Lexical reading in dysfluent readers of German

Gangl, Melanie; Moll, Kristina; Jones, Manon W.; Banfi, Chiara; Schulte-Körne, Gerd; Landerl, Karin

Scientific Studies of Reading

DOI:
10.1080/10888438.2017.1339709

Published: 01/01/2018

Publisher's PDF, also known as Version of record

SCIENTIFIC STUDIES OF READING
2018, VOL. 22, NO. 1, 24–40
https://doi.org/10.1080/10888438.2017.1339709

Lexical Reading in Dysfluent Readers of German

Melanie Gangl[a], Kristina Moll[b], Manon W. Jones[c], Chiara Banfi[a], Gerd Schulte-Körne[b], and Karin Landerl[a]

[a]University of Graz; [b]University Hospital Munich; [c]Bangor University

ABSTRACT

Dyslexia in consistent orthographies like German is characterized by dysfluent reading, which is often assumed to result from failure to build up an orthographic lexicon and overreliance on decoding. However, earlier evidence indicates effects of lexical processing at least in some German dyslexic readers. We investigated variations in reading style in an eye-tracking paradigm with German dysfluent 3rd and 4th graders. Twenty-six TypFix-readers (fixation counts within the range of 47 age-matched typical readers) were compared with 42 HighFix-readers (increased fixation counts). Both groups showed lexical access: Words were read more efficiently than nonwords and pseudohomophones. TypFix-readers showed stronger reliance on lexical reading than HighFix-readers (smaller length effects for number of fixations and total reading time, stronger lexicality effects for gaze duration, stronger word-pseudohomophone effects for mean saccade amplitude). We conclude that in both groups, sublexical and lexical reading processes were impaired due to inefficient visual-verbal integration.

Profound and persistent deficits in reading fluency are a main characteristic of dyslexia across orthographies (Peterson & Pennington, 2012). In the context of the phonological deficit view of dyslexia (Vellutino, Fletcher, Snowling, & Scanlon, 2004), dysfluent reading is interpreted as a consequence of deficient phonological decoding (Rack, Snowling, & Olson, 1992), which in turn hampers the buildup of an orthographic lexicon via self-teaching (Share, 1995). However, although remediating poor readers' phonological awareness and decoding skills can markedly improve their reading accuracy, it does not automatically improve reading fluency as well (Torgesen & Hudson, 2006).

Dysfluent reading is not simply a consequence of an inability to work out correct word pronunciations. This is obvious in orthographies that are phonologically more transparent than English, like German. The relatively consistent letter-sound relationships in German enable even struggling readers to understand the mappings between written and spoken language and to acquire high reading accuracy. Still, reading remains extremely slow and laborious (e.g., Hutzler & Wimmer, 2004; Wimmer, 1993). Wimmer (1993, Wimmer & Schurz, 2010) suggested that dysfluent reading is caused by a deficit in the efficiency of accessing phonology from the visual information of letter strings. This account is in line with the consistent finding that dysfluent reading in German is strongly associated with a marked deficit in rapid automatized naming (RAN; Denckla & Rudel, 1976) of visually presented stimuli, which is evident even before reading instruction starts (Wimmer & Mayringer, 2002).

At the onset of reading development, slow visual-verbal access is assumed to impede the efficient translation of letters into sounds during phonological decoding. Although the output of slow decoding is usually correct in consistent orthographies, the buildup of an orthographic lexicon via

CONTACT Karin Landerl karin.landerl@uni-graz.at Institute of Psychology, University of Graz, Universitätsplatz 2, Graz 8010, Austria.

self-teaching is still impaired. Orthographic representations are based on multiple associations between phonemes and corresponding letters (Ehri, 1992, 2014), and dysfluent readers are perhaps unable to achieve sufficiently fast and simultaneous activation of the phonemes corresponding to the letters in order to form such associations (Jones, Kuipers, & Thierry, 2016).

Dysfluent readers' overreliance on sequential letter processing is reflected in particularly large length effects, that is, a consistent increase in reading times with increasing number of letters (De Luca, Borrelli, Judica, Spinelli, & Zoccolotti, 2002; De Luca, Di Pace, Judica, Spinelli, & Zoccolotti, 1999; Di Filippo, De Luca, Judica, Spinelli, & Zoccolotti, 2006; Hutzler & Wimmer, 2004; Hyönä & Olson, 1995; Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Körne, 2003). However, much less is known on lexico-orthographic processing in dysfluent reading of transparent orthographies. Even if the buildup of orthographic representations is hampered by dysfluent reading, it is still likely that over time at least highly frequent words are stored as orthographic representations. Indeed, this is what Moll, Hutzler, and Wimmer (2005) reported in a single case eye-tracking study of an adult German dyslexic reader: This case demonstrated an extremely large length effect for low-frequency words, confirming that his reading was strongly based on sequential decoding. However, no length effect was evident for high-frequency words, suggesting that he had indeed succeeded in establishing orthographic entries for this word set. Nevertheless, even for those words, his gaze durations and reading onset times were almost twice as high as those of the control group, indicating that the low efficiency of phonological access affected his lexical reading as well. Similarly, Hawelka, Gagl, and Wimmer (2010) reported that two thirds of their word stimuli received only one fixation (or were skipped) by adolescent dyslexic readers of German and no length effect was observed for these words. Thus, it seems that processing of these items was not sequential. But even for this word set, gaze durations were clearly prolonged compared to typical readers. Spinelli et al. (2005) also identified a subgroup of Italian dyslexic children who did not exhibit the typical systematic length effect in vocal reaction times for words up to four letters long.

Thus, the general question arises as to what extent lexical processes are applied during dysfluent word reading and whether for those words that can be stored in the orthographic lexicon access is intact or also impaired. Furthermore, are there differences in the amount of lexical reading among dysfluent readers? The current study aimed to assess these research questions in an eye-tracking paradigm. Eye-tracking presents a wealth of information on children's reading over and above accuracies and latencies, as different parameters on saccade direction (forward vs. backward) and duration of eye-fixations can be measured. Here, we distinguished two groups of dyslexic readers based on eye movement data, both of which showed overall reading dysfluency. We reasoned that the very same overall dysfluent reading time can result from an untypically high number of fixations in combination with relatively short fixation durations in Dysfluent Reader A (hereon referred to as HighFix-readers). However, it can also result from the reverse pattern of an average number of fixations within the range of typical readers but relatively high fixation durations in Dysfluent Reader B (hereon referred to as TypFix-readers). This contrast is particularly interesting for the present study, as we assume that HighFix-readers rely rather on a piecemeal sublexical reading strategy based on a series of short saccades, whereas TypFix-readers are perhaps able to retrieve orthographic representations from memory, but this retrieval process takes unduly long. These eye movement profiles are analogous to the "plodder" and "explorer" distinctions made previously (Olson, Kliegl, Davidson, & Foltz, 1985).

Generally, dyslexic readers are described as showing increased numbers of (rightward) fixations, (leftward) regressions, higher fixation durations, and shorter forward saccades (for a review, see Rayner, 1998). The present study aims to investigate whether it is possible to uncover differential eye-movement profiles within these overall patterns identified in previous group studies of dyslexia in different orthographies. To achieve this, we selected from a large sample of dysfluent readers those who showed a number of fixations during word reading that was within normal ranges and compared their reading profiles with another group of dysfluent readers who had a clearly increased fixation count compared to typical readers.

As we were particularly interested in differences in reliance on lexical access among the dysfluent readers, our analysis focused on experimental effects that are informative in this respect. Thus, we analyzed the length effect specifically for words (but not for sublexical decoding of nonwords), and we contrasted reading of words with nonwords (lexicality effect) and with pseudohomophones (W-PsH effect). The lexicality effect describes the consistent finding that words that can be lexically accessed are read more efficiently than nonwords, which must be sublexically decoded. If dysfluent individuals read words with a similar sublexical reading strategy as nonwords, they should show a reduced or no lexicality effect. However, the typical finding is an increased lexicality effect (e.g., Juphard, Carbonnel, & Valdois, 2004; Rack et al., 1992; Wimmer, 1996; Ziegler et al., 2003). Some studies in the transparent Italian orthography suggest that dysfluent readers' lexicality effect is comparable to typical readers once general differences in reading speed are accounted for (Di Filippo et al., 2006; Paizi, De Luca, Zoccolotti, & Burani, 2013), which the authors interpreted as spared lexical access. However, the interpretation of the lexicality effect is generally problematic, as it is unclear to what extent it reflects direct access of orthographic representations. It is also possible that the phonological lexicon is accessed only after partial or full decoding of the letter sequence based on the assembled phonology, although no orthographic representation is present.

A more direct way to investigate lexical reading is to compare reading latencies for words and derived pseudohomophones. Pseudohomophones are unfamiliar letter strings representing real word pronunciations (e.g., *rane* for *rain*). A sublexical strategy would induce similar reading latencies for words and pseudohomophones, as the very same pronunciation is decoded, and postdecoding access of the phonological lexicon based on the assembled phonology is equally possible for both item types. However, if an orthographic representation is available for a certain word, reading latencies should be lower for the familiar word than for the unfamiliar pseudohomophonic spelling, for which lexical access is not possible. Among typical readers, a W-PsH effect appears as early as first grade after only few presentations of a spelling (Reitsma, 1983). Of interest, Moll and Landerl (2009) found clear W-PsH effects for 10-year-old German-speaking dysfluent readers, in accordance with untypically high reading latencies for all stimulus types. Only children with isolated spelling problems in combination with age adequate reading did not show a significant latency difference between words and pseudohomophones, not even for words they could spell correctly. It seems that these poor spellers had not succeeded in developing a sufficiently large number of orthographic representations that would enable correct spelling and lexical access during reading. Nevertheless, their sublexical decoding strategy was highly efficient, allowing them to read with age-adequate fluency.

In sum, there seem to be individual differences in the extent to which lexical strategies are applied among dysfluent readers. In the current study we were interested whether there are distinct reader profiles that vary in the amount of lexical strategies applied among dysfluent readers in an orthography that can be reliably decoded by sublexical processes. As just explained, we distinguished between a Typical Fixation Count group and a High Fixation Count group of dysfluent readers. In line with earlier findings, we expected that overreliance on sublexical reading, reflected by an untypically high number of fixations (characterized as High Fixation Count pattern), would be the dominant reading style among dysfluent readers of German. Still, we were interested to what extent this reader group nevertheless uses lexical reading strategies. Second, we assumed that dysfluent readers who exhibit unduly high reading latencies in spite of a fixation count that corresponds to typically developing readers (Typical Fixation Count pattern) might be those who apply lexical reading strategies, though obviously with very low efficiency.

We expected the HighFix-readers to show overall prolonged processing times compared to typically developing readers, which should induce increased gaze durations (the sum of fixation durations from the first fixation until a saccade leaves the item, which is called "first pass"), as well as increased total reading times (the sum of all fixation durations on an item, including interword regressions) associated with their higher number of fixations. Furthermore, overreliance on sublexical reading was predicted to induce reduced lexicality and W-PsH effects, as well as increased effects of word length for the HighFix-readers. For TypFix-readers we expected that the prolonged processing times (gaze duration, total reading times) would mainly indicate sluggish lexical access. Such reliance on lexical access should induce lexicality and W-PsH effects that are larger than for HighFix-readers and comparable to controls, and word length effects were predicted to be smaller than for HighFix-readers.

Method

Participants

Groups of TypFix-readers and HighFix-readers were selected from a sample of 75 nine-year-old dysfluent readers who all had a percentile below 20 in a classroom reading fluency test. In addition, they showed performance below percentile 20 in at least one subtest of an individually administered 1-min word and nonword reading test and below 25 on the other subtest (see test descriptions next). An age-matched control group (N = 47) showed performance above percentile 25 and below percentile 75 in the classroom reading and spelling measures. To select a homogeneous group of average readers, we applied a slightly more stringent criterion of percentiles between 30 and 70 for the individually administered 1-min word reading test.

TypFix-readers and HighFix-readers were categorized based on the total number of fixations for words in the eye-tracking paradigm described next. All 26 children who showed a total number of fixations within the typical range of the control group were categorized as TypFix-readers (controls: M = 2.35, SD = .33, range = 1.65–3.06, TypFix-readers: M = 2.75, SD = .23, range = 2.08–3.04). Forty-two children who showed a clearly higher total number of fixations of at least 3 standard deviations above the mean of the controls (M = 4.25, SD = .79, range = 3.34–6.51) were selected as HighFix-readers.

All children had German as first language, a nonverbal IQ at or above 85, and normal or corrected-to-normal vision. Children with a clinical diagnosis of attention deficit hyperactivity disorder or an increased score on a parental questionnaire on attention deficits were not admitted to the study. Consent for the study was given by the ethics committees of the University of Graz and the University of Munich and all parents gave written consent.

Participant characteristics are shown in Table 1. The three groups were well matched on age, attention ratings, and nonverbal and verbal IQ. As expected, the majority of dysfluent readers showed a clearly higher total number of fixations than controls. However, about one third of the poor readers showed a total number of fixation within the typical range of the controls (TypFix-readers). As defined by our selection criteria, HighFix-readers showed reliably more fixations than controls and TypFix-readers. Note that even though TypFix-readers were selected to have fixation counts within the normal range, they still showed a small (0.4 fixations) but significant difference to the control group.

Both groups of dysfluent readers showed seriously impaired reading performance with a mean percentile of 10 on the classroom test and below percentile 15 on the two individual reading tests.[1] Although they performed at similarly low levels in the classroom reading fluency and 1-min nonword reading tests, performance of HighFix-readers was significantly lower compared to TypFix-readers for 1-min word reading and spelling. In the phonological awareness and rapid automatized naming measures described next, TypFix-readers' and HighFix-readers' performance was comparable, with both being lower than controls' performance.

[1]Note that 51 of 55 dysfluent readers who were recruited in Munich fulfilled the criteria of a reading disorder (dyslexia) according to the German diagnostic guidelines (percentile ≤ 16 in one of the subtests of the individually administered standardized reading test; Moll & Landerl, 2010, and IQ ≥ 70) and received formal diagnoses. All 13 dysfluent readers of the second collaborating site in Graz also fulfilled these criteria. However, the Austrian school system does not recognize any formal diagnosis of dyslexia.

Table 1. Participant characteristics for Controls, TypFix-Readers, and HighFix-Readers.

	Group				
	Controls	TypFix-Readers	HighFix-Readers	*F*	*p*
N	47	26	42		
Girls/Boys	24/23	15/11	22/20		
Age (months)	113.28 (3.87)	112.92 (5.60)	113.64 (6.09)	.16	.852
Total number of fixations	2.35 (.33)$_{b,c}$	2.75 (.23)$_{a,c}$	4.25 (.79)$_{a,b}$	149.46	< .001
Reading (% rank)					
Classroom reading	55.82 (13.76)$_{b,c}$	9.44 (5.79)$_{a}$	10.24 (8.31)$_{a}$	259.86	< .001
Word reading	50.47 (10.59)$_{b,c}$	14.71 (7.39)$_{a,c}$	5.74 (4.07)$_{a,b}$	378.47	< .001
Nonword reading	51.47 (16.99)$_{b,c}$	14.71 (7.52)$_{a}$	11.81 (8.98)$_{a}$	128.96	< .001
Spelling (% rank)	55.79 (12.44)$_{b,c}$	30.50 (20.53)$_{a,c}$	16.77 (16.13)$_{a,b}$	68.64	< .001
ADHD score	.49 (.37)	.51 (.29)	.42 (.29)	.96	.386
Nonverbal IQ	106.87 (9.72)	111.58 (13.22)	107.64 (12.65)	1.44	.242
Verbal IQ					
Digit span	10.38 (2.34)	10.04 (1.76)	9.69 (2.32)	1.09	.341
Vocabulary	12.32 (3.06)	12.69 (3.30)	12.40 (2.99)	.13	.882
PA (% correct)	80.96 (12.17)$_{b,c}$	63.55 (17.45)$_{a}$	68.67 (18.73)$_{a}$	11.60	< .001
RAN (items/sec)[a]					
Digits	2.12 (.41)$_{b,c}$	1.85 (.32)$_{a}$	1.76 (.26)$_{a}$	13.57	< .001
Objects	1.09 (.20)$_{b,c}$	.96 (.26)$_{a}$	.94 (.16)$_{a}$	6.88	< .01

Note. Standard deviations are in parentheses. Subscript letters indicate that the mean differs reliably ($p < .05$) from the referred-to mean (post hoc Bonferroni tests): a = controls, b = TypFix-readers, c = HighFix-readers. ADHD = attention deficit hyperactivity disorder; PA = phonological awareness; RAN = rapid automatized naming.

[a]Due to experimental error there are missing values for one control child for RAN digits, for one HighFix-reader and two controls for RAN objects and for one HighFix-reader for both RAN measures.

Materials and procedure

Reading

A standardized reading speed test (Wimmer & Mayringer, 2014; parallel test reliability greater than .86 according to manual) was given as classroom measure. Simple sentences were read silently and were marked as semantically right or wrong (e.g., “Trees can speak.”). After 3 min the task was terminated and the number of correctly marked sentences was determined.

In addition, an individually administered 1-min reading speed task (Moll & Landerl, 2010; parallel test reliability between .90 and .95 according to manual) was given. It contains a word and a nonword reading list with increasing item length and complexity. Children read both lists aloud as fast as possible without making errors. The number of items read correctly within 1 min was scored.

Spelling

The standardized classroom spelling task (Müller, 2004; split-half reliability = .95 according to manual) contained 44 words that had to be written into sentence frames. The experimenter dictated each word, read the full sentence, and repeated the word. Number of incorrect word spellings was scored.

Nonverbal IQ

Series, Classification, Matrices and Topology from the German version of the Culture Fair Intelligence Test (Weiß, 2006; reliability = .92 as described in the manual) were given as estimates of nonverbal IQ.

Verbal IQ

We applied Digit Span and Vocabulary from the German version of the Wechsler Intelligence Scale for Children (Petermann & Petermann, 2011).

Attention rating

Parents were asked to answer a standardized questionnaire (Döpfner, Görtz-Dorten, Lehmkuhl, Breuer, & Goletz, 2008) consisting of 20 items with a 4-point rating scale investigating symptoms of inattention, hyperactivity, and impulsivity.

Phonological awareness

The computerized phoneme deletion task programmed with Presentation 16.3 (Neurobehavioral Systems, Inc., Berkeley, CA) consisted of four practice and 25 test trials (20 mono- and five disyllabic nonwords), which were presented via headphones. Children repeated each nonword first and then pronounced it without a specified phoneme (e.g., "/fɔlt/without /t/"—/fɔl/). The experimenter marked responses for correctness. Any nonword that was not pronounced correctly was replayed up to two times. Items not pronounced correctly by the child were excluded from analysis (0.9%). The ratio of correct responses to the total number of responses was scored. Cronbach's alpha was .79.

Rapid automatized naming

Standard paradigms of RAN-objects and RAN-digits (Denckla & Rudel, 1976) were presented. Both conditions required to name a matrix of 40 items as quickly and accurately as possible. Simple pictured objects and digits were presented on separate sheets in five columns and eight lines. Item order was randomized, and each item was presented once in each line. Children were familiarized with each condition with a 3 × 5 RAN array format. The time needed to name the full item set and any occurring errors were recorded and transformed into items named correctly per second. The correlation between conditions was .46, which corresponds to earlier studies (Van Den Bos, Zijlstra, & Lutje Spelberg, 2002).

Eye-tracking paradigm

Apparatus. Eye-movements of the dominant eye were recorded with an EyeLink 1000 Tower Mount eye-tracker in Graz and an EyeLink 1000 Plus Desktop Mount eye-tracker in Munich (SR Research, Toronto, Canada). The experiment was controlled with Experiment Builder software (RS Research, version 1.10.1241). Children were seated in front of a 20-in. monitor (120-Hz refresh rate, 1024 × 768 resolution) in Graz and a 15.6-in. monitor (120-Hz refresh rate, 1280 × 960 resolution) in Munich at a viewing distance of 65 cm. Stimulus presentation was similar at both collaborating sites with an uppercase letter height of about 0.62° of visual angle. Children put their forehead up against a forehead rest to minimize head movements. A 9-point calibration cycle at the beginning and after each break was used to ensure a spatial resolution of less than 0.5° of visual angle.

Stimuli and design. The item set contained 80 words, 80 pseudohomophones, and 80 nonwords (e.g., *Hand, Hant, Hond*, see the appendix for full item list) and 60 filler items. Forty short (three to five letters) and 40 long words (six to nine letters) with high frequency (mean absolute frequency of 1537.80 for 9- to 10-year-old children according to childLex corpus; Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015) were selected. Pseudohomophones were derived from the words by exchanging one phonologically identical grapheme and nonwords were derived by exchanging one grapheme per syllable. Words, pseudohomophones, and nonwords were matched on number of letters, bigram-frequency, and trigram-frequency according to childLex (Schroeder et al., 2015). Item characteristics are shown in Table 2.

The total item set was divided into three blocks consisting of 80 items each and assembled in four pseudorandomized orders, which were randomly assigned to participants. Two blocks were made up of words and pseudohomophones, and the third block contained nonwords only. Words and pseudohomophones were arranged in a fixed random order with the restriction that no more than two words or pseudohomophones appeared in immediate sequence. Furthermore, a word and the

Table 2. Item characteristics for words (W), pseudohomophones (PsH), and nonwords (NW).

	W		PsH		NW			
	M	*SD*	*M*	*SD*	*M*	*SD*	*F*	*p*
No. of letters	5.54	1.61	5.71	1.74	5.54	1.61	0.30	.742
Log bigram-frequency								
childLex corpus	5.19	5.06	5.16	5.11	5.13	4.99	0.57	.568
childLex lexicon	5.61	5.50	5.58	5.54	5.56	5.43	0.45	.641
Log trigram-frequency								
childLex corpus	4.25	4.28	4.19	4.32	4.12	4.22	1.29	.277
childLex lexicon	4.72	4.79	4.69	4.87	4.59	4.73	0.92	.399

Note. The corpus based bigram- and trigram-frequencies refer to the number of occurrences per million tokens in the corpus of the database childLex, whereas the lexicon based frequencies refer to the number of occurrences per million types.

corresponding pseudohomophone did not occur together in the same block to avoid recognition effects.

Procedure. The experiment was run in a dimly lit room. Items were displayed in single lines in black on a white background in Arial font. Each line contained eight target and two filler items at the beginning and the end of the line. Filler items were not considered for analyses. Presentation started with a word/pseudohomophone block, followed by the nonword and the second word/pseudohomophone block. Short breaks were given after each block and after the first half of the nonwords. Children were familiarized with the task format with four lines of practice items and each subsequent block was introduced with one line of practice items.

A line was displayed after the child had fixated a left-sided yellow smiley linked to a fixation trigger for at least 250 ms. If no fixation on the smiley was detected within 5,000 ms of trial onset, a new 9-point calibration cycle was run and the experiment continued from the point at which it was interrupted. When a fixation was identified a line of 10 items appeared in the center of the computer screen with the first item at the location of the smiley. Children were instructed to read each item aloud at their own speed without making mistakes while their eye-movements were recorded. Reading aloud was chosen in order to control for reading accuracy. Reading errors were noted by the experimenter and corrected only during practice. At the end of each line, children had to immediately look at a small cross in the lower right corner of the screen. The line disappeared as soon as the cross was fixated and the next trial started with the smiley on the left side of the screen. Cronbach's alpha for words, pseudohomophones, and nonwords was high for reading accuracy (.71, .83, .92 respectively) and number of fixations (.94, .94, .92 respectively).

Eye-tracking analysis. Each target item was defined as region of interest. Two fixations within 0.5° of visual angle were pooled and short fixations (< 80 ms) were excluded. Data loss due to problems with calibration accuracy, or because a child did not read the whole line, was 2.27%. Moreover, for each eye-tracking parameter data ±2.5 standard deviations from the individual mean of each item type (words, pseudohomophones, nonwords) by length (short, long) category were removed (2.50%).

We obtained eye-tracking data in first-pass reading pertaining to *number of fixations, mean (rightward) saccade amplitude*, and *gaze duration* (the sum of fixations on the target region before a forward saccade to the next region), as well as *total reading time* (the sum of all fixations on the target region, including interword regressions).

Linear mixed-effects models were implemented with lme4 (Bates, Maechler, & Dai, 2008) and the languageR package in R (R Development Core Team, 2008). Each analysis examined one dependent variable (DV) as a function of the between-subjects factor group (controls, TypFix-readers, HighFix-readers) and one of the within-subjects variables pertaining to stimulus type. These included (a) word length effects (short vs. long words), (b) lexicality effects (words vs. nonwords), and (c) W-PsH effects (words vs. pseudohomophones). Separate a priori hypotheses were made for these within-subjects factors, reflected in our decision to conduct separate analyses. The analyses were based on correctly read items, and for lexicality and W-PsH effects, only corresponding items read correctly in both conditions were included. Fixed effects were obtained for group (controls, TypFix-readers, HighFix-readers) and condition (Level 1, Level 2; see prior specification for each factor). The baseline (intercept) condition comprised controls/Condition 1. Fixed effects of group were determined based on whether (a) TypFix-readers/Condition 1 and (b) HighFix-readers/Condition 1 contributed unique variance to the model relative to the baseline. A fixed effect of condition moreover determined whether controls/Condition 2 contributed unique variance to the model relative to the baseline. Interaction effects were detected based on the contribution of (a) TypFix-readers/Condition 2 to the model, beyond the additive contribution of the group and condition fixed effects, and (b) HighFix-readers/Condition 2 to the model, also beyond the additive contribution of the group and condition fixed effects. For example, if TypFix-readers' performance in Condition 2 contributed unique variance to the model, beyond the contribution of fixed effects (TypFix/Condition 1 + controls/Condition 2), then that interaction would emerge significant. Separate analyses were run in which TypFix-readers/Condition 1 comprised the baseline, in order that direct comparisons could also be made between TypFix-readers and HighFix-readers. In all analyses, participants and item variances were entered as random effects variables, for which maximal intercepts and slopes were specified insofar as this was possible with respect to model convergence (Barr, Levy, Scheepers & Tily, 2013).

Results

Reading accuracy

As expected, reading accuracy (see Table 3) was close to ceiling for controls in all conditions. Dysfluent readers also showed high accuracy rates for short and long words and pseudohomophones (though significantly different from controls). Even for nonwords, accuracy was reasonably high for both groups of dysfluent readers, although the difference in performance was more marked compared to controls (especially for long items).

Eye-tracking paradigm

A Box-Cox power transformation indicated the following transformations for effective normalization of each DV: mean saccade amplitude (log), gaze duration (log), total reading time (square root). No transformation was applied for number of fixations (see Liversedge, Drieghe, Li, Yan, Bai, & Hyönä, 2016). Means and standard deviations corresponding to the original scales are provided in Table 4. Linear mixed-effects results are reported in Table 5 comprising the changes in coefficient values (*b*) along with *t* and *p* values (derived from the normal approximation method; see Barr et al., 2013) for each analysis (DV). In addition, the result patterns are displayed in Figure 1.

Table 3. Reading Accuracy (% Read Correctly) for Words (W), Pseudohomophones (PsH), and Nonwords (NW) Separately for Each Group.

	Group				
	Controls	TypFix-Readers	HighFix-Readers	F	p
W short	98.56 (2.38)$_{b,c}$	95.35 (3.44)$_a$	93.32 (5.58)$_a$	18.92	< .001
W long	98.50 (1.99)$_{b,c}$	93.12 (4.96)$_a$	92.50 (4.89)$_a$	29.33	< .001
PsH short	95.70 (3.89)$_{b,c}$	88.30 (7.73)$_a$	87.37 (8.96)$_a$	18.20	< .001
PsH long	97.42 (3.35)$_{b,c}$	88.98 (8.61)$_a$	91.54 (6.84)$_a$	18.46	< .001
NW short	95.54 (4.56)$_{b,c}$	82.28 (12.57)$_a$	82.75 (11.91)$_a$	24.44	< .001
NW long	91.82 (6.41)$_{b,c}$	73.62 (16.48)$_a$	79.87 (16.53)$_a$	17.87	< .001

Note. Subscript letters indicate that the mean differs reliably ($p < .05$) from the referred-to mean (post hoc Bonferroni tests): a = controls, b = TypFix-readers, c = HighFix-readers.

Table 4. Means (Standard Deviations) of all eye-tracking parameters for Controls, TypFix-Readers, and HighFix-Readers separately for each investigated within-subject factor: Word Length Effect, Lexicality Effect, and W-PsH Effect.

	Word Length Effect				Lexicality Effect				W-PsH Effect			
	Short W		Long W		W		NW		W		PsH	
No. of fixations (N)												
Controls	1.52	(0.67)	2.18	(1.03)	1.84	(0.93)	2.62	(1.59)	1.86	(0.93)	2.23	(1.25)
TypFix-readers	1.71	(0.82)	2.67	(1.38)	2.14	(1.18)	3.08	(1.90)	2.18	(1.21)	2.67	(1.57)
HighFix-readers	2.28	(1.32)	3.52	(2.12)	2.86	(1.85)	3.83	(2.69)	2.90	(1.86)	3.53	(2.36)
Mean saccade amplitude (degrees)												
Controls	0.85	(0.25)	0.98	(0.32)	0.94	(0.31)	0.82	(0.22)	0.94	(0.30)	0.89	(0.31)
TypFix-readers	0.83	(0.25)	0.95	(0.30)	0.91	(0.29)	0.80	(0.20)	0.91	(0.29)	0.87	(0.27)
HighFix-readers	0.80	(0.23)	0.88	(0.25)	0.85	(0.25)	0.76	(0.19)	0.85	(0.25)	0.83	(0.25)
Gaze duration (ms)												
Controls	534	(275)	720	(389)	627	(350)	1119	(727)	628	(350)	848	(510)
TypFix-readers	646	(401)	934	(623)	778	(531)	1509	(1019)	778	(529)	1082	(741)
HighFix-readers	926	(692)	1308	(997)	1102	(878)	1820	(1324)	1113	(881)	1431	(1067)
Total reading time (ms)												
Controls	675	(362)	882	(419)	777	(404)	1421	(787)	779	(405)	1044	(554)
TypFix-readers	813	(474)	1147	(687)	968	(611)	2003	(1127)	969	(594)	1386	(846)
HighFix-readers	1323	(884)	1873	(1099)	1585	(1025)	2746	(1517)	1585	(1024)	2037	(1182)

Note. W = words; PsH = pseudohomophones; NW = nonwords.

Table 5. Results from the linear mixed-effects models analyses for all eye-tracking parameters separately for each investigated within-subject factor: Word Length Effect, Lexicality Effect, and W-PsH Effect.

	No. of Fixations (N)			Mean Saccade Amplitude (log)			Gaze Duration (Log)			Total Reading Time (Square Root)		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
Word length effect												
Group 1 (controls vs. TypFix)[a]	0.19	0.08	2.45*	−0.03	0.03	−1.09	0.15	0.05	3.17**	2.30	0.94	2.45*
Group 2 (controls vs. HighFix)[a]	0.77	0.09	8.74***	−0.05	0.02	−2.02*	0.42	0.04	9.32***	9.62	0.91	10.58***
Group 3 (TypFix vs. HighFix)[b]	0.58	0.09	6.31***	−0.02	0.03	−0.63	0.26	0.05	5.16***	7.32	0.99	7.40***
Condition (short W vs. long W)[a]	0.67	0.07	9.43***	0.14	0.02	5.92***	0.28	0.04	7.26***	3.81	0.58	6.61***
Group 1 × Condition[a]	0.30	0.09	3.45**	−0.01	0.03	−0.36	0.07	0.04	1.66	1.62	0.60	2.68**
Group 2 × Condition[a]	0.60	0.11	5.50***	−0.04	0.02	−1.88†	0.02	0.04	0.53	3.27	0.77	4.25***
Group 3 × Condition[b]	0.30	0.11	2.75**	−0.03	0.03	−1.25	−0.04	0.05	−0.98	1.65	0.70	2.36*
Lexicality effect												
Group 1 (controls vs. TypFix)[a]	0.34	0.09	3.78***	−0.02	0.03	−0.84	0.19	0.05	3.93***	3.14	0.90	3.48**
Group 2 (controls vs. HighFix)[a]	1.05	0.09	11.63***	−0.08	0.02	−3.23**	0.42	0.04	10.17***	11.35	0.83	13.72***
Group 3 (TypFix vs. HighFix)[b]	0.71	0.10	7.19***	−0.05	0.03	−1.90+	0.23	0.05	4.73***	8.21	0.93	8.81***
Condition (W vs. NW)[a]	0.78	0.12	6.32***	−0.11	0.02	−6.33***	0.50	0.05	10.93***	9.53	1.01	9.48***
Group 1 × Condition[a]	0.18	0.14	1.29	0.01	0.02	0.33	0.10	0.05	1.76†	4.02	1.18	3.42**
Group 2 × Condition[a]	0.19	0.14	1.42	0.02	0.02	1.02	−0.04	0.05	−0.95	3.09	1.09	2.85**
Group 3 × Condition[b]	0.01	0.15	0.10	0.01	0.02	0.50	−0.14	0.06	−2.50*	−0.93	1.22	−0.76
W-PsH effect												
Group 1 (controls vs. TypFix)[a]	0.34	0.09	3.69***	−0.04	0.03	−1.52**	0.18	0.05	3.74***	3.07	0.93	3.30**
Group 2 (controls vs. HighFix)[a]	1.06	0.09	11.36***	−0.07	0.02	−3.26***	0.43	0.04	10.43***	11.12	0.86	12.95***
Group 3 (TypFix vs. HighFix)[b]	0.72	0.10	7.41***	−0.03	0.03	−1.30	0.25	0.05	5.18***	8.06	0.95	8.47***
Condition (W vs. PsH)[a]	0.38	0.09	4.21***	−0.06	0.02	−3.10**	0.27	0.04	6.67***	4.36	0.59	7.37***
Group 1 × Condition[a]	0.12	0.09	1.32	0.01	0.02	0.73	0.02	0.04	0.67	1.66	0.64	2.60**
Group 2 × Condition[a]	0.26	0.10	2.51*	0.04	0.02	2.53*	−0.03	0.03	−1.05	1.26	0.69	1.82†
Group 3 × Condition[b]	0.14	0.10	1.44	0.03	0.02	1.47	−0.06	0.04	−1.53	−0.40	0.66	−0.61

Note. W = words; PsH = pseudohomophones; NW = nonwords.
[a]Baseline = controls on Condition Level 1 (primary analysis). [b]Baseline = TypFix-readers on Condition Level 1 (supplementary analysis in order to compare TypFix- and HighFix-readers).
†$p < .08$. *$p < .05$. **$p < .01$. ***$p < .001$.

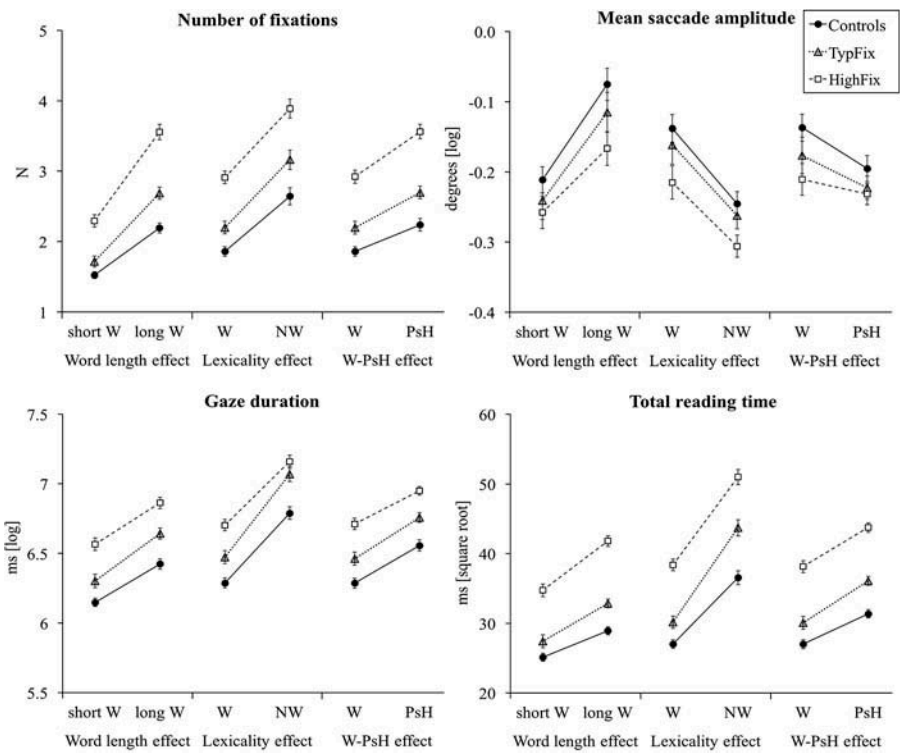

Figure 1. Coefficient values (standard errors) of all eye-tracking parameters for controls, TypFix-readers, and HighFix-readers for (1) short and long words read correctly (word length effect), (2) words and corresponding nonwords read correctly (lexicality effect), and (3) words and corresponding pseudohomophones read correctly (W-PsH effect).

Word length effect

Our main research question was the extent to which dysfluent readers rely on lexical access during reading. Lexical processes are largely irrelevant for nonword reading; therefore, we focused on effects of item length for words only.

Fixed effects of group showed a staircase pattern with clear differences between controls and HighFix-readers, with TypFix-readers performing at a level between these extremes. HighFix-readers exhibited a higher number of fixations, shorter rightward saccades (mean saccade amplitude), and prolonged gaze durations and total reading times compared with controls for short words. TypFix-readers did not differ from controls and HighFix-readers in mean saccade amplitude but showed more fixations and longer gaze durations/total reading times compared with controls (and fewer/shorter durations compared with HighFix-readers). As expected, fixed effects of condition in controls showed higher scores for long than for short words in all dependent variables. Interaction effects moreover revealed word length effects for both groups of dysfluent readers in all variables: TypFix-readers and HighFix-readers showed even larger length effects than controls for number of fixations and total reading time, and HighFix-readers showed even more marked effects on both measures compared with TypFix-readers. Furthermore, HighFix-readers' length effect for mean saccade amplitude tended to be smaller than controls' ($p = .06$). No interactions were revealed for gaze duration.

In summary, the dysfluent reading style of both TypFix-readers and HighFix-readers was reflected by inefficient eye-movements in all eye-tracking parameters (with the exception of mean saccade amplitudes among TypFix-readers). Both groups of dysfluent readers showed an overreliance on sublexical reading processes in terms of increased word length effects for number of fixations and total reading time. In line with predictions, these effects were particularly strong for HighFix-readers.

Lexicality effect

Fixed effects of group confirmed more efficient processing for controls compared to TypFix-readers and HighFix-readers, as well as more efficient processing for TypFix-readers than HighFix-readers across all variables (except saccade amplitude, which was similar for TypFix-readers and controls but lower for HighFix-readers). Lexicality effects were significant for all variables, expressed as higher number of fixations, shorter mean saccade amplitudes, higher gaze durations, and total reading times for nonwords than for words. All three groups exhibited comparable lexicality effects for number of fixations and mean saccade amplitude. TypFix-readers' lexicality effect was marginally larger than controls' (p = .08) and clearly larger than HighFix-readers' for gaze duration, whereas there was no difference between controls and HighFix-readers. Surprisingly, for total reading time the lexicality effect was similar for both groups of dysfluent readers and larger than for controls.

As previously mentioned, reading accuracy for nonwords (particularly the longer ones) was lower among dysfluent than control readers. As we included only item pairs with correct readings of words as well as corresponding nonwords, the analyses of the lexicality effect were based on a smaller and perhaps easier item set for dysfluent readers. To control for such a confound, we tested the lexicality effect for short items only as for these items the group differences in reading accuracy were rather small. The pattern of these analyses was largely the same, with one interesting difference: For number of fixations TypFix-readers showed a significantly larger lexicality effect than controls ($b = .31$, $t = 2.48$, $p < .05$).

Thus, although we had predicted reduced lexicality effects for HighFix-readers who are supposedly over reliant on sublexical decoding, their lexicality effects were largely comparable to controls and even larger for the late measure of total reading time. The particularly large lexicality effects among TypFix-readers are in line with our assumption of overreliance on lexical procedures. Still, note that TypFix-readers' reading style for words was quite inefficient compared to controls for all eye-tracking parameters except saccade amplitudes.

W-PsH effect

Fixed effects of group largely confirmed the pattern of group differences just described: Controls showed more efficient processing than both groups of dysfluent readers, and TypFix-readers showed more efficient eye-movement patterns than HighFix-readers (except for mean saccade amplitude). Fixed effects also indicated that controls read pseudohomophones less efficiently than words, with a higher number of fixations, lower mean saccade amplitudes and higher gaze durations and total reading times. Interaction effects indicated that—as predicted—TypFix-readers' W-PsH effects were largely comparable to controls' and even greater for the late measure of total reading time. Also in line with our predictions was the finding that HighFix-readers' W-PsH effect was smaller than controls' for mean saccade amplitude, but it was still significant ($b = -.05$, $t = -2.03$, $p < .05$). For all other eye-tracking parameters, HighFix-readers too showed clear indications of a word advantage, which was comparable to controls for gaze duration and even larger for number of fixations and marginally larger for total reading time (p = .07).

Thus, although we found overall impaired eye-movement processes among dysfluent readers, our results on the W-PsH effect still revealed clear indications for lexical access for HighFix-readers as well as TypFix-readers.

Discussion

Dysfluent reading in consistent orthographies is generally interpreted as overreliance on a sublexical reading strategy. However, dysfluent reading may also result from very slow lexical access. The present study used eye-tracking to investigate in detail to what extent dysfluent readers in the consistent German orthography show evidence for lexical reading. Based on their eye-movement patterns during word reading, we selected a group of children who showed the expected pattern of unusually high numbers of fixations compared to typically developing readers, indicating strong reliance on piecemeal small-unit processing. Of interest, our sample of dysfluent readers also included children who showed fixation counts within the range of typically developing children, although their reading fluency was similarly impaired as that of the HighFix-readers. We assumed that their fixation patterns might indicate inefficient attempts to access the word lexicon. Note that the parameter for group selection was assessed within the experimental reading paradigm, which might be considered as a methodological limitation of this study. However, in our analysis we focused on experimental effects that are informative on lexical reading, that is, word length, lexicality, and W-PsH effects, which were obviously not part of the group selection, and we analyzed a number of different eye-tracking parameters. It is also important to note that our group distinction is unlikely to constitute a reliable subtype divide (as such subtype categorizations have generally been of limited success; see, e.g., Peterson, Pennington, Olson, & Wadsworth, 2014). However, we reasoned that differentiating dysfluent readers based on this central eye-tracking parameter would allow us to specify differences in the reliance on lexical and sublexical processing within dysfluent readers.

First, it was interesting that among a relatively large sample of dysfluent readers, the HighFix-pattern was clearly more frequent ($N = 42$) than the TypFix-pattern ($N = 26$), confirming that this is the more typical pattern of dysfluent reading in the consistent German orthography. The two groups of dysfluent readers were quite similar with respect to cognitive profiles with age-adequate attention, and IQ, and comparable deficits in phonological awareness and rapid automatized naming. Dysfluent readers showed seriously impaired performance on standardized reading measures with a mean percentile of 10 on a classroom reading test for both groups. On a specific word reading measure and on spelling, HighFix-readers' performance was even lower than TypFix-readers'. This finding is in line with the assumption that TypFix-readers are better at storing word specific knowledge in orthographic memory than HighFix-readers, which they can then use for word recognition and orthographically correct spelling.

Within our eye-tracking paradigm, both groups of dysfluent readers showed marked processing difficulties compared to controls: Overall, TypFix-readers showed higher number of fixations, gaze durations, and total reading times than controls. HighFix-readers' eye-movement patterns were not only less efficient than controls' but also differed from TypFix-readers' (except for mean saccade amplitude). At first glance, it could be argued that our group distinction simply represents two degrees of severity of reading impairment. However, the fine-grained differences we observed in experimental reading effects suggest a more differentiated view.

Prolonged word length effects are generally interpreted as a central indicator of overreliance on sublexical reading and have been reported for dyslexia across orthographies (Wimmer & Schurz, 2010). In line with predictions, we found that the length effect was particularly marked among HighFix-readers: Their length effect was larger compared to controls' and TypFix-readers' for number of fixations and total reading time. It was also marginally smaller compared to controls for mean saccade amplitudes but did not differ from controls for gaze duration. The difference between parameters is probably due to a higher number of refixations among the HighFix-readers: Mean saccade amplitudes, which are based on rightward saccades during first-pass, were not adequately adapted to word length. Number of fixations is based on rightward as well as leftward eye-movements during first pass (before the eye leaves the word) and thus includes the so-called intraword regressions. Therefore, the increased length effect among HighFix-readers on number of fixations seems to result from a higher number of leftward intraword regressions for long words.

Similarly, on gaze duration, which is based on the sum of fixation durations during first-pass, HighFix-readers showed an adequate length effect, whereas the length effect for total reading time, which includes regressive eye-movements after first pass, was unduly large. This difference between parameters thus seems to result from a particularly high number of regressions to long words after first-pass. Of interest, the higher number of refixations during first-pass, particularly for longer words, was not associated with a corresponding increase in processing times reflected in gaze duration, which might suggest that intraword regressions were made if fixations during rightward saccades were too short in duration to recognize the presented word.

Contrary to our expectations, we also found more marked length effects for TypFix-readers compared to controls for number of fixations and total reading time. In summary, prolonged length effects indicating overreliance on sublexical reading were found not only for HighFix-readers but also (though smaller) for TypFix-readers. But to what extent do dysfluent readers in a consistent orthography rely on lexical processes during reading?

All three groups showed clear lexicality effects. Although we had expected to see a reduced effect among HighFix-readers, their lexicality effect was largely comparable to controls. For total reading time, they had an even larger lexicality effect than controls. Again, this effect seems to be associated with a higher number of regressions after the eyes had left the nonword, as the lexicality effect on the first-pass parameters was not larger than for controls. A tentative explanation is that lexical access in this group takes place mostly after a bottom-up decoding process but still constitutes an important validation procedure whether or not the result of the decoding process is correct. In case of nonwords, lexical verification is prohibited, which may induce a higher number of regressions.

For TypFix-readers, who were assumed to rely more strongly on lexical reading than HighFix-readers, we had predicted lexicality effects comparable to controls and larger than for HighFix-readers. Although the contrast to HighFix-readers was confirmed only for gaze duration, TypFix-readers exhibited even larger lexicality effects than controls for gaze duration (marginally) and total reading time. When only short items were considered, the TypFix-control contrast on the lexicality effect was significant for all eye-tracking parameters except mean saccade amplitude.

Thus, in summary both groups of dysfluent readers exhibited strong effects of lexicality, but the origin of this effect is actually not completely clear. Does it result from overreliance on lexical reading (as we had predicted for TypFix-readers) or from particular problems with sublexical reading (the classic nonword reading deficit of dyslexia; Rack et al., 1992; Wimmer, 1996)? Note that in the consistent orthography of German, words can usually be decoded sublexically and the lexicality effect may mostly reflect access to the word entry in the mental lexicon based on the assembled phonology. Thus we consider the W-PsH effect more revealing concerning reliance on lexical reading: As the pronunciation of words and derived pseudohomophones is identical, access to the phonological lexicon is exactly the same. Any difference in processing load must therefore be due to differences in orthographic access.

In line with our expectations there was no indication for a smaller W-PsH effect for TypFix-readers compared to controls, confirming that they differentiated between correct and incorrect spellings. The effect was even more marked for TypFix-readers' total reading time, indicating a particularly high number of regressions to the pseudohomophonic spellings. Surprisingly and contrary to predictions, HighFix-readers showed only a smaller W-PsH effect than controls for saccade amplitude, whereas the effect was even more pronounced than for controls for number of fixations and (marginally) for total reading time. Thus, HighFix-readers did not adapt their saccade amplitudes to item type in the same way as typically developing readers. However, as was already discussed in the context of the lexicality effect, HighFix-readers made more intraword refixations for pseudohomophones, which seems to indicate that they were confounded by the incorrect spelling of the pseudohomophones.

The current study shows that dysfluent reading in German results not only from overreliance on sublexical reading. In fact, lexical as well as sublexical strategies are applied, though both of them seem to be seriously impaired. During typical development, slow and laborious sublexical decoding procedures are over time replaced by more efficient lexical access strategies (e.g., Rau, Moeller, & Landerl, 2014). Dysfluent readers' development is characterized by even slower and more laborious decoding, which is supplemented by slow and laborious lexical access. This evidence confirms the "phonological efficiency" view (Wimmer, 1993), which postulates slow visual-verbal access for both assembled and addressed phonology. Our study found only minor tendencies of TypFix-readers to overrely on lexical reading and HighFix-readers to overrely on sublexical reading. A limitation of our group distinction is that it was based only on one (critical) parameter (total number of fixations), which was assessed within our eye-tracking paradigm. Nevertheless, our data show that there are differences in the (im)balance of reading strategies among dysfluent readers. To provide personalized intervention for dyslexic individuals, it is important to identify any such imbalances between reading strategies as part of a detailed diagnosis. Eye-tracking and the experimental effects investigated in the current study provide a promising means to devise such fine-grained diagnostic tools.

Funding

This research was supported by grants from the German Research Foundation (DFG, grant no. MO 2569/2-1) and the Austrian Science Fund (FWF, grant no. I 1658-G22).

ORCID

Chiara Banfi http://orcid.org/0000-0001-7683-1853

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., & Dai, B. (2008). *lme4: Linear mixed-effects models using S4 classes* (R package version 0.999375-27). Retrieved from http://lme4.r-forge.r-project.org

De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye movement study of developmental dyslexia. *Brain and Language*, *80*, 617–626. doi:10.1006/brln.2001.2637

De Luca, M., Di Pace, E., Judica, A., Spinelli, D., & Zoccolotti, P. (1999). Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia. *Neuropsychologia*, *37*, 1407–1420. doi:10.1016/S0028-3932(99)00038-X

Denckla, M., & Rudel, R. G. (1976). Rapid 'automatized' naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, *14*, 471–479. doi:10.1016/0028-3932(76)90075-0

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. doi:10.1080/09297040500346571

Döpfner, M., Görtz-Dorten, A., Lehmkuhl, G., Breuer, D., & Goletz, H. (2008). *DISYPS-II: Diagnostik-System für psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche–II* [Diagnostic System for Mental Disorders in Childhood and Adolescence]. Göttingen, Germany: Hogrefe.

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–143). Hillsdale, NJ: Erlbaum.

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*, 5–21. doi:10.1080/10888438.2013.819356

Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition*, *115*, 367–379. doi:10.1016/j.cognition.2009.11.004

Hutzler, F., & Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. *Brain and Language*, *89*, 235–242. doi:10.1016/S0093-934X(03)00401-2

Hyönä, J., & Olson, R. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1430–1440. doi:10.1037/0278-7393.21.6.1430

Jones, M. W., Kuipers, J.-R., & Thierry, G. (2016). ERPs reveal the time-course of aberrant visual-phonological binding in developmental dyslexia. *Frontiers in Human Neuroscience*, *10*, 71. doi:10.3389/fnhum.2016.00071

Juphard, A., Carbonnel, S., & Valdois, S. (2004). Length effect in reading and lexical decision: Evidence from skilled readers and a developmental dyslexic participant. *Brain and Cognition*, *55*, 332–340. doi:10.1016/j.bandc.2004.02.035

Liversedge, S. P., Drieghe, D., Li, X., Yan, G., Bai, X., & Hyönä, J. (2016). Universality in eye movements and reading: A trilingual investigation. *Cognition*, *147*, 1–20. doi:10.1016/j.cognition.2015.10.013

Moll, K., Hutzler, F., & Wimmer, H. (2005). Developmental dyslexia in a regular orthography: A single case study. *Neurocase*, *11*, 433–440. doi:10.1080/13554790500263537

Moll, K., & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading*, *13*, 359–382. doi:10.1080/10888430903162878

Moll, K., & Landerl, K. (2010). *SLRT-II: Lese- und Rechtschreibtest* [Reading and Spelling Test]. Bern, Switzerland: Hans Huber.

Müller, R. (2004). *DRT 3: Diagnostischer Rechtschreibtest für 3. Klassen* [Diagnostic Spelling Test for the 3rd Grade]. Göttingen, Germany: Beltz.

Olson, R. K., Kliegl, R., Davidson, B. J., & Foltz, G. (1985). Individual and developmental differences in reading disability. In G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 4, pp. 1–64). New York, NY: Academic Press.

Paizi, D., De Luca, M., Zoccolotti, P., & Burani, C. (2013). A comprehensive evaluation of lexical reading in Italian developmental dyslexics. *Journal of Research in Reading*, *36*, 303–329. doi:10.1111/j.1467-9817.2011.01504.x

Petermann, F., & Petermann, U. (Hrsg.). (2011). *WISC-IV: Wechsler Intelligence Scale for Children-IV*. Frankfurt, Germany: Pearson.

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet*, *379*, 1997–2007. doi:10.1016/S0140-6736(12)60198-6

Peterson, R. L., Pennington, B. F., Olson, R. K., & Wadsworth, S. (2014). Longitudinal stability of phonological and surface subtypes of developmental dyslexia. *Scientific Studies of Reading*, *18*, 347–362. doi:10.1080/10888438.2014.904870

R Development Core Team (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly*, *27*, 29–53. doi:10.2307/747832

Rau, A. K., Moeller, K., & Landerl, K. (2014). The transition from sublexical to lexical processing in a consistent orthography: An eye-tracking study. *Scientific Studies of Reading*, *18*, 224–233. doi:10.1080/10888438.2013.857673

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422. doi:10.1037/0033-2909.124.3.372

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology*, *36*, 321–339. doi:10.1016/0022-0965(83)90036-X

Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior Research Methods*, *47*, 1085–1094. doi:10.3758/s13428-014-0528-1

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*, 151–218. doi:10.1016/0010-0277(94)00645-2

Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., & Zoccolotti, P. (2005). Length effect in word naming in reading: Role of reading experience and reading deficit in Italian readers. *Developmental Neuropsychology*, *27*, 217–235. doi:10.1207/s15326942dn2702_2

Torgesen, J. K., & Hudson, R. (2006). Reading fluency: Critical issues for struggling readers. In S. J. Samuels & A. Farstrup (Eds.), *Reading fluency: The forgotten dimension of reading success* (pp. 130–158). Newark, DE: International Reading Association.

Van Den Bos, K. P., Zijlstra, B. J. H., & Lutje Spelberg, H. C. (2002). Life-span data on continuous-naming speeds of numbers, letters, colors, and pictured objects, and word-reading speed. *Scientific Studies of Reading*, *1*, 25–50. doi:10.1207/S1532799XSSR0601_02

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*, 2–40. doi:10.1046/j.0021-9630.2003.00305.x

Weiß, R. H. (2006). *CFT 20-R: Grundintelligenztest Skala 2 – Revision* [Culture Fair Intelligence Test – Scale 2]. Göttingen, Germany: Hogrefe.

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics*, *14*, 1–33. doi:10.1017/S0142716400010122

Wimmer, H. (1996). The nonword reading deficit in developmental dyslexia: Evidence from children learning to read German. *Journal of Experimental Child Psychology*, *60*, 80–90. doi:10.1006/jecp.1996.0004

Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology*, *94*, 272–277. doi:10.1037/0022-0663.94.2.272

Wimmer, H., & Mayringer, H. (2014). *SLS 2-9: Salzburger Lese-Screening für die Schulstufen 2-9*. Salzburg, Germany: Hans Huber.

Wimmer, H., & Schurz, M. (2010). Dyslexia in regular orthographies: Manifestation and causation. *Dyslexia*, *16*, 283–299. doi:10.1002/dys.411

Ziegler, J. C., Perry, C., Ma-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. doi:10.1016/S0022-0965(03)00139-5

Appendix

List of the Items Used in the Eye-Tracking Paradigm

Words	Pseudohomophones	Nonwords
	Short	
arm	ahrm	orm
Hut	Hud	Lut
nur	nuhr	lur
Rad	Rahd	Rak
Tür	Tühr	Tar
Zoo	Zoh	Zot
Arzt	Artzt	Alzt
Bild	Billd	Buld
Brot	Broht	Krot
Film	Vilm	Folm
froh	fro	groh
Füße	Füse	Fäne
Glas	Glaas	Glap
Gras	Graas	Gres
Hals	Halls	Kals
Hand	Hant	Hond
Hose	Hoße	Kofe
Jahr	Jaar	Mahr
Kind	Kint	Hind
Luft	Lufft	Muft
Mund	Munt	Munk
Obst	Opst	Abst
rund	runt	sund
Saft	Safft	Naft
warm	wahrm	werm[a]
Wind	Wint	Nind[a]
Bäume	Beume	Käune
Blitz	Plitz	Glitz
Farbe	Fabe	Furde
Fuchs	Fux	Fechs
Haare	Hare	Laale
Katze	Katse	Kapfe
Kleid	Klaid	Kleud
König	Köhnig	Käbin
krank	kranck	krink
Stein	Stain	Staum
stolz	stoltz	stelz
teuer	täuer	beuel
Tiere	Tire	Kiene
Vater	Fater	Kaber
	Long	
Butter	Butta	Lutten
fahren	faren	dahnen
falsch	fallsch	fulsch
fehlen	felen	fahmen
Ferien	Fehrien	Delien
fragen	frahgen	flaben
Leiter	Laiter	Meiker
Mutter	Mutta	Sutten
Schnee	Schneh	Schree
Schule	Schuhle	Scheke
selbst	sellbst	solbst
Sommer	Somma	Summel
sparen	schparen	spolen
wählen	wehlen	sähren
Wälder	Welder	Pälger
Wasser	Wassa	Dassel

(*Continued*)

(Continued).

Words	Pseudohomophones	Nonwords
wohnen	wonen	gohlen
zählen	zehlen	zühnen
Zettel	Zättel	Zutten
Zimmer	Zimma	Zommel
Blätter	Bletter	Flätten
Familie	Famielie	Sanilia
Fenster	Fänster	Fonstel
gestern	gesstern	tesbern
Mädchen	Medchen	Müdchel
niemals	nimals	miefals
Schwanz	Schwantz	Schwunz
täglich	teglich	fägnich
Wohnung	Wonung	Kohnang
erfunden	erfunnden	engunsen
Lehrerin	Lererin	Fehrasin
Mittwoch	Mitwoch	Littmoch
schlecht	schlächt	schlocht
sprechen	schprechen	spracker
streiten	schtreiten	streufen
wünschen	wünnschen	tünschan
Geheimnis	Geheimniss	Beleimnes
schwimmen	schwimen	schwillen
unmöglich	unmöklich	urdögnich
verlaufen	ferlaufen	ventauben

[a]Due to experimental error, the nonword *werm* was not presented, and the nonword *Nind* was presented twice. The second occurrence of *Nind* was not included in the analysis.