

Training reading fluency: is it important to practice reading aloud and is generalization possible?

Sini Hintikka · Karin Landerl · Mikko Aro ·
Heikki Lyytinen

Received: 4 September 2007 / Accepted: 14 March 2008 /
Published online: 16 May 2008
© The International Dyslexia Association 2008

Abstract Outcomes of three different types of computerized training in sub-lexical items (word-initial consonant clusters) on reading speed for 39 German-speaking poor readers in Grades 2 and 3 were evaluated. A *phonological–orthographic association group*, a *reading aloud group*, and a *combined group* were compared in performance with an untrained *control group*. During short-term training, the intervention groups showed higher gains than the control group in reading speed of the trained sub-lexical items and of the words containing the trained segments. No differences were found between the intervention groups. In the development of pseudoword reading, the groups did not show differential improvements. The generalization effect to pseudoword reading was similar, whether the pseudowords contained the trained segment as a syllable or as a non-syllabic letter string. The gains induced by training were specific to the materials used in training and did not induce gains in general reading speed.

Keywords Computer-assisted training · Consonant clusters · Dyslexia · Generalization effects · Reading fluency

The main problem children with dyslexia have, at least in more transparent orthographies (like Dutch, German, or Finnish), is poor reading fluency rather than mere inaccuracy (Holopainen et al. 2001; Landerl et al. 1997; Wimmer 1993; Yap and van der Leij 1993;

S. Hintikka (✉) · M. Aro
Niilo Mäki Institute, University of Jyväskylä, P.O. Box 35, Jyväskylä 40014, Finland
e-mail: sini.hintikka@nmi.fi

S. Hintikka · K. Landerl
University of Salzburg, Salzburg, Austria

K. Landerl
University of Tübingen, Tübingen, Germany

H. Lyytinen
University of Jyväskylä, Jyväskylä, Finland

Ziegler et al. 2003; Zoccolotti et al. 1999). Dysfluent reading is often conceptualized as a failure of visual–orthographic word recognition on the basis of the dual-route model (Castles and Coltheart 1993). Thus, especially in the research literature on English, the enhancement of reading fluency is closely related to the concepts of orthographic learning or to the development of visual word recognition skills. It is proposed that orthographic learning occurs through self-teaching (Ehri 1992; Share 1995); that is, the phonological recoding of new words results in the acquisition of the orthographic representations of letter strings in these words.

The main interest of the present study was to evaluate the mechanisms of training reading fluency for German-speaking poor readers, whose reading accuracy is high and reading slow. Slow reading is highly a persistent handicap (see, e.g., Klicpera and Schabmann 1993, Landerl and Wimmer 2008, Thaler et al. 2004; Torgesen et al. 2001). More specifically, the main focus of the present study was to use sub-lexical items as training material and investigate the nature of generalization effects. To our knowledge, training targeted at the sub-lexical, multiletter level has not been widely evaluated. Typically, in reading intervention studies, the sub-lexical level has involved individual letter-sound correspondences (e.g., Defior and Tudela 1994; Hintikka et al. 2005; Hohn and Ehri 1983, and these in combination with practice in phonological awareness, Ehri et al. 2001; Schneider et al. 2000); however, these studies have been mainly conducted with beginning readers with the aim of improving reading accuracy. The most commonly used method for enhancing reading fluency has been *repeated reading*, which typically consists of the repetition of words or a passage. Previous training studies have revealed that repetitions of words or pseudowords can enhance the reading speed of poor readers (Berends and Reitsma 2006a, b; Judica et al. 2002; Lemoine et al. 1993; Levy et al. 1999; Martin-Chang and Levy 2005; Thaler et al. 2004; Wentink et al. 1997). However, when this kind of training has been targeted at the poorest readers, they have not attained the level of average readers (e.g., Thaler et al. 2004), or the generalization effects have been low (e.g., Berends and Reitsma 2006a, b; Lemoine et al. 1993; Lovett et al. 1990; Martin-Chang et al. 2007; Thaler et al. 2004).

Why select multiletter, sub-lexical units instead of words as targets in training reading fluency? First, children with dyslexia seem to have deficiencies in processing larger orthographic units efficiently (Di Filippo et al. 2006; Spinelli et al. 2005; Ziegler et al. 2003). By contrast, Martens and de Jong (2006) showed that the acquisition of orthographic knowledge in the reading of children with typical reading skills stems from the ability to rely on multiletter features. Second, given the above-mentioned item-specific effects of word naming intervention studies (e.g., Thaler et al. 2004), training of sub-lexical units might induce generalization effects that are not limited to single words. Third, there is a lack of studies investigating the issues of different training units and generalization effects.

In another study, Hintikka et al. (submitted) evaluated the outcomes of computerized training aimed at increasing efficiency of access to multiletter sub-lexical units (word-initial consonant clusters) among German-speaking poor readers in Grades 2 and 4. During computer training, children were offered a large number of repetitions in a task associating orthographic units with their corresponding phonological units. The outcomes of the computer-trained group were compared with the performance of children participating in a paired reading program in which they read books with an adult tutor. The results showed that, in a task in which generalization of the computer trained skill was required (reading *transfer words* containing the trained onset consonant clusters), both groups exhibited a similar improvement in accuracy and speed. Therefore, the generalization effects of the computer training program were low. However, as mentioned earlier, generalization seems

to be a persistent problem in training studies (e.g., Berends and Reitsma 2006a, b; Thaler et al. 2004). Further studies are needed to investigate the transfer or generalization issue, as training is of little relevance if children do not learn to apply the trained skills to everyday reading. An additional question that arose from the study by Hintikka et al. (submitted) was related to the practice done by the paired reading group. The children in the paired reading group read words and texts aloud. This practice might have promoted phonological recoding skills that in turn could have enhanced orthographic representations. Is practice in reading aloud or oral reading—the most commonly used method in reading training—crucial in improving reading fluency?

Two lines of evidence have contributed to our understanding of the role of reading aloud: (1) training studies on orthographic learning and (2) neuroimaging studies. Share (2004) suggests that availability of a familiar phonological and/or oral form could form a base for orthographic learning. Only a few studies have examined silent reading or directly compared reading aloud with another type of practice. Recently, Kyte and Johnson (2006) compared the effects of reading aloud with the outcomes of concurrent articulation on orthographic learning for children with typical reading skills. They found that orthographic learning of pseudowords was more enhanced in a reading aloud condition. Bowey and Muller (2005) and De Jong and Share (2007) found that orthographic learning (in English and in Dutch, respectively) occurs both during oral and silent reading among typical readers. However, the results regarding word naming tasks measuring reading speed are conflicting: In one study, silent reading practice was associated with the naming speed advantage (Bowey and Muller 2005), whereas, in other studies, only oral reading enhanced reading speed or no training effects on speed were found (De Jong and Share 2007; Kyte and Johnson 2006, respectively).

An important question is related to the applicability of the findings for children with reading deficits and especially to their performance in reading speed tasks, which is critical in regular orthographies, as accuracy is often close to the ceiling level even among poor readers (Landerl 2001; Wimmer 1993). Berends and Reitsma (2007) found that reading aloud during fluency training among poor readers did not lead to larger improvements in reading speed compared with silent reading. A similar result was reported by Thaler et al. (2004): No differences emerged between a passive (the naming of consonant cluster segments of trained words was done by the computer) and an active naming (a child named the items) condition in the reading fluency development. In summary, the mechanisms, in which reading aloud during training influences reading speed development (especially among poor readers), are not well understood.

Neuroimaging studies on dyslexia indicate that, in participants with dyslexia, the left inferior frontal cortex is overactivated (e.g., Brunswick et al. 1999; Georgiewa et al. 2002; Grünling et al. 2004). These overactivations might show increased reliance on phonological articulatory decoding and pronunciation assembly to compensate for a deficit in fast orthographic word recognition (Pugh et al. 2000; Sandak et al. 2004). What training implications do these findings have? In training reading fluency, is articulatory decoding a necessary compensatory mechanism for poor readers?

In the present study, we directly contrasted practice in reading aloud with training of the associations between phonological and orthographic units without the requirement of overt naming. In the association task, the child heard a phonological unit through headphones and clicked the corresponding orthographic stimulus, selecting from a number of options presented on the screen. During this type of training, the child was provided with a phonological form. This might be important, as the availability of a phonological form could form a base for orthographic learning (Share 2004) and access to the phonemic level might be fragile for children with dyslexia (De Jong and van der Leij 2003). In addition,

this type of training is appealing from an educational point of view, as it enables independent practice, and thus, the continuous presence of an adult tutor is not required. On the other hand, reading aloud could increase the depth of processing between spelling and phonology and promote phonological recoding, which is seen as a critical prerequisite for orthographic learning (De Jong and Share 2007). In addition, we included a group that practiced with both types of training tasks. Providing the phonological form in addition with practice in reading aloud might be more effective than a mere association or reading aloud task. A control group that did not participate in specific training was taken as a comparison group.

Each training group had word-initial consonant clusters (like *kla* or *stre*), which were followed by a vowel to ensure pronounceability, as training items. Consonant clusters occur frequently in German and have been shown to be difficult for children with dyslexia (Bruck and Treiman 1990; Snowling 1981; Treiman 1991). Choosing the cluster onsets as training material enabled us to compare the results of the present study with the outcomes of the study by Hintikka et al. (submitted). Two features were of special importance in the design of the present study. First, in each training condition, the computer practice was designed to emphasize speed of responding. Second, the congruency between training and testing was taken into account. Martin-Chang et al. (2007) pointed out that transfer increases when the same cognitive processes are engaged during training and transfer tasks. Instead of implementing a list-reading task, as we had done in the Hintikka et al. (submitted) study, we applied a computerized item naming task presenting each item individually, thus increasing the congruency between the training program and testing. But note that the training and outcome measures were not identical.

An important question investigated in the present study was the role of syllables in facilitating generalization effects. Empirical evidence shows that at least some of the phonological influences in visual word recognition operate at the level of sub-lexical units (e.g., Cole et al. 1999; Ferrand and Grainger 1992; Perfetti and Bell 1991). This owes to the fact that the phonological system is structured before reading acquisition (Ziegler and Goswami 2005). The precise nature of phonological influence at the sub-lexical level is not well known. In spoken language, the most basic units are syllables (Ziegler and Goswami 2005). However, the critical unit used in visual word recognition may vary across languages along with differences in the forms of reading instruction used as well as in the phonological structure of the language, in the consistency of its orthography and in its phonological and orthographic neighborhood characteristics (Cole et al. 1999; Ziegler and Goswami 2005). In German, the syllable structure is rather complex; however, there is some evidence that syllables play an important role in the visual word recognition of skilled adult readers (Stenneken et al. 2007). We wanted to evaluate whether training effects are facilitated when the trained segments are the initial syllables of pseudowords as compared to the pseudowords containing the trained segments as non-syllabic letter strings. This kind of comparison using identical items across the two conditions and controlling for the word length was only possible with pseudowords.

In the present study, the following research questions were examined: (1) What are the outcomes of practice in the correspondences between phonological and orthographic units compared to reading aloud on reading speed? The results of three intervention groups (association, reading aloud, and combined) will be compared with the results of a control group that does not receive any specific training. (2) Does training aimed at the sub-lexical level induce improvements in the reading of sub-lexical units? We evaluate whether the reading of transfer words and pseudowords containing the trained sub-lexical segments improves during the training period. (3) Is generalization facilitated when the trained

segment (e.g., *fla*) is the first syllable of a pseudoword (*fla-sof*) as compared to a non-syllabic letter string (*flas-to*)? If the syllable plays an important role in transfer effects in reading fluency, the generalization effect of training should be greater to pseudowords in which the trained segment is a syllable. (4) Are gains in reading skills specific to the trained material, or does training induce improvements in general reading speed?

Method

Participants

The participants were selected among 20 classrooms of second and third graders from five elementary schools in Salzburg in Austria. Permissions for all children were obtained both from parents and from schools. The central criterion for selection for the study was a reading speed of below the 25th percentile in an individually administered subtest (a list of high frequency words) of the “Salzburger Lese- und Rechtschreibtest” (Landerl et al. 1997).¹ All poor readers had either been nominated by their classroom teacher or performed poorly (at least 1 SD below the age norm) in a classroom reading fluency test (Mayringer and Wimmer 2003).

After the assessment, 31 children were randomly assigned to three training groups. Ten children were assigned to association and reading aloud groups and 11 children to the combined group. We decided to include each child in the training programs because it was likely that, due to attrition (illness, missing data), we would have to exclude some children from the final analyses. Eventually, the data of one girl (in the association group) had to be excluded because of voice-key malfunction. Reanalyzing the data so that one participant was excluded from the combined group to achieve a more balanced design did not change the results. Thus, we decided to report the results of the original analyses. The *association group* ($n=9$) practiced associations between orthographic and phonological units. The participants in the *reading aloud group* ($n=10$) read the trained items aloud. For the *combined group* ($n=11$), practice in both association and oral reading was provided. We checked that the children in the three groups had a similar pretest level in the reading task of high frequency words and in nonverbal intelligence quotient (IQ) assessed with *Raven's Coloured Progressive Matrices—German version* (Bulheller and Häcker 2002).

The nine *control children* were selected from another project conducted at the same time at the University of Salzburg. These children attended different schools and participated in a comprehensive neuro-cognitive test battery. In the context of the present study, they received all pre- and posttest measures, but did not receive any particular treatment during the training period. We acknowledge that this is not an ideal assignment procedure as it is not fully randomized. However, selecting poor readers as an untrained control group from the very same classrooms as the training groups would have involved serious ethical and practical dilemmas. It would have been difficult to justify to teachers and parents to whom intervention would be provided and to whom not. We decided that matching the initial level between the groups would be sufficient to enable comparisons between the groups. The detailed background characteristics of the groups are presented in Table 1. Importantly, statistical analyses of the pretest

¹ Two children had a reading speed above the 25th percentile in the list of frequent words; however, they were included in the participant group as their performance in two other subtests was low (<25th percentile).

Table 1 Descriptive characteristics (standard deviations in parentheses) of children in the training and control groups

Variable	Association (<i>n</i> =9)		Reading aloud (<i>n</i> =10)		Combined (<i>n</i> =11)		Control (<i>n</i> =9)	
Boys/girls	4/5		8/2		2/9		3/6	
Grade (2/3)	5/4		5/5		6/5		2/7	
Age	8; 7	(0.7)	8; 11	(0.6)	8; 7	(0.7)	8; 11	(0.5)
Raven (%ile)	31.9	(27.2)	36.9	(27.7)	33.2	(27.4)	38.3	(25.7)
Reading speed measures								
Frequent words (%ile)	17.0	(6.3)	17.4	(14.8)	19.9	(14.8)	17.7	(12.8)
Short text (%ile)	16.1	(7.7)	12.5	(9.1)	17.6	(11.6)	18.0	(11.3)
Word-like pseudowords (%ile)	30.4	(16.9)	26.4	(15.3)	31.9	(21.2)	22.7	(15.0)
Word-unlike pseudowords (%ile)	23.8	(15.1)	26.6	(19.0)	28.4	(18.5)	19.7	(16.2)

The groups did not differ significantly ($p>0.05$) at the pretest.

measures revealed no significant differences between the groups in any of the measures (all $ps>0.10$). All children were instructed within regular classroom settings, and all the participants with German as a second language had received all their formal education in German.

Design and materials

The study was carried out from March to May and consisted of a pretest, a training period, and a posttest. After the screening and individual assessment, the participants in the training groups attended training on six successive school days. The posttests were conducted on Day 1 or Day 2 after the final training session. Altogether, training was provided over a period of 8 days (including one weekend). As a few children were unable to attend training on successive school days, their training period was slightly longer (10 days). For the control group, the tests were also administered within a time interval of eight to ten school days corresponding to the time interval of testing for the training groups.

Training program and procedures

The computerized training program was developed at the University of Jyväskylä (see Hintikka et al. 2005; or Lyytinen et al. 2007, for a description of the program), translated into German, and adapted for the purposes of the study. The stimulus material consisted of 28 multiletter, sub-lexical items. These items were formed from high frequency word-initial consonant clusters (*kr-*, *fl-*, *str-*, and *schl-*) with a vowel added (e.g., *kra*, *fle*, *stro*, *schlei*). The same onset clusters were used as training material in previous studies (Hintikka et al. submitted; Thaler et al. 2004). The items were divided into three levels to maintain the attention span of the children. The first level consisted of the onset clusters *kr-* and *fl-* with one of five vowels added (*-a*, *-e*, *-i*, *-o*, *-u*), the second level consisted of the consonant clusters *str-* and *schl-* with one of five vowels, and the third level of these four onset clusters with a diphthong (*-au*, *-ei*, *-eu*). Each level was played in lower case letter formats. On each level, the trained item was practiced six times, and during each session, each level was played through once. Hence, during the training period (6 days), the total number of repetitions of the trained onset cluster syllables was 36. The duration of each training session was 15 to 20 min. The training was conducted by the first author and trained MA students. The children practiced one-on-one with an adult tutor during normal school hours in a separate quiet classroom.

In each training group, the following features were the same. The trained items were always embedded in ‘balls’ falling downward on the screen. This type of task required rapid responding: In the case of a slow response (the ball hit the ground), a competitor snatched the ball. Additionally, the speed of responding was emphasized by dividing the screen into two parts: When the correct item was recognized in the upper part, two points were given, and when in the lower part, only one point was given. The speed at which the balls fell was initially set low. As the game proceeded, a greater challenge was introduced by increasing the speed of presentation. More specifically, during the first three sessions, the maximum presentation time was 9 s (4 s in the upper half), and during the last three sessions, the maximum presentation time was 6 s (3 s in the upper half).

Association group In the program administered to the association group, a single auditory stimulus was delivered (via high quality headphones) concurrently with four orthographic items (target and three distractors) that appeared at the top of the screen. The orthographic items embedded within balls immediately began to drop downward on the computer screen, and the player’s task was to home in on the relevant orthographic item and to ‘catch’ it by clicking the mouse. If the player did not identify the correct spelling before the ball hitting the ground or erroneously clicked on the incorrect spelling, the target item was repeated in the next trial, and the correct response was color-highlighted.

Reading aloud group In this program, four orthographic items appeared at the top of the screen, again embedded within balls, which immediately began to drop downward on the computer screen. The player’s task was to read aloud the item written in the color-highlighted ball. The experimenter marked the response accuracy by clicking the mouse. In the case of an incorrect response, the correct color-highlighted ball was halted for couple of seconds.

Combined group In the combined group, both association and reading aloud tasks were done. During each session in both task types, the participants practiced the trained item three times. The association task was practiced first; then, at a separate level, the same items were read aloud.

Test material and procedures

Experimental task The experimental task consisted of the 28 trained onset cluster syllables (onsets *kr-*, *fl-*, *str-*, and *schl-* with vowels *-a*, *-e*, *-i*, *-o*, *-u* and diphthongs *-au*, *-ei*, *-eu*), 20 words containing the trained segments (as in the word *Kralle*), and 40 pseudowords containing the trained segments (as in the pseudowords *krakal* or *fleinte*). Examples of the tasks and materials are shown in Table 2. The words have been used earlier in a study by Thaler et al. (2004), and they were all mono- or bisyllabic and in the vocabulary of an average 8-year-old child. The pseudowords were all bisyllabic and belonged to one of two categories: (1) pseudowords with the trained segment as the initial syllable (as in *krakal* or *fleitel*), (2) pseudowords in which a trained segment occurred but did not correspond to a syllable (as in *kralka* or *fleinte*).

The items in the experimental task were presented individually in the middle of a laptop screen in the yellow 40-point font that is used in Austrian school books on a black background. The child sat in front of the computer with the experimenter next to him or her. Three practice items were presented first, after which the test items were administered in a random order. The task was to read the items aloud as quickly and accurately as possible. All trials started with a fixation cross to focus the child’s attention. After the fixation cross,

Table 2 Description and examples of the outcome measures

	Materials	Number of items	Task
Trained items	E.g., kra, fle, stro, schlei	28	Naming items on computer screen; voice-key measurement
Words containing the trained segments	E.g., Kralle, flechten, Strick, Schlamm	20	Naming items on computer screen; voice-key measurement
Pseudowords containing the trained segment as the first syllable	E.g., krakal, fleitel, strafon, schlitun	20	Naming items on computer screen; voice-key measurement
Pseudowords containing the trained segment as a non-syllabic letter string	E.g., kralka, fleinte, strafon, schliktu	20	Naming items on computer screen; voice-key measurement
Control task: Frequent words	E.g., Mama, Baum, Katze, Brief	30	List-reading task

there was a pause for 500 ms, after which the stimulus item appeared. The item was remained on-screen until the experimenter marked the response accuracy by clicking the mouse. There was a pause for 700 ms between the experimenter clicking the mouse and the next trial. The responses were recorded for subsequent scoring. The voice key registered responses, measuring both the time from target onset to response onset (onset time) and the time from target onset to mouse clicking (offset time). No feedback was given.

Standardized reading test These subtests were taken from a standardized reading test (Landerl et al. 1997). The test requires a child to read aloud lists of words, pseudowords, and a short text. The reading tasks were preceded by practice sheets. After the practice items, the experimenter presented the lists to the child, simultaneously started a stopwatch and stopped it when the last item had been attempted. The list-completion time and the number of items read incorrectly were recorded. Only the subtest presenting short high frequency words was conducted both in the pre- and posttest. In the posttest, a parallel list version, in which the words were different but controlled for frequency and length, was used. Examples of the materials are shown in Table 2. The other subtests were administered only in the pretest.

Nonverbal abilities Nonverbal IQ was assessed with Raven's Coloured Progressive Matrices—German version (Bulheller and Häcker 2002).

Results

Reading accuracy of the trained items, transfer words, and transfer pseudowords

Reading accuracy was generally high. Percentages of correct readings in reading the trained items, words, and pseudowords with the trained segments at the pretest varied between 88.8% (SD=7.1%) and 96.1% (SD=4.9%). At the posttest, the percentage accuracy was between 89.0% (SD=7.3%) and 95.9% (SD=3.8%). The results are presented separately

Table 3 Reading accuracy percentages (standard deviations in parentheses) of the trained items and words with the trained segments by group

Group	Trained items				Words with the trained segments			
	Pretest		Posttest		Pretest		Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Association	92.8	(5.7)	94.0	(4.0)	95.6	(4.6)	95.0	(5.6)
Reading aloud	91.9	(6.2)	93.9	(6.3)	96.0	(5.2)	94.4	(8.2)
Combined	92.2	(7.4)	92.0	(5.2)	93.2	(4.6)	95.9	(3.8)
Control	92.2	(4.5)	94.3	(3.5)	96.1	(4.9)	93.3	(6.6)

for the item types and groups in Tables 3 and 4. These high accuracy results are typical for German-speaking readers and do not justify further analyses.

Reading speed of the trained items, transfer words, and transfer pseudowords

The experimental task in the present study included measures of the voice onset times (the time between the target onset and response onset) and the offset times (the time from target onset to mouse clicking). Although, typically, onset times are used, we chose to analyze the offset times, which include naming durations. It has been suggested that inclusion of the naming duration could be a more sensitive and reliable measure of decoding efficiency for children (De Jong and Share 2007; Thaler et al. 2004). In the present study, the offset times showed stronger correlations with a standardized list-reading test than did the onset times. The correlation between the standardized reading subtest of frequent words and the offset times of the experimental word reading task in the pretest was $r(39)=0.72$. The correlations between the standardized subtests of reading pseudowords and reading pseudowords in the experimental task varied between 0.79 and 0.85. The onset times were not normally distributed and were therefore log transformed. The correlations that were based on these measures were 0.40 for word and 0.52 and 0.56 for pseudoword reading.

As it was technically impossible to analyze the offset times of the voice recordings, the mouse clicking by the experimenter was used as the offset time. The mouse clicking was considered to be a reliable measure, as (a) the same experimenter administered all the tests, (b) the offset times for the words with the consonant clusters were comparable to the offset times in the study by Thaler et al. (2004) who also had German-speaking poor readers as

Table 4 Reading accuracy percentages (standard deviations in parentheses) of the pseudowords with the trained segments by group

Group	Trained item as a syllable				Trained item as a non-syllabic letter string			
	Pretest		Posttest		Pretest		Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Association	90.9	(4.3)	93.0	(3.2)	91.9	(5.0)	91.0	(5.2)
Reading aloud	90.4	(6.8)	92.3	(6.5)	90.1	(7.4)	91.5	(8.0)
Combined	89.1	(7.9)	90.7	(5.7)	90.3	(9.6)	89.0	(7.3)
Control	89.7	(5.1)	92.3	(3.0)	88.8	(7.1)	92.4	(5.3)

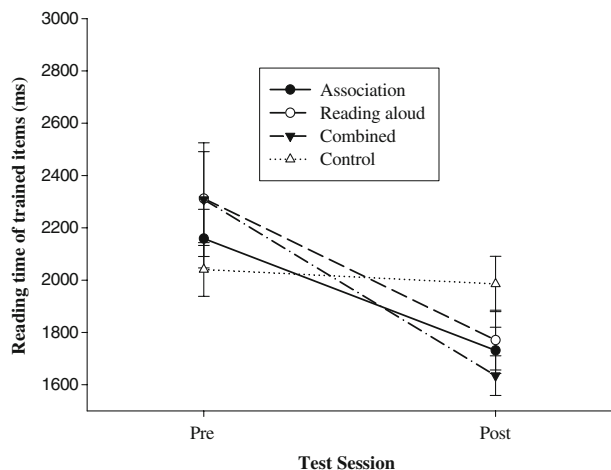
participants and used identical words as stimulus items, and (c) the correlations between the list reading tests and offset times were high.

The response times of incorrect trials and invalid trials (e.g., coughing) were excluded from the analyses. For both pre- and posttest, this was necessary for less than 10% of all responses with no significant differences between groups, all $F_s < 1$. Means and standard deviations were computed separately for each child for the trained onset cluster syllables, the words containing the trained segments, and the two categories of pseudowords. Response times that differed by more than two standard deviations from the child's mean were considered to be outliers and were corrected within the range of two standard deviations.

Trained items The offset times of the trained items were subjected to analysis of variance (ANOVA) with group (association, reading aloud, combined, control) as the between-subjects factor and test session (pre-test, posttest) as the within-subjects factor. The test session effect was significant $F(1, 35) = 49.89, p < 0.001, \eta_p^2 = 0.59$. The groups performed at a similar level, as the main effect of a group was not significant, $F < 1$. The most important result was that the group \times test session interaction effect was significant, $F(3, 35) = 4.86, p < 0.01, \eta_p^2 = 0.29$. Inspection of Fig. 1 shows that children in all three training groups increased their reading speed from pre- to posttest, while in the control group, reading speed remained largely unchanged. This was confirmed by the planned paired sample t -tests, showing that the children in the association group, reading aloud group, and the combined group improved in reading speed, $t(8) = 6.18, p < 0.001; t(9) = 5.35, p < 0.001$; and $t(10) = 3.81, p < 0.01$, respectively. However, the reading speed of the children in the control group did not increase during the training period, $t(8) = 1.09, p > 0.31$. The ANOVA was repeated without the control group. The test session effect was again significant, $F(1, 27) = 51.2, p < 0.001, \eta_p^2 = 0.66$, but now, neither the group effect nor the interaction effect was reliable, $F_s < 1$, indicating that the gains were comparable between the training groups.

Transfer words containing the trained segments Another four (group) by two (test session) ANOVA was carried out for words containing the trained segments. The test session effect was significant, $F(1, 35) = 50.72, p < 0.001, \eta_p^2 = 0.59$, and the main effect of group was not, $F < 1$. The interaction was significant, $F(3, 35) = 4.18, p < 0.05, \eta_p^2 = 0.26$, indicating

Fig. 1 Mean reading times by group at the pre- and posttest for trained sub-lexical items. The vertical lines depict standard errors of the means



differential improvement among the groups. Figure 2 indicates and the planned paired sample *t*-tests confirmed that the three training groups showed an increase in reading speed from pretest to posttest [for the association group, $t(8)=6.63$, $p<0.001$, reading aloud group, $t(9)=3.19$, $p<0.05$, and for the combined group, $t(10)=5.36$, $p<0.001$]. The children in the control group did not improve their reading speed, $t(8)=0.63$, $p>0.54$. The same ANOVA was repeated without the control group. In this analysis, only the test session effect was significant, $F(1, 27)=62.6$, $p<0.001$, $\eta_p^2 = 0.70$, while the group effect and the interaction were not, $F_s<1$, showing that there were no statistically significant differences between the training groups with respect to their reading level and reading speed gains during the training period.

Transfer pseudowords containing the trained segments Two types of pseudowords were included in the analyses: pseudowords containing the trained segments as first syllables and pseudowords containing the trained segments as non-syllabic letter strings. The offset times of both types of pseudowords were subjected to separate ANOVAs with group (association, reading aloud, combined, control) as the between-subjects factor and test session (pre-test, post-test) as the within-subjects factor. The participants improved their reading speed in both types of pseudowords during the training period, for the pseudowords containing the trained segments as first syllables, $F(1, 35)=13.28$, $p=0.001$, $\eta_p^2 = 0.28$, and for the pseudowords containing the trained segments as non-syllabic letter strings, $F(1, 35)=21.60$, $p<0.001$, $\eta_p^2 = 0.38$. Neither the main effect of a group nor the test session \times group interaction effect were significant, $F_s<1$. Although Figs. 3 and 4 indicate more rapid development in reading speed in the training groups compared to the control group, the statistical analyses indicate improvements in all groups. Large standard deviations in the rather small samples may explain this nonsignificance of the interaction effect. In addition, the control group also showed a small improvement. When the same ANOVAs were rerun without the control group, only the effect of test session was significant, for the pseudowords containing the trained segments as first syllables, $F(1, 27)=18.5$, $p<0.001$, $\eta_p^2 = 0.41$, and for the pseudowords containing the trained segments as non-syllabic letter strings, $F(1, 27)=25.9$, $p<0.001$, $\eta_p^2 = 0.49$. As neither the group effect nor the interaction were significant, $F_s<1$, there were no differences between the training groups in their reading level or in the improvement during the training period.

Fig. 2 Mean reading times by group at the pre- and posttest for words containing the trained segments. The vertical lines depict standard errors of the means

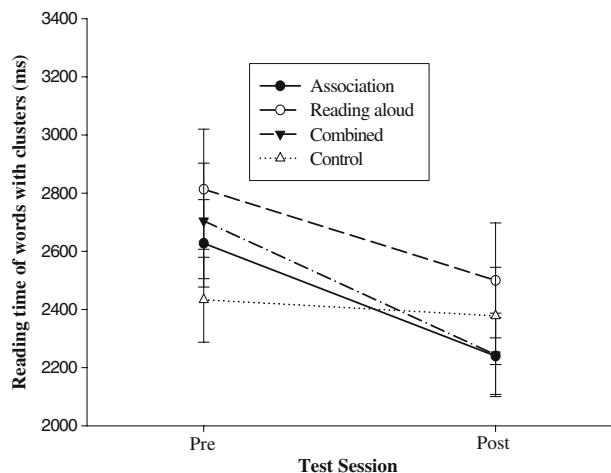
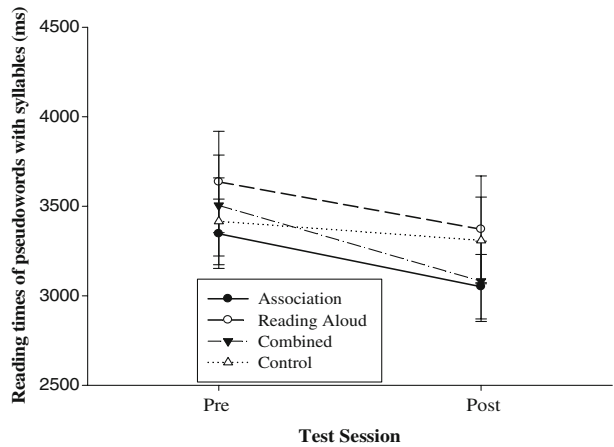


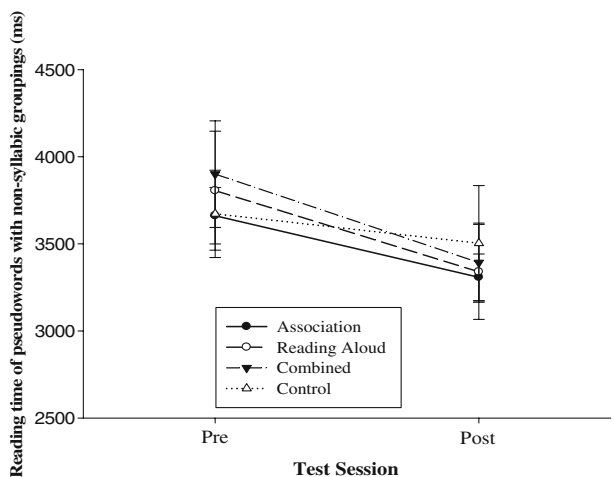
Fig. 3 Mean reading times by group at the pre- and posttest for pseudowords containing the trained segments as syllables. The vertical lines depict standard errors of the means



Was the generalization effect facilitated when the trained segment corresponded to a syllable of the pseudoword? As there were no statistically significant differences between the three trained groups in pseudoword reading, we decided to combine them to form one single training group. We wanted to find out whether the generalization effect to the pseudowords containing the trained segment as the first syllable would differ from that to the pseudowords in which the trained segment was a non-syllabic letter string. The offset times of the pseudowords were subjected to ANOVA with the pseudoword type (trained segments as syllables, trained segments as non-syllabic letter strings) and test session (pre-test, post-test) as within-subjects factors. The analysis disclosed a significant improvement in the reading times of pseudowords, $F(1, 58) = 48.0$, $p < .001$, $\eta_p^2 = .45$. However, differences in the reading speed of the two pseudoword types throughout the study, and more importantly, in the development of reading speed between the pseudoword types were not found, both $F_s = 1$.

Percentages of children showing gains in the groups To analyze the changes in the groups more closely at the individual level, we counted how many children improved their reading

Fig. 4 Mean reading times by group at the pre- and posttest for pseudowords containing the trained segments as non-syllabic letter strings. The vertical lines depict standard errors of the means



speed from pre- to posttest by at least 10% (see Table 5). It was observed that, among the trained groups for the trained sub-lexical items, the percentages of children showing gains were quite high (approximately 80%), whereas in the control group, only one participant showed speed-related gains (11.1%). As expected, the percentages of children benefiting from training were lower for the transfer items: For words, the percentages varied between 60% and 72.7%, for pseudowords between 33.3% and 70%. In word reading in the control group, two out of nine (22.2%) and, in pseudoword reading, three children out of nine (33.3%) showed speed-related gains.

A control task: reading speed of frequent words

To analyze whether the gains induced by training were specific to the material used in the training or whether training produced improvements in general reading speed, we analyzed the development of reading speed in the subtest of high frequency words from the standardized reading test (Landerl et al. 1997). The reading times were subjected to ANOVA with group (association, reading aloud, combined, control) as the between-subjects factor and test session (pretest, posttest) as the within-subjects factor. Neither the test session effect nor the group effect was significant, $F_s \leq 1$. The group \times test session interaction failed to reach the conventional levels of statistical significance, $F(3, 35) = 2.17$, $p = 0.11$, $\eta_p^2 = 0.16$. Thus, the statistical analysis indicates that training did not induce gains in general reading speed. Inspection of Fig. 5 shows a small trend favoring the combined group; however, this improvement was not powerful enough to produce statistically significant effects. When the ANOVA was rerun without the control group, neither the test session effect nor the group effect was significant, $F_s \leq 1$. However, this time, the interaction effect was significant, $F(2, 27) = 5.1$, $p < 0.05$, $\eta_p^2 = 0.28$. Paired sample t -tests showed that there was a reliable improvement in the combined group, $t(10) = 4.91$, $p < 0.01$, while in the association and in the reading aloud group, no improvements were found, $p_s > 0.40$.

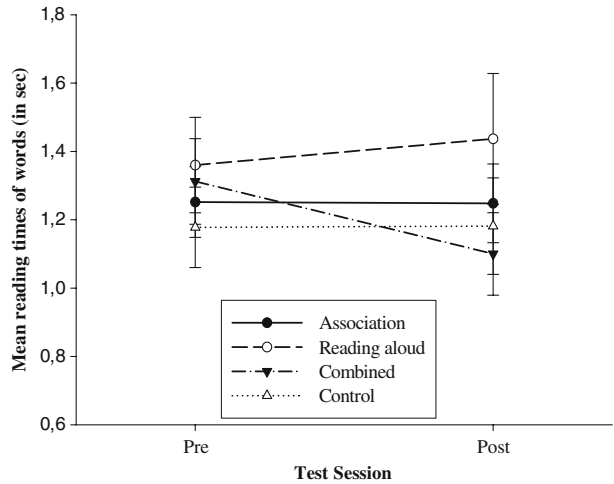
Effect sizes of the training programs

We calculated effect sizes (Cohen's d) for the training conditions by computing the difference between the training and control group at the post-test subtracted by the between-group mean difference at the pre-test, which was corrected for the correlation between the pre- and post-test. The mean difference was then divided by the pooled standard deviation. The effect sizes for each outcome measure by group are shown in Table 6. Using Cohen's (1988) criterion of .80 as a large effect size, .50 as a medium effect size and .20 as a small effect size the training conditions produced large effect sizes on trained items. On transfer

Table 5 Percentages of children showing improvements in reading speed (at least 10% gain compared to the pretest performance) by group

Group	Trained items	Words with trained segments	Pseudowords with syllables	Pseudowords with letter strings
Association	77.8	66.7	55.6	33.3
Reading aloud	80.0	60.0	60.0	70.0
Combined	81.8	72.7	54.5	63.6
Control	11.1	22.2	33.3	33.3

Fig. 5 Mean reading times by group at the pre- and posttest in a standardized reading test of frequent words. The vertical lines depict standard errors of the means



words containing the trained items the effect sizes were mainly medium. On reading transfer pseudowords the effect sizes were small to medium. On the control task of reading high-frequency words, the training conditions produced mainly small effect sizes.

Discussion

The main interest of the present study was to evaluate the outcomes of reading fluency training among German-speaking poor readers in Grades 2 and 3. More specifically, the main aim was to use multiletter, sub-lexical items (onset cluster syllables, like *kra*) as training material to investigate the nature of generalization effects. The performance of an association group (practicing phonological–orthographic correspondences without the requirement of oral articulation), a reading aloud group, and a combined group on the reading speed of the trained sub-lexical items, transfer words, and pseudowords containing the trained segments were compared with the outcomes of an untrained control group. During training, the intervention groups showed higher gains in the trained sub-lexical items and in the transfer words than the control group (effect sizes were between 0.34 and 1.77 for the training groups). No statistically significant differences were found between the three intervention groups. In the development of pseudoword reading, the interventions did not produce higher gains compared to the performance of the control group. The generalization effect of training to pseudoword reading was similar, whether the pseudowords contained the trained segment as a syllable or as a non-syllabic letter string. Finally,

Table 6 Effect sizes (Cohen's *d*) for outcome measures and control task by group

Group	Trained items	Words with trained segments	Pseudowords with syllables	Pseudowords with letter strings	Control: frequent words
Association	1.11	0.65	0.47	0.25	0.15
Reading aloud	1.11	0.34	0.12	0.31	-0.13
Combined	1.77	0.74	0.41	0.36	0.42

reading speed in a standardized subtest did not improve significantly; thus, the gains induced by training were specific to the items containing the trained materials and did not produce improvements in general reading speed.

An important result was that reading aloud practice was as effective as the training of correspondences between the phonological and orthographic units in terms of reading outcomes. With respect to effect sizes, the combined group, which received practice both in phonological–orthographic correspondences and reading aloud, showed the highest gains in most experimental tasks; however, no statistically significant differences between the training groups emerged. Share (2004) has suggested that the availability of a familiar phonological/oral form might be a significant factor in orthographic learning. In the present study, both oral articulation by the child and the phonological form provided by the computer were associated with similar gains in the development of reading speed. The findings of the present small-scale study are consistent with the results of previous experiments: Compared with silent reading or a more passive training method (the computer providing the phonological form), reading aloud does not lead to any additional advantages in reading fluency training (Berends and Reitsma 2007; Thaler et al. 2004). In summary, on the basis of the findings of these short-term intervention studies, the requirement of oral reading is not necessary in training reading fluency. However, in the previously mentioned studies and in the present experiment, the comparison condition for reading aloud involved practice in the connections between print and phonology or exposure to the printed words (silent reading). It has been claimed that, when more able readers encounter printed words, the phonology of those words becomes automatically activated, regardless of whether the words are read aloud or silently (Booth et al. 1999, see also Berends and Reitsma 2007); thus, during reading exercises without the requirement of overt naming, phonology could have been activated, even in the poor readers of the present study. This might explain the lack of differences between the training conditions.

Sandak et al. (2004) compared the behavioral and neurobiological outcomes of three training conditions (orthographic, phonological, or semantic) on learning new words. Phonological training showed advantages over the other conditions, which, according to the authors, indicated that phonological training facilitated phonological assembly, recoding, and rapid mapping of orthography to phonology. Neurobiologically, phonological training contributed to activation in the main reading regions, also in the regions related to fluency in word recognition. Thus, Sandak et al. suggest that, to optimize learning, initial and remedial reading instruction must explicitly draw attention to the phonological properties of words. In the present study, each training condition was related to phonological features. However, some children practiced reading aloud, thereby mapping orthography to phonology, while some children worked in the reverse order, that is, they mapped phonology to orthography without pronunciation. As both types of training resulted in improvements, we conclude that, in practicing reading fluency, the role of rapid mapping between phonology and orthography may be crucial.

Altogether, 36 repetitions during six school days were enough to produce an improvement in the reading speed of the trained sub-lexical items. At the individual level, approximately 80% of the children in the intervention groups showed gains (at least a 10% increase in reading speed compared to the initial level), while in the control group, only one participant (11.1%) was able to increase his or her reading speed. To our knowledge, previous training studies have not directly been targeted at the multiletter, sub-lexical level; thus, it is a new finding that the direct training of these items is effective. The result is important, as children with dyslexia seem to have deficiencies in processing larger orthographic units efficiently (e.g., Di Filippo et al. 2006; Spinelli et al. 2005; Ziegler et al. 2003).

However, training in sub-lexical units is only efficient if children are able to apply this acquired knowledge in word recognition, as faster and more accurate word recognition is a prerequisite for growth in reading fluency (e.g., Perfetti 1992). One of the most important results of the present study was that it was possible to achieve generalization from sub-lexical level to words containing the sub-lexical items, as the interventions were associated with better learning of the transfer words than the control condition. In training programs focused on reading fluency, generalization effects to untrained material have not been widely evaluated (Berends and Reitsma 2006b), and in the few studies in which generalization has been examined, the effects have been low or absent (e.g., Berends and Reitsma 2006a, b; Lemoine et al. 1993; Lovett et al. 1990; Martin-Chang et al. 2007; Thaler et al. 2004). Typically, the results of repeated word naming studies have indicated that the effects of training are item specific; that is, the repetition of words enhances word-specific orthographic representations (e.g., Berends and Reitsma 2006b; Kuhn and Stahl 2003; Thaler et al. 2004). However, from the instructional point of view, this type of training as a remedial one-to-one tutoring program with word-specific outcomes is a time-consuming task. The results of the present study suggest that training at the sub-lexical level can transfer to word reading and lead to generalization effects that are not limited to single words, as the trained sub-lexical units are contained in several words.

Hintikka et al. (submitted) similarly evaluated the outcomes of computerized training aimed at increasing the efficiency of access to multiletter, sub-lexical units among German-speaking poor readers in Grades 2 and 4. The outcomes of the computer-trained group were compared with the performance of children participating in a paired reading program in which they read books with an adult tutor. The computer program used in that study was similar to the training program of the association group in the present study. In the study by Hintikka et al. (submitted) the generalization effects of the computer training program were low, as in a task measuring generalization of the computer trained skill (reading a list of *transfer words* containing the trained segments) both the computer and the paired reading group exhibited a similar improvement. The results were to some degree different from those in the present study in which generalization to the transfer words was found. Three possible reasons for the conflicting results can be suggested. First, the comparison group in the study by Hintikka et al. (submitted) practiced general text-reading skills with an adult tutor. Contrary to our expectations, this type of nonspecific intervention was associated with improvements in the task of reading words with the onset consonant clusters (onset consonant clusters were the focus of the computer program). In the computer group, 55.6% of the participants improved their word reading speed with the trained clusters (at least 10% faster compared to the pretest performance), whereas in the paired reading group, 33.3% showed speed-related gains. In the present study, using the same criteria to evaluate improvements, it was observed that approximately 60% to 73% of the children in the training groups increased their word reading speed as against 22.2% in the untrained control group. These results prompt two conclusions: First, the present interventions seemed to produce higher gains than the computer training in the study by Hintikka et al. (the possible reasons for this result are discussed below), and second, paired reading practice was more effective than no treatment at all. It seems to be possible to enhance the reading skills of poor readers with several types of reading-related practice.

A second possible reason for the fact that we were able to induce generalization in the present but not in our earlier study is that, in the present study, the computer program was modified to encourage the participants to respond quickly; thus, the children may have been more aware of the importance of fast responding. Third, the experimental task in the present study was more congruent with the training program than in the earlier study by Hintikka et al.

(submitted). In this study, a computer task with an isolated item-naming task was used to measure the outcomes of training, whereas in the other study, a list-reading task was administered. Martin-Chang et al. (2007) note that transfer increases as congruency between the training and transfer tasks increases. In the light of this *transfer-appropriate processing hypothesis* (Martin-Chang et al.), the positive outcomes in the present study were probably facilitated because of the higher congruency between the experimental and transfer tasks. However, it must be noted that the experimental task was not identical with the training programs and that the most commonly used task in measuring word reading speed (naming items with a voice key) was employed. Therefore, it is not likely that the effects were merely due to improvements in the computer task itself and not in reading fluency. A transfer task that is congruent with the training program is important to measure the specific outcomes of training; however, it is equally important to evaluate the effects of training on everyday reading away from the training settings (see Olson et al. 1997). Further studies are needed to evaluate the extent to which and under which conditions training can produce transfer effects.

The results of the present study in terms of generalization effects were only partially successful, as in reading pseudowords, no specific training-induced gains were found. Although the trained groups increased their pseudoword reading speed, ANOVAs showed for the control group comparable improvements. Although direct comparison between words and pseudowords was not the main focus of the present study, the results showed differences in transfer effects to word as compared to pseudoword reading. There are three possible reasons for that difference. First, as we were interested in creating two categories of pseudowords (pseudowords containing the trained segment as a syllable or as a non-syllabic letter string), the word and pseudoword items were not matched with respect to stimulus length. The average word length for words was 6.5 letters and for pseudowords 7.3 letters, which may have had an effect on generalization. Second, Share (1999, 2004) has claimed that word meaning per se should not play a significant role in word-specific orthographic learning. On the other hand, in the study by Hintikka et al. (submitted), ‘print exposure’ did not enhance a rapid phonological recoding strategy in reading pseudowords, but helped pupils to gain more rapid access to words. On the basis of the present results, it is not possible to decide whether the difference in terms of transfer effects to words and pseudowords is explained by word meaning per se or by stimulus length. Third, the failure to achieve an interaction effect between test session and group in pseudoword reading might have a statistical explanation: The groups were small, and there were large standard deviations within them. Thus, the difference between the intervention groups and the control group would have had to be large to produce a statistically significant effect. With respect to effect sizes, the training groups showed mainly small effect sizes on pseudoword reading (varying between 0.12 and 0.47).

The generalization effect of training on pseudoword reading was similar for both the pseudowords containing the trained segment as a syllable and those in which it was a non-syllabic letter string; thus, syllable boundaries did not play a significant role in generalization. As noted before, the critical unit used in visual word recognition may vary across languages. In German, the syllable structure is rather complex (e.g., Stenneken et al. 2007; Ziegler and Goswami 2005), which might reduce the possibility that the syllable is an important unit in visual word recognition. However, Stenneken et al. (2007) found that syllables play an important role in visual word processing among German-speaking skilled adult readers. The present results are not incompatible with the previous findings, as in the present study, the role of syllables was evaluated in relation to their effects on enhancing *generalization*. In addition, the present results concern young poor readers who presumably still rely to some extent on a serial grapheme–phoneme decoding strategy in reading pseudowords.

The results showed that no statistically significant improvements among the participants as a whole appeared in reading a list of frequent words. There was a trend indicating that the combined group exhibited gains, whereas other groups did not improve; however, this trend did not reach the level of statistical significance when all the four groups were included in the analysis. This result was not surprising, as it was rather unlikely that, after a training period of only 2 weeks (and using specific training materials), a large improvement in a control task, which was a standardized list-reading test, would be found. An improvement could have occurred, if the training programs would have had effects on general factors, like motivation or global reading strategy (e.g., facilitating a more efficient processing or recognition of all the clusters of a written language). The findings extend the earlier conclusions: Repeated reading of words or sub-lexical items does not enhance the general reading strategy, but rather the orthographic processing or recognition of the trained items. Thus, the improvements induced by repeated reading training can be described as *specific to the trained material*. If significant improvements in general reading speed are aimed for, the intervention duration should be much longer, and the extent of training materials should be greater.

One limitation of the present study is that, due to ethical and practical limitations discussed in “*Method*,” random assignment of participants could only be realized for the three training groups, but not for the control group. Random assignment is especially critical if differences between schools and teachers in their instructional methods and educational context are supposed to play a role in the development of reading skills. In the present study, reading development was evaluated during a time interval of eight to ten schooldays; therefore, it was not expected that the instructional methods would have a large effect on reading development during this short period. Additionally, the groups were matched according to their pretest level. Because of the short time interval between testing and nonsignificant differences between the groups in the pretest, we conclude that it is unlikely that the differential development between the groups would be an effect of school factors rather than of training. Two further limitations of the present study that should be mentioned are a short training period and small sample size, which might have obscured differences between the training groups. To allow more definite conclusions about the effectiveness of the three training types (association, reading aloud, combined) in training reading fluency, future studies should increase both the training duration and the number of participants.

In summary, the present study suggests that, for poor readers, orthographic access to multiletter units without semantics can be enhanced by repeated reading practice. More importantly, training multiletter, sub-lexical units facilitates training effects that are not limited to specific words, as sub-lexical training can lead to improvement at the word level and the trained sub-lexical units are encountered in several words. The improvements induced by repeated reading training seem to be specific to the materials involving the trained segments and do not induce general improvements in reading speed. Further studies are warranted to investigate more carefully what extent and under what conditions training can produce transfer effects. Effective repeated reading exercises can be administered through different types of tasks (training associations between print and phonology, reading aloud); however, in training reading fluency, the role of rapid mapping between phonology and orthography is presumably crucial.

Acknowledgments This paper is part of the first author’s doctoral thesis. This study was supported by the Finnish Center of Excellence Program of the Academy of Finland (44858 and 213486), the Finnish Graduate School of Psychology, and Niilo Mäki Institute. We would like to thank the schools and children who participated in the study. We tender our thanks to Michael Freeman for polishing the language.

References

- Berends, I., & Reitsma, P. (2006a). Addressing semantics promotes the development of reading fluency. *Applied Psycholinguistics*, *27*, 247–265.
- Berends, I., & Reitsma, P. (2006b). Remediation of fluency: Word specific or generalised training effects? *Reading and Writing*, *19*, 221–234.
- Berends, I., & Reitsma, P. (2007). Orthographic analysis of words during fluency training promotes reading of new similar words. *Journal of Research in Reading*, *30*, 129–139.
- Booth, J. R., Perfetti, C. A., & MacWhinney, B. (1999). Quick, automatic, and general activation of orthographic and phonological representations in young readers. *Developmental Psychology*, *35*, 3–19.
- Bowey, J. A., & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading: A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *92*, 203–219.
- Bruck, M., & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics: The case of initial consonant clusters. *Journal of Experimental Psychology*, *50*, 156–178.
- Brunswick, N., McCrory, E., Price, C. J., Frith, C. D., & Frith, U. (1999). Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: A search for Wernicke's Wortschatz? *Brain*, *122*, 1901–1917.
- Bulheller, S., & Häcker, H. (2002). *Coloured progressive matrices—Deutsche Bearbeitung und Normierung* [German Adaptation and Norms]. Frankfurt: Swets and Zeitlinger.
- Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*, 149–180.
- Cohen, J., (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.
- Cole, P., Magnan, A., & Grainger, J. (1999). Syllable-sized units in visual word recognition: Evidence from skilled and beginning readers of French. *Applied Psycholinguistics*, *20*, 507–532.
- Defior, S., & Tudela, P. (1994). Effect of phonological training on reading and writing acquisition. *Reading and Writing*, *6*, 299–320.
- De Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading*, *11*, 55–71.
- De Jong, P. F., & van der Leij, A. (2003). Developmental changes in the manifestation of phonological dyslexic children learning to read a regular orthography. *Journal of Educational Psychology*, *9*, 22–40.
- Di Filippo, G., de Luca, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2006). Lexicality and stimulus length effects in Italian dyslexics: Role of the overadditivity effect. *Child Neuropsychology*, *12*, 141–149.
- Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–144). Hillsdale: Lawrence Erlbaum Associates.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*, 250–287.
- Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology*, *42A*, 353–372.
- Georgiewa, P., Rzanny, R., Gaser, C., Gerhard, U. J., Vieweg, U., Freesmeyer, D., et al. (2002). Phonological processing in dyslexic children: A study combining functional imaging and event related potentials. *Neuroscience Letters*, *318*, 5–8.
- Grünling, C., Ligges, M., Huonker, R., Klinger, M., Mentzel, H. J., Rzanny, R., et al. (2004). Dyslexia: The possible benefit of multimodal integration of fMRI- and EEG-data. *Journal of Neural Transmission*, *111*, 951–969.
- Hintikka, S., Aro, M., & Lyytinen, H. (2005). Computerized training of the correspondences between phonological and orthographic units. *Written Language & Literacy*, *8*, 155–178.
- Hohn, W. E., & Ehri, L. C. (1983). Do alphabet letters help prereaders acquire phonemic segmentation skill? *Journal of Educational Psychology*, *75*, 752–762.
- Holopainen, L., Ahonen, T., & Lyytinen, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities*, *34*, 401–413.
- Judica, A., De Luca, M., Spinelli, D., & Zoccolotti, P. (2002). Training of developmental surface dyslexia improves reading performance and shortens eye fixation duration in reading. *Neuropsychological Rehabilitation*, *12*, 177–198.
- Klicpera, C., & Schabmann, A. (1993). Do German-speaking children have a chance to overcome reading and spelling difficulties? A longitudinal survey from the second until the eighth grade. *European Journal of Psychology of Education*, *8*, 307–334.

- Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*, 3–21.
- Kyte, C. S., & Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology, 93*, 166–185.
- Landerl, K. (2001). Word recognition deficit in German: More evidence from a representative sample. *Dyslexia, 7*, 183–196.
- Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An eight-year follow-up. *Journal of Educational Psychology, 100*, 150–161.
- Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German-English comparison. *Cognition, 63*, 315–334.
- Landerl, K., Wimmer, H., & Moser, E. (1997). *Salzburger Lese- und Rechtschreibtest* [Salzburg Reading and Spelling Test]. Bern: Hans Huber.
- Lemoine, H. E., Levy, B. A., & Hutchinson, A. (1993). Increasing the naming speed of poor readers: Representations formed across repetitions. *Journal of Experimental Child Psychology, 55*, 297–328.
- Levy, B. A., Bourassa, D. C., & Horn, C. (1999). Fast and slow namers: Benefits of segmentation and whole word training. *Journal of Experimental Child Psychology, 73*, 115–138.
- Lovett, M. W., Warren-Chaplin, P. M., Ransby, M. J., & Borden, S. L. (1990). Training word recognition skills of reading disabled children: Treatment and transfer effects. *Journal of Educational Psychology, 82*, 769–780.
- Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, A.-M., & Taanila, M. (2007). Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nordic Psychology, 59*, 109–126.
- Martens, V. E. G., & de Jong, P. F. (2006). The effect of visual word features on the acquisition of orthographic knowledge. *Journal of Experimental Child Psychology, 93*, 337–356.
- Martin-Chang, S. L., & Levy, B. A. (2005). Fluency transfer: Differential gains in reading speed and accuracy following isolated word and context training. *Reading and Writing, 18*, 343–376.
- Martin-Chang, S. L., Levy, B. A., & O’Neil, S. (2007). Word acquisition, retention, and transfer: Findings from contextual and isolated word training. *Journal of Experimental Child Psychology, 96*, 37–56.
- Mayringer, H., & Wimmer, H. (2003). *SLS 1–4: Das Salzburger Lese-Screening für die Klassenstufen 1–4* [Salzburg Reading-Screening for Grade 1–4]. Bern: Verlag Hans Huber.
- Olson, R. K., Wise, B., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the posttraining development of word recognition. *Scientific Studies of Reading, 1*, 235–253.
- Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading Acquisition* (pp. 145–173). Hillsdale: Lawrence Erlbaum Associates.
- Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 30*, 473–485.
- Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation Developmental Disabilities Research Reviews, 6*, 207–213.
- Sandak, R., Mencl, W. E., Frost, S. J., Rueckl, J. G., Katz, L., Moore, D., et al. (2004). The neurobiology of adaptive learning in reading: A contrast of different training conditions. *Cognitive, Affective, & Behavioral Neuroscience, 4*, 67–88.
- Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Educational Psychology, 92*, 284–295.
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151–218.
- Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*, 95–129.
- Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology, 87*, 267–298.
- Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research, 43*, 219–234.
- Spinelli, D., de Luca, M., Di Filippo, G., Mancini, M., Martelli, M., & Zoccolotti, P. (2005). Length effect in word naming latencies: Role of reading experience and reading deficit. *Developmental Neuropsychology, 27*, 217–235.
- Stenneken, P., Conrad, M., & Jacobs, A. M. (2007). Processing of syllables in production and recognition tasks. *Journal of Psycholinguistic Research, 36*, 65–78.

- Thaler, V., Ebner, E. M., Wimmer, H., & Landerl, K. (2004). Training reading fluency in dysfluent readers with high reading accuracy: Word specific effects but low transfer to untrained words. *Annals of Dyslexia*, *54*, 89–113.
- Torgesen, J., Rashotte, C., & Alexander, A. W. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 333–355). Timonium, MD: York Press.
- Treiman, R. (1991). Children's spelling errors on syllable-initial consonant clusters. *Journal of Educational Psychology*, *83*, 346–360.
- Wentink, W. M. H., van Bon, W. H. J., & Schreuder, R. (1997). Training poor readers' phonological decoding skills: Evidence for syllable bound processing. *Reading and Writing: An Interdisciplinary Journal*, *9*, 163–192.
- Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics*, *14*, 1–33.
- Yap, R., & van der Leij, A. (1993). Word processing in dyslexics: An automatic decoding deficit? *Reading and Writing: An Interdisciplinary Journal*, *5*, 261–279.
- Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*, 3–29.
- Ziegler, J. C., Perry, C., My-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology*, *86*, 169–193.
- Zoccolotti, P., De Luca, M., Di Pace, E., Judica, A., Orlandi, M., & Spinelli, D. (1999). Markers of developmental surface dyslexia in a language (Italian) with high grapheme-phoneme correspondence. *Applied Psycholinguistics*, *20*, 191–216.