

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Hautala, Jarkko; Hawelka, Stefan; Aro, Mikko

Title: Dual-stage and dual-deficit? Word recognition processes during text reading across the reading fluency continuum

Year: 2022

Version: Published version

Copyright: © The Author(s) 2021

Rights: _{CC BY 4.0}

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Hautala, J., Hawelka, S., & Aro, M. (2022). Dual-stage and dual-deficit? Word recognition processes during text reading across the reading fluency continuum. Reading and writing, 35(3), 663-686. https://doi.org/10.1007/s11145-021-10201-1



Dual-stage and dual-deficit? Word recognition processes during text reading across the reading fluency continuum

Jarkko Hautala¹10 · Stefan Hawelka²10 · Mikko Aro³10

Accepted: 5 August 2021 © The Author(s) 2021

Abstract

Central questions in the study of visual word recognition and developmental dyslexia are whether early lexical activation precedes and supports decoding (a dualstage view) or not (dual-route view), and the locus of deficits in dysfluent reading. The dual-route view predicts early word frequency and length interaction, whereas the dual-stage view predicts word frequency effect to precede the interaction effect. These predictions were tested on eye movements data collected from (n=152) children aged 9-10 among whom reading dysfluency was overrepresented. In line with the dual-stage view, the results revealed an early word frequency effect in first fixation duration followed by robust word length effect in refixation probability and an interaction of word frequency and word length in summed refixation duration. This progression was advanced in fluent reading to be observable already in first fixation duration. Poor reading fluency was mostly explained by inflated first fixation durations, and to stronger word frequency and length effects in summed refixation duration. This pattern of results suggests deficits in early letter encoding and slowness in serial grapheme-phoneme conversion. In contrast to the widely held belief, the holistic orthographic processing of words seemed to be intact.

Keywords Word recognition \cdot Eye movements \cdot Developmental dyslexia \cdot Word length \cdot Word frequency

Jarkko Hautala jarkko.hautala@nmi.fi; jarkko.v.hautala@jyu.fi

¹ Niilo Mäki Institute, Asemakatu 4, P.O. Box 29, 40101 Jyvaskyla, Finland

² Centre for Cognitive Neuroscience, University of Salzburg, Salzburg, Austria

³ Department of Education, University of Jyväskylä, Jyvaskyla, Finland

Introduction

Research of visual word recognition has led to detailed theories and computational models (Coltheart et al., 2001; Perry et al., 2010; Seidenberg, 2007). These models have been developed predominantly to explain response accuracies and times to individually presented words, leaving the actual time course of cognitive processes involved somewhat underspecified. On the other hand, equally sophisticated models have been developed for the eve movement control of reading, in which rudimentary assumptions about word recognition progression had been made (see e.g., Reichle, 2015). Although theoretical integration of eye movement control and visual word recognition would be pertinent for understanding reading and its development, attempts for such an integration are rare (Hawelka et al., 2010). In fact, visual word recognition theories are commonly used as a general conceptual framework to interpret reader's eye movements especially in developmental studies, in which word recognition cannot be understood simply as an all-or-none perception, but as a gradual decoding of a word (Hutzler & Wimmer, 2004; Rau et al., 2014, 2015; Tiffin-Richards & Schroeder, 2015). Taken together, there is a clear need of conceptual linking readers' eye movements with visual word recognition sub-processes. As a step towards this direction, the present study examines the cognitive architecture of visual word recognition as reflected in eye movements during text reading among 9 to11-year-old readers across the whole continuum of reading fluency.

The prominent dual-route view of (single) word recognition assumes that word recognition is achieved via two parallel and independent pathways handling the visually encoded letter string: an indirect and a direct route (Coltheart et al., 2001; Perry et al., 2010). The indirect route assembles the phonology of a word via serial *grapheme-phoneme conversion* (GPC) producing a word length (WL) effect; the direct route addresses whole-word phonology via the activation of whole-word representations in a mental (i.e., orthographic) lexicon. The speed of such activation is assumed to be dependent mostly on word frequency (WF). Although, in principle, dual-route models predict length effects only for words which are not represented in the orthographic lexicon, such as pseudowords, in reality many low frequency words may be novel to a reader. Therefore, it can be assumed that the dual-route view predicts WF×WL interaction (Balota et al., 2004; Kapnoula et al., 2017). Noteworthy, due to early divergence of the two routes immediately after the letter encoding stage (Perry et al., 2014), this interaction should start to emerge early (Fig. 1).

Fig. 1 Schematic illustration of dual-route and dual-stage views of word recognition. The dotted line represents the connection emerging potentially late in the course of reading development



However, it is not sure whether the traditional dual-route view is still entirely compatible with the current neurocognitive understanding of word recognition (e.g., Jobard et al., 2003; Taylor et al., 2013). Neurocognitively, word recognition begins in the visuo-occipital cortex with letter encoding (Thesen et al., 2012) and continues by feedforward activation towards larger orthographic units in the visual word-form area (Cohen et al., 2000; Dehaene et al., 2005) and ventral occipitotemporal cortex (Price & Devlin, 2011). After this, lexico-semantic processing is assumed to take place in the temporal cortex (Schurz et al., 2010; Wydell et al., 2003). According to the interactive activation model (Price & Devlin, 2011), even the earliest activations of letter combinations are immediately transmitted to relevant frontal and parietal areas and back with the feedback being used for enhancing the still ongoing orthographic processing in the ventral occipitotemporal cortex (Cornelissen et al., 2009; Wheat et al., 2010; Woodhead et al., 2014). The parietal feedback is believed to be associated with the control of serial attention for attaining precise encoding and ordering of letters and thus controlling eye movements (Pugh et al., 2013; Reichle et al., 2003; Richlan, 2014), whereas the frontal feedback may provide phonological and other, higher level predictions (Himmelstoss et al., 2020; Richlan, 2014).

The compatibility of the dual-route architecture with this neurocognitive view boils down to the question whether the encoded letter string is subject to serial processing from the beginning. Another possibility is that parallel orthographic processing largely precedes serial processing (Jobard et al., 2003, 2011), the latter being needed when the former process/mechanism encounters difficulties, such as during reading a nonword. We label this alternative view as a dual-stage view (Fig. 1), in which early but incomplete parallel activation (e.g. **THYNK**, the bolded bigrams representing the rapidly activated ones) is subject to the serial decoding process (TH-**Y**-NK, the bolded letter representing the one requiring most attention; Hautala & Parviainen, 2014). The previous models of early parallel processing of letter strings (Dehaene et al., 2005; Price & Devlin, 2011) and serial predictions of word's sublexical part (Sibley et al., 2010) are promising accounts for possible mechanisms of parallel and serial processing. The core prediction of this view is that the WF effect precedes WL effect.

One potential way of trying to disentangle the early stages of word recognition is to analyze the reader's first-pass eye movements (Hawelka et al., 2010). Apart from the parafoveal preview effects, the earliest foveal measure reliably reflecting word recognition processes is the duration of the very first fixation on a word (i.e., first fixation duration; FFD). Given that word recognition may even complete during a single fixation, special analytic methods—such as quantile regression—is needed to study which effect precedes another effect during FFD, that is, affecting already short fixation durations. From the perspective of eye movement control, a decision to refixate a word must be made during FFD. In the prominent E–Z Reader model of eye movement control during reading (Reichle et al., 2003), the first word recognition stage, the "familiarity check", informs the saccadic system whether the current word will be recognized by a single fixation (i.e., whether whole-word recognition is imminent). Thus, the first fixation duration (FFD) can be assumed to reflect at least coarse word-form activation. However, FFD also reflects other processes: Sometimes, when knowing early on that a refixations will be needed, readers reduce their FFDs (so-called number of fixation and duration trade-off effect) manifesting in a negative influence of WL on FFD (Loberg et al., 2019; Sperlich et al., 2015). In addition, a small and early WL effect in FFD may also stem from early visual encoding processes [about 5 ms per letter as estimated by Hautala and Loberg, (2015); see also Reichle et al., (2003)]. The decision to refixate (refixation probability, RP) is mostly governed by the visual extent of a word (Hautala & Loberg, 2015; Hautala et al., 2011a) and also, to some extent, by WF reflecting lexical influences (e.g., Hawelka et al., 2010), the former suggesting a need to sample more visual information from a word, and the latter the need to resolve lexical recognition. The remaining word recognition processes can be summarized with a summed refixation duration (SRD) measure, which is the sum of all first-pass fixation durations, except the first fixation on a word, and therefore, a late measure complementary to RP.

It is a widely accepted view that developmental dyslexia (DD) stems mostly from a *phonological processing deficit*. However, it is far from clear how this core deficit affects the development of visual word recognition processes. According to the prevalent view, deficient learning of G-P associations impairs phonological decoding and self-teaching of new words resulting in an impoverished orthographic lexicon (Álvarez-Cañizo et al., 2018; Araújo & Faísca, 2019; Blomert, 2011; Conway et al., 2017; Dürrwächter et al., 2010; Hautala et al., 2011b; Hyönä & Olson, 1995; Mehlhase et al., 2019; Perry et al., 2019; Richlan, 2019; Saksida et al., 2016; Share, 2008). In the dual-route framework, such a limited orthographic lexicon should lead to (1) a larger WL effect due to a higher probability of a word being read by the indirect route and (2) a stronger WF×WL interaction due to having orthographic word representations only for the most frequent words. These effects should start to emerge immediately after letter encoding has been completed, which, however, may endure the whole FFD in readers with low reading fluency. In this case, the effects of WF and WL and their interaction may emerge only in later RP and SRD measures.

According to current neurocognitive knowledge, DD is associated with profound difficulties in establishing efficient word recognition circuitry (Perry et al., 2019; Price & Devlin, 2011), manifesting in poorer general connectivity between visual and verbal areas (Schurz et al., 2015) and delayed lateralization of the reading network on the left cerebral cortex (Finn et al., 2014). More specifically, weaker response in the left fusiform gyrus for words and a stronger involvement of parietal areas suggest higher reliance on phonological decoding (Pollack et al., 2015; Richlan, 2014). From the perspective of the dual-stage view, one may assume that the lack of orthographic word representation is not the core deficit, but difficulty in establishing direct connections from orthographic word representations to their phonological counterparts (Boets et al., 2013). In addition, the word decoding may be laborious and poorly automatized. Together these difficulties should lead to normal WF effect in FFD, followed by an—compared to normally reading children inflated WL and WF interaction in later measures (i.e., RP and SRD; Hautala et al., 2011b).

Finally, some authors have suggested that slowness in visual word recognition in DD is largely explained by a single deficit in early visuo-orthographic processing disrupting essentially the letter encoding stage (Boros et al., 2016; Martelli et al., 2009; Moll & Jones, 2013; Paizi et al., 2013; Prado et al., 2007; Sperlich et al., 2015). Such a deficit should be associated with highly inflated FFDs, which would then largely explain all other difficulties in word processing, such as generally inflated WF and WL effects and their interactions.

Linguistic factors in visual word recognition

The cognitive architecture of word recognition and the underlying deficits in developmental dyslexia are believed to be universal (Carioti et al., 2021; Rueckl et al., 2015), but the language is believed to have an important role in determining the cognitive computations (Frost, 2012). Orthographic depth (Katz & Frost, 1992), that is the complexity of mapping between orthography and phonology, is known to strongly determine the speed of reading acquisition (Seymour et al., 2003). In opaque orthographies the pronunciation rules may not be straightforward and a large portion of words are completely irregular, whereas in transparent orthographies serial decoding leads to the correct pronunciation for most of the words. Although there are findings that serial decoding is emphasized in transparent orthographies (Ziegler et al., 2001, 2003), recent studies have complicated this issue (Marinus et al., 2015; Schmalz et al., 2017). To illustrate, morpho-syllabic parsing is necessary for accessing correct word meaning and to generate proper stress assignment. In principle, to segment syllable and morpheme boundaries requires parallel processing of multiple letters. Consistently with this fact, research suggests that the parsing is instantiated on the whole-word input rather than progressing in a strictly serial manner over the letter string (Kuperman et al., 2008; Perry et al., 2010). Studies in transparent orthographies indicate no or small number of syllable effects on word recognition times across reading fluency (Barca et al., 2002; Chetail, 2014; De Luca et al., 2008; Hautala et al., 2013). Furthermore, morphological complexity actually supports word recognition presumably due to richer lexical activation (Burani et al., 2008; Hasenäcker & Schroeder, 2017).

The present study was conducted in the Finnish language, which has a fully transparent orthography with single letter-sound mappings, a rich morphology, and a clear syllable structure with syllabic stress always on the first syllable (Aro, 2017). Similar as in other transparent orthographies, in Finnish developmental dyslexia manifests mainly as slow but fairly accurate reading (Aro, 2017). The fully transparent orthography allows beginning and dysfluent readers to rely even on the highly serial online "sliding" decoding strategy, in which decoding partially overlaps with pronunciation during reading aloud performance (Hautala et al., 2013).

Developmental eye movement studies of reading

Developmental eye movement studies and behavioral studies of WF×WL interaction (Paizi et al., 2013) have produced largely convergent findings. However, concerning the time course of word recognition processes, several previous studies (Huestegge et al., 2009; Hutzler & Wimmer, 2004; Hyönä & Olson, 1995; Joseph et al., 2009, 2013; Rau et al., 2014, 2015; Tiffin-Richards & Schroeder, 2015) have found minimal to non-existing length effects in FFD, followed by strong length effects in RP and in summed duration of first-pass fixations, i.e. gaze duration (GD). The same studies also consistently report a WF×WL interaction in these late measures. However, some studies have reported WF effect emerging in children first in GD, not in FFD (Huestegge et al., 2009; Joseph et al., 2013). Despite the possibility that all word recognition processes in fluent adult readers may occur during a single fixation, there are findings of word frequency effect preceding word length effects. Calvo and Meseguer (2002) reported a frequency effect in FFD, preceding the length effect that they observed in GD. Likewise, Kliegl et al. (2004) did not observe the WL effect in FFD, but in later measures.

There is a paucity of studies on the eye movement of individuals with DD during reading. Dürrwächter et al. (2010) reported larger length and frequency interactions for dyslexic German children as compared to typical readers in various eye movement measures. Hutzler and Wimmer (2004) report a somewhat larger length and frequency effect in FFD for dyslexic readers, but the difference in effect sizes between dyslexic and normal readers was much more pronounced in GD. Hawelka et al. (2010) replicated the findings of Hutzler and Wimmer (2004) and found a much more pronounced WL×WF interaction among dyslexic readers for the number of fixations per word and GD, but they did not assess whether this interaction also occurs in FFD. However, a sizable correlation between dyslexic readers' mean single fixation duration and their (poor) performance in a rapid automatized naming task (Denckla & Rudel, 1974) was suggestive of a speed deficit in accessing the phonology of instantiated orthographic representations. Taken together, the results of eye movement studies on reading seem to lean more in the direction of the dualstage rather than the dual-route view of word recognition.

In Finnish, low reading fluency (RF) in adulthood was associated with longer total fixation duration on words, but not with a larger length effect (Hautala & Loberg, 2015). On the contrary, Hautala et al. (2011b), found that dyslexic children exhibit a substantial WL effect in the number of first-pass fixations even when reading the same items repeatedly, yet show a robust lexicality (words vs. pseudoword) effect on the average fixation durations. Overall, the existing literature, although scarce, supports the notion that DD is associated with slower GPC. However, less is known about the early lexical level of word processing in DD due to a lack of studies analyzing FFD.

The present study

In this study, the word recognition process of developing readers during text reading was studied by analyzing FFD, RP, and SRD with linear mixed models (LMM) in a large group of Finnish third and fourth grade children across a broad continuum of reading fluency.

The first goal of the study was to determine whether the dual-route or dual-stage view provides a better account of word recognition progression in fluent reading. The respective research question was whether the WF effect precedes the WL effect in fluent readers' eye movements? Crucially, the dual-route view predicts the early

and parallel rise of WL, WF, and WF×WL interaction effects, starting in FFD and continuing in RP and SRD measures. The dual-stage view, in contrast, posits that the WF effect precedes the WF×WL interaction. In fluent readers this pattern may appear already within FFD. Thus, in order to tease apart the time course of the influence WF and WL, we will resort to a quantile regression approach (e.g. Yap et al., 2012); the interaction process is then expected to continue through the RP and SRD measures.

Our second research question asked which component processes are mostly affected in low reading fluency and DD? The low-end readers are overrepresented in our sample, allowing us to draw conclusions for DD in transparent orthographies. First, it was assumed that low reading fluency (RF) is associated with generally longer FFDs due to the expected difficulty in the visual letter encoding stage (Paizi et al., 2013). Then, according to the dual-route perspective, word recognition difficulties in DD result mostly due to a small orthographic lexicon (e.g., Bergmann & Wimmer, 2008; Hawelka et al., 2010). This should lead to a pronounced WF×WL interaction emerging as soon as the letter encoding stage has been completed, that is, in FFD and in later measures. The dual-stage view, to the contrary, does not assume that activation of orthographic word representation would guarantee access to its phonological counterpart. This view would predict an intact WF effect in FFD, to be followed by a pronounced WF×WL interaction in later RP and SRD measures. We also studied whether a single deficit in letter encoding provides a sufficient account for low RF. This was examined with a regression analysis testing whether longer FFDs can sufficiently explain RF as predicted by the single-deficit view, or whether WF and WL effects may explain additional variance in RF, as predicted by the dualroute and dual-stage accounts of DD.

Materials and methods

Participants

The participants were 152 third- and fourth-grade students from five schools in central Finland, with a mean age of 10 years and 1 month (SD=7 months). All students followed the standard curriculum with school instruction provided in Finnish. The eye-tracking data of 10 students were excluded for various reasons (participant quitting the participation, inaccurate calibration, or participant just skimming through the texts). This resulted in a final sample of 142 students, consisting of 54 third graders and 88 fourth graders (79 girls, 63 boys). According to the questionnaire completed by the students' caregivers, Finnish was not the first language of three students, but they still reported moderate or good oral language proficiency in Finnish. The caregivers also reported that 19 students had either a suspected or identified reading deficit (DD is rarely formally diagnosed in Finland).

The data reported in the present article concerns screening and pretest assessments of a reading intervention study targeting dysfluent readers. The study was pre-evaluated by the Ethical Committee of the [Anonymized for review], and the research was conducted according to the ethical principles for medical research involving human subjects set forth by the Declaration of Helsinki.

Reading fluency assessment

The RF assessment consisted of two tasks. The first was a standardized readingaloud task involving a word list (Lukilasse 2; Häyrinen et al., 2013), with the raw score being the number of words read accurately within 2 min (M = 78.6, SD = 16.2, range 32-105). Because the norms of this task were collected at the end of each grade, and our assessment was conducted in November, standardized scores were calculated based on the average of previous and current class-level norms. The second reading task required students to read a 124-word text ("Exciting travels") aloud. The number of words read correctly within 1 min was used as a raw score (M=69.1, SD=22.6, range 14-124), which was then standardized into z-scores by grade level with reference to large-scale research data (FirstSteps-study, e.g., Kiuru et al., 2015). The average of the standardized values across the two tasks was used as an index of RF (Cronbach alpha reliability = 0.917). On average, the standardized RF of the participants in this study was relatively low (M = -0.66, SD = 1.06, range)-3.13 to 1.78) but normally distributed (skewness = -0.005, SE = 0.20). To account for the grade-effect, sample-specific instead of age-normative standardized values were used in the analyses.

Materials

Participants read two excerpts from the beginning of abridged versions of classical stories: *Little Heidi* by Johanna Spyri (1881; Finnish abridgment by Kati Weiss) with 457 words, and *Adalmina's Pearl* by Zacharius Topelius (1865) with 403 words. The copyrights of these works have expired, and modernized Finnish abridgements are available at https://iltasatu.org/.

After each story, students completed five four-choice questions about the story and one Yes/No question about whether the story was familiar to the reader. Reading comprehension questions were answered with an accuracy of M=79%, SD=16%, range: 30–100%. Thirty-four children were familiar with one of the stories; eight knew both. Two-gram frequencies of consecutively appearing word pairs were derived from the Finnish N-gram corpus (2014). Word frequency, minimum syllable frequency in a word (an index of sublexical difficulty), and number of syllable indices were derived from the latest published corpora (Table 1; Huovilainen, 2018). All corpus frequency measures were log-transformed occurrences in a million words. Word length and frequency were correlated with r=-0.75.

Apparatus and procedure

Eye movements were recorded with SMI remote eye-tracking devices with a 250 Hz sampling rate installed on laptop computers with a screen size of 34.5×19.5 cm. We used fully adjustable chin rests modified from camera mounts

Story	nr	WL		WF (log)		MinSyl (og)	Twogram (log)		
5		M (SD)	Min–max	M (SD)	Min–max	M (SD)	Min–max	M (SD)	Min-max	
Heidi	457	6.3 (2.6)	2–17	1.7 (1.3)	0-4.5	2.8 (0.7)	0.71-4.52	1.9 (1.1)	1.04-5.83	
Adalmina	403	6.4 (3.0)	2–17	1.9 (1.4)	0–4.5	2.9 (0.8)	0.83-4.31	1.9 (1.2)	1.04-6.45	

Table 1 Means (SD) of word properties of the stimulus texts

WL, word length in letters; WF, word frequency in a million word; MinSyl, minimum syllable frequency in a million word; Twogram, frequency of appearance of two consecutive words

to stabilize the participants' heads while they sat on a non-adjustable chair. The texts were presented with the SMI Experiment Center 3.6 program on 11 five-line screens with no option to return to previous screens (Fig. 2). Arial 28 pt. font was used, corresponding to approximately five letters per degree of visual angle at 60 cm viewing distance. A full-screen 13-point calibration routine was completed prior to both of the stories being read, and a four-point calibration validation routine was completed in the middle of each story.

Two researchers in a dimly lit classroom at each school recorded the eye movements with four eye-tracking devices. Instructions for the task were given simultaneously via on-screen text and through headphones. The experiment began with a calibration and two-screen practice text followed by a multiple-choice comprehension question to introduce participants to the experimental procedure. After the practice text, calibration was repeated, and students proceeded to the actual experimental texts. After reading a screen of text, students proceeded to the next screen by looking at a large gaze-sensitive area centered on a target arrow in the right-bottom corner of the screen. A pause intervened between the two stories,



Fig. 2 An experimental screen overlaid with areas-of-interest and the gaze path of an exemplary participant

allowing the children the opportunity to lift their heads from the chinrest before recalibrating and continuing.

Eye movement data processing

The recorded data was imported into the SMI Begaze 3.6 program for preprocessing. To detect refixation saccades with small amplitudes, sensitive saccade detection parameters of 20 deg/s minimum angular velocity, a saccade duration of 15 ms, and a minimum fixation duration of 50 ms were applied, and blinks were excluded from the data. The vertical boundaries of automatically generated word-specific area-ofinterests were manually extended to a middle position between the lines.

Trained research assistants manually inspected scanpaths of all screen recordings to correct systematic drifts in the data (n=210) and mark occasions where data was of poor quality or fully (n=105) or partially (n=65) missing, affecting 380 of the 1694 screens (22%). The inter-rater agreement as to whether to make a correction on a screen was 94% for all of the 142 recordings of the first text screen.

The first- and second-pass fixations were identified with a custom script in the SPSS 26 program. Being aware that some of the participants only skimmed through parts of the test, only text lines in which more than 60% of words were fixated, were included in the analysis (1573 words excluded). Return-sweep fixations that did not land on the next text line's initial word were excluded. The area-of-interest aggregated data was exported for statistical analysis. After preprocessing, the data consisted of 114,485 word-observations (M=753 per participant) out of 122,120 possible observations.

Analyses

To study the time course of word recognition processes, statistical analyses were conducted separately for FFD, RP, and SRD measures using (generalized) linear mixed models with the Imer-package in R (Bates et al., 2019). The full-factorial fixed-effect structure of the independent variables WL, WF, and RF was analyzed. First, for FFD and SRD, even the simplest models with polynomial terms for WL with random linear slope effect did not converge. Next, in the modeling of linear effects, the control variables of story familiarity, minimum syllable frequency, and two-gram frequency were not significant in the maximum model and were dropped. Maximal random structure resulting in model convergence was applied. Whenever convergence had to be obtained, the correlation of random effects was dropped from the model utilizing the afex-package (Singmann et al., 2015). For the SRD model this consisted of random intercepts and slopes of WL and WF for participants, and random intercept for items. For the FFD model, the random slope of WF was omitted. For the binomial RP model only random intercepts were included. Only instances in which a word was fixated were included in the analyses (i.e., word skips were excluded). For FFD and SRD, log-transformed values were analyzed to fulfill regression model assumptions, including random distribution of residuals and good fit over the entire scale of the dependent variable. The multicollinearity was

inspected with the vif-function (variance inflation factor) of the car-package and was not found to be problematic for any of the analyses (<3.1; De Jongh et al., 2015). The observed statistical power (simr -package; Green & MacLeod, 2016) in the *lmer*-analyses was optimal (90–95%) for testing two-way interactions, but low for testing the three-way interaction (30–40%). Standardized effect sizes are being reported to evaluate the practical relevance of the effects.

To unravel the time course of WF and WL effects and their interaction in fluent reading, a Bayesian quantile regression analysis (brms-package in R; Bürkner, 2017) was conducted for the students belonging to the highest tertile in reading fluency within the sample (sample-specific z-value > 0.56, grade-normative reading fluency M=0.50, SD=0.53, n=47). In the analysis, FFDs (shorter than 800 ms) were analyzed in 0.25, 0.5 and 0.75 quantiles. Informative priors of beta estimates and their standard deviations were derived by a linear mixed model with a formula lmer (FF D~WF*WL+(1lid)+(1litem)). The estimation was run with 2000 iterations for the specified quantiles leading to model convergence (the diagnostic rhat-values were lower than 1.1).

The LMM analyses provided insight into the time course of the word recognition processes, but not their interdependence. For example, longer overall FFD may lead to larger length, frequency, and interaction effects in FFD, RP, and SRD measures, as suggested by the single-deficit view (Paizi et al., 2013). Such possible interdependence of the effects was studied in a separate hierarchical regression analysis. Here, the individual WF, WL, and WF × WL interaction effect coefficients (Carter & Luke, 2018) are used to explain the students' reading fluency in a three-step hierarchical regression analysis: The first step contains the intercept and all WF and WL-related effects found to be significant in the FFD measure. The second step contains all significant WF and WL-related effects for RP, and the third step contains the same for SRD. The recommended backward selection procedure of significant predictors (van Houwelingen & Sauerbrei, 2013) could not be used due to the time-course ordering of our predictors. Instead, a stepwise procedure with a strict inclusion criterion was used (p values < 0.01 for inclusion and > 0.05 for removal). Further, it was checked that the significance level of predictors remained identical in the full model, suggesting that the solution was robust. The observed statistical power of the analysis was 92% to detect effect sizes > 0.15 (GPower; Faul et al., 2009).

Results

The present research questions are answered by inspecting the order of appearance of the WF, WL, RF and their interaction effects across FFD, RP and SRD measures. However, the first research question "Does the WF effect precede the WL effect in fluent readers' eye movements?" required additional quantile regression analysis, and the second question "Which component processes are mostly affected in low reading fluency?" required running also a hierarchical regression analysis.

Descriptive raw data are plotted in Fig. 3 showing how WF effect can be seen already in FFD, weakly on RP and then strongly on GD, being pronounced for long rare words and especially for slow readers. The WL effect manifests strongly



Fig. 3 Means and confidence intervals of FFD, RP and GD as a function of reading fluency (columns), word length (horizontal axis) and word frequency (lines). The percentile groups represent the thirds of observed values in each measure

in RP and then in the interaction with WF in GD, being clearly larger for less fluent readers. Linear change seems to be the dominant pattern, apart from the "bump" at word lengths of 15–16 letter words, which happened to be rare in texts. The standardized beta coefficients (b') and statistical test results for the effects reaching significance are shown in Table 2. Estimated marginal means of LMM results are shown in Fig. 4.

Effect	First fixation duration				Refixation probability				Summed refixation dura- tion			
	b´	SE	t	Sig	b'	SE	z	Sig	b´	SE	t	Sig
RF	22	.02	- 8.98	***	97	0.08	- 12.10	***	32	.02	- 14.32	***
WL	.01	.01	0.99		1.59	0.05	30.9	***	.17	.02	11.04	***
WF	07	.01	-6.76	***	37	0.05	-7.04	***	18	.02	-11.09	***
RF×WL	.01	.01	2.31	*	07	0.03	-2.48	*	04	.01	-4.70	***
RF×WF	.01	.01	2.64	**	.01	0.03	-0.31		.04	.01	4.40	***
WL×WF	01	.01	-0.86		.00	0.04	0.01		10	.01	-9.43	***
RF×WL×WF	01	.00	-1.89	1	04	0.03	-1.41		.00	.00	-0.48	

RF, reading fluency; WL, word length; WF, word frequency; SE, standard error **p < .001, *p < .01, *p < .05, 1 < .1



Fig. 4 Estimated marginal means with their 95% confidence intervals produced from the linear mixed models. Z refers to standardized values, RF=reading fluency, WF_log=logarithmic word frequency, length=word length

First fixation duration

The model log(FFD)~RF*WL*WF+(1+WLllid)+(1litem) revealed highly significant main effects related to RF (b'=-0.22) and WF (b'=-0.07), which were accompanied by weak but significant interactions of RF×WF (b'=0.01) and RF×WL (b'=0.01). The main effects indicate that FFD increased as a function of decreasing RF and WF. The interactions indicate that the WF had a slightly greater influence for less fluent readers, and WL had a slightly greater influence for faster readers. Importantly, the main effect of WL was not significant. These results indicate that the WF effect precedes the WL effect in dysfluent readers. In contrast, among the fluent readers the order of WF and WL effect could not be resolved by this analysis.

The quantile regression analysis provided more fine-grained information about the time course of the effects in fluent readers. The model was run with a formula brm $(bf(FFD \sim WF * WL + (1|id) + (1|item))$. The 0.25 quantile (M = 126 ms, SD = 33 ms, max = 168 ms) analysis resulted in significant (95%) effects for WF, (b = -6.44,CI95% = -8.53 to -4.36, Evidence ratio = Infinite, Posterior probability = 1) and WF×WL-interaction, (b = -0.78, CI95% = 0.43 - 1.13, Evidence ratio = Infinite, Posterior probability = 1). The 0.5 quantile (M = 209 ms, SD = 25 ms, max = 256 ms) analysis resulted in significant effects for WF (b = -6.03, CI95% = -9.11 to -2.99, Evidence ratio = 1332, Posterior probability = 1) and WL (b = 2.03, CI95% = 1.03-3.06, Evidence ratio = 1332, Posterior probability = 1). The 0.75 quantile (M=319 ms,SD=40 ms, max=399 ms) analysis resulted in significant effects for WL (b=4.78, CI95% = 3.25 - 6.41, Evidence ratio = Infinite, Posterior probability = 1) and WF×WL-interaction, (b = -1.97, CI95% = -2.77 to -1.19, Evidence ratio = Infinite,Posterior probability = 1).

Figure 5 shows the nature of these interactions: The WF×WL interaction at 0.25 quantile resulted from a small WL effect for the high-frequency words. In the 0.5 quantile, there was a uniform WL effect across WF range. Finally, the strong WF×WL interaction appeared in the 0.75 quantile, showing a strong WL effect for low frequency words. As will be discussed, this complex pattern of results can be interpreted in terms of parafoveal preview providing a "headstart" for holistic visuo-orthographic processing of frequent words in fluent reading.



Fig. 5 Estimated marginal means with their 95% confidence intervals produced from the quantile regression model

Refixation probability

The model $RP \sim RF * WL * WF + (1lid) + (1litem)$ revealed highly significant main effects related to RF (Odds-ratio [OR]=0.64), WL (OR=2.07) and a weak WF effect (OR=0.84). These main effects were accompanied by a significant but very weak interaction of RF×WL (OR=0.97). The direction of these effects was that infrequent or longer words were refixated more often, and the WL effect was stronger in less fluent readers. These results indicate that following the early WF effect in FFD, WL effect manifests strongly in refixations largely independent from RF. However, dysfluent readers generally made more refixations even to high-frequency short words.

Summed refixation duration

The model log(SFD) ~ RF * WL * WF + (1 + WL + WFllid) + (1litem) revealed highly significant main effects of RF (b' = -0.32), WL (b' = 0.17), and WF (b' = -0.18), which were accompanied by highly significant two-level interactions of RF×WF (b' = 0.04), RF×WL (b' = -0.04) and WL×WF (b' = -0.10). The main effects indicate that SRD increased as a function of decreasing RF and WF and increasing WL. The interactions indicate that the WL effect increased as a function of decreasing WF and that the effects of WL and WF were larger for less fluent readers. These results demonstrate how the decoding speed of a word is strongly dependent on its frequency. Considering high frequency words, low RF was associated with a pronounced WL effect, which was not the case for readers with a high RF.

Hierarchical regression analysis

Table 3 presents the results of the stepwise regression analysis, in which RF was explained by individual coefficients derived from the above reported lmer analyses. The FFD intercept coefficients, that is, the average FFD, explained 36% of RF. In addition, the WL coefficients of FFD explained 4.6% of the additional variance, indicating that individuals who exhibited the WL effect in FFD were actually faster readers. However, the WL and WF coefficients for RP did not explain any additional variance of RF, suggesting that refixations are closely tied with inflated FFDs. In the third step, SRD effects explained an additional 14.1% of RF in the order of WF and

	•	•						
Step	Effect	b'	R^2	Adj. R ²	SE	ΔR^2	$\Delta F(1, > 135$	6)
1	FFD, intercept	49***	.36	.35	.75	.357	77.9***	
1	FFD, length	.17**	.40	.40	.72	.046	10.8**	
3	SRD, word frequency	.30***	.52	.51	.65	.114	32.5**	
3	SRD, word length	27***	.57	.56	.66	.027	7.53**	

Table 3 Results of the regression analysis

FFD, first fixation duration; SRD, summed refixation duration; b', standardized beta coefficient of the full model; Adj., adjusted; Δ , change

****p*<.001, ***p*<.01, **p*<.05

WL coefficients. These SRD findings indicate that RF is partly explained by the efficiency of the late GPC process itself, which is, therefore, not fully dependent on the reader's initial response during FFD.

Discussion

The present study investigated whether a dual-route (Perry et al., 2010; Pritchard et al., 2012) or a dual-stage (Jobard et al., 2003, 2011) view of word recognition is more compatible with eye movement behavior during reading in a sample of 3rd and 4th grade children representing a broad range of reading fluency distribution. We hypothesized that the dual-stage view would predict a sequential manifestation of WF and WL effects, whereas the dual-route view would predict a parallel rise of these effects. In addition, it was studied which processes explain the individual differences in reading fluency.

Dual-route or dual-stage?

Concerning fluent reading, the quantile regression analysis provided novel evidence about the early processes during a word's first fixation: The earliest observed effect was a weak length effect for high-frequency words followed by a length effect for low frequency words. This pattern of results can be understood only by taking parafoveal preview processes into account: Parafoveally activated orthographic word representations facilitated the subsequent foveal visuo-orthographic processing of a word, such as overcoming influences of visual eccentricity (Reichle et al., 2003) and visual crowding (Hautala & Loberg, 2015; Hautala et al., 2011a). A similar advancement of processes in fluent readers has been previously reported for the effect of word predictability (Hawelka et al., 2015). Then at later quantiles of FFDs, the GPC started, as evidenced by a strong WL effect only for low-frequency words. Overall, these findings indicate that activation of orthographic word forms precede decoding through parafoveal preview, providing support for the dual-stage view (Jobard et al., 2003, 2011; see Fig. 6).

Also, in line with the dual-stage predictions, a strong WF x WL interaction was observed in SRDs of fluent readers, paralleling previous GD findings (Hutzler & Wimmer, 2004; Hyönä & Olson, 1995; Rau et al., 2014, 2015; Tiffin-Richards & Schroeder, 2015). According to the dual-stage view, more frequent words are decoded faster, because early orthographic activations provide a facilitative input for the GPC. Although the dual-stage view currently lacks a computational model, the actual cognitive mechanisms supporting it are well understood. In the occipitotemporal axis, visual perception activates exceedingly higher-level representations, e.g., from letter features, letters, bigrams, syllables, and words (Vinckier et al., 2007). All of these activations may constitute an input for GPC—even complete word-forms if they do not activate corresponding phonological representations (consider a frequently seen but never heard word of a foreign language). When the word is highly

activated at orthographic level, little GPC activity is needed to resolve the serial order of G-Ps (Ossmy et al., 2014), resulting in a faster decoding process.

Refixations were almost exclusively determined by WL and little by WF, confirming previous findings that refixations are predominantly made for sampling more visual information about the word's end (Hautala & Loberg, 2015; Hautala et al., 2011a), and to a lesser degree to lexical processing difficulties (Bertram & Hyönä, 2003; Tiffin-Richards & Schroeder, 2015; cf. Hawelka et al., 2010). This is particularly apparent when looking at how highly fluent readers read highly frequent words (Fig. 4). There is virtually no WL effect in FFD and SRD measures, suggesting little involvement of GPC and warranting a conclusion that recognition was based on direct association from orthographic to phonological word representations (Fig. 6, dotted line).

However, the remaining question is why fluent readers refixate frequent long words? We suggest that both direct and indirect routes should be understood as predictive coding of linguistic content (Friston et al., 2012; Sibley et al., 2010). Concerning the indirect route, predictive function would explain why it is much easier for people to decode pseudowords with a predictable internal structure than pseudowords that confront language's internal transitional properties (Perry, 2018). During this process, a strong but false prediction (e.g., "judge") may have either a facilitatory ("judga") or distracting influence ("jugde") on pseudoword or low-frequency word reading. Concerning the direct route, the predictive function would explain the refixation behavior observed in the present study: Whenever a visual perception of sufficient quality has not been obtained from word end, a refixation is needed to confirm or elaborate the perception, that is, for other reasons than conducting GPC.

Being grounded on accuracy and response time findings, it is perhaps no surprise that the dual-route view has some difficulties in predicting the proper time course of processes. Yet, the dual-route view is accurate in the sense that processing of low and high frequency words *can* diverge early, manifesting in the present data already as a consequence of parafoveal preview. We suggest that two modifications to the dual-route view may be required to explain the present findings. First, activation of orthographic word representation should not guarantee direct access to its phonological counterpart, but these associations would also need to be fully automatized. This issue is discussed further below. Second, the feedback mechanism from orthographic word representations to letter encoding may need to be rather strong, and such feedback should somehow be able to facilitate the actual GPC also. To our knowledge, this mechanism has not been explicitly addressed in previous research.

Fig. 6 Schematic illustration of the suggested model of word recognition and the deficits in developmental dyslexia (black crosses) based on the present results. Dotted line represents a connection assumed to appear late in reading development



Word recognition deficits in dysfluent reading

Among readers with low reading fluency the time course of the studied effects was more pronounced: The frequency effect emerged first during the FFD, followed then by a length effect in RP and a marked interaction of WF and WL in SRD. This pattern of results indicate that the activation of orthographic representations clearly precedes and facilitates decoding in less fluent reading. In addition, low reading fluency was characterized by clearly longer fixation durations and increased RP, and a tendency to show stronger WF and WL effects, suggesting overall inefficiency in processing. For studying which of the deficiencies are primary and which are redundant, a regression analysis was conducted in which reading fluency was explained by individual random effect coefficients of each measure. The results indicated that reading fluency is mostly explained by the overall level of FFD, yet SRD coefficients also explained a substantial amount of additional variance in reading fluency. Together these findings suggest a deficiency in early visual letter processing, an intact activation of orthographic word representations, and laborious decoding (Fig. 6).

Overall, the present findings seem to be incompatible with the prevalent learning deficit view that dysfluent readers lack orthographic word representations. Instead the findings suggest a general difficulty in orthographic processing and slowness in connecting activated orthographic word representations to their phonological counterparts directly without considerable decoding involved. This dual-deficit view (Fig. 6) may be a viable alternative to previously established views of single deficit in early visuo-orthographic processing (Paizi et al., 2013; Wimmer & Schurz, 2010) and, on the other hand, views of multiple deficits within the dual-route architecture in DD, including difficulty in establishing orthographic word representation, associating them to their phonological correspondents, laborious decoding, and difficulties in lexical evaluation at phonological level (Bergmann & Wimmer, 2008; Perry et al., 2019). Our results are also in line with the general consensus that the establishment of the direct lexical route takes considerable time and practice to develop (Jobard et al., 2011; Ziegler et al., 2014).

The principal deficiency in early visuo-orthographic processing, i.e. in extracting the orthographic template (Paizi et al., 2013) may involve a limited capacity to process letters in parallel (visual attention span; Hawelka & Wimmer, 2005; Lobier et al., 2012). In line with this explanation, Hautala and Parviainen (2014) found that dysfluent adult readers' fixational eye movement responses to word-end orthographic violations were delayed in comparison to typical readers. Additional factors may include excessive visual crowding (Bertoni et al., 2019; Martelli et al., 2009), which seems to be specific to the processing of symbols (Doron et al., 2015), and less efficient parafoveal processing of the word prior to fixation (Silva et al., 2016).

However, the present results also indicate that whenever a sufficient orthographic template is extracted, possibly even for the beginning of a word, it seems capable of activating orthographic word representations. Such an activation seems to be a fairly compulsive result of letter encoding and not granting direct access to phonological word representations. Therefore, the difficulty in establishing the direct route (Jobard et al., 2011; Ziegler et al., 2014) seem not to result from a deficiency in

forming orthographic word representations, but from a deficiency to associate them to their phonological correspondents (Boets et al., 2013).

Recently, subtypes of dyslexia were modeled by adding noise to specific modules of the connectionist dual-route model (Ziegler et al., 2019). It was found that deficits in accessing phonological word representations may be the largest single factor explaining all subtypes of dyslexia. Our results are in line with this explanation. However, Peterson et al. (2013) suggest that a minority of readers may experience specific difficulties reading irregular words due to deficits in semantic processing.

Limitations

The properties of the studied language may also affect the generalization of the results to some extent. It is possible that reading a fully transparent orthography of a morphologically highly agglutinative language—Finnish—emphasizes sublexical decoding more than, for example, the English orthography, in which accessing phonology requires more parallel processing of letters and where WF information is strongly utilized in reading (see Schmalz et al., 2015; Ziegler et al., 2001). Therefore, the feasibility of the dual-stage view needs yet to be studied in deeper orthographies, despite some compatible evidence obtained in eye movement studies conducted in English orthography (Hyönä & Olson, 1995; see also, Joseph et al., 2009).

The present study did not address morpho-syllabic parsing, other than controlling for sublexical difficulty. The lack of WL effect in fluent readers' fixation durations show that fluent readers are remarkably efficient in reading morpho-syllabically complex (i.e. long) words. It is possible that morpho-syllabic complexity explains to some extent the word length effect observed in summed refixation duration for less fluent readers. However, previous studies in Italian (another highly transparent orthography) suggest that morphological complexity actually supports word recognition in DD (Burani et al., 2008), possibly because morphemes may provide richer lexical activation for readers with limited ability for parallel letter processing.

In line with previous multiline text–reading studies (Hutzler & Wimmer, 2004; Hyönä & Olson, 1995), we observed a "normal" lengthening of FFD as a function of WL, rather than the trade-off effects observed in some single sentence-reading studies (Loberg et al., 2019; Sperlich et al., 2015; Vitu et al., 2001). It may be that single-sentence tasks induce a shallower type of first-pass processing by allowing the reader to reread the most relevant part of the sentence immediately. In text reading, such a strategy would distract the encoding of story narrative and the flow of reading.

Finally, it is possible that continuous reading generally advances the activation of orthographic word representations both via contextual predictions and parafoveal preview, which both are absent when reading single words. Therefore, it is possible that the here-introduced dual-stage view is emphasized in continuous reading. It is thus an interesting question whether evidence for WF effect preceding the WL effect can be obtained by analyzing fixation durations of single word recognition.

Implications and conclusions

One possible practical implication of the dual-stage view would be that decoding skill of at least regularly pronounced words might be improved by learning to make better use of early orthographic activations. Such a process may be practiced by showing a word with brief exposure time and asking students first to guess the word before engaging into decoding activity. However, experimental training studies are needed to test this possibility.

Taken together, the present study provides novel evidence for the view that readers' visual word recognition during continuous reading may be best understood as a dual-stage process in which early holistic lexical processing of a word, whenever needed, is followed and complemented by a more fine-grained G-P conversion. In this framework, low RF seems to be associated with impaired pre-lexical processing of letter strings and slower G-P conversion, whereas early lexical processing may be intact. However, due to their central theoretical importance, the present findings need to be replicated in different orthographies and with methods which are sensitive to the time course of word recognition.

Acknowledgements The authors would like to thank the participating students and teachers, research team, research assistants and collaborators for their contributions to the ReadDrama -research project.

Authors' contribution JH, SW and MA contributed on the designing and writing of the study. JH designed, executed and reported the analyses.

Funding Open access funding provided by University of Jyväskylä (JYU). This research was supported by Grants 319911, 352020 and 317030 to JH from Academy of Finland.

Data availability The materials and code used in this research are available from the corresponding author, JH, upon reasonable request. The research data is currently pseudonymized and not shared until it is anonymized.

Declarations

Conflict of interest The authors report no conflict of interests.

Consent to participate Informed written consent was derived from students and their caregivers prior the study.

Consent for publication All authors have approved the publication of this article.

Ethical approval The study was pre-evaluated by the Ethical Committee of the University of Jyväskylä.

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

References

- Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2018). The role of sublexical variables in reading fluency development among Spanish children. *Journal of Child Language*, 45(4), 858–877.
- Araújo, S., & Faísca, L. (2019). A meta-analytic review of naming-speed deficits in developmental dyslexia. Scientific Studies of Reading, 23(5), 349–368.
- Aro, M. (2017). Learning to read finnish. In L. Verhoeven & C. Perfetti (Eds.), *Learning to read across languages and writing systems* (pp. 416–436). Cambridge University Press.
- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133(2), 283.
- Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers, 34*(3), 424–434.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2019). lme4: Linear mixed effects models using Eigen and S4. R package v. 1.1–21.
- Bergmann, J., & Wimmer, H. (2008). A dual-route perspective on poor reading in a regular orthography: Evidence from phonological and orthographic lexical decisions. *Cognitive Neuropsychol*ogy, 25(5), 653–676.
- Bertoni, S., Franceschini, S., Ronconi, L., Gori, S., & Facoetti, A. (2019). Is excessive visual crowding causally linked to developmental dyslexia? *Neuropsychologia*, 130, 107–117.
- Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language*, 48(3), 615–634.
- Blomert, L. (2011). The neural signature of orthographic-phonological binding in successful and failing reading development. *NeuroImage*, 57(3), 695–703.
- Boets, B., de Beeck, H. P. O., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., et al. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science*, 342(6163), 1251–1254.
- Boros, M., Anton, J. L., Pech-Georgel, C., Grainger, J., Szwed, M., & Ziegler, J. C. (2016). Orthographic processing deficits in developmental dyslexia: Beyond the ventral visual stream. *Neuro-Image*, 128, 316–327.
- Burani, C., Marcolini, S., De Luca, M., & Zoccolotti, P. (2008). Morpheme-based reading aloud: Evidence from dyslexic and skilled Italian readers. *Cognition*, 108(1), 243–262.
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software, 80(1), 1–28.
- Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical, and contextual factors. *The Spanish Journal of Psychology*, 5(1), 66–77.
- Carioti, D., Masia, M. F., Travellini, S., & Berlingeri, M. (2021). Orthographic depth and developmental dyslexia: A meta-analytic study. Annals of Dyslexia. https://doi.org/10.1007/s11881-021-00226-0
- Carter, B. T., & Luke, S. G. (2018). Individuals' eye movements in reading are highly consistent across time and trial. *Journal of Experimental Psychology: Human Perception and Performance*, 44(3), 482.
- Chetail, F. (2014). Effect of number of syllables in visual word recognition: New insights from the lexical decision task. *Language, Cognition and Neuroscience, 29*(10), 1249–1256.
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, 123(2), 291–307.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204.
- Conway, A., Brady, N., & Misra, K. (2017). Holistic word processing in dyslexia. *PLoS ONE*, 12(11), e0187326.
- Cornelissen, P. L., Kringelbach, M. L., Ellis, A. W., Whitney, C., Holliday, I. E., & Hansen, P. C. (2009). Activation of the left inferior frontal gyrus in the first 200 ms of reading: Evidence from magnetoencephalography (MEG). *PLoS ONE*, 4(4), e5359.

- De Jongh, P. J., De Jongh, E., Pienaar, M. P., Gordon-Grant, H., Oberholzer, M., & Santana, L. (2015). The impact of pre-selected variance in ation factor thresholds on the stability and predictive power of logistic regression models in credit scoring. ORiON, 31(1), 17–37.
- De Luca, M., Barca, L., Burani, C., & Zoccolotti, P. (2008). The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Cognitive and Behavioral Neurology*, 21(4), 227–235.
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, 9(7), 335–341.
- Denckla, M. B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters and numbers by normal children. *Cortex*, 10(2), 186–202.
- Doron, A., Manassi, M., Herzog, M. H., & Ahissar, M. (2015). Intact crowding and temporal masking in dyslexia. *Journal of Vision*, 15(14), 13–13.
- Dürrwächter, U., Sokolov, A. N., Reinhard, J., Klosinski, G., & Trauzettel-Klosinski, S. (2010). Word length and word frequency affect eye movements in dyslexic children reading in a regular (German) orthography. *Annals of Dyslexia*, 60(1), 86–101.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160.
- Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., et al. (2014). Disruption of functional networks in dyslexia: A whole-brain, data-driven analysis of connectivity. *Biological Psychiatry*, 76(5), 397–404.
- Friston, K., Adams, R., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology*, 3, 151.
- Frost, R. (2012). A universal approach to modeling visual word recognition and reading: Not only possible, but also inevitable. *The Behavioral and Brain Sciences*, 35(5), 310.
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7(4), 493–498.
- Hasenäcker, J., & Schroeder, S. (2017). Syllables and morphemes in German reading development: Evidence from second graders, fourth graders, and adults. *Applied Psycholinguistics*, 38(3), 733–753.
- Hautala, J., Aro, M., Eklund, K., Lerkkanen, M. K., & Lyytinen, H. (2013). The role of letters and syllables in typical and dysfluent reading in a transparent orthography. *Reading and Writing*, 26(6), 845–864.
- Hautala, J., Hyönä, J., & Aro, M. (2011a). Dissociating spatial and letter-based word length effects observed in readers' eye movement patterns. *Vision Research*, 51(15), 1719–1727.
- Hautala, J., Hyönä, J., Aro, M., & Lyytinen, H. (2011b). Sublexical effects on eye movements during repeated reading of words and pseudowords in Finnish. *Psychology of Language and Communication*, 15(2), 129–149.
- Hautala, J., & Loberg, O. (2015). Breaking down the word length effect on readers' eye movements. Language, Cognition and Neuroscience, 30(8), 993–1007.
- Hautala, J., & Parviainen, T. (2014). Gaze position reveals impaired attentional shift during visual word recognition in dysfluent readers. *PLoS ONE*, 9(9), e108937.
- Hawelka, S., Gagl, B., & Wimmer, H. (2010). A dual-route perspective on eye movements of dyslexic readers. *Cognition*, 115(3), 367–379.
- Hawelka, S., Schuster, S., Gagl, B., & Hutzler, F. (2015). On forward inferences of fast and slow readers. An eye movement study. *Scientific Reports*, 5(1), 1–8.
- Hawelka, S., & Wimmer, H. (2005). Impaired visual processing of multi-element arrays is associated with increased number of eye movements in dyslexic reading. *Vision Research*, 45(7), 855–863.
- Häyrinen, T., Serenius-Sirve, S., & Korkman, M. (2013). Lukilasse 2. Lukemisen, kirjoittamisen ja laskemisen seulontatesti 1.-6. vuosiluokille [Screening test for reading, spelling and counting for the grades 1-6]. Helsinki: Hogrefe Psykologien Kustannus Oy.
- Himmelstoss, N. A., Schuster, S., Hutzler, F., Moran, R., & Hawelka, S. (2020). Co-registration of eye movements and neuroimaging for studying contextual predictions in natural reading. *Language*, *Cognition and Neuroscience*, 35(5), 595–612.
- Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, 49(24), 2948–2959.
- Huovilainen, T. M. (2018). Psycholinguistic descriptives. The Language Bank of Finland. Retrieved from: http://urn.fi/urn:nbn:fi:lb-2018081601.
- Hutzler, F., & Wimmer, H. (2004). Eye movements of dyslexic children when reading in a regular orthography. Brain and Language, 89(1), 235–242.

- Hyönä, J., & Olson, R. K. (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*(6), 1430.
- Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *NeuroImage*, 20(2), 693–712.
- Jobard, G., Vigneau, M., Simon, G., & Tzourio-Mazoyer, N. (2011). The weight of skill: Interindividual variability of reading related brain activation patterns in fluent readers. *Journal of Neurolinguistics*, 24(1), 113–132.
- Joseph, H. S., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, 49(16), 2078–2086.
- Joseph, H. S., Nation, K., & Liversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. School Psychology Review, 42(2), 207–222.
- Kapnoula, E. C., Protopapas, A., Saunders, S. J., & Coltheart, M. (2017). Lexical and sublexical effects on visual word recognition in Greek: Comparing human behavior to the dual route cascaded model. *Language, Cognition and Neuroscience, 32*(10), 1290–1304.
- Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 45–66). Amsterdam: Elsevier.
- Kiuru, N., Aunola, K., Lerkkanen, M. K., Pakarinen, E., Poskiparta, E., Ahonen, T., et al. (2015). Positive teacher and peer relations combine to predict primary school students' academic skill development. *Developmental Psychology*, 51(4), 434.
- Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *The European Journal of Cognitive Psychology*, 16, 262–284.
- Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. Language and Cognitive Processes, 23(7–8), 1089–1132.
- Loberg, O., Hautala, J., Hämäläinen, J. A., & Leppänen, P. H. (2019). Influence of reading skill and word length on fixation-related brain activity in school-aged children during natural reading. *Vision Research*, 165, 109–122.
- Lobier, M., Zoubrinetzky, R., & Valdois, S. (2012). The visual attention span deficit in dyslexia is visual and not verbal. *Cortex*, 48(6), 768–773.
- Marinus, E., Nation, K., & de Jong, P. F. (2015). Density and length in the neighborhood: Explaining cross-linguistic differences in learning to read in English and Dutch. *Journal of Experimental Child Psychology*, 139, 127–147.
- Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading, and developmental dyslexia. *Journal of Vision*, 9(4), 14–14.
- Mehlhase, H., Bakos, S., Landerl, K., Schulte-Körne, G., & Moll, K. (2019). Orthographic learning in children with isolated and combined reading and spelling deficits. *Child Neuropsychology*, 25(3), 370–393.
- Moll, K., & Jones, M. (2013). Naming fluency in dyslexic and nondyslexic readers: Differential effects of visual crowding in foveal, parafoveal, and peripheral vision. *The Quarterly Journal of Experimen*tal Psychology, 66(11), 2085–2091.
- Ossmy, O., Ben-Shachar, M., & Mukamel, R. (2014). Decoding letter position in word reading. *Cortex*, 59, 74–83.
- 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(3), 303–329.
- Perry, C. (2018). Reading orthographically strange nonwords: Modelling backup strategies in reading. Scientific Studies of Reading, 22(3), 264–272.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cognitive Psychology*, 61(2), 106–151.
- Perry, C., Ziegler, J. C., & Zorzi, M. (2014). CDP++ Italian: Modelling sublexical and supralexical inconsistency in a shallow orthography. *PLoS ONE*, 9(4), 94291.
- Perry, C., Zorzi, M., & Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. *Psychological Science*, 30(3), 386–395.
- Peterson, R. L., Pennington, B. F., & Olson, R. K. (2013). Subtypes of developmental dyslexia: Testing the predictions of the dual-route and connectionist frameworks. *Cognition*, 126(1), 20–38.
- Pollack, C., Luk, G., & Christodoulou, J. A. (2015). A meta-analysis of functional reading systems in typically developing and struggling readers across different alphabetic languages. *Frontiers in Psychology*, 6, 191.

- Prado, C., Dubois, M., & Valdois, S. (2007). The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span. *Vision Research*, 47(19), 2521–2530.
- Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, 15(6), 246–253.
- Pritchard, S. C., Coltheart, M., Palethorpe, S., & Castles, A. (2012). Nonword reading: Comparing dual-route cascaded and connectionist dual-process models with human data. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5), 1268.
- Pugh, K. R., Landi, N., Preston, J. L., Mencl, W. E., Austin, A. C., Sibley, D., et al. (2013). The relationship between phonological and auditory processing and brain organization in beginning readers. *Brain and Language*, 125(2), 173–183.
- 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(3), 224–233.
- Rau, A. K., Moll, K., Snowling, M. J., & Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *Journal of Experimental Child Psychology*, 130, 92–105.
- Reichle, E. D. (2015). Computational models of reading: A primer. Language and Linguistics Compass, 9(7), 271–284.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, 26(4), 445.
- Richlan, F. (2014). Functional neuroanatomy of developmental dyslexia: The role of orthographic depth. *Frontiers in Human Neuroscience*, *8*, 347.
- Richlan, F. (2019). The functional neuroanatomy of letter-speech sound integration and its relation to brain abnormalities in developmental dyslexia. *Frontiers in Human Neuroscience*, 13, 21.
- Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W. J., Bick, A., Frost, S. J., et al. (2015). Universal brain signature of proficient reading: Evidence from four contrasting languages. Proceedings of the National Academy of Sciences, 112(50), 15510–15515.
- Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J. F., Bricout, L., et al. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental Psychology*, 52(10), 1503.
- Schmalz, X., Marinus, E., Coltheart, M., & Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomic Bulletin & Review*, 22(6), 1614–1629.
- Schmalz, X., Robidoux, S., Castles, A., Coltheart, M., & Marinus, E. (2017). German and English bodies: No evidence for cross-linguistic differences in preferred orthographic grain size. *Collabra: Psychology*, 3(1), 5.
- Schurz, M., Sturm, D., Richlan, F., Kronbichler, M., Ladurner, G., & Wimmer, H. (2010). A dualroute perspective on brain activation in response to visual words: Evidence for a length by lexicality interaction in the visual word form area (VWFA). *NeuroImage*, 49(3), 2649–2661.
- Schurz, M., Wimmer, H., Richlan, F., Ludersdorfer, P., Klackl, J., & Kronbichler, M. (2015). Restingstate and task-based functional brain connectivity in developmental dyslexia. *Cerebral Cortex*, 25(10), 3502–3514.
- Seidenberg, M. S. (2007). Connectionist models of reading. In M. G. Gaskell (Ed.), *The Oxford hand-book of psycholinguistics* (pp. 235–250). New York, NY: Oxford University Press.
- Seymour, P. H., Aro, M., Erskine, J. M., & Collaboration with COST Action A8 Network. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94(2), 143–174.
- Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an" outlier" orthography. *Psychological Bulletin*, 134(4), 584.
- Sibley, D. E., Kello, C. T., & Seidenberg, M. S. (2010). Learning orthographic and phonological representations in models of monosyllabic and bisyllabic naming. *European Journal of Cognitive Psychology*, 22(5), 650–668.
- Silva, S., Faísca, L., Araújo, S., Casaca, L., Carvalho, L., Petersson, K. M., & Reis, A. (2016). Too little or too much? Parafoveal preview benefits and parafoveal load costs in dyslexic adults. *Annals of Dyslexia*, 66(2), 187–201.
- Sperlich, A., Schad, D. J., & Laubrock, J. (2015). When preview information starts to matter: Development of the perceptual span in German beginning readers. *Journal of Cognitive Psychology*, 27(5), 511–530.

- Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin*, 139(4), 766.
- The Finnish N-gram Corpus, version 1 (FNC1). (2014). Distributed by the University of Helsinki on behalf of the FIN-CLARIN consortium. Retrieved from: http://urn.fi/urn.ibn:fi:lb-2014073038.
- Thesen, T., McDonald, C. R., Carlson, C., Doyle, W., Cash, S., Sherfey, J., et al. (2012). Sequential then interactive processing of letters and words in the left fusiform gyrus. *Nature Communications*, 3(1), 1–8.
- Tiffin-Richards, S. P., & Schroeder, S. (2015). Word length and frequency effects on children's eye movements during silent reading. *Vision Research*, 113, 33–43.
- van Houwelingen, H. C., & Sauerbrei, W. (2013). Cross-validation, shrinkage and variable selection in linear regression revisited. Open Journal of Statistics, 3(2), 79–102.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron*, 55(1), 143–156.
- Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research*, 41(25–26), 3513–3533.
- Wheat, K. L., Cornelissen, P. L., Frost, S. J., & Hansen, P. C. (2010). During visual word recognition, phonology is accessed within 100 ms and may be mediated by a speech production code: Evidence from magnetoencephalography. *Journal of Neuroscience*, 30(15), 5229–5233.
- Wimmer, H., & Schurz, M. (2010). Dyslexia in regular orthographies: Manifestation and causation. Dyslexia, 16(4), 283–299.
- Woodhead, Z. V. J., Barnes, G. R., Penny, W., Moran, R., Teki, S., Price, C. J., & Leff, A. P. (2014). Reading front to back: MEG evidence for early feedback effects during word recognition. *Cerebral Cortex*, 24(3), 817–825.
- Wydell, T. N., Vuorinen, T., Helenius, P., & Salmelin, R. (2003). Neural correlates of letter-string length and lexicality during reading in a regular orthography. *Journal of Cognitive Neuroscience*, 15(7), 1052–1062.
- Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance*, 38(1), 53.
- Ziegler, J. C., Bertrand, D., Lété, B., & Grainger, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmen*tal Psychology, 50(4), 1026.
- Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science*, 12(5), 379–384.
- 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(3), 169–193.
- Ziegler, J. C., Perry, C., & Zorzi, M. (2019). Modeling the variability of developmental dyslexia. In L. Verhoeven, C. Perfetti, & K. Pugh (Eds.), *Developmental dyslexia across languages and writing* systems (pp. 350–371). Cambridge University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.