Effects of a Whole-Class Reading Program Designed for Different Reading Levels and the Learning Needs of L1 and L2 Children

Susanne Seifert^{1*}, Susanne Schwab¹ & Barbara Gasteiger-Klicpera¹

¹*Karl-Franzens-University of Graz, Merangasse 70/2, 8010 Graz - Austria*

* corresponding author

E-Mail addresses:

susanne.seifert@uni-graz.at, susanne.schwab@uni-graz.at, barbara.gasteiger@uni-graz.at

DOI: 10.1080/10573569.2015.1029176

Running Head: *Differentiating Reading Program*

Abstract

This article aims to evaluate a whole-class reading program (LARS) and its effects on decoding, reading comprehension, and language abilities of L1 and L2 German second graders. The program consisted of teacher training and differentiated reading materials offered adapted to the needs of both L1 and L2 learners. The materials were divided into four different reading levels. For the evaluation, a pre-post-test design with a comparison group was used. Decoding, reading comprehension, and language abilities of all students (n=377) were assessed before and after the implementation of the program. 159 students received reading instruction twice a week with the LARS program during regular school hours for a period of seven months. 218 students received conventional reading instruction without special materials (e.g. no differentiated materials) provided by the research team. The results indicate that the development of reading comprehension and expressive vocabulary (trained words) was significantly improved by the program.

Keywords: evaluation, reading instruction, reading comprehension, German as a second language, differentiation

Introduction

Reading Abilities

Proficiency in reading remains one of the main aims of schooling as reading abilities act as the foundation for the whole education process (Education, Audiovisual and Culture Executive Agency, 2011). International studies such as PIRLS (Progress International Reading Literacy Study), however, revealed that in German-speaking countries many children show an alarmingly poor performance in reading competence at the end of primary school (fourth grade). About 20% of the fourth graders in Austria (Vogtenhuber et al., 2012), for example, and 15% of those in Germany (Bos, Bremerich-Vos, Tarelli, & Valtin, 2012) lack the ability to pick simple information out of written texts. Thus, those children are far away from reaching the goal of reading proficiency. For children with German as a second language (L2) the risk for reading deficits is even higher than for first language (L1) German children, due to the L2 children's insufficient language proficiency in the German language (Schabmann, Landerl, Bruneforth, & Schmidt, 2012).

Proficient reading means that an individual can successfully manage the demands of different kinds of text- and reading-related tasks (Artelt et al., 2007). These demands include above all the comprehension of written texts or tasks. The foundation for becoming a proficient reader lies in learning to decode within the first school years (Perfetti, Landi, & Oakhill, 2005). Decoding involves learning initially to identify letter-sound-correspondences and more and more the efficient decoding of whole words (lexical reading). During this learning process, a continually growing number of words are stored in the lexicon and can thus be retrieved increasingly automatically (Ehri, 1999; Klicpera, Schabmann, & Gasteiger-Klicpera, 2013). This automatic retrieval of written words from the lexicon represents a prerequisite for efficient reading comprehension (Klicpera et al., 2013). In a consequence, decoding and reading comprehension are closely related to each other. The better a child is able to decode, the more processing resources are available for reading comprehension (Perfetti et al., 2005).

The Relationship Between Reading and Language Abilities

Once the alarming results in reading proficiency of fourth graders were published, an intensive process of reflection in search of specific reasons for these findings was initiated by the German-speaking countries, for instance. One possible reason for these poor results is the connection to deficits in language development. Internationally, scholars, who focused on children with impairments, have proven the interrelation of reading development and language abilities. These studies examined, on the one hand, children with pre-school language impairments and revealed their high risk of developing reading problems in school (e.g. Bishop & Snowling, 2004; Catts, Fey, Tomblin, & Zhang, 2002). On the other hand, these studies also looked at children with reading impairments and showed that these children quite often have deficits in vocabulary and grammar (Catts, Adlof, & Weismer, 2006; Nation & Snowling, 1998; Snowling, Gallagher, & Frith, 2003). Besides research on impaired children, longitudinal studies on reading development represent further evidence for the interrelation of reading and language skills. Lee (2011), amongst others, illustrated that early vocabulary skills can serve as a predictor for later reading comprehension development up to 5th grade. Moreover, intervention studies also revealed the relevance of language abilities for reading skills. In such studies, vocabulary skills, for instance, had an impact on the success of the reading interventions and programs (e.g. Duff et al., 2008; Vadasy, Sanders, & Abbott, 2008; Whiteley, Smith, & Connors, 2007).

The comparison of L1 and L2 learners' reading development seem to be an additional proof for the interrelation of reading and language abilities. As Verhoeven (2011) summarized, L2 learners when learning to read in their L2 often show more problems in decoding and reading comprehension than L1 learners. Quite often, reading comprehension is especially impaired in L2 learners (Melby-Lervag & Lervag, 2013; Verhoeven & van Leeuwe, 2012). Furthermore, particularly for L2 learners, scholars observed deficient vocabulary skills as predictors for deficits in reading comprehension (Hutchinson, Whiteley, Smith, & Connors, 2003).

Most of the reported evidence for the interrelation of reading and language abilities comes from Anglo-American studies and are not directly transferable to German-speaking countries and schools. In contrast to the English language, the German language consists of a phonologically transparent orthography (Seymour, Aro, & Erskine, 2003). Therefore, decoding skills are less

influenced by phonological awareness in German reading children than in English readers (Wimmer, Mayringer, & Landerl, 2000). However, some German studies also analyzed the relationship between reading and language abilities. For instance, the vocabulary and grammar skills in German pre-school children were proven to influence their later reading comprehension (Ennemoser, Marx, Weber, & Schneider, 2012).

The Challenge of Reading Instruction in Classes with L1 and L2 Children

The way of reading instruction plays an important role for the development of reading skills and the prevention of reading problems of certain children (Schabmann, Schmidt, Klicpera, Gasteiger-Klicpera, & Klingebiel, 2009; Snow & Juel, 2005). The challenge that teachers face in reading classes is to meet the needs of all learners since individual reading levels vary. Some students might still struggle with certain letter-sound-correspondences, while others already manage to read short texts or even books fluently and independently. Such heterogeneous reading abilities are even more challenging for teachers when many students in class are struggling with the language of instruction as L2 learners often do. L2 learners are a growing population at elementary schools worldwide. In the U.S., for example, about 9% of all children enrolled in public schools speak English as L2 (Kena, Aud, Johnston et al., 2014). In some states, the percentage is more than twice as big (e.g. California where English Language Learners constitute approximately 23% of all children enrolled in public schools, Kena et al., 2014). The situation in German-speaking elementary schools is quite similar. In Austria, for instance, one out of four children at the end of fourth grade speaks German (the language of instruction in school) as L2 (Statistik Austria, 2014). As mentioned before, L2 learners often show problems in learning to read in the L2 which is often associated with deficits in their L2 language abilities (e.g. Esser, 2006).

Hence, teachers must consider these children's learning needs during reading instruction as well as those of the L1 learners. The individual needs of all children in heterogeneous reading classes should be addressed through individualized support taking the different ability levels into consideration (Schabmann et al., 2012; Woodward & Talbert-Johnson, 2009). Differentiation represents one of the ways to meet the needs of L2 learners, but may also benefit L1 learners of

different ability levels (Baecher, Artiglieri, Patterson, & Spatzer, 2012). Research has shown that undifferentiated instruction leads to moderate gains only in high achieving students, while average and low achieving students lag behind (McGill-Franzen, Zmach, Solic, & Zeig, 2006; Schumm, Moody, & Vaughn, 2000). Therefore, especially for L2 learners who often happen to be average or rather low achieving students in reading, differentiation of instruction is indispensable.

The American RAND Reading Study Group (2002) highlighted the importance of providing students with practice and reading materials appropriate for their individual level as one of the most effective teacher behaviors. Thus, differentiation in the U.S. is widely achieved through individualized reading lessons where children read different books or magazines that fit their reading level. The challenge in classrooms with this kind of differentiation is that students do not work on the same content. Therefore, post-reading communication in either group or classroom settings is difficult. Expanded communication settings, however, ought to be used to enrich children's vocabulary knowledge, facilitate, and deepen their comprehension of the texts and topics (Artelt et al., 2007). This proves especially important for L2 learners. Within the classroom, when working with the entire class, teachers can realize differentiation, for instance, by offering differentiated materials or instructions (Ankrum & Bean, 2007; Baecher et al., 2012). The benefit of differentiating the same reading text into different ability levels is that all students are able to work on the same topic, though with differing degrees of difficulty. Thus, post-reading exchanges and discussions are possible within the whole class independently of the heterogeneity in reading levels. Differentiated reading texts can be of use not only in reading classes but also in content area classes, e.g. social science classes. Either way, working on the very same topic with all students benefits overall classroom communication.

In schools in German-speaking countries, conventional reading materials provided for reading instruction are not differentiated. Teachers have the choice to either 1) let all students work on the same texts independent of their ability level and adapt the reading time if necessary or 2) let the students read different texts designed for different ability levels. Neither of these possibilities allows the joint learning on the same content and post-reading discussions about the topic. On the one hand, if teachers decide for the first option, students are ready at different time points. On the other hand, if

teachers decide for the second option, students have read totally different texts and are not able to talk in the whole class about the contents.

Evidence for Effective Reading Instruction

The educational system in German-speaking countries has so far not been very successful in providing equal opportunities for children with different learning preconditions as well as children with reading difficulties, as the results in the PIRLS study demonstrate. Therefore, in the last decade, scholars have published numerous recommendations for reading instruction in German-speaking countries (e.g. Artelt et al., 2007; Philipp, 2013). These publications are primarily meta-analyses based on research conducted in English speaking countries (e.g. Berkeley, Scruggs, & Mastropieri, 2009; Edmonds, Vaughn, Wexler et al., 2009; Slavin, Lake, Chambers, Cheung, & Davis, 2010) as only a small number of national studies tried to evaluate best practice models for reading instruction. These meta-analyses revealed that different kinds of interventions can be effective in educating children with reading problems. Yet the effectiveness of these interventions seems to be dependent on various variables such as the setting of the intervention (e.g. one-to-one tutoring, small group, parenting instruction), length and intensity of the intervention (e.g. length and number of remedial lessons, number of remedial lessons per week), and the group of children (e.g. children with difficulties in reading, children with general learning difficulties). Moreover, a meta-analysis conducted by Nelson, Benner, and Gonzalez (2003) reported that the treatment responsiveness of children with reading deficits to early reading interventions does not depend on the deficit level and on the ethnicity of the students, showing that both children with more and less severe deficits as well as both children with L1 and L2 can profit equally from interventions. In contrast, the review of Al Otaiba and Fuchs (2002) revealed that the proficiency in the language of instruction may influence the treatment responsiveness.

Although additional support is important for children with reading problems, teachers in the reading class should offer adequate in-class support to these students as well since individual support by specialized teachers is not an indispensable requirement (Rose, 2009). The key factor for supporting children, however, is the quality of teaching (see also Bell, McPhillips, & Doveston, 2011).

To find out which way of reading instruction is effective in supporting every child in the reading class, several studies attempted to evaluate reading programs that are implemented in the classroom. Above all, Hattie (2009) revealed in his meta-analyses of 50 studies based on reading research that vocabulary programs, repeated reading, phonics instruction, and reading comprehension programs are the most effective ones. The National Reading Panel (National Institute of Child Health and Human Development, 2000) came to quite similar results. In line with the report from the NRP, effective instruction for beginning reading needs to consider the areas of phonemic awareness, phonics, fluency, vocabulary, and comprehension. These five areas should be directly instructed to the children with respect to their individual learning needs (Rupley, Blair, & Nichols, 2009). According to a meta-analysis by Slavin et al. (2010), programs that helped to alter teacher-students interactions on a daily basis had a higher impact on students' reading performance than programs only focusing on curricula changes. Furthermore, the authors showed that programs implemented in earlier school years (preschool to first grade) tend to be more effective than programs implemented from second grade on. Another meta-analysis (Guthrie, McRae, & Klauda, 2007) evaluated the reading program CORI (Concept Oriented Reading Instruction, see Guthrie, Wigfield, & Perencevich, 2004). They showed that a multilayered program that combines different aspects relevant for reading (in this case motivation and reading strategies) can be very effective for the later elementary grades. In addition, a meta-analysis focusing on vocabulary programs (Elleman, Lindo, Morphy, & Compton, 2009) revealed moderate effects of vocabulary trainings on reading comprehension. Nevertheless, vocabulary instruction remains an indispensable component of reading instruction (e.g. National Institute of Child Health and Human Development, 2000), especially for struggling readers (Taylor, Mraz, Nichols, Rickelman, & Wood, 2009). But as Wanzek (2013) has shown, vocabulary instruction is often under-represented in the reading class and even more in supplemental reading interventions.

Concerning effective reading instruction for L2 learners, Goldenberg (2010) emphasizes that the knowledge of how reading instruction should be designed for L1 learners most probably also holds true for L2 learners. But due to their limited L2 proficiency, L2 learners need additional instructional support.

As previously mentioned, in Germany and Austria, there have been only a few systematic investigations of whole-class reading programs. Frequently, the programs or parts of them are adapted from English programs. But English programs must be cautiously adapted for German-speaking children. For German children other programs than the ones that proved to be effective for English readers might lead to learning gains. One of the adapted programs that proved to be effective for German children is the German *Leseprofi-Training* for third-graders (Munser-Kiefer, & Kirschhock, 2009). This program integrates aspects of the above-mentioned CORI program (Guthrie, Wigfield, & Perencevich, 2004) and the paraphrase-strategy RAP (Schumaker, Denton, & Deshler, 1984). The *Leseprofi-Training* is recommended to use only from third grade on (Munser-Kiefer, 2012). Fischer (2012) designed a different German program, the *Leseförderung im Anfangsunterricht*, for the first year of schooling. In particular, decoding abilities were fostered by a phonics-based approach including work on the letter-sound-correspondences, repeated reading of words that include the already learned letter-sound-correspondences, and phonological awareness training. Although this program focused on weak readers, the work in the whole class was possible through differentiated materials adapted to the different student's needs. However, the weak readers benefitted most from this program which was implemented twice a week for 10 weeks.

Currently, scholars have yet to evaluate German programs for the use in the whole class in second grade—when reading comprehension becomes even more critical. Furthermore, the evaluated reading programs in German-speaking countries have this far not focused on the needs of second language learners. We know of only one German program which considers the learning needs of L2 German children in particular and tries to foster their reading abilities in a whole-class program with L1 and L2 children (Förster, Niebuhr-Siebert, & Hellmich, 2012). However, at this point in time, an evaluation of the program remains unavailable. The lack of reading programs that are effective—both for children with L1 and L2 German—lead to the conceptualization, implementation, and piloting of a reading program which takes the needs of L2 learners into account. This program—developed by the authors of this article—which was named LARS (Language And Reading Skills), offers materials in different levels of difficulty and therefore allows an implementation into the daily schedule of classes with heterogeneous learning needs. The program that combines vocabulary work with intensive

reading is especially designed for enhancing reading comprehension, as this ability often is limited in L2 learners while decoding is relatively intact (Melby-Lervag & Lervag, 2013). The previous work of the present authors (Schwab & Gasteiger-Klicpera, 2014; Schwab, Seifert, & Gasteiger-Klicpera, 2014) evaluated the pilot study of this reading program. In this study, the reading materials were offered in three ability levels (above average level, average level, below average level). The effects on the reading and also the language skills of a small sample (55 students were supported using the LARS reading program, 50 students were in a comparison group receiving conventional instruction) were investigated. The LARS-group showed higher gains in decoding and comprehension than the comparison group but there were hardly any differentiating effects for sub-groups (e.g. students with different basic abilities or students with L2 or L1 German). Only on the level of word decoding was the program's beneficial to students with higher abilities. The shortcomings of this study lie in the implementation design. Because a research member implemented the program, no realistic schooling conditions were retained. In addition, the children with the lowest reading abilities were separated from the class and taught in small groups during the program (see Schwab, Seifert, & Gasteiger-Klicpera, 2013). In the light of the aims of inclusive class settings (Feyerer & Prammer, 2003), these children should be fostered in the whole class as well. The inclusion of the children with quite weak reading abilities in the LARS program required the offer of another differentiating level. Thus, the materials were adapted for four varying levels.

Purpose and Research Questions

This study explores the effects of the whole-class differentiating reading program LARS when teachers implement it twice a week during a whole school year under realistic schooling conditions. The purpose was to analyze if students who received the LARS program show a better development in decoding and reading comprehension compared to second graders who were taught the conventional way. As reading abilities are closely related to language skills, and the program is designed to also address language skills, the effects on vocabulary and grammar knowledge are examined in a second step.

This specific differentiating reading program contains components that aimed at supporting especially the group of L2 German learners. But the program should also be effective for children with L1 German. Therefore, in a third step, it needs to be examined if both groups of students can profit from the reading program.

The LARS program provides materials for four different reading levels. All students in a class should benefit from the program in the same way, since it was not designed to foster solely weak readers. However, in the differentiated reading program *Leseförderung im Anfangsunterricht* for first graders evaluated by Fischer (2012), especially weak readers profited most from the program. Therefore, in a fourth step, it has to be examined, if the children in each level of difficulty can profit equally from the reading program LARS.

Method

Procedure

The study was organized in a quasi-experimental pre-post-test-design with a comparison group. At the beginning of October 2012 (pretest, first measurement point, Time 1) 400 second graders from 21 classes of eight primary schools in Styria, a province in the Southeast of Austria, were tested. In the participating schools, 30% or more students with German as L2 were enrolled. From November 2012 to May 2013 nine classes from three schools received the LARS reading program (LARS group). For this group, the program was implemented into regular school lessons. The remaining twelve classes of the other five schools formed the comparison group that received conventional instruction (C-group). At the end of June 2013 (posttest, second measurement point, Time 2) reading and language achievement of all students was assessed again.

Sample

The study to be evaluated included 404 students (198 boys, 206 girls). On average, the students were seven years and eight months old ($SD=0.44$). For nearly half of the sample (52.5%), German was the L1. For the remaining 47.5%, German was the L2. The L2-group spoke 25 different first languages.

The allocation of schools (with two or three second grade classes) to either the LARS-group or the C-group was random. As the analyses are based on longitudinal data, only students who completed the assessments at both measurement points (pre- and posttest) were included in the final sample and therefore also in the subsequent analyses. Hence, the final sample consisted of 377 students (159 students in the LARS-group and 218 students in the comparison group). The students of the LARS-group (50.9% boys, 49.1% girls) were 7.84 years old ($SD=0.45$), 45.9% had L2 German. The comparison group (C-group) receiving conventional reading instruction consisted of students (45.4% boys, 54.6% girls) that were on average 7.74 years old ($SD=0.42$). 49.5% had L2 German. The age difference between the LARS- and C-group was significant ($t_{375}=-2.26, p<.05$). The gender distribution ($\chi^2_{1, 377}= 1.127$, n.s.) and also the distribution of L2 learners ($\chi^2_{1, 377}= 0.485$, n.s.) were similar in both groups.

To determine whether students who failed to complete the tests at both measurement points differ from those in the final sample, independent t-tests were calculated on some relevant variables (L2, reading abilities, and language abilities). The results showed no systematic bias in the final sample in any of these variables.

As far as the distribution (based on their initial reading abilities) of the students into the four different ability levels in reading is concerned, 28.9% of the students of the LARS-group belonged to the above average ability level (level 1), 44% to the average ability level (level 2), 13.8% to the below average ability level (level 3) and 13.2% to the lowest ability level (level 4). In the C-group, 30.3% of the students belonged to the above average ability level, 47.7% to the average ability level, 13.3% to the below average ability level and 8.7% to the lowest ability level. The Mann-Whitney-U-Test showed no differences in the ability levels of both groups ($U = 18215.5$, n.s.).

The study included 9 teachers that implemented the LARS program and 12 teachers that organized their reading lessons conventionally without any modifications. As typical for elementary school teachers, all of the teachers, both the ones in the LARS- and in the C-group, teach all subjects in the elementary school, along with “German language arts”. Within this subject, teachers instruct reading, spelling, and writing, while also teaching their students about grammar. Concerning the reading lessons, in the first grade teachers usually use traditional phonics instruction, teaching letter-sound

correspondences. They help the students to master the alphabetic principle and to understand underlying letter-sound correspondences. In the second grade, the reading abilities should be expanded step by step. Students should read almost independently and comprehend more and more different kinds of texts.

Interviews with the teachers that implemented LARS were carried out before the program was implemented. These LARS-teachers were on average 40.68 years old (minimum: 25 years; maximum: 59 years). On average, they started with their teaching practice 14.6 years ago (minimum: 1 month; maximum: 37 years). Most of the teachers have been teaching in the class in which they implemented the LARS program since the children entered school (six teachers). One teacher began to teach the class in the second half of the first grade. Two teachers just started to teach the respective classes at the beginning of second grade.

Description of the Differentiating Reading Program

The reading program was conducted twice a week for seven months. Every LARS-lesson took as long as a regular lesson (50 minutes). Thus, every child received on average 50 LARS-lessons.

Differentiating materials form one of the main pillars of the LARS reading program. As the materials are designed in four different reading ability levels, they allow joint learning of students with different ability levels on the same subject in one classroom (see also Feyerer & Prammer, 2003). To offer the students this opportunity, the materials were adjusted to the mainstream curriculum of class instruction. This ensured that the texts and the vocabulary they contained were relevant to all students. The second grade curriculum for German language arts is structured around various topics (e.g. autumn, Christmas, farm animals), following the cycle of the year. Nine of these topics were selected to frame the vocabulary and texts offered throughout the year in the program.

Every topic encompasses one sequence of vocabulary instruction and several sequences of intensive reading. In the course of the program, the sequences of vocabulary instruction marked the beginning of new topics. Since research showed evidence that vocabulary work has a positive effect on reading comprehension skills (e.g. Elleman et al., 2009; Rupley, Logan, & Nichols, 1999; Souvignier & Antoniou, 2007; Fischer & Gasteiger-Klicpera, 2011), the sequences of vocabulary instruction as pre-reading tasks are considered essential. For these sequences, 15 relevant words were

pre-selected for each topic area. During the vocabulary lessons, teachers introduced the topic and the topic-related vocabulary by activities where all students were actively involved and practiced the words in various contexts. These words were then represented frequently in the texts to be read by the students in the sequences of intensive reading.

The 15 topic-related words are controlled in terms of relevance (highly relevant for the topic and everyday life), frequency (highly frequent in primary school students' vocabulary), orthographical depths (regular words) and prosodic features (typical German word stress). In order to stimulate vocabulary development, the students worked on vocabulary pictures and so called "Wimmelbilder." These pictures display a vast amount of different objects and people in action. The objects and the actions are not only related to the topic but also represent the 15 vocabulary words. Vocabulary growth is further fostered through the use of individualized dictionaries that progressively were filled with the introduced words by the children themselves.

The second sequences of LARS-lessons are reading lessons that build the main part of the program. During these lessons, the students worked independently with texts (narrative texts as well as non-fictional texts) and corresponding reading-comprehension tasks. The texts and tasks come in four different levels. Each student was assigned to one of the four levels based on his/her individual reading skills (the percent rank of word decoding was used as the criterion of level assignment as it marks the number of words a child is able to read in a minute). Before and after the independent reading of the texts, where teachers gave support when needed, teachers were proposed to open and close the reading lessons with pre- and post-reading activities in the whole class. Discussions about the texts were encouraged.

The texts of the four different levels primarily differ in the complexity of the text. In a first step, information about the number of words read per minute by second grade students (at the beginning and in the middle of second grade) enrolled in Austrian schools was taken from the norms of a standardized test (*Salzburg reading and spelling test*, SLRT II: Moll & Landerl, 2010). Based on the norms, the range of numbers of words read per minute by students of different ability levels (level 1: above average reading ability, percentile rank (PR) above 70; level 2: average readers, PR between

30 and 70; level 3: below average reading ability, PR between 15 and 30; and level 4: lowest reading ability, PR below 15) were identified.

In a second step, the appropriate text length for texts that could be read in approximately 15 to 20 minutes adequately by students of each ability level was extrapolated from the amount of words per minute. Since longitudinal studies in German-speaking countries (Klicpera & Gasteiger-Klicpera, 1993; Landerl & Wimmer, 2008) showed that the development of reading abilities of good readers is much faster than the development of weak readers, these expected developmental differences were respected during the construction of the reading texts throughout the school year.

In a third step, the texts were written considering the topic as well as the topic-related vocabulary. However, the texts in the different ability levels did not only differ in text length. For the classification of reading texts, Björnsson (1968) examined 15 of the factors that make a text more or less easy to read. He identified only two factors that contribute explicitly to the difficulty of a text, which are also present in the German language: the word and sentence length. Therefore, in the LARS materials, the complexity of every text differed from one ability level to the other in these two factors. The word and sentence length gradually decreased from level 1 to level 4. For example, texts on ability level 1 contain longer words and sentences than texts on ability level 2. All texts in all ability levels were afterwards classified by the index LIX proposed by Björnsson. This index is calculated as the sum of the average sentence length and the percentage of long words (number of words that consist of more than six letters) in a text. As Nickel (2011) stressed, the index LIX can help to differentiate reading texts according to the needs of the learners. Thus, the LARS materials were created in accordance to this index. For every text, the average sentence length and percentage of long words decreased from level 1 to level 4 and so does the LIX as well (cf. table 1). Bamberger and Vanecek (1984) identified the average LIX for German second grade reading texts (LIX=24). At the beginning of the intervention—at the beginning of second grade—the texts of the first topic, for instance, show at ability level 1 an average LIX of 23.03, at ability level 2 an average LIX of 22.49, at ability level 3 an average LIX of 21.45, and at ability level 4 an average LIX of 20.24. Because the abilities of every reader are expected to increase, the complexity of texts in every ability level should

also increase. Thus, from the beginning of second grade to the end of the school year, the texts gained more and more difficulty depending on the learning gains that can be expected for each ability level.

Table 1

Average Number of Words, Sentences, Average Sentence Length, Percentage of Long Words (More than 6 Letters) and LIX in all LARS-Texts to the Topics 1 (First Topic) and 9 (Last Topic) in the Second Grade

Topic	Ability Level	Average Number of Words	Average Number of Sentences	Average Sentence Length (SL)	Average Percentage of Long Words (LW)	LIX = SL+LW
1	1	311.5	37.17	8.46	14.57	23.03
	2	213.5	29.33	7.37	15.12	22.49
	3	106.17	16.67	6.49	14.96	21.45
	4	60	12.5	4.84	15.39	20.24
9	1	514.75	59.75	8.67	18.65	27.32
	2	361	42.5	8.66	18.54	27.19
	3	225.75	30.5	7.44	17.34	24.78
	4	136.25	20.25	6.82	17.20	24.02

After and while reading, students should be engaged in reading comprehension tasks to ensure their understanding of the text. This also allows further work with and discussion of the content. For the differentiated LARS materials, any written instruction was adapted to the ability levels, following the same procedures and principles as described before. Correspondingly to the texts, the reading comprehension tasks varied according to level in their complexity (sentence and word length) and the number of tasks and sub-tasks that needed to be completed by the students.

For more information about the reading program, see also Seifert, Schwab and Gasteiger-Klicpera (2014).

Implementation of the Reading Program LARS

The regular teacher taught the LARS-lessons. Before the implementation of the program, all of the teachers were trained how to apply LARS. During the first two months of the reading program, one member of the project team intensively supported each teacher. The project member was present at all LARS-lessons and provided support if questions were raised. After two months the teachers were implementing the program alone in the classroom. In this way, realistic classroom conditions were present. The project member continued to visit the teacher once a month in order to bring new materials for the following topic, to explain the materials, or to support the teacher if required, and to monitor the lesson. In order to make the implementation of the LARS program transparent to the researches, teachers were asked to answer some questions about the LARS-lessons (the topic discussed in a lesson, the text read, etc.) and the experiences they made in these lessons in a log book after every LARS-lesson (e.g. what worked well, what went wrong, which children needed support, etc.).

Assessments

A broad test battery was used to examine the reading and language abilities of all children. The reading part of the *Salzburg reading and spelling test* (SLRT II: Moll & Landerl, 2010) was used to measure decoding competencies. The SLRT is an individual reading test that assesses decoding in

two sub-scales: the sub-scale pseudo-words (assessing non-lexical decoding) and the sub-scale words (assessing lexical decoding). Pseudo-words and words presented in a list are read out by the students in one minute, respectively. According to the manual, the retest reliability is high ($r_{tt} = .90-.98$).

Reading comprehension was tested with the *reading comprehension test for first to sixth graders* (Elfe 1-6: Lenhard & Schneider, 2006). The Elfe 1-6 is used to ascertain the reading comprehension on the word-, sentence- and text-level with three sub-scales. The internal consistency of the sub-scales ranges from .92 to .97. In the present study, only the sub-scales of word and sentence comprehension were assessed.

Vocabulary knowledge was assessed with a modified version of the receptive vocabulary sub-scale of the *vocabulary- and word finding-test for 6 to 10 year olds* (WWT 6-10: Glück, 2007; $\alpha = .84$), which was adapted for the use as a group test. This test consists of 18 items. The task of the students is to choose one out of four pictures that matches best the word the instructor says. The internal consistency of this test is .79 for Time 1 as well as for Time 2. In addition, at Time 1, the original short form of the WWT was used to control whether the modified version represents a reliable and valid tool for vocabulary assessment. The WWT is divided into the sub-scales expressive and receptive vocabulary. The correlation between the modified receptive vocabulary test and the original WWT by Glück (2007) is .71 with the expressive and .73 with the receptive vocabulary sub-scale.

A test to measure the learning gains of the vocabulary to be taught in the reading program was also conducted. This expressive vocabulary test consisted of 27 items of the words included in the vocabulary sequences. The reliability was .87 at Time 1 and .88 at Time 2.

The grammar skills (expressive knowledge, not receptive) were measured with a modified version of the sub-test grammar of the *Potsdam-Illinois Test for Psycholinguistic Abilities* (P-ITPA: Esser, Wyschkon, Ballaschk, & Hänsch, 2010). The reliability at Time 1 was .89 and .9 at Time 2.

Furthermore, overall cognitive abilities were assessed with the *Cultural Fair Intelligence Test* (CFT-1; Cattell, Weiß, & Osterland, 1997) at Time 1. Information about the children's first language and age was collected through teacher questionnaires.

Statistical Analysis

To compare the initial ability levels of both groups (LARS-group vs. C-group), multivariate analyses of variance and t-tests were calculated. To analyze the effects of the differentiating reading program, multivariate variance analyses with repeated measurement as well as t-tests were calculated. The initial ability level (level 1 to 4) as well as the language (L1 or L2 German) were used as group factors. The age and the intelligence were used as control variables in these analyses. As dependent variables, the achievement of the students (word decoding, pseudo-word decoding, word reading comprehension, sentence reading comprehension, receptive vocabulary—of untrained words only, and expressive vocabulary—of trained words only, and grammar) were inserted as raw scores in the analyses. Concerning the intelligence, IQ-scores were used for the calculation. The two sub-scales of decoding as well as the two sub-scales of reading comprehension were analyzed within multivariate analyses with repeated measurement.

Results

Group Differences at the First Measurement Point (Time 1)

Regarding the initial abilities of the LARS- and the C-group in decoding ($F_{2,375} = 0.80$, n.s.), no significant group difference could be detected. For reading comprehension, however, the multivariate analysis showed differences between the LARS- group and the C- group ($F_{2,375} = 6.34$, $p < .01$), but no significant differences at univariate level could be found for reading comprehension (cf. table 2). Furthermore, no group differences became apparent neither in general intelligence nor in language skills (cf. table 3).

Table 2

Initial Mean-Differences, Standard Deviations, F-Values and Significances in the Reading Abilities of the LARS-Group Receiving the Program and C-Group Receiving Conventional Reading Instruction

Sub-Test	Time 1		<i>F</i>	<i>p</i>
	<i>M (SD)</i>			
	LARS-Group	C-Group		
Word Decoding				
(Number of Correct Read Words per Minute)	30.15 (13.97)	31.76 (14.48)	1.17	.28
Pseudo-Word Decoding				
(Number of Correct Read Pseudo-Words per Minute)	25.77 (8.67)	26.35 (8.33)	0.43	.51
Reading Comprehension on Word-Level	16.15 (8.00)	15.32 (7.37)	1.10	.30
Reading Comprehension on Sentence-Level	5.67 (4.15)	6.19 (3.98)	1.52	.22

Table 3

Initial Mean-Differences, Standard Deviations, T-Values and Significances in IQ and the Assessed Language Abilities of the LARS-Group Receiving the Program and C-Group Receiving Conventional Reading Instruction

Sub-Test	Time 1		<i>t</i>	<i>p</i>
	<i>M (SD)</i>			
	LARS-Group	C-Group		
Intelligence (IQ)°	96.12 (12.23)	96.74 (12.06)	0.49	.62
Expressive Vocabulary (WWT)°	13.62 (8.66)	13.98 (9.57)	0.38	.70
Receptive Vocabulary (WWT)°	32.31 (6.58)	31.31 (6.68)	-1.46	.15
Expressive Vocabulary (Trained Words)	16.59 (4.95)	16.22 (5.70)	-0.68	.50
Receptive Vocabulary (Adapted Group-Test)	13.33 (3.31)	12.72 (3.70)	-1.65	.10
Grammar (Adapted P-ITPA Sub-Test)	19.11 (5.87)	19.41 (6.46)	0.45	.65

Note. [°]Assessed as control variable only at Time 1.

Group Differences in the Development of Reading Skills

The variance analysis with repeated measurement (IQ and age was controlled) showed a significant effect for the interaction between time and the implementation of the LARS reading program in reading comprehension on sentence-level. Students of the LARS-group showed significantly higher increases than students of the C-group ($F_{1, 370}=7.99, p<.01, \eta^2=.02$; cf. table 4).

This effect can be interpreted as a small effect. There were no significant interaction effects between time and LARS reading program for word decoding, pseudo-word decoding and reading comprehension on word-level. All further interactions, in particular with the group variables initial ability level (level 1 to 4) and language (L1 or L2), were not significant. Therefore, those F-values are not presented.

Table 4

Mean-Differences, Standard Deviations, F-Values and Significances in Reading Abilities of the LARS-Group Receiving the Program and C-Group Receiving Conventional Reading Instruction

Sub-Test	LARS-Group		C-Group		F°	Eta^2
	Time 1	Time 2	Time 1	Time 2		
	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$		
Word Decoding	30.15	46.87	31.85	49.30	0.63	.00
(Number of Correct Read Words per Minute)	(13.97)	(16.25)	(14.54)	(17.68)		
Pseudo-Word Decoding	25.77	33.86	26.37	34.58	0.06	.00
(Number of Correct Read Pseudo-Words per Minute)	(8.67)	(9.80)	(8.36)	(10.58)		
Reading Comprehension on Word-Level	16.15	30.68	15.32	29.27	1.65	.00
	(8.00)	(10.50)	(7.37)	(10.28)		
Reading Comprehension on Sentence-Level	5.67	13.06	6.19	12.82	7.99**	.02
	(4.15)	(5.20)	(3.98)	(5.17)		

Notes. $^{\circ}$ F-values of the interaction between time and LARS reading program. ** $p < .01$.

Group Differences in the Development of Language Skills

The variance analysis with repeated measurement (IQ and age was controlled) showed significant effects for the interaction between time and LARS reading program in expressive vocabulary (trained words) ($F_{1, 369} = 7.42$, $p < .01$, $Eta^2 = .02$). Students of the LARS-group showed

significantly larger increases than students of the comparison group (cf. table 5). This effect can be interpreted as a small effect. There were no significant interaction effects between time and LARS reading program for receptive vocabulary and grammar.

For expressive vocabulary (trained words), i.e. the vocabulary included in the program, a significant interaction between time x LARS reading program x language (German as L1 or L2) was revealed ($F_{1,367}=8.08, p<.01, \eta^2=.02$). This effect was due to differences between the LARS-group and the C-group when considering L1 children. L1 children of the LARS-group showed significantly larger increases than L1 children of the C-group (cf. figure 1). Nevertheless, students with L2 and students with L1 of the LARS-group show nearly the same performance gains in expressive vocabulary (trained words). Thus, with the help of the intervention, students with L2 gain nearly as much vocabulary knowledge from the LARS reading program as students with L1.

Moreover, the analyses showed no further interactions. In particular, interactions with the group variable initial ability level (level 1 to 4) remained absent. Hence, the F-values are not presented.

Table 5

Mean-Differences, Standard Deviations, F-Values and Significances in Language Skills of the LARS-Group Receiving the Program and C-Group Receiving Conventional Reading Instruction

Sub-Test	LARS-Group		C-Group		F°	Eta^2
	Time 1	Time 2	Time 1	Time 2		
	$M (SD)$	$M (SD)$	$M (SD)$	$M (SD)$		
Receptive Vocabulary (Adapted Group-Test)	13.33 (3.31)	14.99 (2.67)	12.72 (3.70)	14.29 (3.33)	0.20	.00
Expressive Vocabulary (Trained Words)	16.59 (4.95)	20.69 (5.19)	16.32 (5.63)	19.60 (5.02)	7.42**	.02
Grammar (Adapted P-ITPA Sub-Test)	19.07 (5.88)	21.95 (4.59)	19.46 (6.38)	22.07 (5.19)	0.27	.00

Notes. $^{\circ}$ F-values of the interaction between time and LARS reading program. ** $p < .01$.

Figure 1. Performance gains in expressive vocabulary of the words included in the LARS program (trained words) of students with L1 and L2 in the LARS-group and the C-group.

Discussion

This study's primary purpose was to examine the effects of a whole-class differentiating reading program (LARS) that comes in four different reading levels and involves vocabulary work. Firstly, this study attempted to investigate the effects of the program on the reading abilities of second grade students with German as L1 and German as L2. The comparison of the group receiving the LARS reading program (LARS-group) with the group receiving conventional reading instruction (C-group) showed that instruction with the LARS reading program twice a week lead to higher increases in reading comprehension (on sentence-level). Because the work concentrated on reading comprehension, it was expected that gains would be revealed both in word and sentence comprehension. However, only the effect on sentence-level was shown and this one was moderately small. In contrast, the increases in decoding from the beginning to the end of the second grade were similar in both groups. First of all, the different reading outcomes for reading comprehension and

decoding can be attributed to the focus of the intervention. As the program aimed at improving reading comprehension and not decoding, the absence of effects for decoding was mostly expected. Secondly, the decoding skills of all students (the LARS-group as well as the C-group) were already average at the beginning of the study. Thus, at the beginning of second grade, the students did not have problems in decoding but they showed problems in reading comprehension. This difference in the reading sub-skills can partially be attributed to the relatively high ratio of second language learners in the sample since it is known that they often have deficits in reading comprehension but not in decoding (Melby-Lervag & Lervag, 2013; Verhoeven & van Leeuwe, 2012).

The study's second aim was to examine whether the reading program also shows effects on the language skills of the students. In particular, as other scholars have illustrated, vocabulary knowledge and grammar skills are related to reading comprehension (e.g. Catts et al., 2006; Ennemoser et al., 2012; Snowling et al., 2003). Therefore, this study scrutinized those two components of language ability.

Regarding vocabulary knowledge, the LARS-group showed significantly larger increases in expressive vocabulary of the trained words compared to the C-group. This small effect reflects a training or learning effect. Admittedly, this study did not assess whether trained words were also known better receptively after the intervention. But since all those words included in the vocabulary work of LARS are high rather than low frequency words, the chance that they are known receptively as well is quite high (Waring, 1996; Webb, 2008). Different performance gains between the LARS-group and the C-group in receptive vocabulary (of untrained words) could not be found. The absence of a difference in receptive vocabulary gains of untrained vocabulary shows that there has been no transfer of the trained vocabulary to untrained vocabulary. However, the lack of transfer effects was shown before in other vocabulary interventions (Elleman et al., 2009).

As regards grammar skills, the LARS-group showed comparable gains as the C-group. The difference between the outcomes for vocabulary and grammar skills can be attributed to the program's focus. While vocabulary work was part of the program, no explicit training for grammar skills was included. Thus, the absence of effects for grammar skills was mostly expected.

Thirdly, this study investigated whether the LARS program was equally effective for L1 and L2 German learners. As was shown, both students with L1 and L2 German benefitted equally from the reading program regarding comprehension skills. Thus, this result reconfirms the outcomes of Nelson et al. (2003) whose meta-analysis revealed that the treatment responsiveness of children to reading interventions is not influenced by the first language. In contrast, the result does not support the findings of Al Otaiba and Fuchs (2002), since in the present study the proficiency in the language of instruction did not influence the treatment responsiveness. In addition, concerning language abilities, L1 and L2 children benefitted both from the LARS program. More precisely, this study saw no difference in the grammar gains of L1 and L2 learners. Similarly, L1 and L2 learners of the LARS-group showed nearly the same vocabulary gains (expressive vocabulary of trained words). Thus, the program offers learning opportunities for both L1 and L2 learners. A difference in the vocabulary gains between the LARS-group and the C-group was found concerning L1 children. L1 children gained more knowledge of these words when taught with the LARS program. This indicates that not only L1 but also L2 children can benefit from vocabulary work, as also discussed by Rupley et al. (1999).

Because the LARS reading program is a reading program offering materials in four different ability levels, children worked on different reading texts and tasks based on their reading ability levels. Therefore, the study's fourth aim was to show whether the students of the different ability levels profit equally from the program. As was shown, this assumption could be proven as no interaction effect for the ability level group was revealed. This result is consistent with previous research (e.g. Nelson et al., 2003) and supports the assumption that students' response to intervention is independent of their ability level in reading. Thus, it can be concluded that the systematically and differentiating reading program does not only foster the development of reading competence in students with difficult starting conditions, but also in students with average or above average abilities.

In general, the effects of the reading program are limited. Nonetheless, the presented results can be interpreted as a valuable success of the LARS reading program because during the program's duration the LARS-lessons replaced two regular school lessons per week. Therefore, the LARS-group and the C-group received a comparable amount of school lessons. However, the study did not assess

the overall amount of reading lessons in the LARS-group (as the LARS program was not designed to completely replace all reading lessons in second grade) nor in the C-group. Since the number of reading lessons per week is not prescribed in the second grade classes in Austria, it is possible that some classes in the C-group had more reading lessons than LARS-classes. Assessing all of the reading lessons and the contents of these lessons during the whole school year would have amounted to an enormous cost that the authors were not able to afford. Consequently, it remains unclear what exactly the C-group accomplished during reading instruction. Still, conventional reading materials in Austrian schools are not differentiated. In contrast, the LARS-group received and worked on such differentiated materials. Thus, pre- and post-reading activities in the whole class on the same topic and text were possible. It can be assumed that the differentiated materials ensured that the students could be actively involved in the classroom and increased the possibility for the students to get a better understanding of the content as well as to improve their reading abilities. The effects show that the students taught with the LARS reading program increased their skills more considerably than students receiving conventional reading instruction, even though the effects were of course limited. We transformed the revealed effects into Cohen's d corrected for different sample sizes and pre-test differences (Klauer, Willmes, & Phye, 2002). Thus, the effect for reading comprehension on sentence-level was $d = 0.175$. Compared to the effects shown by Slavin et al. (2010) for reading programs executed from the second grade on ($d = 0.13$), the revealed effects of the LARS program are quite similar.

Nevertheless, the study's limited effects must be discussed, particularly in regard to previous findings. Compared to the pilot study (see Schwab & Gasteiger-Klicpera, 2013; Schwab, Seifert, & Gasteiger-Klicpera, 2014), where effects on both decoding sub-scales and all sub-scales of reading comprehension could be revealed, it is astonishing that in the present study no effects in decoding and only small and limited effects in reading comprehension were found. As the LARS reading program period in the previous study was shorter and the materials were only divided in three different ability levels, greater effects were expected for the present study. Obviously, the absence of higher effects has to be ascribed to the fact that in the previous study a project member conducted the implementation of the LARS reading program whereas in the present study the teacher led the program. Hence, during the pilot study one extra person in the classroom supported the students during the LARS-lessons and

assured high implementation quality. This is not surprising, as scholars have shown that an additional person can raise students' engagement (Stephen, Ellis, & Martlew, 2010). Furthermore, the inconsistency in implementation among the classes could have been reduced through the presence of a project member. Without a project member, the fidelity of the program could not be guaranteed. After the first two months of intense support, the fidelity was only controlled by monthly visits and monitoring, as well as logbooks filled out by the teachers after every LARS-lesson. As first evaluations of this qualitative data show, the teachers implemented LARS in quite different ways. Some of the teachers enthusiastically integrated the LARS-lessons and the materials into their regular schooling routines and used the materials for extended discussions on the texts, vocabulary, and topics in other subjects as well. On the contrary, others seemed to experience LARS-lessons as an obligation and conducted these lessons rather superficially. Perhaps, then, the amount of experience a teacher already possessed is related to the study's fidelity. Further analysis of the data collected from the logbooks, monitoring, and interviews will help understand how the LARS program has to be implemented to lead to positive outcomes. As already assumed by Snow and Juel (2006), the quality of implementation of reading programs has much more impact on the program's effects on reading than the nature of the program itself. They concluded that teachers should have profound knowledge about the learning subject, the (different) learning preconditions of the students, and the knowledge about adequate ways to guide learning effectively for each student.

For future research, it is recommended to select the participating teachers more carefully in order to avoid the pitfall that they are participating in the study only because they are told to do so by their principal. In addition, teachers should be trained intensely in implementing the reading program with the intention to promote understanding of reading development, to foster instructional quality, and avoid differences in fidelity. This also includes the specification and clarification of the program's procedures and guidelines in order to ensure implementation of the program as originally designed (Lane, Bocian, MacMillan, & Gresham, 2004). Moreover, qualitative data should be used to observe what actually occurs during the lessons when a program is implemented.

Scholars have to interpret the results of the current study carefully, because some limitations must be considered. The first limitation concerns the sample. The results are not generalizable for

average second grade students because the examined students (both in the LARS- and in the C-group) contained an obviously higher amount of second language learners than typical in Styrian schools. This is based on the decision to focus on the needs of L2 learners. Furthermore, the methodological design of the study exhibits limitations. As the LARS reading program links different elements to foster reading abilities (e.g. systematic vocabulary work, different tasks to improve reading comprehension, the different levels of texts), it is not possible to conclude which part of the LARS reading program led to its moderate success. Basically, research indicates that multiple components of reading supported in an intervention lead to positive effects in reading abilities (Guthrie et al., 2007). Perhaps, the combination of the components of the LARS program can be considered as one reason for the effects. For the current study, the effectiveness of the reading program, although limited, might be attributed to the quantitatively and qualitatively differentiated materials adapted to the initial skills of the students that enabled the students to work steadily and independently on their reading level following their own pace and time.

To reveal whether the effects on sentence comprehension and expressive vocabulary knowledge of trained words persist over time, follow-up data should be assessed. Furthermore, it should be examined whether expanding the program would lead to better results. For instance, simply adjusting the reading level was possibly not sufficient to foster all children adequately. In addition, the vocabulary work included in the reading program was probably not enough to foster the L2 learners effectively. Undoubtedly, as Goldberg concluded (2012), L2 learners need additional instructional support. Perhaps, then, advices for such additional support should be added to the program.

As the experiences of the research group with teachers in the sample show, reading instruction is a very complex and challenging task for teachers. Thus, they mostly appreciate highly to receive systematic and motivating materials supplementing their work. Nevertheless, it is obvious that for successful reading instruction, continuous training of the teachers is needed. Perhaps a reading-expert should support the teacher during the time the instruction is implemented (Helf & Cooke, 2012; Kretlow & Bartholomew, 2010) in order to ensure high quality reading instruction for every pupil in the class. However, successful reading instruction has to be adapted to the different learning needs of

the students resulting in the change of methods and materials, learning settings, and classroom organization.

Acknowledgements. We would like to thank the Styrian government (department of education) for supporting the current research project financially. We are grateful to Norbert Holzer, Kerstin Waldmüller, Petra Watko, Kathrin Ortner and Mike Trauntschnig for their assistance during the whole project, especially with respect to the preparation of the reading instruction and data collection.

References

- Al Otaiba, S., & Fuchs, D. (2002). Characteristics of Children Who Are Unresponsive to Early Literacy Intervention. *Remedial and Special Education, 23*(5), 300-316.
- Ankrum, J. W., & Bean, R. M. (2007). Differentiated Reading Instruction: What and How. *Reading Horizons, 48* (1), 133-146.
- Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., Schneider, W., Stanat, P., Ostermeier, C., Schiefele, U., Valtin, R., & Ring, K. (2007). *Förderung von Lesekompetenz. Expertise*. Berlin: Bundesministerium für Bildung und Forschung.
- Bamberger, R., & Vanecek, E. (1984). *Lesen – Verstehen – Lernen – Schreiben. Die Schwierigkeitsstufen von Texten in deutscher Sprache*. Wien: Jugend und Volk.
- Baecher, L., Artiglieri, M., Patterson, D., & Spatzer, A. (2012). Differentiated Instruction for English Language Learners as „Variations on a Theme“. *Middle School Journal, 43* (3), 14-21.
- Bell, S., McPhillips, T., & Doveston, M. (2011). How do teachers in Ireland and England conceptualise dyslexia? *Journal of Research in Reading, 34*, 171-192.
- Berkeley, S., Scruggs, T. E., & Mastropieri, M. A. (2009). Reading Comprehension Instruction for Students with Learning Disabilities, 1995–2006: A Meta-Analysis. *Remedial and Special Education, 31*, 423-436.
- Bishop, D. V. M., & Snowling, M. J. (2004). Developmental Dyslexia and Specific Language Impairment: Same or Different? *Psychological Bulletin, 130*, 858-886.
- Björnsson, C. H. (1968). *Läsbarhet*. Stockholm: Liber.
- Bos, W., Bremerich-Vos, A., Tarelli, I., & Valtin, R. (2012). Lesekompetenzen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos, & K. Schwippert (Eds.), *IGLU 2011*.

- Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (pp. 91-135). Münster u.a.: Waxmann.
- Cattell, R. B., Weiß, R. H. & Osterland, J. (1997). *Grundintelligenztest Skala 1 (CFT 1)* (5. Auflage). Göttingen: Westermann.
- Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language Deficits in Poor Comprehenders: A Case for the Simple View of Reading. *Journal of Speech, Language and Hearing Research*, 49, 278-293.
- Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A Longitudinal Investigation of Reading Outcomes in Children with Language Impairments. *Journal of Speech, Language, and Hearing Research*, 45, 1142-1157.
- Duff, F. J., Fieldsend, E., Bowyer-Crane, C., Hulme, C., Smith, G., Gibbs, S., & Snowling, M. J. (2008). Reading with vocabulary intervention: evaluation of an instruction for children with poor response to reading intervention. *Journal of Research in Reading*, 31, 319-336.
- Education, Audiovisual and Culture Executive Agency (2011). Teaching reading in Europe. Contexts, Policies and Practices. Retrieved from: <http://eacea.ec.europa.eu/education/eurydice>
- Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Tackett, K. K. et al. (2009). A Synthesis of Reading Interventions and Effects on Reading Comprehension Outcomes for Older Struggling Readers. *Review of Educational Research*, 79, 262-300.
- Ehri, L. C. (1999). Phases of development in learning to read words. In J. Oakhill, & R. Beard (Eds.), *Reading Development and the Teaching of Reading: A Psychological Perspective* (pp. 79-108). Oxford, UK: Blackwell Publishers.
- Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic Phonics Instruction Helps Students Learn to Read: Evidence From the National Reading Panel's Meta-Analysis. *Review of Educational Research*, 71, 393-447.
- Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The Impact of Vocabulary Instruction on Passage-Level Comprehension of School-Age Children: A Meta-Analysis. *Journal of Research on Educational Effectiveness*, 2 (1), 1-44.
- Ennemoser, M., Marx, P., Weber, J., & Schneider, W. (2012). Spezifische Vorläuferfertigkeiten der Lesegeschwindigkeit, des Leseverständnisses und des Rechtschreibens. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44 (2), 53-67.
- Esser, H. (2006). *Migration, Sprache und Integration*. Berlin: Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration, Wissenschaftszentrum Berlin für Sozialforschung.
- Esser, G., Wyschkon, A., Ballaschk, K. & Hänsch, S. (2010). *Potsdam-Illinois Test für Psycholinguistische Fähigkeiten - P-ITPA*. Göttingen: Hogrefe, Testzentrale.
- Feyerer, E. & Prammer, W. (2003). *Gemeinsamer Unterricht in der Sekundarstufe I. Anregungen für eine integrative Praxis*. Weinheim; Basel; Berlin: Beltz.
- Fischer, U. (2012). *Leseförderung im Anfangsunterricht*. Duisburg: Universitätsverlag Rhein-Ruhr.

- Fischer, U., & Gasteiger-Klicpera, B. (2011). *Leseförderung durch Wortschatzarbeit: Ergebnisse einer frühen Interventionsstudie zur individuellen Leseförderung*. In B. Lütje-Klose, M. T. Langer, B. Serke, M. Urban (Eds.), *Inklusion in Bildungsinstitutionen - eine Herausforderung an die Heil- und Sonderpädagogik* (pp. 272-277). Bad Heilbrunn: Verlag Julius Klinkhardt
- Foorman, B. R. (2007). Primary Prevention in Classroom Reading Instruction. *Teaching Exceptional Children*, 39, 24-30.
- Förster, S., Niebuhr-Siebert, S., & Hellmich, F. (2012). Förderung der Lesekompetenz durch ein wortschatzbasiertes Lesestrategietraining bei Kindern mit Deutsch als Zweitsprache. In F. Hellmich, S. Förster, & F. Hoya (Eds.), *Bedingungen des Lehrens und Lernens in der Grundschule. Bilanz und Perspektiven* (pp. 199-202). Wiesbaden: Springer.
- Glück, C. W. (2007). *Wortschatz- und Wortfindungstest für 6- bis 10-Jährige (WWT 6-10)*. München: Elsevier.
- Goldenberg, C. (2010). Reading instruction for English language learners. In M. Kamil, P. D. Pearson, E. Moje, & P. Afflerbach (Eds.), *Handbook of Reading Research*, Vol. IV (pp. 684-710). Newark, DE: International Reading Association.
- Guthrie, J. T., McRae, A., & Klauda, S. L. (2007). Contributions of Concept-Oriented Reading Instruction to Knowledge About Interventions for Motivations in Reading. *Educational Psychologist*, 42 (4), 237-250.
- Guthrie, J. T., Wigfield, A., & Perencevich, K. C. (2004). *Motivating reading comprehension: Concept-Oriented Reading Instruction*. Mahwah, NJ: Erlbaum.
- Hattie, J. (2009). *Visible Learning*. London: Routledge.
- Helf, S., & Cooke, N. L. (2012). Reading Specialist: Key to a Systematic Schoolwide Reading Model. *Preventing School Failure*, 55, 140-147.
- Hurry, J., & Sylva, K. (2007). Long-term outcomes of early reading intervention. *Journal of Research in Reading*, 30, 227-258.
- Hutchinson, J. M., Whiteley, H. E., Smith, C. D., & Connors, L. (2003). The developmental progression of comprehension-related skills in children learning EAL. *Journal of Research in Reading*, 26 (1), 19-32.
- Kena, G., Aud, S., Johnson, F., Wang, X., Zhang, J., Rathbun, A., Wilkinson-Flicker, S., & Kristapovich, P. (2014). *The Condition of Education 2014* (NCES 2014-083). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved from: <http://nces.ed.gov/pubsearch>
- Klicpera, C., & Gasteiger-Klicpera, B. (1993). *Lesen und Schreiben – Entwicklung und Schwierigkeiten: Die Wiener Längsschnittuntersuchungen über die Entwicklung, den Verlauf und die Ursachen von Lese- und Schreibschwierigkeiten in der Pflichtschulzeit*. Bern: Huber Verlag.

- Klicpera, C., Schabmann, A., & Gasteiger-Klicpera, B. (2013). *Legasthenie – LRS. Modelle, Diagnose, Therapie und Förderung* (4th Edition). München: Reinhardt.
- Kretlow, A. G., & Bartholomew, C. C. (2010). Using Coaching to Improve the Fidelity of Evidence-Based Practices: A Review of Studies. *Teacher Education and Special Education*, 33, 279-299.
- Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment Integrity: An Essential –But Often Forgotten–Component of School-Based Interventions. *Preventing School Failure*, 48 (3), 36-43.
- Lee, J. (2011). Size matters: Early vocabulary as a predictor of language. *Applied Psycholinguistics*, 32, 69-92.
- Lenhard, W., & Schneider, W. (2006). *Ein Leseverständnistest für Erst- bis Sechstklässler (ELFE 1-6)*. Weinheim: Beltz.
- McAlenney, A. L., & Coyne, M. D. (2011). Identifying At-Risk Students for Early Reading Intervention: Challenges and Possible Solutions. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 27, 306–323.
- McGill-Franzen, A., Zmach, C., Solic, K., & Zeig, J.L. (2006). The Confluence of Two Policy Mandates: Core Reading Programs and Third Grade Retention in Florida. *Elementary School Journal*, 107, 67-91.
- Melby-Lervåg, M., & Lervåg, A. (2013). 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. DOI: 10.1037/a0033890.
- Moll, K., & Landerl, K. (2010). *SLRT-II – Verfahren zur Differentialdiagnose von Störungen der Teilkomponenten des Lesens und Schreibens*. Bern: Huber.
- Munser-Kiefer, M. (2012). Lesen im Leseteam trainieren. In M. Philipp, & A. Schilcher (Eds.), *Selbstreguliertes Lesen. Ein Überblick über wirksame Leseförderansätze* (pp. 100-115). Seelze-Velber: Klett/Kallmeyer.
- Munser-Kiefer, M., & Kirschhock, E.-M. (2009). Entwicklung und Evaluation eines Lesekompetenztrainings – Eine Vorstudie zur Frage nach der geeigneten Sozialform. In C. Röhner, C. Henrichwark, & M. Hopf, (Eds.), *Europäisierung der Bildung - Konsequenzen und Herausforderungen für die Grundschulpädagogik* (pp. 166-171). Wiesbaden: VS Verlag.
- Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, Rimes, Vocabulary, and Grammatical Skills as Foundations of Early Reading Development: Evidence From a Longitudinal Study. *Developmental psychology*, 40 (5), 665-681.
- Nation, K., & Snowling, M. J. (1998). Semantic Processing and the Development of Word Recognition Skills: Evidence from Children with Reading Comprehension Difficulties. *Journal of Memory and Language*, 39, 85-101.
- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research*

- Literature on Reading and Its Implications for Reading Instruction* (National Institutes of Health Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.
- Nelson, J. R., Benner, G. J., & Gonzalez, J. (2003). Learner Characteristics that Influence the Treatment Effectiveness of Early Literacy Interventions: A Meta-Analytic Review. *Learning Disabilities Research and Practice, 18*, 255-267.
- Nickel, S. (2011). Textschwierigkeit objektivieren: Der Lesbarkeitsindex LIX. *Alfa-Forum, 76*, 30-32.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The Acquisition of Reading Comprehension Skill. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading* (pp. 227-247). Malden: Blackwell.
- Philipp, M. (2013). *Lese- und Schreibunterricht*. Tübingen: Francke.
- RAND (Reading Study Group) (2002). *Reading for Understanding: Towards an R&D Program in Reading Comprehension*. Santa Monica, CA: RAND.
- Rose, J. (2009). *Identifying and Teaching Children and Young People with Dyslexia and Literacy Difficulties*. London: Department for Children Schools and Families.
- Rupley, W. H., Logan, J. W., & Nichols, W. D. (1999). Vocabulary Instruction in a Balanced Reading Program. *The Reading Teacher, 52*, 336-346.
- Rupley, W. H., Blair, T. R., & Nichols, W. D. (2009). Effective Reading Instruction for Struggling Readers: The Role of Direct/ Explicit Teaching. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 25* (2-3), 125-138.
- Schabmann, A., Landerl, K., Bruneforth, M., & Schmidt, B. M. (2012). Lesekompetenz, Leseunterricht und Leseförderung im österreichischen Schulsystem. Analysen zur pädagogischen Förderung der Lesekompetenz. In B. Herzog-Punzenberger (Eds.), *Nationaler Bildungsbericht Österreich 2012, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (pp. 17-70). Graz: Leykam.
- Schabmann, A., Schmidt, B. M., Klicpera, C., Gasteiger-Klicpera, B., & Klingebiel, K. (2009) Does systematic reading instruction impede prediction of reading a shallow orthography? *Psychology Science Quarterly, 51*, 315-338.
- Schumm, J. S., Moody, S. W., & Vaughn, S. (2000). Grouping for Reading Instruction: Does One Size Fit All? *Journal of Learning Disabilities, 33*, 477-488.
- Schumaker, J. B., Denton, P. H., & Deshler, D. D. (1984). *The paraphrasing strategy*. Lawrence: University of Kansas.
- Schwab, S., & Gasteiger-Klicpera, B. (2014). Förderung der Lesekompetenzen bei Kindern der zweiten Schulstufe – Evaluierung eines differenzierten Sprach- und Leseförderprogramms im Rahmen des Grundschulunterrichts. *Zeitschrift für Bildungsforschung, 4* (1), 63-79.
- Schwab, S., Seifert, S., & Gasteiger-Klicpera, B. (2013). Reading intervention in second-grade children with poor reading abilities. *Procedia - Social and Behavioral Sciences, 106*, 2205-2216.

- Schwab, S., Seifert, S., & Gasteiger-Klicpera, B. (2014). Improving Reading in Children With German as a First or Second Language. *Journal of Early Years Education*. DOI: 10.1080/09669760.2014.909308.
- Seifert, S., Schwab, S., & Gasteiger-Klicpera, B. (2014). Ein Konzept für differenzierten Leseunterricht in heterogenen Klassen. *Sprache Stimme Gehör*, 38, e21-e27.
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174.
- Slavin, R.E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2010). *Effective Reading Programs for the Elementary Grades: A Best-Evidence Synthesis*. Baltimore, MD: Johns Hopkins University, Center for Data-Driven Reform in Education.
- Snow, C. E., & Juel, C. (2005). Teaching Children to Read: What Do We Know about How to Do It? In M. J. Snowling, & C. Hulme (Eds.), *The Science of Reading* (pp. 501-520). Malden: Blackwell.
- Snowling, M., & Hulme, C. (2005). Learning to Read with a Language Impairment. In M. J. Snowling, & C. Hulme (Eds.), *The Science of Reading* (pp. 397-412). Malden: Blackwell.
- Souvignier, E., & Antoniou, F. (2007). Förderung des Leseverständnisses bei Schülerinnen und Schülern mit Lernschwierigkeiten – eine Metaanalyse. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 76, 46-62.
- Statistik Austria (2014). *Bildung in Zahlen 2012/2013 – Tabellenband*. Retrieved from: http://www.statistik.at/web_de/services/publikationen/5/index.html
- Stephen, C., Ellis, J., & Martlew, J. (2010). Taking active learning into the primary school: a matter of new practices? *Journal of Early Years Education*, 18, 315-329.
- Swanson, H. L. (1999). *Interventions for Students With Learning Disabilities: A Meta-Analysis of Treatment Outcomes*. New York: The Guilford Press.
- Taylor, D. B., Mraz, M., Nichols, W. D., Rickelman, R. J., & Wood, K. D. (2009). Using Explicit Instruction to Promote Vocabulary Learning for Struggling Readers. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 25 (2-3), 205-220.
- Torgesen, J. K. (2005). Recent Discoveries on Remedial Interventions for Children with Dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The Science of Reading* (pp. 521-537). Malden: Blackwell.
- Vadasy, P. F., & Sanders, E. A. (2008). Repeated Reading Intervention: Outcomes and Interactions with Readers' Skills and Classroom Instruction. *Journal of Educational Psychology*, 100 (2), 272-290.
- Vadasy, P. F., Sanders, E. A., & Abbott, R. D. (2008). Effects of Supplemental Early Reading Intervention at 2-Year Follow Up: Reading Skill Growth Patterns and Predictors. *Scientific Studies of Reading*, 12, 51-89.

- Verhoeven, L. (2011). Second Language Reading Acquisition. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of Reading Research* (pp. 661-683). New York: Routledge.
- Verhoeven, L., & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing*, 25, 1805-1818.
- Vogtenhuber, S., Lassnigg, L., Gumpoldsberger, H., Schwantner, U., Suchaň, B., Bruneforth, M. et al. (2012). Indikatoren D: Output – Ergebnisse des Schulsystems. In M. Bruneforth, & L. Lassnigg (Eds.), *Nationaler Bildungsbericht Österreich 2012, Band 1: Das Schulsystem im Spiegel von Daten und Indikatoren* (pp. 111-164). Graz: Leykam.
- Wanzek, J. (2013). Building Word Knowledge: Opportunities for Direct Vocabulary Instruction in General Education for Students With Reading Difficulties. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 30 (2), 139-164.
- Wanzek, J., & Vaughn, S. (2007). Research-Based Implications From Extensive Early Reading Interventions. *School Psychology Review*, 36, 541-561.
- Waring, R. A. (1996). Comparison of the Receptive and Productive Vocabulary Sizes of Some Second Language Learners. *Immaculata - occasional papers of Notre Dame Seishin University*, 1. Retrieved from: <http://www.robwaring.org/papers/various/vocsize.html>
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30 (1), 79-95.
- Whiteley, H. E., Smith, D., & Connors, L. (2007). Young children at risk of literacy difficulties: Factors predicting recovery from risk following phonologically-based intervention. *Journal of Research in Reading*, 30, 249-269.
- Wigfield, A., Guthrie, J. T., Perencevich, K. C., Taboada, A., Klauda, S. L., McCrae, A., & Barbosa, P. (2008). Role of Reading Engagement in Mediating Effects of Reading Comprehension Instruction on Reading Outcomes. *Psychology in the Schools*, 45, 432-445.
- Wimmer, H., Mayringer, H., & Landerl, K. (2000). The Double-Deficit Hypothesis and Difficulties in Learning to Read a Regular Orthography. *Journal of Educational Psychology*, 92, 668-680.
- Woodward, M.M., & Talbert-Johnson, C. (2009). Reading Intervention Models: Challenges of Classroom Support and Separated Instruction. *The Reading Teacher*, 63 (3), 190-200.