Scientific Studies of Reading

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hssr20

Rapid Coding of Syllable Structure by Dysfluent Developing Readers

Lisa Hintermeier, Jarkko Hautala & Mikko Aro

To cite this article: Lisa Hintermeier, Jarkko Hautala & Mikko Aro (2023): Rapid Coding of Syllable Structure by Dysfluent Developing Readers, Scientific Studies of Reading, DOI: 10.1080/10888438.2023.2203863

To link to this article: https://doi.org/10.1080/10888438.2023.2203863

Published online: 25 Apr 2023.

SCIENTIFIC STUDIES OF READING
https://doi.org/10.1080/10888438.2023.2203863

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Rapid Coding of Syllable Structure by Dysfluent Developing Readers

Lisa Hintermeier [a], Jarkko Hautala [b], and Mikko Aro [a]

[a]Department of Education, University of Jyväskylä, Jyväskylä, Finland; [b] Niilo Mäki Institute, Jyväskylä, Finland

ABSTRACT

Purpose: The present study investigated whether the number of syllables affects developing readers' word recognition when controlling for word length and word frequency and, if so, whether the effect is dependent on reading fluency. The target language was Finnish, a language with a transparent orthography and a simple syllable structure.

Method: Eye movements of 142 third and fourth graders were recorded during silent reading of two stories. Reading fluency was assessed separately. For analyses, a data subset containing words of a certain length (6,7,9 letters) and varying syllable number (2,3,4 syllables) was extracted from the data set. Using linear mixed-effects modeling, the effect of the syllable number on various eye-tracking measures across different levels of reading fluency was studied.

Results: Results revealed a statistically significant, impeding number of syllables effect in first fixation duration but non-significant effects in the later reading measures. Furthermore, fluent and dysfluent readers did not differ regarding the number of syllables effect.

Conclusion: These findings suggest that in Finnish developing readers, syllabic parsing is a highly rapid and automatized process, which predominantly takes place during the early holistic orthographic processing of a word, and that qualitatively similar orthographic processing occurs in fluent and dysfluent beginning readers.

Introduction

Learning to read is one of the most important milestones that young children need to achieve during their first years in primary school. However, what exactly happens during the progression from novice to skilled reader with regard to developing solid reading fluency and accuracy is highly dependent on the characteristics of the respective orthography and is currently not fully understood (see Share, 2008).

In alphabetic writing systems, the first step toward successful reading acquisition is to learn and apply grapheme-phoneme correspondences. In this way, children become able to sound out words letter-by-letter and finally read their first written words (see, for example, Ehri, 1995). Nevertheless, this serial decoding is a relatively slow process and requires conscious cognitive effort; therefore, word recognition needs to be accelerated in order to achieve reading fluency. When beginning readers reach the so-called *"consolidated alphabetic phase"* (Ehri, 1995, p. 117), they have already encountered a set of new words through their growing reading experience and are therefore also able to establish orthographic representations based on frequently recurring letter combinations, such as syllables or morphemes. These sublexical units could help beginning readers facilitate the decoding process by

CONTACT Lisa Hintermeier lisa.m.hintermeier@jyu.fi Department of Education, University of Jyväskylä, Jyväskylä 40014, Finland

decreasing the need for serial decoding of single graphemes, which can ultimately speed up word recognition (Hautala et al., 2021).

However, as stated in Ziegler and Goswami's *Psycholinguistic Grain Size Theory* (Ziegler & Goswami, 2005), specific language characteristics, such as orthographic transparency or the complexity of a language's syllabic structure, might determine the role and use of these sublexical units in reading. Orthographically opaque languages like English are characterized by inconsistent mappings between graphemes and phonemes (Seymour et al., 2003). The combination of inconsistent letter (cluster) pronunciations and a relatively simple morphological system might favor reading via a lexical rather than a sublexical strategy, especially in developing readers (see Ziegler & Goswami, 2005). On the other hand, Finnish, the language of the present study, is characterized by a complex morphological system. Due to its agglutinative inflectional morphology, as well as productive derivation and compounding, Finnish words tend to be long, with the average length being between seven and eight letters, and any noun can have more than 2000 different forms (Aro, 2017). In addition, since grapheme-phoneme correspondences are almost perfectly regular in Finnish, and the syllable structure is simple, developing readers might favor processing smaller units during reading.

On the basis of the aforementioned linguistic features, there are at least two sublexical units larger than the grapheme-phoneme level that might serve as helpful processing units when beginning readers proceed toward reading fluency in Finnish: morphemes and syllables. Morphemes are the smallest units that carry meaning in a language, and especially in morphologically rich languages like Finnish, where words can have numerous agglutinated affixes, morphemes could serve as suitable processing units. Some previous studies have found that developing readers utilize morphemic information in both naming (in *French*, see Colé et al., 2012) and lexical decision tasks (in *German*, see Hasenäcker & Schroeder, 2017). Moreover, results from an eye-tracking study conducted with Finnish preschoolers indicated that typically developing first grade readers are already sensitive to morphemes (Häikiö & Vainio, 2018).

Apart from morphemes, the syllable may also serve as a reasonable processing unit in Finnish, considering that words are almost without exception multisyllabic (Aro, 2017). Furthermore, the syllable is already familiar to developing readers in Finnish, as multisyllabic words are hyphenated at the syllable boundary in early reading materials. Previous research from Häikiö et al. (2015) for example has shown that Finnish readers, as early as in the first grade, were distracted by syllable-incongruent hyphenations in polysyllabic words which indicates that already very early in their reading development, children progress from a mainly serial letter-by-letter decoding to the use of larger units, at least in Finnish. Interestingly, however, more advanced readers seemed to be slowed down by hyphenated words regardless of the hyphen position when compared to unhyphenated control items, which suggests a capability of whole-word processing or at least parallel processing of syllabic units (Häikiö et al., 2015).

Other earlier results regarding the role of the syllable have mainly been drawn from tasks focusing either on lexical access by using lexical decision tasks (e.g., Alvarez et al., 2001; Carreiras et al., 1993) or on speech production by using naming tasks (e.g., Carreiras & Perea, 2004; Perea & Carreiras, 1998). In studies where the frequency of the first syllable was manipulated, words with a highly frequent first syllable were recognized more slowly than words with less frequent syllables in lexical decision tasks (Alvarez et al., 2001; Carreiras et al., 1993; Conrad et al., 2006; Luque et al., 2013; Perea & Carreiras, 1998), but produced faster in naming tasks (Carreiras & Perea, 2004; Perea & Carreiras, 1998). These contradictory findings can be explained by differences in task demands, as high lexical competition slows down lexical decision processes, while language production seems to be facilitated by a highly frequent first syllable (Luque et al., 2013).

Another line of research has focused on the number of syllables effect on word recognition – that is, on the question of whether the syllable is being used as a processing unit in word recognition. More specifically, the number of syllables effect has been investigated in studies that implemented tasks on speech production (Eriksen et al., 1970; Ferrand, 2000; Klapp et al., 1973; Mason, 1978) or lexical access (Ferrand & New, 2003, in *French*; Stenneken et al., 2007, in *German*). The earliest results regarding the number of syllables effect mainly arise from studies investigating the English language, where a number of syllables effect has been both found (Klapp et al., 1973; Mason, 1978) and not found (Forster & Chambers, 1973; Frederiksen & Kroll, 1976) during naming. However, as highlighted by Jared and Seidenberg (1990) and more recently by Chetail (2014), prolonged pronunciation times with increasing number of syllables in English might also be subject to grapheme-phoneme irregularities, which mostly concern vowels: increasing the number of syllables necessarily increases the number of vowels, which in turn can account for longer response times in the naming task. But also in orthographies other than English, like German (Stenneken et al., 2007) or French (Ferrand & New, 2003; Ferrand, 2000), it has been shown that the number of syllables affects word recognition significantly. For instance, Chetail (2014) found the number of syllables to influence word recognition speed in French in a lexical decision task, where they found trisyllabic words to be processed more slowly than bisyllabic words. This effect was present for both high- and low-frequency words. Some previous studies similarly point to the role of the syllable in the context of Finnish word recognition processes: Hautala et al. (2013) conducted both lexical decision and naming tasks to examine word length and syllable effects in fluent and dysfluent second graders. Their results implicated that the number of syllables affected only the response times in the lexical decision task of the dysfluent readers, while such an effect was not found in the naming task, and the number of syllables did not affect the response times of fluent readers in either task.

Despite offering interesting insights into the underlying mechanisms of visual word recognition, previous results regarding syllabic processing, and more specifically the number of syllables effect, exclusively stem from methods that might not necessarily reflect the natural reading process, as these tasks involve additional processing steps like speech production (in naming tasks) or decision making (in lexical decision tasks), which are not present in natural reading situations. Likewise, words are usually presented in isolation during such tasks, which poorly reflects natural reading situations. A method that allows for studying reading in more natural reading situations is eye-tracking. Another advantage of eye-tracking is that we can gain insight into the time course of stages of visual word recognition by measuring various indices of eye movements tapping into earlier stages, such as first-pass eye movements (e.g., *first fixation duration*), and later stages, such as *refixation probability* or *summed refixation duration* (Hautala et al., 2021). This is particularly interesting, as models of eye movement control during reading, such as the prevalent *E-Z Reader model* (Reichle et al., 2003), assume early and later stages in word identification. More specifically, the *E-Z Reader model* presumes two stages: the so-called *familiarity check*, during which the orthographic form of the word is being identified; and the second stage, where the phonological and/or semantic word form is being identified and can thus be referred to as the actual lexical access stage. More importantly, the familiarity check informs whether a word will be identified during a single fixation or whether a refixation is needed.

There are also well-established computational models of visual word recognition that specifically take sublexical units into account, although they have mostly been developed for the English language. The *Connectionist Multiple-Trace Memory Model* (*MTM model;* Ans et al., 1998) proposes that pseudowords/nonwords are sequentially decomposed into phonological segments (i.e., syllabic units) since they are read via the analytic route, which in turn leads to increased naming latencies when the number of syllables increases. Finally, the *Connectionist Dual-Process Model* (*CDP++ model*; Perry et al., 2010) of reading English words aloud suggests a serial graphemic parsing process across an attentional window of three neighboring letters. If the graphemic parser encounters two vowel graphemes, the word is identified as bisyllabic. In addition, the number of syllables effect was reported to be larger for low-frequency words than for high-frequency words in the *CDP++ model* simulations.

Crucially, however, the aforementioned models are exclusively based on skilled reading and can thus not be applied to the reading processes of developing readers, not only because of potential differences in reading strategies, but also due to the variation within their reading levels. In order for reading researchers to model developing but also compromised reading computationally, the first step is to understand the underlying processes.

In summary, previous research started to study the role of the syllable in reading more intensively in the past years, although beginning and dysfluent reading remains an area where more insights are needed, especially in terms of underlying processes in visual word recognition. To achieve this, more sensitive methods are required to study natural reading situations, as the majority of earlier results are based on methods such as lexical decision and naming tasks (e.g., Ferrand & New, 2003, in *French*; Stenneken et al., 2007, in *German*). To fill these gaps, the present study will try to expand on the knowledge of word identification in children on their way to reading fluency by investigating how the number of syllables affects both the early and later stages of visual word recognition in Finnish third and fourth graders using eye movement recordings in continuous reading. More specifically, we want to investigate whether the number of syllables has an effect on developing readers' word recognition and, if so, whether the effect is dependent on reading fluency. Based on previous research on syllabic processing in Finnish, as well as other language contexts, we predict to only find small effects of the number of syllables in general. However, we expect the effect to differ between different levels of reading fluency, i.e., to find a stronger effect in the group of dysfluent readers who presumably read most words by decoding. Furthermore, previous research on the number of syllables effect found different results for high- vs. low-frequency words, and computational models of reading predict a number of syllables effect for low-frequency words only (Ans et al., 1998; Perry et al., 2010). Therefore, we studied whether the number of syllables effect is dependent on word frequency.

An exploratory aspect of the present analysis is to investigate how different eye-tracking measures are affected by the number of syllables while reading. To our knowledge, previous studies that investigated the number of syllables effect have not used eye-tracking. Thus, the use of eye-tracking technology in the present study will provide novel insights into the time-course of visual word recognition processes during reading.

Methods

Participants

The present study is a reanalysis of a data set reported by Hautala et al. (2021), and the data was collected as part of the screening and pretest assessments within an intervention study designed to support struggling readers (Hautala et al., 2022). The study was pre-evaluated by the Ethical Committee of the University of Jyväskylä, and the research was conducted according to the ethical principles for medical research involving human subjects set forth by the Declaration of Helsinki. Prior to the study, informed written consent was obtained from both the children and their caregivers.

In order to identify possible participants for the reading interventions, classroom teachers were instructed to distribute invitation letters to third/fourth graders deemed to require targeted support in reading fluency (25 per school) and to 30 typical readers in each school. Then, reading fluency of students with a consent for participating was assessed in a separate session. All students involved in this study followed the standard curriculum, with school instruction provided in Finnish. For three students, caregivers reported that Finnish was not the child's native language; however, as they demonstrated solid Finnish language proficiency (both in written and spoken language), they were included in the study.

In half of the schools participating in the reading intervention study, additional eye-tracking data collection sessions were conducted, which served as the basis for the present analysis. Subsequently, 152 third and fourth graders from Central Finland were included in the present study (mean age = 10 years and 1 month, *SD* = 7 months). Due to technical difficulties during the eye-tracking measurements and experiment discontinuation, 10 students were excluded, resulting in a final sample of 142 students (79 girls and 63 boys) from Grade 3 ($N = 54$) and Grade 4 ($N = 88$) for the present analysis.

Reading fluency assessment

In a separate session prior to the eye-tracking experiment, students' reading fluency was assessed in two steps: first, a standardized reading test was used (*Lukilasse 2*; Häyrinen et al., 2013), where the students were asked to read aloud a list of words with increasing length as rapidly and accurately as they could. The time limit was two minutes, and the number of correctly read words within the time limit was used as the raw score. The raw scores were transformed into standard scores separately for both grade levels on the basis of grade-specific normative data published in the test manual.

Second, students were given a reading task in which they were asked to read aloud a short text (*"Exciting travels"*, 124 words), again as rapidly and accurately as possible. The number of words read accurately within a one-minute time limit was used as the raw score. According to grade-specific large-scale research data (data from *First Steps study*, Lerkkanen et al., 2006), the raw scores were subsequently transformed into standard scores.

To obtain an index of reading fluency for the present analysis, the average of the standardized values across both tasks was calculated (Cronbach's alpha reliability =.917). On average, the standardized reading fluency of the participants in this study was relatively low ($M = -0.66$, $SD = 1.06$, range: −3.13–1.78), but normally distributed (skewness = −0.005, $SE = 0.20$). 28,87% of the participants scored below the 10th percentile relative to the normative samples used for the fluency assessments; poor readers were thus overrepresented in the present sample. Sample-specific *z*-scores were used in the analyses to exclude any grade effects.

Procedure

Subsequently, the students participated in the actual eye-tracking recording session. Data collection was conducted in dimly lit classrooms in the schools of the participants by two researchers using SMI remote eye-tracking devices. The students were asked to sit down on a non-adjustable chair and to put their chin on an adjustable custom chinrest to minimize head movements. The participants' eye movements during the two text-reading tasks were recorded at a 250 Hz sampling rate. Prior to both tasks, the eye tracker was calibrated using a 13-point full-screen calibration procedure. After half of each story was read, a four-point recalibration-validation procedure was repeated. To familiarize themselves with the procedure, the participants completed a short practice session. Another calibration followed the practice trials before the actual experiment started.

The materials were presented on a 19.5 × 34.5 cm screen, which was placed at a 60 cm distance to the participant, on eleven five-line screens using the *SMI Experiment Center 3.6* program. To proceed with the next screen, participants were required to look at a gaze-sensitive area on the right-bottom corner of the screen. Between the two texts, a break was included to offer the participants the possibility to recover, followed by a recalibration procedure. Both stories were followed by five four-choice comprehension questions (in sum, 10 questions), which were answered with an accuracy of $M = 79\%$, $SD = 16\%$, range: 30%–100%. Task instructions were delivered simultaneously, both as written text on screen and auditorily via headphones.

Reading materials

During the eye movements recording, participants were asked to read two slightly adapted versions of classic short stories, "*Little Heidi*" (Johanna Spyri, 1881/2015, 457 words) and "*Adalmina's Pearl*" (Zacharias Topelius, 1865/2017, 403 words)[1] during the experiment. Additionally, they were asked whether the stories were familiar to them. Thirty-four children were familiar with one of the stories; eight knew both. The stories were each followed by five

Table 1. Frequency and examples of the items included in the present analysis (syllable boundaries are marked with a hyphen).

6 Letters		7 Letters		9 Letters	
2 Syllables	3 Syllables	2 Syllables	3 Syllables	3 Syllables	4 Syllables
N=71	N=24	N=20	N=55	N=26	N=19
het-ken	ha-lu-an	jol-loin	e-nem-pää	kii-tä-mään	he-lo-kis-ta
kas-voi	hei-ni-ä	saa-tiin	i-loi-nen	ko-dis-saan	ko-ho-si-vat
mu-kaan	mai-to-a	tai-vaan	ko-li-naa	vuo-hi-neen	va-el-si-vat

multiple-choice comprehension questions. In order to measure the number of syllables effect in words of the same length (i.e., number of letters), a data subset containing only words with six, seven, or nine letters and two, three, or four syllables[2] of different types (e.g., CVVC-CV-V or CVC-CV-CVVC) was extracted from the whole data set, resulting in a final data set of 215 different words per participant (note that some words occurred more than once in the texts). Comparability between the selected items in terms of frequency and two-gram predictability was measured and ensured using an analysis of variance (ANOVA); the results of the post hoc Tukey test indicated that words of the same length, but varying number of syllables did not differ significantly from each other with regards to frequency and two-gram predictability. An overview about the frequency of the items with varying combinations with regard to the number of letters and syllables, as well as some example stimuli for each letter-syllable category can be found in Table 1.

Eye-tracking measures

To gain deeper insight into the time course of the reading processes, both early and later eye movement variables were analyzed for the selected target words. Temporally strictly sequential process measures were *first fixation duration*, *first pass refixation probability*, and *summed refixation duration*. Summative measures were *gaze duration* and *total fixation duration*.

Data processing

The data processing procedures were identical to those of a previous study of this data set (Hautala et al., 2021). After data collection, the eye-tracking data was preprocessed using the *SMI Begaze 3.6* program. Blinks were excluded from the data, a saccade duration of 15 ms, a minimum fixation duration of 50 ms, and sensitive saccade detection parameters of 20 deg/s minimum angular velocity were applied in order to detect refixation saccades with small amplitudes. Word-specific interest areas were automatically generated, but their vertical boundaries were manually extended to the middle position between the lines. Subsequently, data quality was assessed by trained research assistants, who manually inspected the scanpaths of the screen recordings. The inspection showed that 22% (i.e., 380 out of 1694) of the screens were affected by either partially ($n = 65$) or fully ($n = 105$) missing data. In addition, systematic drift corrections were applied ($n = 210$). For all 142 recordings, the decisions of whether or not to apply corrections to the screen were 94% consistent among the independent raters. A custom *SPSS 26* script was used to identify first- and second-pass fixations. In the final analysis, only lines where at least 60% of the words were fixated were included to remove occurrences where participants only skimmed the text, which led to an exclusion of 1573 words. In addition, we excluded return-sweep fixations whenever they did not land on the first word of the next line. The pre-processing procedure led to a final data set including 114,485 word observations ($M = 753$ per participant) out of 122,120 possible observations, and the area-of-interest aggregated data was finally exported for statistical analysis.

Statistical analysis

To normalize the data, log-transformation was applied to all measures. Outliers in the eye-tracking variables of each item category (i.e., letter-syllable combinations) were removed from the data by first standardizing it and removing values that were more than three standard deviations from the mean. In addition, we included only instances in which the word was fixated in the analysis. In the case where a word occurred multiple times in the text, only the first occurrence was considered.

The statistical analysis was conducted in two steps: In the main analysis which focused on the overall number of syllables effect, linear mixed-effects modeling was conducted in *R* (R Core Team, 2021) using the *afex* package (Singmann et al., 2021) to measure the number of syllables effect for all eye movement variables, which served as the dependent variables. We started with a maximum model structure: a three-way interaction between the level of *number of syllables* (i.e., 2/3/4 syllables), *number of letters* (i.e., 6/7/9 letters) and *reading fluency*, and a two-way interaction between *word frequency* and *reading fluency* as a fixed effects structure. The random effects structure existed of random intercepts and slopes of *number of syllables*, *number of letters* and *reading fluency* for participants and random intercepts for *items*, and whenever the maximum model failed to converge, we gradually simplified the model structure by removing non-significant effects or correlations between random effects from the model structure, non-significant main effects were retained. Finally, the goodness-of-fit between the models was estimated using a likelihood ratio test.

Descriptive analyses of the data indicated variation in the present sample with regard to reading strategies across the reading fluency continuum. More specifically, dysfluent readers seemed to a) make more fixations and b) be slower readers, as indicated by longer average gaze durations and total fixation durations (see Appendix A-C), compared to the more fluent readers in the present sample. In a secondary analysis, we were thus also interested in possible differences between fluent and dysfluent readers regarding the syllable as a processing unit. More specifically, we wanted to see whether the children at the higher end of the reading fluency continuum differed from the children with the lowest fluency scores regarding the number of syllables effect. Therefore, we calculated the percentile ranks for the reading fluency scores within the sample and created two groups: a low reading fluency group with a fluency rank below the 10th percentile and a high reading fluency group with a fluency rank above the 90th percentile. After calculating these, 17 children were thus assigned to the high reading fluency group and 16 children to the low reading fluency group.

We again applied linear mixed-effects modeling using the same eye-tracking measures as in the main analysis (*first fixation duration*, *summed refixation duration*, *gaze duration*, *first pass refixation probability*, *total fixation duration*). However, this time, a categorical reading fluency variable was included as an additional main effect to compare the number of syllables effect across the low- and high-fluency groups. As in the main analysis, we started with the maximum model structure and gradually removed non-significant effects from the formula, and non-significant main effects were retained. For this analysis, the maximum model structure consisted of a three-way interaction between the level of *number of syllables*, *word length*, and *reading fluency category* (i.e., low vs. high), and a two-way interaction between *word frequency* and *reading fluency category* as fixed effects. For the random effect structure, we applied random intercepts and slopes of the *number of syllables*, *number of letters* and *reading fluency* for participants, and random intercepts for *items*.

The statistically significant results of the individual models will be described in the following sections, and the whole model outputs can be found in the Appendix (see Appendix D-M). It should be noted that as its default, the *mixed* function in *afex* uses sum-to-zero contrasts for interpreting categorical variable coefficients in linear mixed models; thus, the intercept refers to the grand mean (Singmann et al., 2021) and the baseline category for the contrasts referred to the middle category in the respective categorical variables (i.e., seven letters and three syllables).

Results

Primary statistical analysis: overall number of syllables effect

First fixation duration

For *first fixation duration*, there was a statistically significant effect of reading fluency ($p < .001$) and word frequency ($p < .001$); that is, *first fixation duration* decreased with increasing fluency and word frequency. Additionally, we found a statistically significant interaction of word length and reading fluency at the second contrast from seven to nine letters ($p = .0284$), indicating that the difference between seven- and nine-letter words was larger for less fluent readers.

The effect of number of syllables on *first fixation duration* turned out to be statistically significant ($p = .0308$), indicating that *first fixation duration* increased with increasing the number of syllables. More specifically, the effect of the number of syllables on *first fixation duration* was statistically significant at the contrast between two- and three-syllabic items ($p = .0308$), but not at the contrast between the three-and four-syllabic items.

Summed refixation duration

Similar to the findings for *first fixation duration*, there was also a statistically significant effect of word frequency ($p < .001$) and reading fluency ($p < .001$) in the *summed refixation duration*, indicating shorter *summed refixation durations* with increasing word frequency and increasing reading fluency. In addition to this, the interaction between word frequency and reading fluency turned out to be statistically significant ($p < .001$), indicating that the word frequency effect is larger for less fluent readers. However, in contrast to the *first fixation duration* results, no statistically significant number of syllables effects were found in the *summed refixation duration*.

Gaze duration

Again, the effects of reading fluency ($p < .001$) and word frequency ($p < .001$) were statistically significant, i.e., *gaze duration* was significantly higher in low-frequency words and less fluent readers. We also found a statistically significant effect of word length; that is, the contrast for six-letter compared to seven-letter words was statistically significant ($p = .0196$), while the contrast between seven- and nine-letter words was not. These results indicate that seven- and nine-letter words were read equally fast, while six-letter words were read significantly faster than seven-letter words. In addition, the interaction between reading fluency and word length was statistically significant for the contrast between six- and seven-letter words ($p = .0168$), but not for the contrast between seven- and nine-letter words, indicating that the length effect was smaller for faster readers. The interaction between reading fluency and word frequency again turned out to be highly statistically significant ($p < .001$), i.e., the effect of word frequency was larger for less fluent readers. No statistically significant number of syllables effects were found in the *gaze duration*.

First pass refixation probability

For *first pass refixation probability*, the length effect turned out to be statistically significant again; that is, the contrast for six-letter compared to seven-letter words was statistically significant ($p < .001$), showing that six-letter items were read with fewer refixations than seven-letter items ($p < .001$), while the contrast between seven- and nine-letter words was not statistically significant. In addition, we found a statistically significant effect of word frequency ($p < .001$) and reading fluency ($p < .001$), indicating a lower *first pass refixation probability* with increasing word frequency and increasing reading fluency. In addition, the interaction between reading fluency and word frequency was statistically significant ($p = .00974$); the word frequency effect was larger for less fluent readers. The number of syllables effect was not found to be statistically significant for *first pass refixation probability*.

Total fixation duration

A statistically significant length effect could again be observed in *total fixation duration*, showing that the contrast between six-letter items and seven-letter items was statistically significant ($p = .0107$), while this was not the case for the contrast between seven- and nine-letter items. Also, we found a statistically significant interaction of word length and reading fluency again at the contrast between six-letter versus seven-letter words ($p = .0202$), which indicates that the difference between seven- and six- letter words was larger for less fluent readers.

In addition, the main effects of reading fluency ($p < .001$) and word frequency ($p < .001$), as well as their interaction ($p < .001$), were statistically significant, indicating a larger effect of word frequency on *total fixation duration* in less fluent readers. Again, we could not find a statistically significant effect on the number of syllables.

Secondary statistical analysis: number of syllables effect across different reading fluency levels

To summarize, the interactions between the number of syllables and fluency category were not statistically significant for any of the eye-tracking measures; that is, highly fluent and dysfluent readers did not differ regarding their number of syllables effect. The individual model outputs can be found in the Appendix (see Appendix I-M).

Discussion

The present study aimed to investigate syllabic processing in Finnish word recognition in children in the third and fourth grades. To do so, we explored if and how the number of syllables affected both early and late eye-tracking measures in developing readers and whether the effect is dependent on reading fluency and word frequency. Moreover, we also tested if the effect differed across different levels of reading fluency (high versus low reading fluency). We found a number of syllables effect only in the earliest eye movement measure of *first fixation duration*, with the three-syllabic words being fixated longer than the bisyllabic words, whereas no such difference was observed between three- and four-syllabic words. These effects were not dependent on reading fluency or word frequency. Conversely, subsequent eye-tracking measures (*summed refixation duration*, *gaze duration*, *first pass refixation probability*, *total fixation duration*) were not significantly affected by the number of syllables.

The lack of a statistically significant number of syllables effect in the summative measures of word recognition (i.e., *gaze duration*, *total fixation duration*) seems to be in line with previous studies investigating syllabic processing in reading development. For example, Hautala et al. (2013) found that, while more proficient second graders were not significantly influenced by the number of syllables, dysfluent Finnish second graders still showed a syllable number effect. In the present study, eye movements of Finnish third- and fourth-grade readers no longer seem to be influenced by the number of syllables, and this was true even for children at the lowest level of reading fluency. From the previous literature, it is already known that the relevance of different units might change throughout the course of reading development (e.g., Ziegler & Goswami, 2005). Our results could thus be seen as a continuation of Hautala et al. (2013) findings, indicating that less proficient readers in Grades three and four, while still showing fluency problems, seem to have developed regarding their sublexical processing in word recognition. This finding also seems to be in line with previous work from Häikiö et al. (2015) who showed that already very early in their reading development, Finnish readers become capable of parallel syllabic processing or even whole-word recognition to some extent.

However, the present finding regarding the significant number of syllables effect in the *first fixation duration* measure irrespective of word frequency also seems to align with previous results from Stenneken et al. (2007). In their study, they found a significant syllable number effect during a lexical decision task in skilled German readers. From their results, they concluded that the orthographic input is being segmented into syllabic units, which, according to the authors, relates to prelexical processing mechanisms since this effect was true for both word and nonword stimuli. The present findings thus seem to complement Stenneken et al. (2007) results by demonstrating that this pattern can also be observed in children at a relatively early stage of literacy development acquiring Finnish, a language with a simpler syllable structure and an even more perfect orthographic transparency than German.

Implications to word recognition models

Both, the *CDP++ model* by Perry et al. (2010) and *MTM model* by Ans et al. (1998) make explicit predictions about a number of syllables effect, that is, they predict a number of syllables effect only for low frequency words which are subject to serial decoding. In the *CDP++ model*, the number of syllables effect stems from a serial graphemic parsing of letter strings and is thus suggested to emerge simultaneously with the word length effect. In the *MTM model*, words that have no orthographic word representations would be read aloud syllable by syllable. Finally, within the *E-Z Reader* framework (Reichle et al., 2003), the present results might reflect that the number of syllables plays a role in identifying the orthographic form, but not in activating phonological representations, as the number of syllables effect was only found in the very early eye-tracking measure (i.e., *first fixation duration*). This view is in contrast with the *MTM model* (Ans et al., 1998), in which unrepresented words are phonologically decoded syllable by syllable. In the case of Finnish, however, it should be noted that the inflected word form is unlikely to be identified in the initial stages, as Finnish words often consist of multiple morphemes. Rather, the identification of the first syllable might support the recognition of the word stem in agglutinative languages, such as Finnish.

The serial graphemic parsing process proposed by Perry et al. (2010) takes place earlier than the syllabic decoding mechanism suggested by *MTM model*, and therefore our finding of an early number of syllables effect is more in line with *CDP++ model*. However, in contrast to the *CDP++ model* predictions, we did find a number of syllables effect also for frequent words. In principle, it may be possible that our developing readers lacked orthographic representations for most of the studied words and therefore read the words mainly by decoding. Yet, this explanation seems unlikely as in a previous study, our fluent readers were found to read high frequency words with almost no signs of serial decoding (Hautala et al., 2021). On these grounds, one would have expected a larger number of syllables effect for low frequency words among fluent readers, which was not observed. The present results therefore seem to suggest that the number of syllables effect stems from rather early stages of orthographic processing. It is noteworthy that the number of syllables effect was not linear, but three-syllabic words were fixated longer than bi- and four-syllabic items. This result suggests that the observed effect may stem from a more general process of coding a word's orthographic structure including its consonant-vowel structure. In addition, the orthographic processing may be seen as a complex process of activations induced by a letter string, and these activations may include bigrams, syllables, morphemes, or words (cf. Dehaene et al., 2005).

Developmental considerations

We further found significant main effects of reading fluency in *first fixation duration*, which may reflect overall slowness in early orthographic processing such as letter encoding (Hautala et al., 2021). There were also interactions between word frequency and reading fluency in all but *first fixation duration*, i.e., the word frequency effect was shown to be higher for less fluent readers. According to Hautala et al. (2021), this result reflects the developmental status of decoding skills: Poor readers seem to have partial access to orthographic representations, as reflected by the larger effect as compared to fluent readers. However, as evidenced by the subsequent length effect (Hautala et al., 2021), they still need to decode even the frequent words, presumably due to a lack of direct connections between orthographic and phonological word representations or an overreliance on a serial decoding strategy, which is strongly emphasized in early reading instruction of (highly transparent) Finnish.

In the second part of the present analysis, which focused on comparing highly fluent and dysfluent readers regarding their number of syllables effect, it was shown that the groups did not differ significantly from one another, which might indicate qualitatively similar orthographic processing. For example, it is possible that both fluent and dysfluent readers use smaller units such as syllables, morphemes, or other letter combinations during reading, but fluent readers' processing skills are simply more automatized, which results in shorter reading times. This finding also seems to be in line with Hautala et al. (2021), who aimed at investigating word recognition processes in developing readers of Finnish. They were especially interested in finding out which underlying processes are compromised in slow readers. Their results contradict the prevailing opinion that the main deficit in dysfluent reading concerns a lack of orthographic representations. Rather, Hautala et al. (2021) conclude that the deficits lie in connecting the orthographic representation to its phonological representation. In such a scenario, the present result suggests that the syllabic parsing automatizes into an orthographic process during the early grades. It should be noted that the sample size in the secondary analysis was relatively small ($N = 34$). Yet, following the approach described by Judd et al. (2017), we found that the analysis was powerful enough for detecting minimum effect sizes between 0.07 and 0.13 for both analyses. However, to understand the developmental role of the number of syllables in reading Finnish, more studies are needed, especially in younger age groups and with varying reading skill levels. It should also be noted that due to the fact that students were initially recruited to participate in a reading intervention study (Hautala et al., 2022), poor readers were overrepresented in the current data set, which has to be considered when interpreting the present results.

As mentioned in the introduction, prevalent models of visual word recognition are mostly targeted at skilled reading only, although there seem to be efforts to extend these computational models to developing and compromised reading (see Perry et al., 2019; Ziegler et al., 2020). To this end, more empirical studies clarifying the time course of processing sublexical units during developing visual word recognition are still needed, as it is likely that the reliance on decoding, as well as the role of sublexical units and more specifically the role of the syllable, differs between experienced and developing reading. Thus, future research should also aim at simulating sublexical effects from a developmental point of view. Another aspect to consider when developing new computational models of word recognition or adaptations of existing versions should be the language context, as the underlying processes of word recognition are likely to differ between languages.

Limitations

A possible limitation of the present study is the fact that, due to investigating visual word recognition processes within a context as opposed to reading isolated words, the items in the present study might not be as strictly controlled for as in the case of single-word studies, which often feature uninflected words. Thus, we cannot rule out the potential impact other units might have on word recognition processes. For example, as mentioned earlier, Finnish is characterized by a very rich morphology; with most words being inflected rather than uninflected. It is therefore possible that emerging readers of Finnish might find recognizing morphemes rather than syllables a more helpful strategy for achieving solid reading fluency after the initial phase of reading acquisition. Results from Hasenäcker and Schroeder (2017), where the use of syllables and morphemes in word recognition was compared between German second and fourth graders using a lexical decision task, showed that while second graders preferred syllables, German fourth graders were already sensitive to morphemic units. Similar results were found in French, where Colé et al. (2012) showed that second graders were already sensitive to morphemes. However, morphemes might be even more relevant in Finnish, since the morphology is in fact even more complex than in German or French (Aro, 2017). Some more recent results from Häikiö and Vainio (2018) also give insight into morphological processing in Finnish early reading, showing that Finnish first graders already seem to be sensitive to morphemic information during reading. Especially when a context is given, readers have various information available that could guide word recognition as well as help predict morphological information. In some situations, processing units other than the syllable might be more beneficial. Therefore, future research should focus on multiple processing units within a comparable setting to disentangle possible inter-relations between them. In addition, subsequent studies could also investigate how pseudowords are being decoded during reading development, as children might process units larger than the syllable or even whole word forms (the latter being likelier in languages with a simpler morphological system than Finnish). Using pseudowords instead of real words could thus help to exclude lexical activation a priori, which helps in understanding how unknown words are being read.

Finally, however, it should be noted that the number of syllables effect turned out to be relatively weak in the present study. The finding that the number of syllables effect was not present in re-reading times, such as *total fixation duration* could suggest that due to higher variances, late measures tend to have lower statistical power than early measures (Von der Malsburg & Angele, 2017). Thus, the number of syllables effect certainly requires further replication in upcoming research, perhaps in different languages and with different items in more controlled sentence reading studies.

Summary and conclusion

In summary, the present results indicate that syllabic parsing is a highly rapid and automatized process that predominantly takes place during the early holistic orthographic processing of a word. The fact that dysfluent readers are also capable of rapid syllabic parsing implies relatively preserved orthographic holistic processing even in readers who lag behind in the development of reading skills according to the age expectations. However, it is quite well established that developmental dyslexia is associated with slow visuo-orthographic processing, such as slow letter encoding (e.g., Paizi et al., 2013), and it is possible that there are even more deficient processing levels involved than generally thought (e.g., difficulties in mapping orthographic activations onto phonological counterparts, see Hautala et al., 2021). Thus, the results suggest an increasing need to examine the fine time course of visuo-orthographic and higher-level orthographic processing in dyslexia.

The current study also provides important initial insights into the time course of syllable processing in a language that notably differs from the dominant languages that reading research has focused on in the past. Therefore, it is beneficial to consider different language contexts, as reading development is influenced by various linguistic features, such as orthographic transparency or morphological complexity. Although investigating the processes involved in reading development cross-linguistically is always a complicated endeavor (see Papadopoulos et al., 2021), reading researchers should aim for this line of research in order to understand how reading develops cross-linguistically and to ultimately use this knowledge to develop efficient training programs to support developing readers.

Note

1. The copyrights of both texts have expired, and modernized Finnish abridgements can be accessed at https://iltasatu.org/.
2. It should be noted that in Finnish, eight-letter words are usually three-syllabic and thus, not enough stimuli were available in the present materials to compare multiple eight-letter words with a varying number of syllables.

Acknowledgments

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

Disclosure statement

No potential conflict of interest was reported by the authors.

Ethics approval statement

The study was approved by the Ethical Committee of the University of Jyväskylä, and the research was conducted in accordance with the ethical principles for medical research involving human subjects set forth by the Declaration of Helsinki. Informed written consent was obtained from both the children and their caregivers prior to the study.

Funding

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement [813546] and the Academy of Finland [317030,319911].

ORCID

Lisa Hintermeier http://orcid.org/0000-0002-5219-069X
Jarkko Hautala http://orcid.org/0000-0002-7402-6364
Mikko Aro http://orcid.org/0000-0002-0545-0591

References

Alvarez, C. J., Carreiras, M., & Taft, M. (2001). Syllables and morphemes: Contrasting frequency effects in Spanish. *Journal of Experimental Psychology Learning, Memory, and Cognition*, *27*(2), 545–555. https://doi.org/10.1037/0278-7393.27.2.545

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, *105*(4), 678–723. https://doi.org/10.1037/0033-295X.105.4.678-723

Aro, M. (2017). Learning to read finnish. In *Learning to read across languages and writing systems*. Cambridge University Press.

Carreiras, M., Alvarez, C. J., & Devega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language*, *32*(6), 766–780. https://doi.org/10.1006/jmla.1993.1038

Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain and Language*, *90*(1), 393–400. https://doi.org/10.1016/j.bandl.2003.12.003

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. https://doi.org/10.1080/23273798.2013.876504

Colé, P., Bouton, S., Leuwers, C., Casalis, S., & Sprenger-Charolles, L. (2012). Stem and derivational-suffix processing during reading by French second and third graders. *Applied Psycholinguistics*, *33*(1), 97–120. https://doi.org/10.1017/S0142716411000282

Conrad, M., Stenneken, P., & Jacobs, A. (2006). Associated or dissociated effects of syllable frequency in lexical decision and naming. *Psychonomic Bulletin & Review*, *13*(2), 339–345. https://doi.org/10.3758/BF03193854

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. https://doi.org/10.1016/j.tics.2005.05.004

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, *18*(2), 116–125. https://doi.org/10.1111/j.1467-9817.1995.tb00077.x

Eriksen, C. W., Pollack, M. D., & Montague, W. E. (1970). Implicit speech: Mechanism in perceptual encoding? *Journal of Experimental Psychology*, *84*(3), 502–507. https://doi.org/10.1037/h0029274

Ferrand, L. (2000). Reading aloud polysyllabic words and nonwords: The syllabic length effect reexamined. *Psychonomic Bulletin & Review*, *7*(1), 142–148. https://doi.org/10.3758/BF03210733

Ferrand, L., & New, B. (2003). Syllabic length effects in visual word recognition and naming. *Resource Policy*, *113*(2), 167–183. https://doi.org/10.1016/s0001-6918(03)00031-3

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 627–635. https://doi.org/10.1016/S0022-5371(73)80042-8

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology Human Perception and Performance*, *2*(3), 361–379. https://doi.org/10.1037/0096-1523.2.3.361

Häikiö, T., Hyönä, J., & Bertram, R. (2015). The role of syllables in word recognition among beginning Finnish readers: Evidence from eye movements during reading. *Journal of Cognitive Psychology*, *27*(5), 562–577. https://doi.org/10.1080/20445911.2014.982126

Häikiö, T., & Vainio, S. (2018). Syllables and inflectional morphemes in early Finnish readers: Evidence from eye-movements. *Journal of Child Language*, *45*(5), 1227–1245. https://doi.org/10.1017/S0305000918000132

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. https://doi.org/10.1017/S0142716416000412

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 & Writing*, *26*(6), 845–864. https://doi.org/10.1007/s11145-012-9394-3

Hautala, J., Hawelka, S., & Aro, M. (2021). 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

Hautala, J., Ronimus, M., & Junttila, E. (2022). Readers' theater projects for special education: A randomized controlled study. *Scandinavian Journal of Educational Research*, 1–16. https://doi.org/10.1080/00313831.2022.2042846

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]*. Hogrefe Psykologien Kustannus Oy.

Jared, D., & Seidenberg, M. S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology Human Perception and Performance*, *16*(1), 92–105. https://doi.org/10.1037/0096-1523.16.1.92

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, *68*(1), 601–625. https://doi.org/10.1146/annurev-psych-122414-033702

Klapp, S. T., Anderson, W. G., & Berrian, R. W. (1973). Implicit speech in reading: Reconsidered. *Journal of Experimental Psychology*, *100*(2), 368–374. https://doi.org/10.1037/h0035471

Lerkkanen, M. K., Niemi, P., Poikkeus, A. M., Poskiparta, M., Siekkinen, M., & Nurmi, J. E. (2006). The first steps study. *Unpublished data, Department of Psychology and Department of Teacher Education, University of Jyväskylä, Jyväskylä, Finland*.

Luque, J. L., López-Zamora, M., Álvarez, C. J., & Bordoy, S. (2013). Beyond decoding deficit: Inhibitory effect of positional syllable frequency in dyslexic Spanish children. *Annals of Dyslexia*, *63*(3–4), 239–252. https://doi.org/10.1007/s11881-013-0082-z

Mason, M. (1978). From print to sound in mature readers as a function of reader ability and two forms of orthographic regularity. *Memory & Cognition*, *6*(5), 568–581. https://doi.org/10.3758/BF03198246

Paizi, D., De Luca, M., Zoccolotti, P., & Burani, C. (2013). A comprehensive evaluation of lexical reading in Italian developmental dyslexics: Lexical reading in Italian developmental dyslexics. *Journal of Research in Reading*, *36*(3), 303–329. https://doi.org/10.1111/j.1467-9817.2011.01504.x

Papadopoulos, T. C., Csépe, V., Aro, M., Caravolas, M., Diakidoy, I. -A., & Olive, T. (2021). Methodological issues in literacy research across languages: Evidence from alphabetic orthographies. *Reading Research Quarterly*, *56*(S1). https://doi.org/10.1002/rrq.407

Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology Human Perception and Performance*, *24*(1), 134. https://doi.org/10.1037/0096-1523.24.1.134

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. https://doi.org/10.1016/j.cogpsych.2010.04.001

Perry, C., Zorzi, M., & Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. *Psychological Science*, *30*(3), 386–395. https://doi.org/10.1177/0956797618823540

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL https://www.R-project.org/

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z reader model of eye-movement control in reading: Comparisons to other models. *The Behavioral and Brain Sciences*, *26*(4), 445–476. https://doi.org/10.1017/S0140525X03000104

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. https://doi.org/10.1348/000712603321661859

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), 584–615. https://doi.org/10.1037/0033-2909.134.4.584

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2021). Afex: Analysis of factorial experiments. R package version 0.28-1. https://CRAN.R-project.org/package=afex

Spyri, J. (2015). *Little Heidi (K. Weiss, Trans.)*. (Original work published 1881). https://iltasatu.org/satu/pikku-heidi/

Stenneken, P., Conrad, M., & Jacobs, A. M. (2007). Processing of syllables in production and recognition tasks. *Journal of Psycholinguistic Research*, *36*(1), 65–78. https://doi.org/10.1007/s10936-006-9033-8

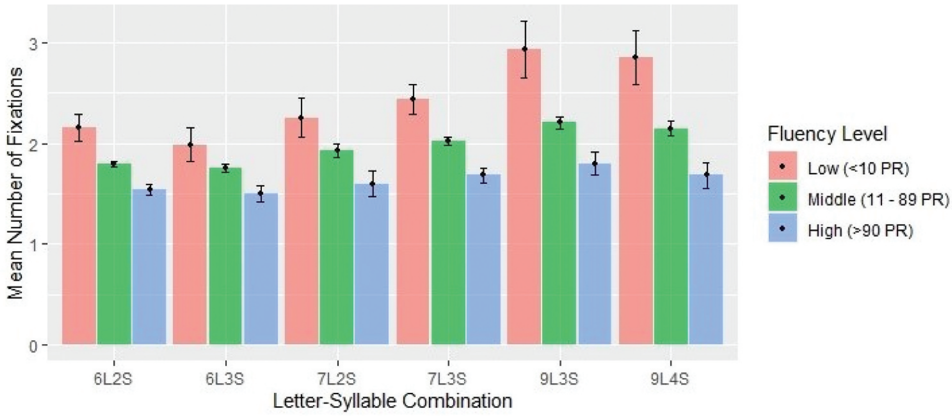
Topelius, Z. (2017). *Adalmina's Pearl*. (Original work published 1865). https://iltasatu.org/satu/adalminan-helmi/

Von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of Memory and Language*, *94*, 119–133. https://doi.org/10.1016/j.jml.2016.10.003

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

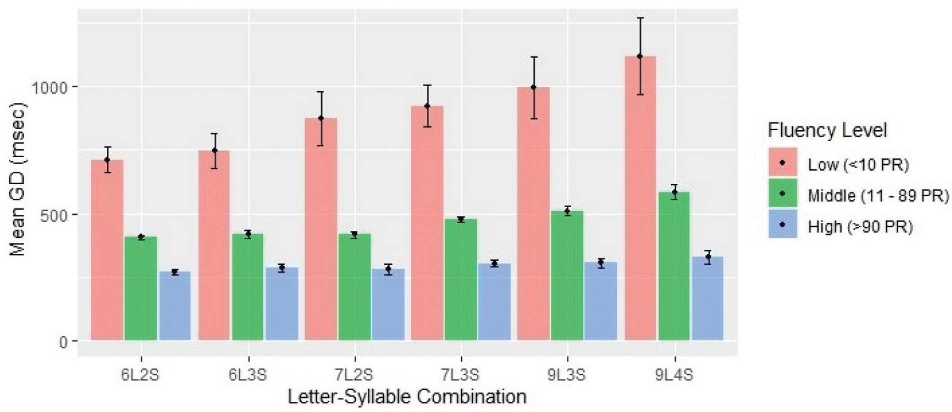
Ziegler, J. C., Perry, C., & Zorzi, M. (2020). Learning to read and dyslexia: From theory to intervention through personalized computational models. *Current Directions in Psychological Science*, *29*(3), 293–300. https://doi.org/10.1177/0963721420915873

Appendix A. Mean Number of Fixations per Item Category and Fluency Level



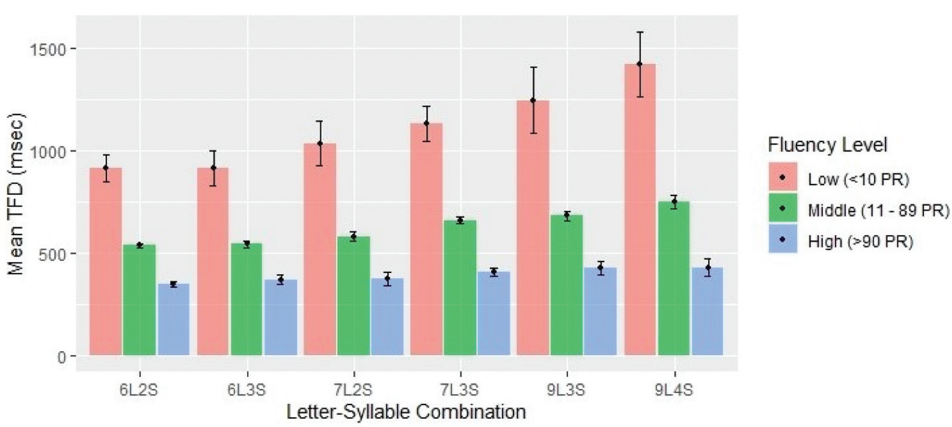
Note. 6L2S = words with 6 letters and 2 syllables; PR = percentile rank.

Appendix B. Mean Gaze Duration (GD) per Item Category and Fluency Level



Note. 6L2S = words with 6 letters and 2 syllables; PR = percentile rank.

Appendix C. Mean Total Fixation Duration (TFD) per Item Category and Fluency Level



Note. 6L2S = words with 6 letters and 2 syllables; PR = percentile rank.

Appendix D. First Fixation Duration (FFD): Results of the Linear-Mixed Model in the Main Analysis

Effect	Estimate	SE	df	t-value	Pr(>|t|)
Intercept	5.597	0.019	226.4	295.492	<.001***
NoS_2	−0.038	0.018	206.5	−2.175	0.031 *
NoS_4	0.003	0.013	193.9	0.219	0.827
WordLength_6	0.012	0.014	206.2	0.863	0.389
WordLength_9	0.012	0.012	204.9	0.980	0.328
ReadingFluency	−0.136	0.016	140.3	−8.628	<.001***
WordFrequency	−0.047	0.008	202.5	−5.581	<.001***
WordLength_6 × ReadingFluency	−0.001	0.005	1090	−0.162	0.871
WordLength_9 × ReadingFluency	−0.012	0.005	6596	−2.193	0.028 *

Note. Model specification: mixed(FFD) ~ NoS+ WordLength* ReadingFluency+ WordFrequency + (1 + NoS + WordFrequency || subject) + (1 | item), data). NoS = Number of Syllables; SE = Standard Error; df = Degrees of Freedom.

Appendix E. Summed Refixation Duration (GD_FFD): Results of the Linear-Mixed Model in the Main Analysis

Effect	Estimate	SE	df	t-value	Pr(>|t|)
Intercept	5.660	0.029	268.025	196.382	<.001***
NoS_2	−0.011	0.037	212.016	−0.290	0.772
NoS_4	−0.033	0.027	190.326	−1.208	0.229
WordLength_6	−0.049	0.030	214.652	−1.658	0.099.
WordLength_9	0.005	0.025	202.322	0.214	0.831
WordFrequency	−0.135	0.018	206.317	−7.653	<.001***
ReadingFluency	−0.275	0.020	148.850	−13.684	<.001***
WordFrequency x ReadingFluency	0.040	0.009	138.605	4.355	<.001***

Note. Model specification: mixed(GD_FFD) ~ NoS + WordLength + WordFrequency * ReadingFluency+ (1 + NoS + WordFrequency || subject) + (1 | item), data). NoS = Number of Syllables; SE = Standard Error; df = Degrees of Freedom.

Appendix F. Gaze Duration (GD): Results of the Linear-Mixed Model in the Main Analysis

Effect	Estimate	SE	df	t-value	Pr(>|t|)
Intercept	5.887	0.023	286,3	261.034	<.001***
NoS_2	−0.035	0.025	207.1	−1.414	0.159
NoS_4	0.011	0.018	205.9	0.623	0.534
WordLength_6	−0.047	0.020	207.4	−2.352	0.020 *
WordLength_9	−0.009	0.017	206.8	−0.541	0.589
ReadingFluency	−0.243	0.017	141.3	−14.390	<.001***
WordFrequency	−0.093	0.012	234.5	−7.606	<.001***
WordLength_6 × ReadingFluency	0.013	0.005	24010	2.391	0.017 *
WordLength_9 × ReadingFluency	0.002	0.005	24030	0.341	0.733
ReadingFluency × WordFrequency	0.024	0.005	161.7	4.513	<.001***

Note. Model specification: mixed(GD) ~ NoS + WordLength * ReadingFluency+ WordLength * ReadingFluency+ (1 + WordFrequency || subject) + (1 | item), data). NoS = Number of Syllables; SE = Standard Error; df = Degrees of Freedom.

Appendix G. Total Fixation Duration (TFD): Results of the Linear-Mixed Model in the Main Analysis

Effect	Estimate	SE	df	t-value	Pr(>|t|)
Intercept	6.158	0.027	329.1	231.441	<.001***
NoS_2	−0.012	0.033	208.2	−0.356	0.722
NoS_4	0.019	0.025	207.5	0.771	0.442
WordLength_6	−0.069	0.027	208.3	−2.574	0.011 *
WordLength_9	0.010	0.023	208.0	0.430	0.668
ReadingFluency	−0.271	0.018	141.4	−15.402	<.001***
WordFrequency	−0.130	0.016	230.3	−8.111	<.001***
WordLength_6 × ReadingFluency	0.013	0.005	24040	2.323	0.020 *
WordLength_9 × ReadingFluency	−0.001	0.005	24060	−0.262	0.793
ReadingFluency × WordFrequency	0.032	0.005	165.7	5.806	<.001***

Note. Model specification: mixed(TFD) ~ NoS + WordLength * ReadingFluency+ WordFrequency * ReadingFluency+ (1+WordFrequency || subject) + (1 | item), data). NoS = Number of Syllables; SE = Standard Error; df = Degrees of Freedom.

Appendix H. First Pass Refixation Probability (RefixProb): Results of the Linear-Mixed Model in the Main Analysis

Effect	Estimate	SE	z-value	Pr(>|z|)
Intercept	−0.904	0.054	−16.862	<.001***
NoS_2	−0.024	0.066	−0.364	0.715
NoS_4	0.069	0.049	1.408	0.159
WordLength_6	−0.274	0.054	−5.072	<.001***
WordLength_9	−0.085	0.045	−1.868	0.062.
WordFrequency	−0.159	0.032	−5.049	<.001***
ReadingFluency	−0.443	0.038	−11.548	<.001***
WordFrequency x ReadingFluency	0.040	0.015	2.585	0.010 **

Note. Model specification: RefixProb ~ NoS + WordLength + WordFrequency * ReadingFluency+ (1 +NoS || subject) + (1 | item), data). NoS = Number of Syllables; SE = Standard Error.

Appendix I. First Fixation Duration (FFD): Results of the Linear-Mixed Model in the Fluency Analysis

Effect	Estimate	SE	df	t-value	Pr(>|t|)
Intercept	518.527	36.216	94.383	14.318	<.001***
NoS_2	−38.465	17.383	637.062	−2.213	0.027 *
NoS_4	5.750	24.109	641.665	0.239	0.812
WordLength_6	16.604	17.498	610.580	0.949	0.343
WordLength_9	43.176	14.422	621.778	2.994	0.003 **
FluencyCategoryhigh	−263.853	48.662	88.220	−5.422	<.001***
WordFrequency	−54.576	13.240	37.799	−4.122	<.001***
NoS_2 × WordLength_6	19.530	39.199	639.617	0.498	0.618
NoS_2 × FluencyCategoryhigh	24.228	22.385	4972.315	1.082	0.279
NoS_4 × FluencyCategoryhigh	−9.232	31.300	4975.447	−0.295	0.768
WordLength_6 × FluencyCategoryhigh	−9.273	22.631	4945.993	−0.410	0.682
WordLength_9 × FluencyCategoryhigh	−43.850	18.630	4970.492	−2.354	0.019 *
FluencyCategoryhigh × WordFrequency	41.045	17.937	34.952	2.288	0.028 *
NoS_2 × WordLength_6 × FluencyCategoryhigh	−30.171	51.202	4967.667	−0.589	0.556

Note. Model specification: mixed(FFD ~ NoS * WordLength * FluencyCategory + WordFrequency * FluencyCategory + (1 + WordFrequency || subject) + (1 | item), data). NoS = Number of Syllables; SE = Standard Error; df = Degrees of Freedom.

Appendix J. Summed Refixation Duration (GD_FFD): Results of the Linear-Mixed Model in the Fluency Analysis

Effect	Estimate	SE	df	t-value	Pr(>|t|)
Intercept	679.491	96.632	189.075	7.032	<.001***
NoS_2	−7.299	57.842	98.152	−0.126	0.900
NoS_4	−5.224	76.960	256.976	−0.068	0.946
WordLength_6	−120.595	57.746	380.883	−2.088	0.037 *
WordLength_9	15.933	45.555	341.339	0.350	0.727
FluencyCategoryhigh	−447.808	151.127	344.278	−2.963	0.003 **
WordFrequency	−118.617	33.025	46.048	−3.592	0.001***
NoS_2 × WordLength_6	83.232	128.077	379.900	0.650	0.516
NoS_2 × FluencyCategoryhigh	−10.600	88.258	150.683	−0.120	0.905
NoS_4 × FluencyCategoryhigh	−4.873	120.631	694.127	−0.040	0.968
WordLength_6 × FluencyCategoryhigh	120.488	92.835	1494.930	1.298	0.195
WordLength_9 × FluencyCategoryhigh	−18.898	68.504	1488.744	−0.276	0.783
FluencyCategoryhigh × WordFrequency	110.428	48.339	56.134	2.284	0.026 *
NoS_2 × WordLength_6 × FluencyCategoryhigh	−70.626	205.852	1510.973	−0.343	0.732

Note. Model specification: mixed(GD_FFD ~ NoS * WordLength * FluencyCategory + WordFrequency * FluencyCategory + (1 + NoS + WordFrequency || subject) + (1 | item), data). NoS = Number of Syllables; SE = Standard Error; df = Degrees of Freedom.

Appendix K. Gaze Duration (GD): Results of the Linear-Mixed Model in the Fluency Analysis

Effect	Estimate	SE	df	t-value	Pr(>|t|)
Intercept	869.606	71.372	86.114	12.184	<.001***
NoS_2	−18.815	32.508	464.035	−0.579	0.563
NoS_4	14.474	45.061	466.046	0.321	0.748
WordLength_6	−95.394	35.751	198.660	−2.668	0.008 **
WordLength_9	25.372	29.177	122.268	0.870	0.386
FluencyCategoryhigh	−584.178	90.886	65.305	−6.428	<.001***
WordFrequency	−142.290	24.470	44.652	−5.815	<.001***
NoS_2 × WordLength_6	25.710	73.300	463.721	0.351	0.726
NoS_2 × FluencyCategoryhigh	1.920	35.920	4899.830	0.053	0.957
NoS_4 × FluencyCategoryhigh	−3.948	50.227	4902.106	−0.079	0.937
WordLength_6 × FluencyCategoryhigh	87.447	41.242	137.791	2.120	0.036 *
WordLength_9 × FluencyCategoryhigh	−24.020	33.529	73.492	−0.716	0.476
FluencyCategoryhigh × WordFrequency	125.619	31.499	33.892	3.988	<.001***
NoS_2 × WordLength_6 × FluencyCategoryhigh	−14.463	82.157	4897.777	−0.176	0.860

Note. Model specification: mixed(GD ~ NoS * WordLength * FluencyCategory + WordFrequency * FluencyCategory + (1 + WordLength WordFrequency || subject) + (1|item), data). NoS = Number of Syllables; SE = Standard Error; df = Degrees of Freedom.

Appendix L. Total Fixation Duration (TFD): Results of the Linear-Mixed Model in the Fluency Analysis

Effects	Estimate	SE	df	t-value	Pr(>|t|)
Intercept	1046.30	87.62	87.70	11.941	<.001***
NoS_2	−29.59	39.99	421.00	−0.740	0.460
NoS_4	44.60	55.43	422.65	0.805	0.421
WordLength_6	−117.09	40.43	406.80	−2.897	0.004 **
WordLength_9	20.94	33.26	413.38	0.630	0.529
FluencyCategoryhigh	−672.10	109.63	62.32	−6.131	<.001***
WordFrequency	−193.15	29.97	49.83	−6.444	<.001***
NoS_2 × WordLength_6	104.52	90.17	420.38	1.159	0.247
NoS_2 × FluencyCategoryhigh	19.87	41.87	4932.32	0.475	0.635
NoS_4 × FluencyCategoryhigh	−20.70	58.55	4933.90	−0.354	0.724
WordLength_6 × FluencyCategoryhigh	92.76	42.29	4916.54	2.193	0.028 *
WordLength_9 × FluencyCategoryhigh	−18.11	34.84	4931.87	−0.520	0.603
FluencyCategoryhigh × WordFrequency	161.45	37.96	35.65	4.254	<.001***
NoS_2 × WordLength_6 × FluencyCategoryhigh	−86.92	95.75	4930.03	−0.908	0.364

Note. Model specification: mixed(TFD ~ NoS * WordLength * FluencyCategory + WordFrequency * FluencyCategory + (1 + WordFrequency || subject) + (1 | item), data). NoS = Number of Syllables; SE = Standard Error; df = Degrees of Freedom.

Appendix M. First Pass Refixation Probability (RefixProb): Results of the Linear-Mixed Model in the Fluency Analysis

Effects	Estimate	SE	z-value	Pr(>|z|)
Intercept	-0.289	0.137	-2.112	0.035*
NoS_2	0.087	0.094	0.921	0.357
NoS_4	0.113	0.069	1.633	0.103
WordLength_6	-0.362	0.078	-4.660	< .001***
WordLength_9	-0.035	0.065	-0.546	0.585
FluencyCategoryhigh	-1.354	0.182	-7.432	< .001***
WordFrequency	-0.253	0.055	-4.576	< .001***
FluencyCategoryhigh x WordFrequency	0.178	0.070	2.532	0.011*

Note. Model specification: mixed(RefixProb ~ NoS+ WordLength + FluencyCategory + WordFrequency * FluencyCategory + (1 + WordFrequency || subject) + (1 | item), data). NoS= Number of Syllables; SE = Standard Error.