The hyphen as a syllabification cue in reading bisyllabic and multisyllabic words among Finnish 1st and 2nd graders

[Post-Print version, as published in Reading and Writing, doi: 10.1007/s11145-015-9584-x]

Tuomo Häikiö, Raymond Bertram & Jukka Hyönä

University of Turku

Finland

Address all correspondence to:

Tuomo Häikiö, Department of Psychology, University of Turku, FIN-20014 Turku, Finland

Tel: +358-2-3338501, Fax: +358-2-3335060, E-mail: tuilha@utu.fi

Acknowledgements

This study was financially supported by the Finnish Graduate School of Psychology, Oskar Öflunds Stiftelse [Oskar Öflund Foundation] and Suomen Kulttuurirahaston Varsinais-Suomen rahasto [Finnish Culture Foundation, Varsinais-Suomi Regional Fund] to the first author and the Academy of Finland (grant 118404) to the second author. We would like to thank Petar Milin for an advice on an earlier version of this paper. Correspondence concerning this article should be addressed to Tuomo Häikiö, Department of Psychology, University of Turku, FIN-20014 Turku, Finland. Email may be sent to tuilha@utu.fi.

The hyphen as a syllabification cue in reading bisyllabic and multisyllabic words among Finnish 1st and 2nd graders

Abstract

Finnish ABC books present words with hyphens inserted at syllable boundaries. Syllabification by hyphens is abandoned in the 2nd grade for bisyllabic words, but continues for words with three or more syllables. The current eye movement study investigated how and to what extent syllable hyphens in bisyllabic (*kah-vi* 'cof-fee') and multisyllabic words (*haa-ruk-ka* 'fork', *ap-pel-sii-ni* 'orange') affect eye movement behavior and reading speed of Finnish 1st and 2nd graders. Experiment 1 showed that 2nd graders had longer gaze durations, needed more fixations and had longer selective regression path durations for hyphenated than concatenated words. This implies that hyphenated words were difficult to process when first encountered, but also hard to integrate with prior sentence context. The effects were modified by number of syllables and reading skill. That is, the hyphenation effects were larger for multisyllabic than bisyllabic words and larger for more than less proficient readers. Experiment 2 showed the same hyphenation effect for 1st graders reading long multisyllabic words, even with a hyphen that was smaller in size and hence visually less salient. We argue that syllable hyphens prevent reasonably proficient readers from using the most efficient processing route for bi- and multisyllabic words and discuss the possible implications of the results for early Finnish reading instruction.

Keywords: hyphenation, syllables, lexical access, reading development, eye movements, Finnish

The hyphen as a syllable cue in reading bisyllabic and multisyllabic words among Finnish 1st and 2nd graders

Several developmental reading theories emphasize the importance of making grapheme-phoneme connections in the initial stages of word reading (e.g., Ehri, 1987, 1989, 1995, 2005; Ehri & McCormick, 1998; Frith, 1985). According to Ehri's phase theory (2005), children move after a pre-alphabetic phase into the *partial-alphabetic* phase, during which they form some connections between letters seen in spellings of words and sounds detected in their pronunciations. That is, the connections pertain to a few letters only and typically these letters are in salient positions like word beginnings or words ends (e.g., recognizing the word SPOON on the basis of the word-initial S and the word-final N). In the *full-alphabetic phase* children can make complete connections between graphemes and phonemes. In English, this means that for a word like SPOON they can make the correspondence between the five letters of the word and its four phonemes, including the correspondence between OO and /u/. The ability to make all the connections also enables them to decode words never read before and to extend their sight vocabulary rapidly. In the next phase in Ehri's (1995) framework, the *consolidated-alphabetic* phase, recurring letter patterns like rimes, syllables and morphemes become consolidated or unitized. Ehri (2005) argues that chunking letters is valuable for remembering and reading words with more than one syllable. A word such as *interesting* would then be learned more easily, because fewer connections would be required to store its representation in memory, the number being reduced from 10 grapheme–phonemes connections to four syllabic chunks.

The most generally used reading instruction method in Finland is the KÄTS method (kirjain-äänne-tavu-sana = 'letter-sound-syllable-word', developed by Karppi, 1983). This method is in line with Ehri's theory, but includes a more prominent role for the syllable at an earlier stage. That is, this method emphasizes that – already after the first letters are taught – word recognition should be mediated via syllables. Thus early reading instruction is based on decoding syllables and assembling whole word pronunciation from the syllables. The semantics of the whole word is then retrieved via the whole-word phonological representation. Especially for longer, morphologically complex words (e.g., talossanikin = 'also in my house') – which are quite common in Finnish due to its morphological productivity – the assembly via syllables is thought to be crucial, as phonological recoding merely via grapheme-phoneme conversions would burden working memory too much. The method not only teaches children how to read, but also how to spell correctly. Nevertheless, it has also received some criticism. More specifically, in her textbook on reading instruction methods, Lerkkanen (2006) describes some possible setbacks of the KÄTS method. She notes that the method is especially suitable for children at risk in reading development, but that for normally developing children the method may be too repetitive and the continuous focus on sounding out syllables may undermine their motivation. Moreover, she even states that – in the end – the development of reading comprehension skills may be delayed.

There is – to our knowledge – no clear evidence that the development of reading comprehension would indeed be compromised by this kind of method, but it can be hypothesized that the strong reliance on the phonological pathway in early Finnish reading instruction has certain consequences. One of the consequences may be a delay in the development of the orthographic pathway. The multiple-route model of silent word reading development of Grainger, Lété, Bertand, Dufau and Ziegler (2012) for instance holds that the orthographic pathway becomes increasingly important during the development of word recognition skills. In support of their model, they found that the impact of phonological manipulations decreased and the impact of orthographic manipulations increased significantly during elementary school years, especially between grade 1 and 2. Häikiö, Hyönä and Bertram (2014) presented a model derived from the Grainger and Ziegler (2011) model that also aims to explain the shift from phonological to orthographic pathways during development in word recognition. An upgraded version of this model can be found in Figure 1.

INSERT FIGURE 1 ABOUT HERE

The model includes multiple routes (Routes 1a to 3b) to word recognition and holds that at the early stages of reading development children recognize words via phonological recoding (Route 1), either via grapheme-phoneme conversions (Route 1a) or – as in Finnish – via mediation of syllabic assembly (Route 1b). As reading skill develops, words will be more often recognized via orthographic units, first via smaller orthographic units (Route 2) feeding into phonological units (Route 2a) or activating the orthographic word directly (Route 2b), later via direct access of the whole orthographic word (Route 3a) or via non-contingent letter clusters (Route 3b) (see Häikiö et al., 2014, for a more detailed explanation). These suggestions are in line with the self-teaching hypothesis of Share (1995), who proposed that exposure to printed words not only invites beginning readers to perform phonological recoding, but also provides them with the opportunity to learn word-specific print-to-meaning connections.

The KÄTS method is supported by presenting hyphens at syllable boundaries in Finnish school books during the first two elementary school years. Hyphens are inserted in words with more than one syllable. Thus, words like *kahvi* = 'coffee' or *haarukka* = 'fork' are presented in hyphenated format (*kah-vi*, *haa-ruk-ka*). The hyphens indicate which letters are to be merged to form syllables. Subsequently the syllables can be assembled to assess the phonological word form, via which the meaning of the word can be retrieved. Since the Finnish word stock includes only few monosyllabic words[1], Finnish children are exposed to many hyphenated words during their first school year (see Figure 2). Hyphenation is gradually removed from words during the 2nd grade, but is preserved in most long words. By the 3rd grade, all textbooks are presented without hyphenation.

It is of interest to know whether and to what extent hyphens at syllable boundaries affect reading behavior and reading development. Do hyphens actually accelerate or slow down word recognition? Do they have the same impact on reading behavior in short and long words? Do hyphens affect reading behavior similarly for more and less fluent readers? Do they come to aid in reading comprehension? The current eye movement study was designed to answer these questions for Finnish children during early reading development.

INSERT FIGURE 2 ABOUT HERE

The study is a follow-up study of a recent study of Häikiö et al. (2014), who investigated to what extent intrasyllabic hyphens (e.g., *ka-hvi*) and intersyllabic hyphens (e.g., *kah-vi*) affect word recognition speed of Finnish 1st and 2nd graders. They found that intrasyllabic hyphenation (e.g., *ka-hvi*) generated slower reading times than intersyllabic hyphenation (*kah-vi*). This led them to conclude that pure grapheme-phoneme recoding without making use of syllabic units is a less efficient reading strategy than syllable-based reading. However, at the same time they found that words with hyphens at syllable boundaries (*kah-vi*, *haa-ruk-ka*) were recognized slower than words in concatenated format (*kahvi*, *haarukka*). Sentences of around six words in length took about one second longer to read when the words included intersyllabic hyphens in comparison to when the words were concatenated. This indicates that intersyllabic hyphens enforce a slower reading strategy upon the early elementary school readers[2].

One possibly is that the hyphen encourages phonological mediation to a larger extent than needed for reasonably proficient 1st and 2nd graders. For instance, it is possible that a whole-word orthographic representation is addressed while bypassing phonological syllable representations in case there are no intersyllabic hyphens, but that the hyphen by separating the word into syllabic units forces sequential syllable processing upon the readers during which the phonological representations of syllables are addressed as well. In line with this idea, Häikiö et al. (2014) argues that the hyphen affects the way visual attention is allocated within the word. More precisely, when processing bisyllabic or multisyllabic hyphenated words, the initial attentional focus may be predominantly allocated to the first syllable (i.e., to a visually distinct object). Without hyphenation visual attention is more likely to stretch out over the whole word. In the latter case, word recognition may be predominantly orthographic in nature (goes by via Routes 2b, 3a, or 3b), which would be faster than using phonological mediation.

The current study investigates to what extent intersyllabic hyphens (e.g., *kah-vi*) affect word recognition speed of Finnish 1st and 2nd graders with a new group of 1st and 2nd graders (in comparison to Häikiö et al., 2014). Four specific issues are addressed. First, it is investigated to what extent the role of hyphens in word reading is modified by the number of syllables. Typically, children need more fixations and more reading time for long multisyllabic than for short bisyllabic words (Hyönä & Olson, 1995). One reason for this may be that for long multisyllabic words they more often resort to a sublexical processing strategy involving phonological mediation than for short bisyllabic words. If this is the case, hyphenation may be more beneficial when reading long multisyllabic than short bisyllabic words. The fact that Finnish ABC book writers preserve hyphenation only in long multisyllabic words until the end of the 2nd grade (e.g., haa-ruk-ka = ‘fork’, ap-pel-sii-ni = ‘orange’) is – at least in part – inspired by this belief. That is, one would think that especially for long multisyllabic words, whole-word phonological representations should be assembled from phonological syllables and this is more easily achieved when hyphens clearly indicate the syllable boundaries. In the current study, we therefore directly tested whether 1st and 2nd graders benefit more from (or are disrupted less by) hyphenation in long multisyllabic words than in short bisyllabic words.

The second issue is the role of proficiency in reading hyphenated words. Earlier studies hinted at the possibility that less proficient readers are less disrupted by hyphenation (Häikiö et al., 2011; Häikiö et al., 2014). In fact, given that it is more likely that less proficient readers resort to syllable-based processing including phonological mediation (Grainger et al., 2012), one may even expect that hyphens would facilitate bisyllabic and multisyllabic word reading for this group. Here we directly assessed this issue by investigating the role of hyphenation among more and less proficient early elementary school children.

Third, it can be hypothesized that hyphens, even though they may slow down word recognition, come to aid in word comprehension. That is, a stronger engagement of the phonological route – as stimulated by the hyphen – may lead to a deeper understanding of the words than recognition via a predominantly orthographic route, as connections between phonological and semantic representations are much stronger for children in their early stages of reading acquisition than connections between orthographic and semantic representations. Thus, we hypothesized that if hyphenated words are understood better during initial decoding than concatenated words, they can be better integrated in the unfolding sentence representation and the need to regress in the text and subsequently reread words is less urgent. In the present study, we made use of a specific eye movement measure (so called selective regression path duration) in order to investigate this issue.

Finally, since hyphens lengthen the spatial extent of words and thereby push word-external letters further away from the point of high visual acuity during initial fixation(s), it may well be that disruptive hyphenation effects are nothing more but visual acuity effects. In order to investigate this issue and make the hyphen condition more comparable to the control condition, we used smaller-sized hyphens in Experiment 2. In fact, the syllable-boundary hyphens used in Häikiö et al. (2014) and in Experiment 1 are longer than the ones typically used in schoolbooks (see Figure 2). This motivated the use of the almost dot-like schoolbook hyphens in Experiment 2.

Experiment 1

In Experiment 1 we investigated whether syllabification by hyphens affects word recognition speed of Finnish 2nd graders reading short bisyllabic words (kah-vi vs. kahvi) and long multisyllabic words (haa-ruk-ka vs. haarukka; ap-pel-sii-ni vs. appelsiini). To this end, the children read sentences that incorporated target words that were either concatenated or syllabified by hyphen, while their eye movements were registered. The children in Experiment 1 were in the 2nd month of the 2nd elementary school year. Each child's reading skill was assessed by the standardized reading test ALLU (Lindeman, 1998). We hypothesized that if early 2nd graders still restrict themselves to phonologically mediated syllable-by-syllable reading, syllabification by hyphens should speed up reading. If, on the other hand, early 2nd graders can make use of letter information across syllable boundaries and mainly engage orthographic pathways, hyphenation will slow down reading. We anticipated that the role of hyphenation is modulated by word length but also by reading skill. Less skilled readers and long multisyllabic words were expected to benefit more from visual marking of syllable boundaries than more skilled readers and short bisyllabic words. Finally, assuming that hyphenation leads to more profound word comprehension, we hypothesized that a measure indexing rereading of the target words would show an advantage for the hyphenated in comparison to the concatenated words.

Method

Participants

Twenty-three 2nd grade children (on average 8:4 years of age) were recruited from an elementary school in Lieto (South-West Finland). At the time of testing, they had received approx. 1 year and 2 months of reading instruction. All participants had normal or corrected-to-normal vision. Permission from the children's parents was acquired prior to the test. The participants received candy or stickers as reward for participation.

Prior to the experiment proper, the children's reading skill was assessed by a sentence comprehension test included in ALLU (Lindeman, 1998), a standardized Finnish reading test for elementary-school children. Children saw 20 pictures of scenes (e.g., people sailing) and had to choose from four sentence alternatives which sentence corresponded to the picture best (e.g., "They swim", "They dive", "They sail", and "They dance"). Children responded to as many picture-sentence combinations as possible within 120 seconds. The average score was 10, with scores ranging from 7 to 16. These values translate to normative levels ranging from 2 to 8 with an average level of 4 (levels 4-6 corresponding to average readers).

Apparatus

Eye movements were recorded monocularly with a remote table-mounted model of EyeLink 1000 (SR Research). The eyetracker is an infrared video-based tracking system with hyperacuity image processing; its spatial accuracy is 0.4 degrees. An infrared LED for illuminating the eye is positioned next to the eye movement camera. A chin-and-forehead rest was used to minimize head movements. The texts were presented on a 20-inch ViewSonic G225f computer screen (refresh rate 150 Hz).

Materials

Twenty-one employees at Turku University rated the age of acquisition (AoA) and familiarity of a list of potential target words. AoA was rated on a 7-point scale (1 = 0-2 years, 2 = 3-4 years, 3 = 5-6 years, 4 = 7-8 years, 5 = 9-10 years, 6 = 11-12 years, 7 > 12 years). Familiarity was rated on a 5-point scale (“How often do you encounter or use the word?”; 1 = very seldom, 2 = seldom, 3 = sometimes, 4 = often, 5 = very often). On the basis of these ratings, we selected 24 short bisyllabic and 24 multisyllabic words (trisyllabic + quadrosyllabic words). We assured that all the selected words were early acquired words (average AoA rating of 2.7, range 1.9-4.0). The bisyllabic words were matched with the multisyllabic words (see Table 1) on word frequency, bigram frequency, family size (i.e., the number of derivations and compounds starting with the target word), AoA, and familiarity. All values, apart from AoA and familiarity, were extracted from a newspaper corpus containing 22.7 million word forms (Laine & Virtanen, 1999).

INSERT TABLE 1 ABOUT HERE

The target words were embedded in sentences. Each short word was paired with a long word, and a sentence identical up to word N+1 was constructed for each pair. The target word never appeared in the beginning or end of the sentence. Five Turku University employees rated whether sentences containing long targets were equally natural as the sentences including short ones (sentences equally natural, long target word sentence more natural, short target word sentence more natural). Whenever one of the sentences was rated as more natural than the other, the sentence frames were replaced, until they were rated as equally natural by the 5 test subjects.

The experiment proper was preceded by six practice sentences and a block of 30 sentences from another experiment. The experimental materials included 72 sentences divided into two blocks, each of which contained 24 experimental and 12 filler sentences. Two versions of both blocks were created, one with all words hyphenated at syllable boundaries and another without hyphenation. Each participant saw one hyphenated and one non-hyphenated block. The block order was counterbalanced across participants. The order of sentences was pseudo-randomized within a block. Each participant saw each target word only once, either in hyphenated or concatenated format. An example sentence pair (with the target word in bold) is presented below.

Short - hyphenated: I-so-i-sä ker-toi, et-tä **aar-re** o-li pii-lo-tet-tu met-sään.

Short - concatenated: Isoisä kertoi, että **aarre** oli piilotettu metsään.

(Grandfather told the **treasure** was hidden in the forest.)

Long - hyphenated: I-so-i-sä ker-toi, et-tä **kynt-ti-lä** o-li pa-la-nut pöy-däl-lä ko-ko yön.

Long - concatenated: Isoisä kertoi, että **kynttilä** oli palanut pöydällä koko yön.

(Grandfather told the **candle** had burned on the table all night.)

The sentences were presented in Courier font with equal width for each character. With a viewing distance of 60 cm one character subtended approximately 0.3 degrees of visual angle. Sentences were all single-line sentences with a maximum of 85 characters per line presented vertically between the top and the middle of the screen.

Procedure

Prior to the experiment, the eye-tracker was calibrated using a three-point calibration grid extending horizontally over the computer screen. Before each sentence appeared on the screen, the participant fixated on a calibration point at the left side of the screen. When the participant fixated the calibration point, the experimenter pushed a button, which made the sentence appear to the right from the position where the calibration point was presented. Participants were instructed to read sentences silently for comprehension at their own pace. They were further told that after varying intervals they would get a statement (16 in total) about the content of the sentence they just read, for which they had to indicate whether it was true or false. Almost all participants responded to the statements correctly in at least 75% of the cases, but the four participants that scored the lowest in the ALLU test (4 participants with score 7) scored below 75% on the comprehension questions as well. For this reason, they were excluded from the initial analysis.

Dependent variables and predictors

We used four common eye movement measures as dependent variables, namely first fixation duration, gaze duration, the number of first-pass fixations (i.e., the number of fixations on the target word before exiting it for the first time) and selective regression path duration. The first three measures are first-pass reading measures, reflecting eye movement behaviour during the initial encounter of the word. Figure 3 provides a depiction of all four measures for the example word *girlfriend*. First fixation duration is used to assess initial word processing, in Figure 3 it is Fixation #3. Gaze duration is the summed duration of fixations on the target word before exiting it for the first time (in the example it is the summed duration of Fixations #3 and #4). Gaze duration informs about how quickly the whole word can be processed. Short and/or high-frequency words tend to be processed very rapidly; they often only require one fixation. In that case first fixation duration equals gaze duration. A relatively long word like *girlfriend* regularly requires more than one fixation, certainly when read by 8-year-old 2nd graders.

When words are more difficult to process and/or hard to integrate into the sentence, a reader may have to regress to an earlier word in order to reactivate the preceding context for extra support. In our example the reader regresses back to the word *lovely* (Fixation #5), before the word *girlfriend* is refixated (Fixation #6). The selective regression path duration is the summed duration of fixations on the target word after entering the target word and before exiting the target word to the right (in the example the selective regression path duration includes Fixations #3, #4, and #6). If the word is comprehended well during the initial encounter, it can be integrated into the unfolding sentence representation immediately and regressive fixations will not be needed. In this case selective regression path duration equals gaze duration. As noted in the Introduction, we expect that this will be more often the case for the hyphenated than for the concatenated word presentation format.

INSERT FIGURE 3 ABOUT HERE

The critical independent variables in the lme-analyses (see below for further details) were hyphenation (*Hyph*, hyphenated vs. concatenated) and number of syllables (*NSyl*, short bisyllabic vs. long multisyllabic words). Continuous variables included in the statistical models were log word frequency (*LFreq*) and age of acquisition (*AoA*). Word familiarity rating (*Fmlrty*) and family size (*FSize*) were also included in the models, but did not improve explanatory power in any of them[3]. Finally, we tested whether hyphenation interacted with the ALLU test score (*ALLU*, ranging from 8 to 16 after exclusion of the participants with a test score of 7). All the other interactions with hyphenation were also entered into the initial models.

Statistical considerations

The duration measures were left-skewed and therefore transformed by a power smaller than 1 as calculated by the Box-Cox system of the PowerTransform function in the library car of

R statistical software (R Development Core Team, 2012). Furthermore, values 2.5 SDs smaller or larger than the grand mean were excluded. Finally, the continuous predictor variables were centred. We used multiple regression mixed-effects modelling with participants and items as crossed random effects. Furthermore, other variables were added to the random effect structure when they improved the model's explanatory value as indicated by significantly higher values of the maximum log likelihood estimate of the model when compared to the model without the random effect (all ps $\leq .023$). All analyses were carried out using R (r Development Core Team, 2012) and the lme4 package, version 1.0.+ (Bates, Maechler, & Bolker, 2012). Specifications for all models include estimates of the regression coefficients. Standard errors and t-values were obtained with the t-test for fixed effects using the difference between the number of observations and the number of fixed effects as the upper bound for the degrees of freedom (see Baayen, 2008; Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000). In the following, an effect or interaction is significant for $|t| > 1.96$.

Results

Trials in which the target word was initially skipped were excluded before analyses (0.6% of trials). No further data trimming was done before the main analysis. The non-transformed means of each measure as a function of hyphenation and number of syllables are presented in Table 2. All the models of the main analyses are presented in Appendix 1.

INSERT TABLE 2 ABOUT HERE

First fixation duration. None of the main effects or interactions was significant.

Gaze duration. This measure showed main effects of *Hyph*, $t = -2.52$ and *AoA*, $t = -3.38$. Gaze durations were on average 55 ms longer for hyphenated than for concatenated words and late acquired words elicited longer gaze durations than early acquired words. None of the interactions with *Hyph* were significant, not even the interactions with *NSyl* or *ALLU*. However, there was a significant interaction between *LFreq* and *NSyl*, $t = 2.87$. Separate analyses for short and long words showed a frequency effect for long multisyllabic words, $t = 4.61$, with less frequent multisyllabic words eliciting longer gaze durations than more frequent ones, whereas there was no frequency effect for short bisyllabic words, $t = 0.30$.

Number of fixations. This measure revealed main effects of *Hyph*, $t = 3.38$, and *AoA*, $t = 2.47$. Hyphenated words elicited on average 0.34 more fixations than concatenated words and late acquired words elicited more fixations than early acquired words. None of the interactions with *Hyph* was significant, including the interactions with *NSyl* or *ALLU*. There was a significant interaction between *LFreq* and *NSyl*, $t < -2.98$. Separate analyses for short and long words showed a frequency effect for long multisyllabic words, $t = -3.05$, with less frequent multisyllabic words eliciting more fixations than more frequent ones, whereas no frequency effect was found for short bisyllabic words, $t = 0.56$.

Selective regression path duration. This measure revealed an interaction between *NSyl* and *LFreq*, $t = 2.87$, and a tendency for an interaction between *NSyl* and *Hyph*, $t = -1.89$. The model with the interactions included predicted more variance than the model with the main effects for *LFreq*, *Hyph* and *NSyl* only, $p = .06$. To follow up the interactions, we conducted separate analyses for bisyllabic and multisyllabic words. These analyses showed that the 23 ms hyphenation effect for short bisyllabic words was not significant, $t = -1.21$, whereas the 177 ms hyphenation effect for long multisyllabic words clearly was, $t = -3.17$[4]. In addition, the frequency effect turned out to be significant for long multisyllabic words, $t = 4.08$, but not for short bisyllabic words, $t = -0.17$.

Post-hoc analyses

The observed frequency effects were mainly restricted to long multisyllabic words. As the group of long words was somewhat heterogeneous in terms of word length and number of syllables, we wanted to assure whether the frequency effects were not word length/number of syllable effects in disguise. More specifically, the 24 long words that were selected included 16 trisyllabic words and 8 quadrosyllabic words (which were also on average one letter longer than the trisyllabic words). It turned out that the quadrosyllabic words elicited on average longer fixation times than trisyllabic words (gaze duration: quadrosyllabic 1003 ms vs. trisyllabic 766 ms; selective regression path duration: quadrosyllabic 1481 ms vs. trisyllabic 1032 ms). The quadrosyllabic words were at the lower end of the frequency distribution (quadrosyllabic: average frequency of 27 per million words; trisyllabic: average frequency of 146 per million words). In order to disentangle the effect of frequency and word length/number of syllables for the long multisyllabic words, we reanalyzed the data with *NSyl* as a 3-level factor (two syllables vs. three syllables vs. four syllables) instead of a two-level factor (bisyllabic vs. multisyllabic). These analyses eliminated all significant frequency effects, but showed significant differences between all levels of *NSyl* in gaze duration, number of fixations and selective regression path duration (*t*s > 3) with bisyllabic words eliciting shorter gaze and selective regression path durations and fewer fixations than trisyllabic words, which in turn were processed faster and with fewer fixations than quadrosyllabic words. In other words, the frequency effects in the main analyses turned out to be word length effects in disguise. The longer the words (or the more syllables there were in the word), the longer the fixation durations and the more numerous the fixations.

We also wanted to assess whether the disruptive role of hyphenation was modulated by reading skill. Therefore we reanalyzed the data by including the data of the four participants with the lowest ALLU scores and who had low comprehension scores in our experiment as well (lower than 75%). For gaze duration and number of fixations a significant interaction between *Hyph* and *ALLU* appeared, $|ts| > 2.7$. The interactions reflected that hyphens slowed down less proficient readers less than more proficient readers. The inclusion of the four least proficient readers reduced the disruptive effect of hyphen in gaze duration from 55 ms to 20 ms and in the number of fixations effect from 0.34 to 0.22. Gaze duration for the four least proficient readers was on average 126 ms shorter for hyphenated words than for concatenated words (H: 1346 ms; C: 1480 ms). However, the inclusion of the four participants did not alter the results for first fixation duration and selective regression path duration; for these measures no significant interactions between *Hyph* and *ALLU* could be observed $|ts| < 1$.

Discussion

The results of Experiment 1 showed that Finnish 2nd grade readers needed more fixations and longer gaze durations for hyphenated than for concatenated words; only in first fixation duration there was no hyphenation effect. These effects of hyphenation were observed in gaze duration and number of fixations for both short bisyllabic words (*kahvi* vs. *kah-vi)* and for long trisyllabic and quadrisyllabic words (*haarukka* vs. *haa-ruk-ka*, *appelsiini* vs. *ap-pel-sii-ni*). When also the least skilled readers were included in the analyses, hyphenation interacted with reading skill in both gaze duration and number of fixations. The results hinted at a beneficial effect of hyphenation for the less proficient 2nd graders in first-pass reading.

The results showed that hyphenation slows down word recognition speed for proficient Finnish 2nd graders. This is in line with the notion that for these children phonological pathways are less involved in word recognition. Instead they are more likely to make use of an orthographic pathway, either via syllables (Route 2b in Figure 1) or via a more direct route

(Route 3a or 3b). For less proficient 2nd graders, the phonological pathway (Route 2a) is a more likely option and intersyllabic hyphens seem to come to aid in using this pathway.

However, the results on selective regression path duration were not affected by reading skill. Children needed on average 180 ms more time to get past a long hyphenated word (e.g., *haa-ruk-ka*) than past its long concatenated counterpart (*haarukka*), no matter whether the least proficient readers were included or not. The 25 ms effect for short words was in the same direction, but did not reach significance in the separate analyses. In both instances, however, it does not seem to be the case that hyphenated words were comprehended better in the sense that they would have been easier to integrate in the unfolding sentence representation. Instead, one could say that for short words the slowing down by hyphenation during first-pass reading is not compensated by facilitated integrative processes and that for long hyphenated words additional second-pass reading is often required to understand them well enough to be integrated in the sentence before proceeding to the subsequent words.

Experiment 2

The insertion of hyphens causes the final letters of words to be further away from the area of sharp visual acuity when making the initial fixation on the word. This implies that the final letters of hyphenated words are visually more degraded during the initial fixation on the word than the final letters in concatenated words. This is likely to induce more effortful processing, for example in the form of making an additional fixation on the word. The visual acuity problem is exacerbated when using hyphens of equal width as the letters, as was the case in Experiment 1. In Experiment 2 we used hyphens that are approximately 1/3 of the width of letters, similarly to the dot-like hyphens used in schoolbooks. Three inserted hyphens correspond to the width of one letter, which significantly diminishes the potential visual acuity problem. In addition, smaller hyphens also make visually less salient segmentation cues. This may entail that for words with smaller-sized hyphens orthography-dominant processing including parallel syllable processing or whole-word processing becomes more viable. In addition to the change in visual manipulation, we recruited 1st graders who were able to read before entering school, but whose proficiency level was on average (naturally) lower than that of the 2nd graders of Experiment 1. Both smaller-sized hyphens and lower-proficiency readers made us expect that the disruptive effect of hyphenation would be attenuated.

Method

Participants

Sixteen 1st grade children (on average 7:9 years of age) were recruited from an elementary school in Turku (South-West Finland). At the time of testing, they had received approx. 6 months of reading instruction. All of the children were able to read when entering school, as assessed by the special education teacher at the beginning of school year using the LukiMat test (a test designed to evaluate the reading skills at the 1st grade, see Salmi, Eklund, Järvisalo, & Aro, 2011). All participants had normal or corrected-to-normal vision. Permission from the children's parents was acquired prior to the test. The participants received candy or stickers as reward for participation.

Prior to the experiment proper, the children's reading skill was assessed by a word recognition test included in ALLU (Lindeman, 1998). Children saw 80 pictures of objects (e.g., pencil = *kynä*) and had to choose which of the four words, all phonologically similar to the correct word (e.g., kylä = village, kyllä = yes, kynä = pencil, and kylmä = cold), corresponds to the picture. Children responded to as many picture-word combinations as possible within five minutes. The average score was 35, with scores ranging from 19 to 58. These values translate to normative levels ranging from 3 to 7 with an average level of 5 (levels 4-6 corresponding to average readers).

Apparatus

The apparatus was identical to that of Experiment 1.

Materials

The words were taken from a list of potential target words for which AoA and familiarity ratings had been acquired earlier. On the basis of these ratings, a total of 90 words was selected. We assured that all the selected words were early acquired (average AoA rating of 2.5, range 1.3-4.0). For each participant, 45 out of 90 words were used as target items and 45 as filler items. Of the 45 target words, 15 were hyphenated and 30 were concatenated. The hyphenated and concatenated target words were embedded in similar sentence frames. Of the 15 hyphenated words, 8 words were long multisyllabic words with 3 or 4 syllables and 7 were short, bisyllabic words. Of the 30 concatenated words, 16 were multisyllabic words with 3 or 4 syllables and 14 were bisyllabic words. The 4 conditions (long hyphenated, long concatenated, short hyphenated and short concatenated) were matched on word frequency, bigram frequency, family size, AoA, and familiarity (see Table 3). All values, apart from AoA and familiarity, were extracted from a newspaper corpus containing 22.7 million word forms (Laine & Virtanen, 1999). The target word never appeared in the beginning or end of the sentence. The experiment proper was preceded by six practice sentences. Altogether, the experiment included 90 sentences divided into two blocks of equal size. The order of sentences was randomized within a block. Each participant saw each target word only once, either in hyphenated or concatenated format. The sentences were presented in Calibri Light font with size 36. The hyphens were font size 18 with the baseline shift of 4 pixels upwards. With a viewing distance of 60 cm one character subtended from approximately 0.1 (a hyphen) to 0.25 degrees of visual angle. Sentences were all single-line sentences with a maximum of 67 characters per line presented between the top and the middle of the screen.

INSERT TABLE 3 ABOUT HERE

Procedure

The procedure was identical to that of Experiment 1 with the exception that there were 15 oral comprehension questions instead of 16 statements. All participants responded to the questions correctly in at least 87% of the cases.

Dependent variables, predictors, and statistical considerations

The dependent variables, predictors, and statistical considerations were identical to those in Experiment 1.

Results

Trials in which the target word was initially skipped were excluded before analyses (0.2% of trials). No further data trimming was done before the main analysis. The non-transformed means of each measure as a function of hyphenation and number of syllables are presented in Table 3. All the models of the main analyses are presented in Appendix 2. *ALLU* was significant in each measure, $|ts| > 4.1$. More proficient readers had longer first fixations, shorter gaze and selective regression path durations, and made less fixations than less proficient ones.

INSERT TABLE 4 ABOUT HERE

First fixation duration. There was a significant effect for *Nsyl*, $t = 2.60$. Longer words elicited shorter first fixation durations than short words. All other main effects and interactions did not approach significance.

Gaze duration. This measure showed an interaction between *Nsyl* and *Hyph*, $t = 2.33$. The separate analyses showed that there was – despite the numerical difference in favour of hyphenated words – no significant effect of hyphenation for short bisyllabic words, $t < 1$. However, for long multisyllabic words hyphenation elicited significantly longer (143 ms) gaze durations, $t = 3.00$. Furthermore, there were significant main effects of AoA and Fmlrty, $|ts| > 2$. Early acquired and familiar words elicited shorter gaze durations than later acquired and less familiar words. None of the other main effects and interactions was significant.

Number of fixations. This measure revealed main effects of *Nsyl* and *AoA*, $ts > 3.6$. Multisyllabic words elicited more fixations than bisyllabic words, and late acquired words elicited more fixations than early acquired words. None of the interactions with *Hyph* were significant.

Selective regression path duration. This measure showed a significant interaction between *Nsyl* and *Hyph*, $t = 2.21$. The interaction reflected the finding that there was – as in gaze duration - no significant hyphenation effect for short bisyllabic words, $t < 1$, but a significant effect for long multisyllabic words (122 ms longer selective regression path durations for hyphenated than concatenated words, $t = 2.63$). Furthermore, there were significant main effects of *AoA* and *Fmlrty*, $|ts| > 2.5$. Selective regression path durations were longer for late acquired and less familiar words than early acquired and familiar words. There were no other main effects or interactions.

Discussion

The pattern of results was somewhat different from that of Experiment 1. In particular, unlike for the 2nd graders in Experiment 1, hyphenation did not lengthen 1st graders' gaze durations on bisyllabic words. Similarly to Experiment 1, hyphenation did not increase selective regression path durations for bisyllabic words. However, hyphenation did not speed up word processing either. For long words, the observed effects were similar in Experiment 1 and 2. In both experiments, long hyphenated words elicited longer gaze and selective regression path durations than long concatenated words, even though for the 1st graders the hyphens were smaller in size. In general, the results of Experiment 2 testify that the hyphenation effects cannot solely be ascribed to visual acuity, as the dot-like hyphens only minimally increased the spatial extent of the words.

General Discussion

The results of the present study clearly demonstrate that hyphenation at syllable boundaries slows down word recognition of early Finnish readers already mastering phonological decoding. It should be noted though that the level of disruption depended on readers' proficiency: 1st graders were less disrupted by hyphenation than 2nd graders, and there was some indication that less proficient 2nd graders may even benefit from hyphenation during the first-pass reading of words. The level of disruption also depended on the number of syllables. For short bisyllabic words there was a disruption effect only in gaze duration for 2nd graders, whereas for long multisyllabic words a disruption effect was obtained both in gaze duration and selective regression path duration for both 1st and 2nd graders. Finally, the size of the hyphen only partly affected the hyphenation effect. An effect observed in gaze duration for short words in Experiment 1 (longer hyphens) did not appear in Experiment 2 (shorter hyphens). This may have been caused by the smaller-sized hyphens but also by the lower level of reading proficiency of the participants of Experiment 2 in comparison to those in Experiment 1. Most importantly, the smaller-sized hyphens in Experiment 2 did not eliminate the disruption effect of hyphenation for long multisyllabic words, as even smaller-sized hyphens slowed down reading of long words with more than 100 ms in gaze duration and selective regression path duration.

The present results are in line with those of Häikiö et al. (2014) by showing that intersyllabic hyphens enforce a processing strategy upon children that is suboptimal in terms of word recognition speed. The more proficient children are in reading, the larger the disruption effect of hyphenation. Following Häikiö et al., we argue that the disruption is associated with the hyphen restricting the initial attention to the first syllable. Subsequently, the letters constituting the syllable will be used to access the phonological representation of the syllable. After this the subsequent syllable(s) will be accessed in the same manner, after which the syllables will be assembled to arrive at the phonology of the whole word. The more syllables there are, the more demanding the assembly process will be, as keeping initial syllable(s) active in working memory while simultaneously accessing subsequent syllables and eventually integrating all the syllables will require cognitive resources. This would explain the correlation between the size of the hyphenation effect and the number of syllables. For less proficient readers the above described decoding process will take place anyway, regardless whether there are hyphens inserted at syllable boundaries or not; in this case hyphens may be even beneficial, as they clearly signal syllable structure.

The question remains how more proficient readers process bisyllabic and multisyllabic words without hyphens at syllable boundaries. Häikiö et al. (2014) suggest that they tend to recognize words via an orthographic route by which syllables are identified in parallel (rather than sequentially). Subsequently, the accessed syllables would address the whole orthographic word form without mediation of phonology (Route 2b in Figure 1). Alternatively, words can be identified via a holistic orthographic route (Route 3a in Figure 1) or via coarse-grained orthography (cf. Grainger et al, 2012; Route 3b in Figure 1). The number-of-syllables effect in combination with the lack of whole-word frequency effects may be taken to indicate that the early readers of the present study prefer syllable-based access. However, with respect to the lack of frequency effect, it should be noted that the frequency range from which the words were chosen was restricted. All target words were familiar to the children and were at least reasonably frequent. This reduced range may have obscured possible frequency effects. Moreover, we did observe whole-word-related effects. We found an effect for age of acquisition in gaze duration and in the number of fixations in both experiments as well as in selective regression path duration in Experiment 2. Similar age-of-acquisition (and also word familiarity) effects were found by Häikiö et al. (2014), suggesting that early Finnish readers can also access bisyllabic and multisyllabic words via a more direct route.

Finally, we hypothesized in the Introduction that intersyllabic hyphens may lead to a better level of comprehension and may make it easier to integrate words in the sentence. However, the results of selective regression path duration indexing integrative processing indicate that, if anything, hyphenation complicates the construction of the sentence representation. More specifically, in case of multisyllabic words, children regressed in the text more and reread words far more often in case of hyphenated than concatenated words.

Implications for reading instruction

In Finnish ABC books hyphens are inserted at syllable boundaries to signal the syllable structure of words with more than one syllable. Hyphenation is preserved for bisyllabic words until the end of the 1st grade and for long multisyllabic words until the end of the 2nd grade. This practice follows the KÄTS method, which is based on decoding syllables and assembling the whole word pronunciation from the individual syllables. Our results should not be taken to imply that the KÄTS-method does not facilitate initial reading development or reading accuracy. It is possible that syllabification by hyphens helps early readers in the initial stages (and struggling readers for a longer period of time) to phonologically recode the word and to arrive at the full phonological representation more swiftly. To establish the general usefulness of hyphenation more firmly, a controlled experiment where beginning readers would be subjected to reading instruction with or without hyphens would be needed. However, the present results do imply that after the initial stages of reading instruction have been completed, hyphens at syllable boundaries slow down reading of normally developing children, especially when inserted in long multisyllabic words. Yet, this slowdown may be considered a positive rather than a negative consequence of hyphenation. That is, slower reading pace and more sequential syllable processing could make children more aware of how words are to be written. Since spelling errors increase with increasing word length, this would be especially important for long words[5]. Thus, the substantial disruption effect of hyphenation for long words may actually serve the beginning readers well, as it may have a positive impact on spelling development, especially given the agglutinative nature of the Finnish language[6] (M-K. Lerkkanen, personal communication, March 30, 2010). These arguments are appealing although at present there is – as far as we know – no empirical evidence to support it.

Lerkkanen (2006) notes, however, that the KÄTS method may be too repetitive for normally developing children and that too heavy reliance on this method may delay the development of reading comprehension skills. Hyphens may for instance prevent or delay the use of other, more subtle segmentation cues like bigram troughs or they may delay the development of orthographic routes and stable orthographic whole-word representations (Häikiö et al., 2014). Therefore, it is important for reading development that teachers also frequently expose children to texts without intersyllabic hyphens. This is already a common practice among elementary-school teachers in Finland. Another practical implication for early Finnish reading instruction that may be drawn from the present results is to limit syllabification by hyphens in ABC books to a shorter period of time. Especially in case of long multisyllabic words hyphens may be eliminated at an earlier stage or replaced by visually less prominent and intrusive cues (e.g., alternate syllable coloring, see Häikiö et al., 2014), at least for pupils following the normal trajectory in reading development.

Conclusions

Generally speaking, it can be concluded that the present results are in line with the multiple-route model of word recognition (Grainger et al., 2012), which incorporates the idea that words can be recognized via different pathways that are at work simultaneously in order to retrieve word meanings as quickly as possible. Grainger et al (2012) showed that in the course of reading development the balance shifts from recognizing words via small units (letters) and via phonological mediation to more direct orthographic access routes involving multi-letter clusters. Our results indicate that early Finnish 1st and 2nd graders use either syllable-based (Route 2b in Figure 1) or more direct (Route3a or 3b) orthographic pathways. How much exactly each route contributes to word decoding depends on factors such as reading skill, word length and word frequency. At the practical level, it can be concluded that in terms of hyphenation the present findings indicate that reading materials should be individualized after the initial stages of reading acquisition, a practice that is widely used in Finland already, but could still be extended.

Footnotes

1. In support of this statement, only 0.1% of the 70,000 words in the reverse dictionary of modern standard Finnish are monosyllabic (Tuomi, 1980).
2. A similar slowing down of hyphenation was found by Häikiö, Bertram, and Hyönä (2011) at the morpheme level. They found that compounds with a hyphen at the constituent boundary (e.g., *ulko-ovi* = ‘front door’) elicited longer reading times for proficient 2nd graders, 4th graders and 6th graders than concatenated compounds (e.g., *sivuovi* = ‘side door’).
3. Both of these variables had high correlations with word frequency, so we used their residuals in the models.
4. Note that exactly the same pattern of results was found for the go-past time, which is the sum of all fixations after entering the target word before exiting it to the right (in Figure 2 the summed duration of Fixations #3 to #6). For this measure the hyphenation effect was 36 ms for short words and 194 ms for long words.
5. We would like to thank an anonymous reviewer for pointing this out.
6. Furthermore, the present results relate to reading of monomorphemic, uninflected words. Since, for example, any Finnish noun can appear in approximately 2,000 possible orthographic forms (Karlsson & Koskenniemi, 1985), most of which have never been encountered before, hyphenation may support the decoding of these typically long morphologically complex words in the early stages of reading development.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge, UK: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. doi:10.1016/j.jml.2007.12.005

Bates, D., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes.* Retrieved from http://CRAN.R-project.org/package=lme4 (R package version 1.0+)

Ehri, L. C. (1987). Learning to read and spell words. *Journal of Reading Behavior,* 19, 5-31. doi:10.1080/10862968709547585

Ehri, L. C. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Reading Disabilities, 22*, 356-365. doi:10.1177/002221948902200606

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

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*, 167-188. doi:10.1207/s1532799xssr0902_4

Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly, 14*, 135-163. doi:10.1080/1057356980140202

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 301–330). London, UK: Erlbaum.

Grainger, J., Lété, B, Bertand, D., Sufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition, 123*, 280-292. doi:10.1016/j.cognition.2012.01.003

Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology, 2*:54. doi:10.3389/fpsyg.2011.00054

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*, 1430-1440. doi:10.1037/0278-7393.21.6.1430

Häikiö, T., Bertram, R., & Hyönä, J. (2011). The development of whole-word representations in compound word processing: Evidence from eye fixation patterns of elementary school children. *Applied Psycholinguistics, 32*, 533-551. doi:10.1017/S0142716411000208

Häikiö, T., Hyönä, J., & Bertram, R. (2014). The role of syllables in word recognition among beginning Finnish readers: Evidence from eye movements during reading. *Journal of Cognitive Psychology*. doi:10.1080/20445911.2014.982126 [Manuscript accepted for publication.]

Karlsson, F., & Koskenniemi, K. (1985). A process model of morphology and lexicon. *Folia Linguistica*, 29, 207-231. doi:10.1515/flin.1985.19.1-2.207. Karppi, S. (1983). *Lukutaidon ABC. Johdatus lukemisen ja kirjoittamisen perustekniikan opetukseen. [The ABC of reading skill. Introduction to teaching the basic techniques of reading and writing.]* Espoo, Finland: Weilin+Göös.

Laine, M., & Virtanen, P. (1999). *WordMill lexical* search *program*. Turku, Finland: University of Turku, Center for Cognitive Neuroscience.

Lerkkanen, M.-K. (2006). *Lukemaan oppiminen ja opettaminen esi- ja alkuopetuksessa. [Learning and teaching reading in preschool and early education.]* Helsinki, Finland: WSOY Oppimateriaalit.

Lindeman, J. (1998). *ALLU: Ala-asteen Lukutesti [ALLU: Reading test for elementary school]*. Turku, Finland: University of Turku, Center for Learning Research.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-PLUS.* New York, NY: Springer-Verlag.

R Development Core Team (2012). *R: A language and environment for statistical computing*. Vienna, Austria; http://www.R-project.org

Salmi, P., Eklund, K., Järvisalo, E., & Aro, M. (2011). *LukiMat - Oppimisen arviointi: Lukemisen ja kirjoittamisen tuen tarpeen tunnistamisen välineet 1. luokalle. [LukiMat – Evaluating learning: Tools for recognizing the need of support in reading and writing for the 1st grade.]* Jyväskylä, Finland: Niilo Mäki Instituutti.

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

Tuomi, T. (1980). *Suomen kielen käänteissanakirja* [Reverse dictionary of modern standard Finnish] (2nd rev. ed.). Hämeenlinna, Finland: Finnish Literature Society.

Wäre, M., Lerkkanen, M.-K., Suonranta-Hollo, L., Korolainen, T., Parkkinen, J., Kirkkopelto, K., & Ketonen R. (2014). *Pikkumetsän aapinen* [ABC book of Tiny Forest]. Helsinki, Finland: Sanoma Pro.

Table 1

Lexical Statistics of the Target Words in Experiment 1.

	Short	Long
Frequency[a]	75	107
Length in characters	4.92	8.33
Length in syllables	2.00	3.33
Bigram frequency[b]	7.48	7.85
Family size	110	129
Age of Acquisition[c]	2.60	2.75
Familiarity[d]	2.94	3.18

Note. [a] Per million; [b] Per thousand; [c] On scale 1-7;[d] On scale 1-5

Table 2

Means for Each Eye Movement Measure in Experiment 1, as a Function of Hyphenation and Word Length (Standard Deviations in Parentheses).

	Long words		Short words	
	Hyphenated	Concatenated	Hyphenated	Concatenated
First Fixation (ms)	282 (115)	294 (124)	288 (128)	299 (125)
Gaze Duration (ms)	886 (687)	803 (648)	641 (537)	614 (580)
Number of Fixations	3.20 (1.92)	2.83 (1.92)	2.22 (1.41)	1.91 (1.33)
Selective Regression Path Duration (ms)	1070 (786)	893 (698)	704 (568)	679 (632)

Table 3

Lexical Statistics of the Target Words in Experiment 2.

	Short	Long
Frequency[a]	116	55
Length in characters	5.00	7.46
Length in syllables	2.00	3.13
Bigram frequency[b]	7.79	7.82
Family size	114	97
Age of Acquisition[c]	2.42	2.52
Familiarity[d]	2.95	3.25

Note. [a] Per million; [b] Per thousand; [c] On scale 1-7;[d] On scale 1-5

Table 4

Means for Each Eye Movement Measure in Experiment 2, as a Function of Hyphenation and Word Length (Standard Deviations in Parentheses).

	Long words		Short words	
	Hyphenated	Concatenated	Hyphenated	Concatenated
First Fixation (ms)	288 (165)	302 (160)	334 (186)	332 (191)
Gaze Duration (ms)	1138 (805)	995 (821)	733 (542)	794 (655)
Number of Fixations	3.80 (2.08)	3.23 (1.97)	2.32 (1.50)	2.44 (1.48)
Selective Regression Path Duration (ms)	1191 (809)	1069 (899)	779 (567)	842 (684)

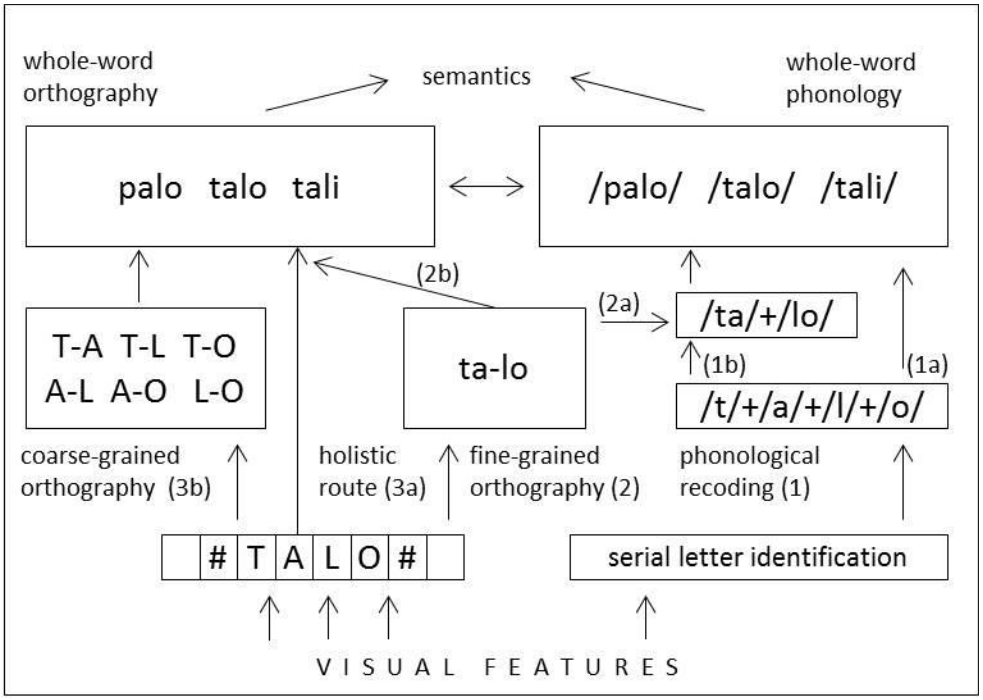
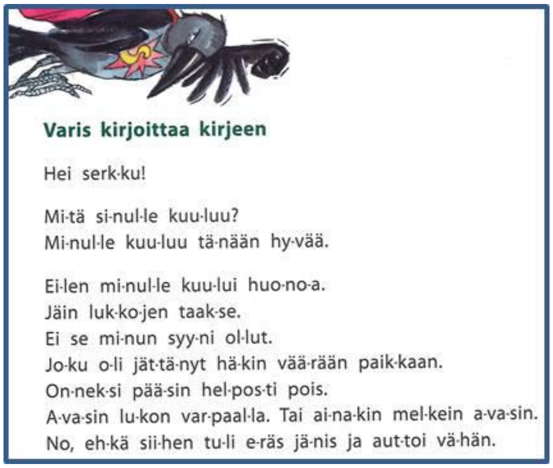


Figure 1. Model of Grainger and Ziegler (2011) adapted to elucidate the reading development of Finnish children when reading polysyllabic words like talo (=house). Adapted with permission from Figure 6 of Grainger and Ziegler (2011, p. 8).



Varis kirjoittaa kirjeen

Hei serk·ku!

Mi·tä si·nul·le kuu·luu?
Mi·nul·le kuu·luu tä·nään hy·vää.

Ei·len mi·nul·le kuu·lui huo·no·a.
Jäin luk·ko·jen taak·se.
Ei se mi·nun syy·ni ol·lut.
Jo·ku o·li jät·tä·nyt hä·kin vää·rään paik·kaan.
On·nek·si pää·sin hel·pos·ti pois.
A·va·sin lu·kon var·paal·la. Tai ai·na·kin mel·kein a·va·sin.
No, eh·kä sii·hen tu·li e·räs jä·nis ja aut·toi vä·hän.

Figure 2. An excerpt from a first-year ABC book (Wäre et al., 2014). Hyphens are inserted at every syllable boundary. Only 15% of the text is monosyllabic. (Reprinted with permission.)

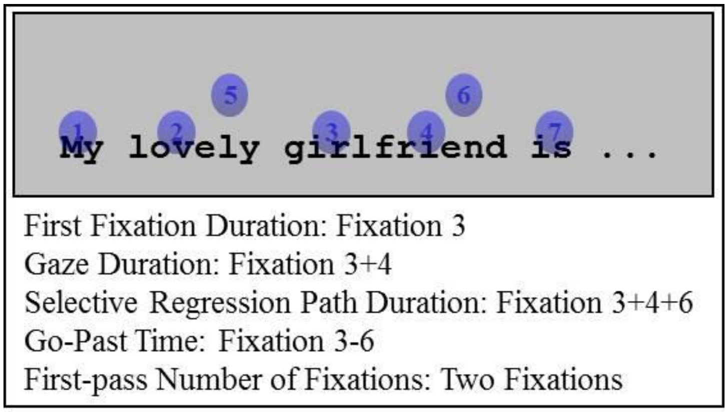


Figure 3. Eye movements of a hypothetical reader reading the phrase ‘My lovely girlfriend is…’; below the phrase the eye movement measures used in the analyses of the current experiment.

Appendix 1

Final models for each measure in the Experiment 1 with fixed effects presented for each measure separately. If an interaction was significant, its main effects are reported as well. Please note that in this case the main effects are not independently interpretable in the lmer() output. Also note that due to the transformations the directions of the effects are reversed. An effect is significant when $|t| > 2$.

(Insert Tables A1-A4 about here.)

Appendix 2

Final models for each measure in the Experiment 2a with fixed effects presented for each measure separately. If an interaction was significant, its main effects are reported as well. Please note that in this case the main effects are not independently interpretable in the lmer() output. Also note that due to the transformations the directions of the effects are reversed for First Fixation Duration. An effect is significant when $|t| > 2$.

(Insert Tables A5-A8 about here.)

Table A1.

First Fixation Duration: Fixed Effects

	Estimate	*SE*	*t*-value
(Intercept)	0.7516	0.0011	682.0
cLFreq	0.0001	0.0006	1.70

Table A2.

Gaze Duration: Fixed Effects

	Estimate	*SE*	*t*-value
(Intercept)	0.5206	0.0054	95.65
HyphPresent	-0.0064	0.0025	-2.52
NSylLong	-0.0148	0.0023	-6.41
cLFreq	-0.0002	0.0023	-0.1
cAoA	-0.0087	0.0026	-3.38
cAllu	0.0079	0.0021	3.75
NSylLong:cLFreq	0.0087	0.0030	2.87

Table A3.

Number of Fixations: Fixed Effects

	Estimate	*SE*	*t*-value
(Intercept)	1.6265	0.3879	4.19
HyphPresent	0.1742	0.0518	3.38
NSylLong	0.3521	0.0510	6.91
cLFreq	0.0420	0.0491	0.86
cAoA	0.1226	0.0497	2.47
cAllu	-0.1003	0.0359	-2.79
NSylLong:cLFreq	-0.1805	0.0606	-2.98

Table A4.

Selective Regression Path Duration: Fixed Effects

	Estimate	*SE*	*t*-value
(Intercept)	0.6244	0.0049	128.27
HyphPresent	-0.0027	0.0028	-0.94
NSylLong	-0.0127	0.0028	-4.47
cLFreq	-0.0003	0.0019	-0.15
cAllu	0.0093	0.0021	4.47
NSylLong:HyphPresent	-0.0070	0.0037	-1.89
NSylLong:cLFreq	0.0068	0.0024	2.87

Table A5.

First Fixation Duration: Fixed Effects

	Estimate	*SE*	*t*-value
(Intercept)	0.2541	0.0028	89.39
NSylLong	0.0054	0.0021	2.60
cAllu	0.0007	0.0002	4.02

Table A6.

Gaze Duration: Fixed Effects

	Estimate	*SE*	*t*-value
(Intercept)	1.775	0.0168	105.88
HyphPresent	-0.0050	0.0080	-0.62
NSylLong	0.0353	0.0089	3.97
cAllu	-0.0053	0.0010	-5.01
cAoA	0.0227	0.0068	3.33
cFmlrty	-0.0108	0.0052	-2.09
HyphPresent:NSylLong	0.0256	0.0110	2.33

Table A7.

Number of Fixations: Fixed Effects

	Estimate	*SE*	*t*-value
(Intercept)	1.0070	0.0845	11.92
NSylLong	0.3391	0.0529	6.41
cAllu	-0.0253	0.0051	-4.94
cAoA	0.1541	0.0423	3.64

Table A8.

Selective Regression Path Duration: Fixed Effects

	Estimate	*SE*	*t*-value
(Intercept)	1.7150	0.0156	109.64
HyphPresent	-0.0038	0.0064	-0.59
NSylLong	0.0354	0.0073	4.87
cAllu	-0.0048	0.0010	-4.81
cAoA	0.0149	0.0055	2.72
cFmlrty	-0.0105	0.0041	-2.55
HyphPresent:NSylLong	0.0194	0.0088	2.21