 | 11.64 | 37.57 | 18.07 | 49.25 | 11.14 | 44.86 | 16.69 | F(1, 23) = 1.66 | 0.211 | – |
| Decoding, pseudowords | 30.83 | 9.91 | 29.43 | 12.19 | 33.67 | 8.18 | 32.71 | 10.02 | F(1, 23) = 0.01 | 0.935 | – |
| Spelling           |       |     |   |    |   |    |   |     |       |
| Correct words      | 14.17 | 4.65 | 15.29 | 5.50 | 18.33 | 4.25 | 16.71 | 6.65 | F(1, 24) = 4.26b | 0.019 | 0.50 |
| Correct graphemes  | 126.83 | 7.85 | 126.43 | 11.11 | 132.83 | 6.01 | 129.21 | 11.19 | F(1, 24) = 5.88b | 0.023 | 0.35 |
| Alphabetic strategy| 17.50 | 2.43 | 17.79 | 3.09 | 18.67 | 1.78 | 17.71 | 3.15 | F(1, 24) = 5.33b | 0.030 | 0.47 |
| Orthographic strategy | 9.42 | 5.25 | 10.14 | 6.29 | 12.67 | 4.85 | 12.50 | 6.54 | F(1, 23) = 0.78 | 0.387 | – |
| Vocabulary         | 26.75 | 4.20 | 28.07 | 3.69 | 30.92 | 4.21 | 29.14 | 4.93 | F(1, 23) = 9.35b | 0.006 | 0.72 |

a Assessed only at pretest
b F-values representing time × group-interaction in repeated measurement ANOVAs
The advantages of the training group endured until follow-up. For application of orthographic rules, a significant training effect of medium effect size emerged at follow-up. Effect sizes were in the medium range ($d_{corr} = 0.35–d_{corr} = 0.71$). The results were confirmed through MANOVAs on the four spelling variables by significant time $\times$ group interactions (see Table 4 in the appendix).

Phonological awareness At posttest, significant and marginally significant ($p = 0.053$) effects of strong effect size in favor of the training group were found for sound deletion and sound substitution, respectively. However, these effects were no longer significant at follow-up. MANOVAs on the three awareness measures (see Table 4 in the appendix) revealed non-significant time $\times$ group interactions when three measurement times were included, and when pretest and follow-up were included ($p > 0.10$ in both cases), whereas a MANOVA including pretest and posttest data confirmed a significant time $\times$ group interaction ($p = 0.025$).

Vocabulary Significant beneficial training effects of moderate effect size were found at both posttest ($d_{corr} = 0.72$) and follow-up ($d_{corr} = 0.62$).

Discussion

In the current study, the effectiveness of a computerized phonological training program was assessed in children with an immigrant background receiving literacy instruction in German as L2. Training effects were assessed by means of a pretest–posttest–follow-up design. Training was administered via the program Lautarium, which has proved effective in prior studies with native German children with and without literacy disorders. In the current sample of L2 learners, we found significant and enduring training gains of
medium effect size in spelling and vocabulary knowledge. Concerning phonological awareness, beneficial training effects of strong effect sizes were found in 2 of 3 subtests at posttest. For reading, no training effects were found, neither for decoding (reading aloud) of words and pseudowords, nor for comprehension of words read silently.

Pretest scores confirmed substantial deficits in German vocabulary knowledge in the current sample of second-graders instructed in German as L2. Average scores were more than 1.5 SD below the mean of the norm sample. Inspection of individual data revealed that vocabulary scores were at least 1 SD below the norm for 23 children (88%), and at least 2 SDs below the norm for 9 children (35%). In contrast, concerning basic literacy skills, the children performed within the average range for word and pseudoword decoding, and only slightly (i.e., less than one half of an SD) below the norm for spelling and for comprehension of words read silently. This pattern of results indicates that acquisition of basic reading skills in L2 does not depend on L2 vocabulary knowledge. In line with this argument, vocabulary scores at pretest were uncorrelated to reading, spelling, and phonological awareness (all ps > 0.24). These findings replicate prior results showing average or near-to-average decoding skills despite deficient vocabulary knowledge in children instructed in L2 (Babayigit 2015; Knell 2018; Limbird et al. 2014; Melby-Lervåg and Lervåg 2014; Prevoo et al. 2016).

Concerning instruction, the results of the current study add further evidence against the argument that teaching basic literacy skills in L2 learners should be postponed until a basic vocabulary is established (Goldenberg 2011; Knell 2018).

Analyses of the training effects revealed significant and marginally significant (p = 0.053) effects of strong effect size in two of three measures of phonological awareness, i.e., deletion and substitution of speech sounds in words. However, these effects were no longer significant at follow-up.

Sound deletion and substitution are not explicitly trained in Lautarium. The latter focusses on blending of phonemes into words, and segmentation of words into the constituent phonemes. These abilities are a precondition for reading and spelling, respectively (Schuele and Boudreau 2008). Deletion and substitution are more complex tasks, as they require not only segmentation, but also manipulation of sounds in words (Pufpaff 2009; Schuele and Boudreau 2008).

Inspection of Tables 2 and 3 indicates that the children from the training group made strong gains during the training period (Table 2), but no further progress after termination of the training, i.e., between posttest and follow-up (Table 3). The controls, in contrast, improved continuously between pretest and follow-up, where they caught up with the training group. We may thus conclude that phonological awareness training in Lautarium accelerates the acquisition of complex, non-trained phonological awareness skills. Unfortunately, training gains on segmentation and blending could not be assessed, as respective tasks were not included in the phonological awareness test. However, we may assume that the enduring effects on spelling described below are accompanied by significant gains in segmentation.

Lautarium-training evoked significant and enduring effects of medium effect size on spelling accuracy (number of words and graphemes correctly spelled; correct application of letter-sound-mappings). Thus, the beneficial effects of Lautarium-training on spelling found in monolingual German children (Klatte et al. 2016, 2017a, b) and in classes of second-graders with and without a migration background (Klatte et al. 2018) were replicated in a sample of children instructed in German as L2. This finding is in line with prior studies demonstrating beneficial effects of phonic-based literacy instruction in both L1 and L2 learners (Ginns et al. 2019; Goldenberg 2011; Ludwig et al. 2019).

In addition to overall spelling accuracy, the spelling test used in the current study allowed analyses of orthographic skills (“orthographic strategy”, i.e., correct application of spelling rules and morpheme knowledge when spelling non-transparent words). Even though Lautarium-training does not include rule- or morpheme-based spelling strategies, a significant beneficial effect on these skills was found at follow-up. Different mechanisms might contribute to this transfer effect. First, it might be attributed to the intensive training of vowel-length perception. As described in the introduction, vowel length is phonemic in German, and orthographically marked. In order to understand and successfully apply the respective orthographic rules, children need to be able to classify a vowel as short or long. Therefore, effective training of vowel-length perception eliminates a major barrier of spelling rule acquisition. Second, Lautarium-
training includes rapid word recognition (word-to-picture matching) of non-transparent words. This might foster the acquisition of orthographic representations of these words, and also of novel words that share orthographical or morphological features (Pacton et al. 2018; Tucker et al. 2016). Third, the significant improvement in spelling transparent words (“alphabetic strategy”) might enhance the “self-teaching mechanism” for spellings of novel, non-transparent words. It has been shown that, in addition to phonological recoding (Share, 1995), spelling acts as a powerful self-teaching tool in orthographic learning (Conrad et al. 2019; Shahar-Yames and Share 2008).

Concerning reading, no effects of Lautarium-training were found. A possible explanation for this is that, for transparent orthographies such as German, phonological skills are more closely related to spelling when compared to reading. Respective studies in German children revealed that the variance explained by phonological awareness declines rapidly after the first school year for reading, but remains reliable in later grades for spelling (Ennemoser et al. 2012; Landerl and Wimmer 2008; Wimmer and Mayringer 2002). This is because grapheme-phoneme-relations are highly regular, whereas phoneme-grapheme-relationships are not (Wimmer and Mayringer 2002). Due to their regularity, letter-sound mappings are easily learned, and word reading accuracy reaches ceiling at the end of the first school year (Seymour et al. 2003). Therefore, learning to read in German requires only basic phonological awareness skills. As phonological awareness training is a key aspect in Lautarium, and the current study was performed in second-graders, stronger effects on spelling when compared to reading were to be expected. Furthermore, studies with children instructed in German language indicate that the predictive strength of phonological awareness for reading is still lower in L2 than in L1 learners (Duzy et al. 2013a, b).

The most important finding from the current study is that computerized grapho-phonological training evokes enduring beneficial effects on vocabulary acquisition in L2 learners. This finding confirms our assumption that phonics-based instruction in L2 learners effectively fosters both written and oral language skills.

We may attribute the significant gains in vocabulary to the integration of auditory-phonological and orthographic training. As outlined in the introduction, phonological awareness skills predict vocabulary acquisition in L1 and L2 learners (Kalia et al. 2018; Marecka et al. 2018), and phonological awareness training fosters word learning in children (de Jong et al. 2000; Melby-Lervåg and Hulme 2010). Concerning L2 learners, phonological training boosts the acquisition of long-term phonological knowledge of the L2, e.g., sensitivity for the phoneme structure of words, and for phonemic features and phonotactic regularities that might be non-existent in L1. The increasing knowledge on the phonological characteristics of the L2 acts as a scaffold when phonological representations of novel words are built up. Especially, when encountering a novel word, existing phonological representations are activated and used to create a representation of the new word in short-term memory. This representation can later be encoded in long-term memory. The efficiency of this process depends on the range and quality of phonological representations already available to the learner (Marecka et al. 2018). Studies confirmed that pre-existing phonological knowledge supports vocabulary acquisition in L2 learners (Majerus et al. 2008; Masoura and Gathercole 2005).

Acquisition of phonological representations is further supported by practicing grapheme-phoneme-mappings and reading and spelling of words. The beneficial effect of print exposure on novel word learning (orthographic facilitation) described in the introduction has proved reliable in a recent systematic review of 23 studies (Colenbrander et al. 2019). In alphabetic systems, orthographic word forms are well specified and systematically related to pronunciations. Therefore, orthographic representations help learners to specify and clarify phonological forms, leading to distinctive and robust phonological representations. Recent evidence suggests that the orthographic facilitation effect is still more pronounced in L2 when compared to L1 learners (Qu et al. 2018). A possible explanation is that, in L2 learners, language representations are often acquired through spoken and written language input in parallel. Therefore, the interlinks between phonological and orthographic representations are presumably stronger when compared to L1 learners. In Lautarium, word meaning is provided by pictures accompanying the target words. The program thus enables the learner to establish high-quality lexical representations consisting of well-integrated...
Despite the encouraging findings, the current study has a number of limitations. First, the sample size is small, and—for reasons of data protection regularities—little is known concerning the language backgrounds of the participants (e.g., characteristics of native language and native language proficiency, main language spoken at home, age at onset of L2 acquisition). Analyses of potential moderating effects of the heritage language was therefore impossible. Second, allocation of participants to treatment conditions (training vs. control) was not randomized, but based on school classes. Even though the groups proved highly comparable at pretest, we cannot rule out differences in regular class instruction during the study period. Third, due to the complexity of the training program, the effectiveness of specific components cannot be evaluated. However, the literature reviewed above lets us assume that the success of the program results from its complexity, as this allows integration of phonological, orthographic, and semantic constituents (Perfetti and Hart 2002).

Despite the encouraging findings, the current study has a number of limitations. First, the sample size is small, and—for reasons of data protection regularities—little is known concerning the language backgrounds of the participants (e.g., characteristics of native language and native language proficiency, main language spoken at home, age at onset of L2 acquisition). Analyses of potential moderating effects of the heritage language was therefore impossible. Second, allocation of participants to treatment conditions (training vs. control) was not randomized, but based on school classes. Even though the groups proved highly comparable at pretest, we cannot rule out differences in regular class instruction during the study period. Third, due to the complexity of the training program, the effectiveness of specific components cannot be evaluated. However, the literature reviewed above lets us assume that the success of the program results from its complexity, as this allows integration of phonological, orthographic, and semantic constituents. Finally, we do not know whether and in what extent the current results can be generalized to other forms of phonics-based instruction. It might be possible that the effects result from the specific training format (computerized training), schedule (intensive, daily training for a period of 9 weeks), and/or specific features of materials and implementation. For example, in Lautarium, each speech item is represented by 4 high-quality recordings produced by 2 professional speakers (a male and a female). This variability might contribute to the beneficial training effect on L2 learners’ vocabulary (Sommers and Barcroft 2011).

In line with prior studies, the current results suggest that teaching basic literacy skills (i.e., phonological awareness, letter-sound mappings, and word-level reading and spelling) is effective in immigrant children with deficient vocabulary knowledge in the language of instruction. Furthermore, becoming literate in L2 enhances L2 vocabulary acquisition. Consequently, in courses targeting vocabulary knowledge, teachers should include written words in addition to phonology and semantics as soon as basic reading skills are established.

**Acknowledgements** Many thanks are due to the children who participated in this study and to their teachers for permanent support, as well as to the student assistants for their help in organization and data acquisition. This project was supported by the German Federal Ministry of Educational Research (BMBF 01GJ1003/01 GJ 1401).

| Table 4 Multivariate analysis for the outcome variables in the domains phonological awareness, reading, spelling with the factors group and time measurement |
|---------------------------------------------------------------|
| Effect | Group (2: Training group and Control group) × Time (3: Pretest, Posttest, Follow-up) MANOVA | Group (2: Training group and Control group) × Time (2: Pretest, Posttest) MANOVA | Group (2: Training group and Control group) × Time (2: Pretest, Follow-up) MANOVA |
| df | F | p | η² | df | F | p | η² | df | F | p | η² |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Phonological awareness | | | | | | | | | | | | |
| Time | 6, 18 | 16.82 | < 0.001 | 0.849 | 3, 22 | 19.94 | < 0.001 | 0.340 | 3, 21 | 24.47 | < 0.001 | 0.778 |
| Group | 3, 21 | 0.39 | 0.765 | 0.052 | 3, 22 | 0.36 | 0.800 | 0.044 | 3, 21 | 1.11 | 0.366 | 0.137 |
| Time × Group | 6, 18 | 2.11 | 0.103 | 0.413 | 3, 22 | 3.77 | 0.025 | 0.340 | 3, 21 | 0.02 | 0.995 | 0.003 |
| Reading | | | | | | | | | | | | |
| Time | 6, 18 | 40.29 | < 0.001 | 0.931 | 3, 22 | 41.24 | < 0.001 | 0.849 | 3, 21 | 77.44 | < 0.001 | 0.917 |
| Group | 3, 21 | 1.15 | 0.354 | 0.141 | 3, 22 | 0.97 | 0.425 | 0.117 | 3, 21 | 1.11 | 0.366 | 0.137 |
| Time × Group | 6, 18 | 0.61 | 0.718 | 0.170 | 3, 22 | 0.53 | 0.668 | 0.067 | 3, 21 | 0.43 | 0.737 | 0.057 |
| Spelling | | | | | | | | | | | | |
| Time | 8, 16 | 38.94 | < 0.001 | 0.951 | 4, 21 | 12.21 | < 0.001 | 0.699 | 4, 21 | 51.23 | < 0.001 | 0.911 |
| Group | 4, 20 | 1.01 | 0.426 | 0.168 | 4, 21 | 0.86 | 0.506 | 0.140 | 4, 20 | 0.63 | 0.648 | 0.112 |
| Time × Group | 8, 16 | 3.17 | 0.025 | 0.611 | 4, 21 | 2.82 | 0.051 | 0.349 | 4, 20 | 4.31 | 0.011 | 0.463 |

semantic, phonological, and orthographic constituents (Perfetti and Hart 2002).
Authors contributions All authors contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript. Data collection was organized by Marita Konerding and Kirstin Bergström.

Funding Open Access funding provided by Projekt DEAL. This project was supported by the German Federal Ministry of Educational Research (BMBF 01GJ1003/01 GJ 1401).

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Availability of data and material Archiving is not mandated but data will be made available on reasonable request.

Ethics approval This study was performed in line with the principles of the Declaration of Helsinki. Regarding ethical standards, the study was approved by the local State education authority of the State of Bavaria and by the responsible data protection officer (Neu-Ulm, Bavaria, Germany).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix

See Table 4.

References

Autorenguppe Bildungsberichterstattung. (2018). Bildung in Deutschland 2018: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration. wbv. https://doi.org/10.3278/6001820fw.

Babayigit, S. (2015). The relations between word reading, oral language, and reading comprehension in children who speak English as a first (L1) and second language (L2): A multigroup structural analysis. Reading and Writing, 28(4), 527–544. https://doi.org/10.1007/s11145-014-9536-x.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. Psychological Review, 105(1), 158–173. https://doi.org/10.1037/0033-295X.105.1.158.

Blatter, K., Faust, V., Jäger, D., Schöppe, D., Artelt, C., Schneider, W., et al. (2013). Vorschulische Förderung der phonologischen Bewusstheit und der Buchstaben-Laut-Zuordnung: Profitieren auch Kinder nichtdeutscher Herkunftssprache? In A. Redder & S. Weinert (Eds.), Sprachförderung und Sprachdiagnostik: Interdisziplinäre Perspektiven (pp. 218–238). Münster: Waxmann.

Boets, B., Vandermosten, M., Poelmans, H., Luts, H., Wouters, J., & Ghesquière, P. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. Research in Developmental Disabilities, 32(2), 560–570. https://doi.org/10.1016/j.ridd.2010.12.020.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read—A causal connection. Nature, 301(5899), 419–421. https://doi.org/10.1038/301419a0.

Brunner, M. (2012). Heidelberger Vokallängendifferenzierungstest (HVT). Winswangen: Westra.

Burnham, D. (2003). Language specific speech perception and the onset of reading. Reading and Writing, 16(6), 573–609. https://doi.org/10.1023/A:1025593911070.

Camorota, S. A., & Zeigler, K. (2015). One in five US residents speaks foreign language at home, Record 63.2 million. Center for Immigration Studies. October. https://cis.org/sites/default/files/camarotalam]>=language-15.pdf. Accessed 27 July 2020.

Chambré, S. J., Ehri, L. C., & Ness, M. (2020). Phonological decoding enhances orthographic facilitation of vocabulary learning in first graders. Reading and Writing, 33(5), 1133–1162. https://doi.org/10.1007/s11145-019-09997-w.

Colenbrander, D., Miles, K. P., & Rickets, J. (2019). To see or not to see: How does seeing spellings support vocabulary learning? Language, Speech, and Hearing Services in Schools, 50(4), 609–628. https://doi.org/10.1044/2019-LSHSS-VOIA-18-0135.

Conrad, N. J., Kennedy, K., Saoud, W., Scallion, L., & Hanusiak, L. (2019). Establishing word representations through reading and spelling: Comparing degree of orthographic learning. Journal of Research in Reading, 42(1), 162–177. https://doi.org/10.1111/1467-9817.12256.

Czapka, S., Klassert, A., & Festman, J. (2019). Executive functions and language: Their differential influence on mono- vs. Multilingual spelling in primary school. Frontiers in Psychology, 10, 97. https://doi.org/10.3389/fpsyg.2019.00097.

Darcy, I., & Krüger, F. (2012). Vowel perception and production in Turkish children acquiring L2 German. Journal of Phonetics, 40(4), 568–581. https://doi.org/10.1016/j.jphon.2012.05.001.

de Jong, P. F., Seveke, M. J., & van Veen, M. (2000). Phonetic sensitivity and the acquisition of new words in children. Journal of Experimental Child Psychology, 76(4), 275–301. https://doi.org/10.1006/jecp.1999.2549.

Demoulin, C., & Kolinsky, R. (2016). Does learning to read shape verbal working memory? Psychonomic Bulletin and Review, 23(3), 703–722. https://doi.org/10.3758/s13423-015-0956-7.

Duzy, D., Ehm, J.-H., Souvignier, E., Schneider, W., & Gold, A. (2013a). Prädiktoren der Lesekompetenz bei Kindern mit

C Springer
Klatte, M., Steinbrinck, C., Bergström, K., & Lachmann, T. (2016). Lautarium: Ein computerbasiertes Trainingsprogramm für Grundschulkinder mit Lese-Rechtschreibschwärmigkeiten. Göttingen: Hogrefe.

Klatte, M., Steinbrinck, C., Pröß, A., Estner, B., Christmann, C. A., & Lachmann, T. (2014). Effekte des computerbasierten Trainingsprogramms “Lautarium” auf die phonologische Verarbeitung und die Lese-Rechtschreibleistungen bei Grundschulkindern. In G. Schulte-Körne (Ed.), Legaesthete und Dyskalukie (pp. 127–144). Stuttgart: Winkler.

Knell, E. (2018). Role of oral language in the development of L2 literacy skills. In J. I. Liontas & M. DelliCarpini (Eds.), The TESOL Encyclopedia of English language teaching (Vol. 25, pp. 1–8). Oxford: Wiley. https://doi.org/10.1002/9781118784235.eelt0238.

Koschollek, C., Bartig, S., Rommel, A., Santos-Hövene, C., & Lampert, T. (2019). Die Gesundheit von Kindern und Jugendlichen mit Migrationshintergrund in Deutschland—Querschnittsergebnisse aus KiGGS Welle 2. https://doi.org/10.25646/6070.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. Nature Reviews Neuroscience, 5(11), 831–843. https://doi.org/10.1038/nrn1533.

Landerl, K. (2003). Categorization of vowel length in German poor spellers: An orthographically relevant phonological distinction. Applied Psycholinguistics, 24(4), 523–538. https://doi.org/10.1023/A:10142716403000262.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. Journal of Educational Psychology, 100(1), 150–161. https://doi.org/10.1037/0022-0663.100.1.150.

Lenhard, W., & Schneider, W. (2006). Ein Leseverständnistest für Erst- bis Sechsstklässler: (ELFE 1–6). Deutsche Schultests: Hogrefe.

Limbird, C. K., Maluch, J. T., Rjosk, C., Stanat, P., & Merkens, H. (2014). Differential growth patterns in emerging reading skills of Turkish-German bilingual and German monolingual primary school students. Reading and Writing, 27(5), 945–968. https://doi.org/10.1007/s11145-013-9477-9.

Limbird, C. K., & Stanat, P. (2006). Prädiktoren von Leseverständnis bei Kindern deutscher und türkischer Herkunftssprache: Ergebnisse einer Längsschnittstudie. In A. Ittel & H. Merkens (Eds.), Veränderungs- und Längsschnittstudien in der Empirischen Erziehungswissenschaft (pp. 93–123). Berlin: VS Verlag für Sozialwissenschaften.

Ludwig, C., Guo, K., & Georgiou, G. K. (2019). Are reading interventions for English language learners effective? A meta-analysis. Journal of Learning Disabilities, 52(3), 220–231. https://doi.org/10.1177/0022219419825855.

Majerus, S., Poncelet, M., van der Linden, M., & Weekes, B. S. (2008). Lexical learning in bilingual adults: The relative importance of short-term memory for serial order and phonological knowledge. Cognition, 107(2), 395–419. https://doi.org/10.1016/j.cognition.2007.10.003.

Manis, F. R., Lindsey, K. A., & Bailey, C. E. (2004). Development of reading in grades K–2 in Spanish-speaking English-language learners. Learning Disabilities Research and Practice, 19(4), 214–224. https://doi.org/10.1111/j.1540-5826.2004.00107.x.

May, P. (2012a). Hamburger Schreib-Prob 1 + (HSP 1 +): Hinweise zur Durchführung und Auswertung. Klett: VP Medien.

May, P. (2012b). Hamburger Schreib-Prob 1-10 (HSP 1-10): Manual/Handbuch. Klett: VP Medien.

May, P. (2012c). Hamburger Schreib-Prob 2 (HSP 2): Hinweise zur Durchführung und Auswertung. Klett: VP Medien.

McArthur, G., Eve, P. M., Jones, K., Banales, E., Kohnen, S., & Doi, L. M., Munson, B., et al. (1997). Are speech perception deficits associated with development dyslexia? Journal of Experimental Child Psychology, 66(2), 211–235. https://doi.org/10.1006/jecp.1997.2383.

Marecka, M., Szewczyk, J., Jelec, A., Janiszewska, D., Rajat, K., & Dziubalska-Kolaczky, K. (2018). Different phonological mechanisms facilitate vocabulary learning at early and late stages of language acquisition: Evidence from Polish 9-year-olds learning English. Applied Psycholinguistics, 39(1), 1–35. https://doi.org/10.1017/S0142716417000455.

Masoura, E. V., & Gathercole, S. E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. Memory, 13(3–4), 422–429. https://doi.org/10.1080/09658210344000323.

Matos, R. (2003). Kaiserslauterner Gruppentests: Hogrefe.

McArthur, G., Eve, P. M., Jones, K., Banales, E., Kohnen, S., & Doi, L. M., Munson, B., et al. (1997). Are speech perception deficits associated with development dyslexia? Journal of Experimental Child Psychology, 66(2), 211–235. https://doi.org/10.1006/jecp.1997.2383.

McArthur, G., Sheehan, Y., Badcock, N. A., Francis, D. A., Wang, H. C., Kohnen, S., et al. (2018). Phonics training for English-speaking poor readers. The Cochrane Database of Systematic Reviews, 12, CD009115. https://doi.org/10.1002/14651858.CD009115.pub2.

McArthur, G., Sheehan, Y., Badcock, N. A., Francis, D. A., Wang, H. C., Kohnen, S., et al. (2018). Phonics training for English-speaking poor readers. The Cochrane Database of Systematic Reviews, 12, CD009115. https://doi.org/10.1002/14651858.CD009115.pub3.

McBride-Chang, C. (1995). Phonological processing, speech perception, and reading disability: An integrative review. Educational Psychologist, 30(3), 109–121. https://doi.org/10.1207/s15326985ep3003_2.
McCarthy, K. M., Mahon, M., Rosen, S., & Evans, B. G. (2014). Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Development, 85*(5), 1965–1980. https://doi.org/10.1111/cdev.12275.

Melby-Lervåg, M., & Hulme, C. (2010). Serial and free recall in children can be improved by training: Evidence for the importance of phonological and semantic representations in immediate memory tasks. *Psychological Science, 21*(11), 1694–1700. https://doi.org/10.1177/0956797610385555.

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin, 140*(2), 409–433. https://doi.org/10.1037/a0033890.

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*(2), 322–352. https://doi.org/10.1037/a0026744.

Moll, K., & Landerl, K. (2014). *Lese- und Rechtschreibtest: SLRT-II: Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT)*. Göttingen: Huber Hogrefe.

Muneaux, M., & Ziegler, J. (2004). Locus of orthographic relatedness independently of orthographic relatedness when they learn to spell new words. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30*(3), 467–486. https://doi.org/10.1037/0269-9505.30.3.467.

Noordenbos, M. W., & Serniclaes, W. (2015). The categorical perception deficit in dyslexia: A meta-analysis. *Scientific Studies of Reading, 19*(5), 340–359. https://doi.org/10.1080/10888638.2015.1052455.

O’Connell, N. S., Dai, L., Jiang, Y., Speiser, J. L., Ward, R., Wei, W., et al. (2017). Methods for analysis of pre-post data in clinical research: A comparison of five common methods. *Journal of Biometrics & Biostatistics, 8*(1), 1–8. https://doi.org/10.4172/2155-6180.1000334.

Pacton, S., Afonso Jaco, A., Nys, M., Foulin, J. N., Treiman, R., & Peereman, R. (2018). Children benefit from morphological relatedness independently of orthographic relatedness when they learn to spell new words. *Journal of Experimental Child Psychology, 171*, 71–83. https://doi.org/10.1016/j.jjecp.2018.02.003.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Studies in written language and literacy* (Vol. 11, pp. 189–213). Precursors of functional literacy Amsterdam: Benjamins.

Petermann, F., Fröhlich, L. P., & Metz, D. (2010). *SET 5–10: Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren*. Göttingen: Hogrefe.

Pfoist, M. (2015). Children’s phonological awareness as a predictor of reading and spelling. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie, 47*(3), 123–138. https://doi.org/10.1026/0049-8637/a000141.

Prevo, M. J. L., Malda, M., Mesman, J., & van Ijzendoorn, M. H. (2016). Within- and cross-language relations between oral language proficiency and school outcomes in bilingual children with an immigrant background. *Review of Educational Research, 86*(1), 237–276. https://doi.org/10.3102/0034654315584685.

Prössoldt, M. V., Michálek, A., Schneider, W., Duzy, D., Glück, D., Souvignet, E., et al. (2013). Effekte kombinierter Förderprogramme zur phonologischen Bewusstheit und zum Sprachverstehen auf die Entwicklung der phonologischen Bewusstheit von Kindergartenkindern mit und ohne Migrationshintergrund. *Friede Bildung, 2*(3), 122–132. https://doi.org/10.1024/2191-9186/a000099.

Pufpaff, L. A. (2009). A developmental continuum of phonological sensitivity skills. *Psychology in the Schools, 46*(7), 679–691. https://doi.org/10.1002/pits.20407.

Qu, Q., Cui, Z., & Damian, M. F. (2018). Orthographic effects in second-language spoken-word recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 44*(8), 1325–1332. https://doi.org/10.1037/xlm0000520.

Quiroga, T., Lemos-Britton, Z., Mostafapour, E., Abbott, R. D., & Berninger, V. W. (2002). Phonological awareness and beginning reading in Spanish-speaking ESL first graders. *Journal of School Psychology, 40*(1), 85–111. https://doi.org/10.1016/S0022-0663(01)00095-4.

Rausch, J. R., Maxwell, S. E., & Kelley, K. (2003). Analytic methods for questions pertaining to a randomized pretest, posttest, follow-up design. *Journal of Clinical Child and Adolescent Psychology, 32*(3), 467–486. https://doi.org/10.1207/S15374424JCCP3203_15.

Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The effectiveness of reading interventions for English learners. *Exceptional Children, 82*(2), 144–169. https://doi.org/10.1177/0014429115585483.

Rickets, J., Bishop, D. V. M., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Quarterly Journal of Experimental Psychology, 62*(10), 1948–1966. https://doi.org/10.1080/17470210802696104.

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics, 27*(3), 465–485. https://doi.org/10.1017/S01424606060036X.

Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology, 100*(1), 175–191. https://doi.org/10.1037/0022-0663.100.1.175.

Schaars, M. M. H., Segers, E., & Verhoeven, L. (2019). Cognitive and linguistic precursors of early first and second language reading development. *Learning and Individual Differences, 72*, 1–14. https://doi.org/10.1016/j.lindif.2019.03.008.

Schöppe, D., Blatter, K., Faust, V., Jäger, D., Stanat, P., Artelt, C., et al. (2013). Effekte eines vorschulischen Trainings der phonologischen Bewusstheit bei Vorschulkindern mit unterschiedlichem Sprachhintergrund. *Zeitschrift Für
