Word-recognition processes in normal and dyslexic readers

Marinus, E.

Citation for published version (APA):
Chapter 4

Variability in the word-reading performance of dyslexic readers: Effects of letter length, phoneme length and digraph presence

Abstract

The marked word-length effect in dyslexic children suggests the use of a letter-by-letter reading strategy. Such a strategy should make it more difficult to infer the sound of digraphs. Our main aim was to disentangle length and digraph-presence effects in word and pseudoword reading. In addition, we examined differences in intra-individual variability between dyslexic and normal readers. Naming tasks were administered to 24 dyslexic readers individually matched to chronological-age and reading-age controls. As expected, dyslexic and younger children showed stronger length effects. In contrast to our expectations, the dyslexic and younger children were faster in reading (pseudo)words with a digraph than in reading (pseudo)word of similar letter length but without a digraph. Normal readers were equally fast on both types of (pseudo)words. However, considering phoneme-length effects, digraph-presence caused an additional delay in all reading groups, but only for pseudowords. In addition, this effect was stronger for the dyslexic and younger readers. Finally, dyslexic readers’ intra-individual variability in reading was larger than the variability in normal readers, especially for short words.

CHAPTER 4. VARIABILITY IN WORD-READING PERFORMANCE

4.1 Introduction

Earlier studies have demonstrated important differences in the word-recognition processes of children with and without dyslexia. Obviously, children with dyslexia are much slower and make more errors than their normal reading counterparts (Vellutino, Fletcher, Snowling, & Scanlon, 2004). However, more specific differences in word-recognition processes have also been observed. The most noted difference is the stronger lexicality effect that is commonly found in dyslexic children. The lexicality effect refers to the finding that words are generally read faster and more accurately than comparable pseudowords. Interestingly, it has been repeatedly found that this effect is more pronounced in children with dyslexia (Marinus & de Jong, 2008; Martens & de Jong, 2006; Rack, Snowling, & Olson, 1992; Ziegler et al., 2003; Zoccolotti et al., 1999; Barca, Burani, di Filippo & Zoccolotti, 2006, but see di Filippo, de Luca, Judica, Spinelli, & Zoccolotti, 2006; Zoccolotti, de Luca, Judica, & Spinelli, 2008). The difficulty that dyslexic children experience in the reading of pseudowords has typically been interpreted as a marker of a phonological deficit (Rack et al., 1992).

Another notable difference between normal reading and dyslexic children is their sensitivity to the length of both words and pseudowords. Several studies found that dyslexic children respond relatively slower to longer words and pseudowords than to shorter ones. (Martens & de Jong, 2006; Ziegler et al., 2003; Zoccolotti et al., 2005). The stronger length effect in dyslexic children is often hypothesized to reflect the dyslexic readers’ impairment to apply lexical reading strategies (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) or “use of orthographic knowledge” to process more or even all the letters of a target word in parallel. Instead, dyslexic readers are thought to persist in using a more laborious serial, letter-by-letter, decoding strategy (Zoccolotti et al., 2005). Evidence for dyslexic children’s persistent use of a letter-by-letter reading strategy is also provided by studies in which dyslexic children were found to make more eye-movements per word than normal readers (de Luca, Borelli, Judica, Spinelli, & Zoccolotti, 2002; Hutzler & Wimmer, 2004).

The purpose of the current study is twofold. Firstly, the nature of the word-length effect in dyslexic and normal reading children is studied in more detail. More specifically, we examined whether the presence of a digraph causes an additional delay on top of the letter-length effect. Digraphs are two-letter graphemes, such as, the **ou** in *soup*, and are a common phenomenon in Germanic languages like English, German and Dutch. Up until now most studies examining the length effect in normal reading and dyslexic children, have not controlled for the presence of digraphs (Martens & de Jong, 2006; Ziegler et al., 2003). As a result it is difficult to conclude that the stronger length effect in younger and dyslexic children reflects a pure letter-by-letter reading strategy. It might as well reflect a grapheme-by-grapheme reading strategy in which the processing of a digraph
causes an additional delay. Disentangling length and digraph effects is also of theoretical relevance as theories of visual word recognition differ in their implementation of how digraphs are processed during reading. Theories and research regarding length effects and the processing of digraphs will be discussed in the next section of the introduction.

The second aim of this study is to examine and explicitly model the variability in the naming latencies of normal reading, dyslexic and younger readers. Thus far, most studies have been concerned with mean differences in reading speed and accuracy between normal and dyslexic readers. However, everyone who has ever examined word recognition skills of normal readers and dyslexics will agree that the reading performance of dyslexics is much more variable than that of normal readers. Besides that the word reading of dyslexics is more variable in very easy tasks, like for instance three letter word naming (Marinus & de Jong, 2008; Martens & de Jong, 2006; Ziegler et al., 2003), its variability also seems to increase more in response to word characteristics (such as lexicality (pseudowords or words), orthographic complexity and word length). Why and how to model the higher variability of dyslexic children will be discussed in more detail after we have elaborated on word length and digraph effects.

4.1.1 Word length and digraph effects

It is often assumed that length effects reflect the serial processing of words and pseudowords (Zoccolotti et al., 2005). Length effects are found in both children and adult readers in naming tasks and lexical decision (Balota, Cortese, Sergent-Marshall, & Spieler, 2004; Martens & de Jong, 2006). Serial processing is one of the core assumptions of the Dual Route Cascaded (DRC) Model (Coltheart et al., 2001) and is also incorporated in the Connectionist Dual Route Processing (CDP+) Model (Perry, Ziegler, & Zorzi, 2007). According to these models every letter string enters the reading system letter-by-letter, starting from the first. Shortly after the serial processing of letters has started, the lexical route (or lexical network) is activated. The lexical route processes letter strings in parallel and is able to map the orthographic input directly to a code in the phonological lexicon that directly activates the pronunciation of whole words. As a result, this route considerably speeds up the word recognition process. It follows from the DRC and CDP+ models that the length effect will be larger for pseudowords than for words. A lot of studies indeed found this lexicality by length interaction effect in both naming and lexical decision tasks (Balota et al., 2004; Weekes, 1997, Spinelli et al, 2005).

Whereas the DRC and CDP+ postulate that the length effect results from the serial processing of the letters of the stimuli some connectionist models (Seidenberg & McClelland, 1989) assume that the effect of word length is a mere consequence of neighbourhood effects. In general shorter words and pseudowords
have a larger neighbourhood size than longer ones, which might form an alternative explanation for the finding that shorter words are read faster than their longer counterparts. However, this is not in line with the recent findings of Balota et al. (2004) that length and neighbourhood seem to have an independent influence on (pseudo) word recognition speed in adult readers.

Most alphabetic orthographies consist of more phonemes than there are letters to denote the phonemes. Therefore, languages often use two or more letters to express a single phoneme (Borgwaldt, Hellwig, & de Groot, 2004). The term “digraph” refers to a two-letter grapheme. Digraphs are an interesting phenomenon in the context of the length effect because it is not possible for a reader to tackle the digraph ambiguity by using a mere serial letter-by-letter reading strategy. According to the DRC Model (Coltheart et al., 2001), digraphs are not processed as perceptual units during the first level of processing. The nonlexical route of the model initially activates an incorrect phoneme for the first letter of the digraph. When the second letter of the digraph comes into view the correct phoneme, corresponding to the digraph, begins to be activated. However, initially the wrong phoneme has been activated. As a result more time is needed to reach the threshold for the pronunciation of the correct digraph phoneme. Therefore, the DRC model predicts that, given an equal number of letters, it will take longer to read words with digraphs than words without (Jackson & Coltheart, 2001b).

Rastle and Coltheart (1998) investigated the role of digraphs in pseudoword reading by asking adult readers of English to read pseudowords consisting of five letters. Half of the pseudowords did not have a digraph (e.g., *breps*), the other half of the words contained two digraphs (e.g., *doaph*). Rastle and Coltheart found that the adult readers were faster on the pseudowords without than on the pseudowords with digraphs. The finding that five letter pseudowords with two digraphs were read slower than five letter pseudowords without digraphs was hypothesized to be a result of the delay that is caused by the letter-by-letter processing of the two digraphs within the pseudowords. Rastle and Coltheart denoted this delay, due to a digraph, as the “whammy effect”. The effect has also been found for words (Rey, Jacobs, Schmidt-Weigand, & Ziegler, 1998). However, just like in the pseudowords of Rastle and Coltheart, a significant whammy effect was only found when five letter words with two digraphs were contrasted with five letter words without digraphs. One digraph did increase the response latencies, however it did not cause enough delay to yield a significant difference.

In the CDP+ Model (Perry et al., 2007), the letters of a (pseudo)word also enter the system one by one, but in contrast to the DRC Model, the letters are visually parsed and segmented into graphemes before they are mapped onto the corresponding phonemes. The contention that graphemes and not letters are the perceptual reading units is supported by the results of letter detection experiments (Rey, Ziegler, & Jacobs, 2000). The reasoning behind these experiments is that if the reading system processes digraphs as a reading unit, then it should be harder
to detect a target letter when it is embedded in a digraph (e.g., the \( a \) in peach) than when it corresponds to a single letter (e.g., the \( a \) in class). Rey et al. indeed found that letter detection latencies were larger if the target letter was embedded within a digraph. Other support that the letters of digraphs are processed in parallel before the mapping onto the phoneme occurs, comes from a phoneme detection study of Peereman, Brand and Rey (2006). Peereman et al. argued that if the sound of the first letter of a digraph is activated before the parsing of the digraph, that it should be easier to make decisions about the presence of the phoneme of the first than about the presence of the phoneme of the second letter of the digraph. Because, according to the DRC model, the phoneme of the second letter will not be activated at all. In contrast, however, Peereman et al. found that it was equally difficult to decide that the target phoneme was not present when it matched the phonological code of the first letter of the digraph as when it matched the phonological code of the second letter. This finding also contradicts the assumption that the sound of the first letter is activated before the digraph is parsed.

To conclude, the DRC and CDP+ models differ in their assumptions about the initial processing of digraphs during reading. According to the DRC model the separate letters will enter the system one-by-one and hence the presence of a digraph will cause a delay in the word recognition process. Empirical data for this are provided by pseudoword and word naming experiments in which skilled adult readers have to read five letter (pseudo)words with at least two digraphs (Rastle & Coltheart, 1998). In contrast the CDP+ model features a graphemic buffer in which digraphs are visually parsed and directly mapped onto their corresponding sound without further delays. Indirect support for this contention is found in letter and phoneme detection experiments (Rey et al., 2000, Peereman et al., 2006).

Although the DRC and CDP+ models are developed to represent skilled adult reading, they are also useful as a framework for making predictions about the processing of digraphs in normal reading, dyslexic and younger reading children. We reasoned that the serial letter-by-letter decoding strategies of dyslexic and younger readers might lead to difficulties in the processing of digraphs. We predict that this effect will cause an additional delay or an “enlarged whammy effect”, as described in the DRC model, when children, especially younger and dyslexic readers, read words and pseudowords with digraphs. If this is not the case, the conclusion might be that digraphs are visually parsed in an earlier stage, as described by the CDP+ model. Theories on reading development, like Ehri’s phase theory (1992; 1998) and Elbro’s reading acquisition model (2005) agree that complex associations like digraphs are learned later than reliable grapheme-phoneme associations. However, both theories do not describe how children process digraphs during word recognition and whether there is development in how digraphs are processed. It is not surprising that developmental theories do not make more specific predictions, because research investigating the processing of digraphs in children, especially in dyslexic children, is scarce.
4.1.2 Variability in reading

The influence of word characteristics, such as lexicality and word length, on the word recognition of normal reading and dyslexic children is usually investigated in terms of mean differences in accuracy and reading speed between these groups. However, especially with respect to speed, the variance of the dyslexics is usually larger, even in the reading of relatively simple three-letter words, and tends to increases even more in response to word characteristics such as word length and lexicality (e.g., Marinus & de Jong, 2008; Martens & de Jong, 2006, Ziegler et al., 2003). One reason for a larger variability among dyslexic readers is probably due to the selection procedure. A cut-off criterion, for example at least one standard deviation below the mean, is often used to select the dyslexic children, whereas the performance of the control groups on reading selection tasks typically falls within a restricted ability range encompassing often minus one to plus one standard deviation from the mean. Thus, the performance range is less restricted for the dyslexic children and, as a result, there can be more variability among children in a dyslexic sample than among children in a control sample. This kind of variability is commonly referred to as between subjects or inter-individual variability. However, the focus in the present study is on the variability of word reading at the individual level of normal and dyslexic readers. Williams, Hultsch, Strauss, Hunter, and Tannock (2005) refer to this phenomenon as intra-individual variability. Intra-individual variability has been examined in some other developmental disorders, such as in ADHD (Geurts et al., 2008) and in aging (Williams et al., 2005). However, to our knowledge, intra-individual variability of reading in dyslexic readers has not yet been explicitly addressed.

Given a variety of words (e.g., words that vary in length and orthographic complexity), it is likely that the intra-individual variability of dyslexic readers will be larger than that of normal readers. As dyslexic children respond stronger to word characteristics such as word length, there will be automatically more variability in their word reading speed of a set words of varying length. However, the question remains whether these differences in intra-individual variability are fully a function of the response of dyslexic children to longer and more complex words or whether their reading is already more variable at the shortest three-letter words.

In the current study we explicitly examined and modeled differences in the intra-individual variability of normal reading and dyslexic children. We expected this variability to be influenced by word characteristics. To study intra-individual variability, we used hierarchical-regression models or multilevel models (Richter, 2006). In these models responses to individual words (or pseudowords) instead of mean scores over words per condition are analyzed. As the individual words are considered as repeated measures within a person, inter-individual and intra-individual variability are easy to compute. An additional advantage of using this type of modeling is that the influence of word and pseudoword characteristics,
that are hard to match for the different length and digraphs conditions, can be controlled. In the current study we controlled for the influence of phonological-neighbourhood size as differences in neighbourhood size can, according to some computational models of reading (Seidenberg & McClelland, 1989), account for the effects of length on word reading.

4.1.3 Outline of the study
To examine the effect of digraph presence on the length effect and to examine the variability of the different reading groups, we administered a word and a pseudoword naming task to grade-4 normal reading and dyslexic children and to younger grade-2 normal readers. The reading performance of the younger readers, the reading age control group, was at a similar level as the dyslexic children. The children read words and pseudowords of varying length with and without digraphs. We investigated the contribution of word length and digraph presence on the mean reading speed and whether these effects were influenced by phonological neighbourhood size and word frequency. Finally, we compared the response variability of the normal reading, younger and dyslexic readers.

4.2 Method

4.2.1 Participants
Twenty-four dyslexic grade-4 children (9 boys, 15 girls), 24 normal reading grade-4 children and 24 normal reading grade-2 children were selected for participation. All children attended regular education and had normal or corrected to normal vision. The dyslexic children were individually matched in receptive vocabulary, non-verbal intelligence, age and gender to the normal reading children from grade 4 (the chronological age controls) and in gender and reading level to the normal reading children from grade 2 (the reading age controls).

The children were selected from a group of 240 grade 2 and 501 grade 4 children of 16 different schools in the area of Purmerend and Zaanstad (The Netherlands). Normal word reading ability was defined to range from three months below to three months above the average reading level. For the dyslexic group, children with a reading lag of at least 1.5 years were selected. As a consequence, their reading level equalled the level of normal reading children in the younger or halfway grade 2. The descriptive statistics of the three reading groups can be found in Table 4.1.

Word-reading ability was assessed with the B version of the Dutch One-minute test (Brus & Voeten, 1995), a speeded test for single-word reading. This test is commonly used to determine the reading level of children in Dutch primary schools. The test consists of 116 words of increasing difficulty. The children
were instructed to read the words aloud as quickly as possible, without making
errors. The score is the number of words read correctly within one minute. On
the basis of this raw score, a standardized score was computed, with a mean of
100 and a standard deviation of 15 (van den Bos, lutje Spelberg, Scheepstra, & de
Vries, 1994). Word reading ability was measured twice, before the matching of
the groups and approximately four weeks later, just before the experimental test
session. Table 4.1 displays the reading level at the time of the administration of
the experimental reading tasks.

Receptive vocabulary of the children was measured with the subtest Vocab-
ulary of the RAKIT, a Dutch intelligence test battery for children (Bleichrodt,
Drenth, Zaal, & Resing, 1987). This test consists of 60 words of increasing dif-
culty. For each word the children had to choose the corresponding picture out of
four alternatives. When a child made four errors in a row, the administration of
the test was stopped. The score was the number of correct answers.

Finally, nonverbal reasoning was assessed with the Raven Standard Progres-
sive Matrices (Raven, Court, & Raven, 1986). This test consists of 60 items. On
each item the children had to choose a pattern from a set of answer options to
complete a series of patterns. The score was the number of correct answers.

### 4.2.2 Materials and design

The stimulus set included 105 high frequent words and 105 pseudowords. The
items varied in length from three (CVC) to five (CCVCC) letters. In addition,
there were four and five letter words without digraphs (CCVC, CCVCC), with
a homogenous vowel digraph and with a heterogeneous vowel digraph (CVVC,
CCVVC). As a result there were seven different conditions, each consisting of
15 words and 15 pseudowords. The words were selected from a word corpus of
child literature (Schrooten & Vermeer, 1994) and were all high frequent. The
conditions were matched on mean word frequency and, to avoid voice key bias
(Kessler, Treiman, & Mullennix, 2002), on the first letter. The CCVC, CCVVC
and CCVCC condition words were also matched on consonant onset cluster. Be-
cause there are only a few (words with) consonantal digraphs in Dutch, we decided
to use words with vowel digraphs only.

As explained in the previous paragraph, we selected words with two different types of vowel digraphs: homogeneous (e.g., the ee in leek) and heterogeneous (e.g., the ou in soup). However, for the current study we only analyzed the words with heterogeneous digraphs. In Dutch the name of the first (and second) letter of all homogeneous digraphs is the same as the pronunciation of the digraph itself. Since one of our main research questions concerned the effect of serial processing of digraphs we reasoned that it made no sense, to include words with digraphs of which the first letter gives enough information to produce the pronunciation of the digraph. Besides we were concerned that the inclusion of an additional digraph condition would make the analyses unnecessarily complex.

In addition, we would like to underline that the pronunciation of Dutch vowel digraphs is very consistent, in contrast to English vowel digraphs like ea (compare for instance leak and bread), they are always pronounced the same. In contrast to Rastle and Coltheart (1998) we did not include words or pseudowords with two digraphs. Dutch monosyllabic words with two digraphs are scarce, especially high frequent words, making it impossible to match for frequency, length and number of digraphs. In addition, our study also concerned the reading of dyslexic and younger grade-2 children. We expected that the reading of monosyllabic words, and especially pseudowords, with two digraphs would result in too many errors in these groups.

Pseudowords were constructed by mixing the onsets, nuclei and codas of the words. With a few exceptions it was possible to form all the pseudowords by this procedure (sometimes one letter had to be substituted to produce a pseudoword). Overall, the pseudowords contained the same graphemes, digraphs and consonantal onset clusters as the words. All words and pseudowords are presented in Appendix C.

As mentioned in the Introduction section it is often difficult, if not impossible, to match experimental conditions on both word frequency and neighbourhood size. As described above we matched the different word conditions for frequency, but as a consequence, not for neighbourhood size. This is a serious problem, because, for example, an alternative explanation for length effects might be the fact that smaller words have in general more neighbours than larger words. However, in the statistical model that we applied to analyze the data, differences in neighbourhood size across conditions differing in length and the presence of a digraph can be controlled statistically (see below).

In reading research two kinds of neighbourhood sizes are distinguished. Phonological neighbourhood size refers to the number of words that can be formed from a word or pseudoword by changing only one phoneme. Orthographic neighbourhood size is defined as the number of words that may be formed from a word or pseudoword by replacing one letter (Coltheart, Davelaar, Jonasson, & Besner, 1977). Recently, Adelman and Brown (2007) found unique facilitatory phono-
logical neighbourhood effects in word naming data, along with an absence of facilitatory orthographic neighbourhood effects. Keeping this in mind, especially because 40% of our stimuli consisted of words and pseudowords with digraphs, we decided to statistically control for phonological and not for orthographic neighbourhood size in our analyses.

4.2.3 Procedure

The words and pseudowords were administered in two separate blocks and within a block the items were presented in random order. Half of the children first read the word block; the other half first read the pseudoword block. In each block a short break (less than one minute) was given after every 40th trial. Each block was preceded by six practice trials. Between the two blocks was a break of ten minutes, in which the children performed an unrelated computer task.

The tasks were programmed in E-Prime 1.0 (Schneider, Eschman, & Zuccolotto, 2002). The words and pseudowords were presented one by one in the middle of a 14.1-inch XGA LCD screen of a D600 Pentium-M 1.3-GHz computer. The stimuli were printed in 46-point lower-case black Arial font, on a white background. A fixation point (+) was projected in the middle of the screen and 750 ms later the stimulus appeared.

The computer registered latencies and responses. The latencies were defined as the time between the appearance of the stimulus on the screen and the onset of the voice key. The (non)words disappeared as soon as the voice key was triggered.

4.2.4 Method of data analysis

The data were analyzed with hierarchical-regression or multilevel models. In word-recognition experiments like the current study, reaction times can be considered as embedded in a two-level hierarchical structure (Richter, 2006). The repeated (pseudo)word reading observations (level 1) are considered as nested under individuals (level 2). Several tutorials on the use of multilevel models for the analysis of repeated measures like this are currently available (Hoffman & Rovine, 2007; Maas & Snijders, 2003; Quené & van den Bergh, 2004). In the present study we closely followed the model specification for repeated measures by Quené and van den Bergh (2004). On the first level we modeled the word recognition speed for separate words and pseudowords within a child (intra-individual or within-subjects variance). Variation among the scores across the children (inter or between-subjects variance) was modeled at the second level, the level of the child.

In contrast to standard multivariate analyses for repeated measures, which are conducted on the mean latency scores per condition, multilevel regression analyses are performed on the latency scores of all separate items. Such an approach
4.2. METHOD

offers four important advantages. As mentioned in the Introduction section, it enables control for the influence of word characteristics, like phonological neighbourhood size and word frequency. By adding these two variables as level-1 predictors to the model they function as a covariate as the latency score of each word (or pseudoword) is corrected for the effects of word frequency and phonological neighbourhood size.

In addition, multilevel analysis offers a more elegant solution when there is no orthogonal manipulation of, in this case, length and absence/presence of a digraph. To examine the letter-length and phoneme-length effect, two separate multivariate analyses for repeated measures would have been needed. However, with multilevel analyses both effects can be tested within one model.

Third, multilevel models can also be applied if the data are not fully complete. As some missing data are inevitable in reaction time experiments with children, mean reaction times being usually based on correct and valid trials only, this is an important advantage.

Finally, one of the research aims was to examine differences in intra(within)-subjects variances between the dyslexic and normal-reading children. Therefore, the variances needed to be split into between and within-subjects variances, which is only possible with multilevel analysis. Moreover, the multilevel approach enables us to test whether the differences in variances between reading groups are significant (Quené & van den Berg, 2004).

The analyses were conducted with MLwiN 2.02 (Rasbash, Steele, Browne, & Prosser, 2004). In this program differences between parameter estimates can be tested with a Chi-square test. In addition, the inverse fit or deviance (defined as -2 log likelihood, Snijders & Bosker, 1999) of each model can be defined. An improvement in fit of a more complex model indicates that the model can explain more of the variance of the data in comparison to a simpler model (Quené & van den Berg, 2004).

To examine the contribution of number of letters and the presence of a digraph on the mean latency scores and variances we specified a “Length-Digraph” model. In this model, there were five variables per reading group, representing the mean latency scores (in ms) of all length and digraph conditions. As a result the model consisted of fifteen variables. Several contrasts were specified to test mean differences in latency scores between the groups. The contrasts will be described in more detail in the separate sections. As the five condition variables within each reading group were highly correlated, we also estimated the covariances to obtain more reliable estimates of the mean reaction times, and the between-subjects and within-subjects variability. The estimated between-subjects variances, within-subjects variances and covariances can be obtained from the first author.

The results are presented in three sections. In the first section we present the results concerning the mean latency scores in the different length and digraph-presentation conditions. Next, the first (item) level variables frequency and phonolo-
gical neighbourhood size were added to the “Length-Digraph” model to examine whether the length and digraph-presence results were influenced by these variables. In addition, it was tested whether the frequency and neighbourhood effects differed for the three reading groups. The final section focuses on differences in variability between normal and dyslexic readers in response to the different length and digraph-presence conditions.

4.3 Results

4.3.1 Data cleaning and error percentages

The analyses were conducted on valid and correct trials only. Latencies were considered invalid when the response was premature (< 325 ms), longer than 6000 ms, in case of a voice key error and when the response deviated more than three standard deviations from a child’s individual mean score. The deviation scores were calculated per length and lexicality condition.

The percentage of invalid latencies was 5.1% (4.4% words, 5.8% pseudowords) for the normal reading children, 8.3% (8.1% words, 8.6% pseudowords) for the dyslexic children and 9.8% (9.4% words, 10.2% pseudowords) for the younger readers. The error percentages over the valid trials per condition for each reading group are presented in Table 4.2. In general the normal reading children made fewer errors than the dyslexic and younger readers. All children made more errors on longer (pseudo)words than on shorter ones and more errors on pseudowords than on words. As can be seen in Table 4.2, the standard deviations of the mean error percentages were very high, sometimes even higher than the mean error percentage itself. In addition there were floor effects for the normal reading readers. Because of the high variability and these floor effects we considered it inappropriate and unreliable to analyze the error percentages and therefore we will conduct analyses on the mean reaction times only.

4.3.2 Disentangling length and digraph effects

As there were 15 variables per lexicality condition, we specified separate models for words and pseudowords to simplify the analyses and presentation of the results. The mean latency scores per length and digraph-presence condition for each reading group are presented in Figure 4.1 and 4.2.

Before we focused on the length and digraph-presence effects, we considered whether there was an overall main effect of reading group. To this end two orthogonal contrasts were specified. First, we compared the normal with the younger and dyslexic readers. We found a significant difference in mean reading speed for both words and pseudowords, $\chi^2 (1) = 63.22, p < .001$, $\chi^2 (1) = 81.76, p < .001$. A second contrast was specified to compare the dyslexic and younger readers. The
mean reading speed for words and for pseudowords did not differ significantly. Next, we specified three different contrasts to test for length and digraph-presence effects and their interactions with reading group.

Results Length Contrast

The first contrast concerned the pure length effect, denoted as the “Length” contrast. With this contrast we tested for the length effect on the words and pseudowords without digraphs. The Length contrast was further divided into the length effect between three and four letter (pseudo)words and the length effect between four and five letter (pseudo)words.

The main effect of length from three to four letters was significant for both words, $\chi^2 (1) = 47.06, p < .001$ and pseudowords, $\chi^2 (1) = 45.48, p < .001$. The main effect of length from four to five letters was also significant for both words, $\chi^2 (1) = 8.99, p < .01$ and pseudowords, $\chi^2 (1) = 34.95, p < .001$. These main effects were qualified by a number of interaction effects with reading group. The length effect from three to four letters was larger for the younger and dyslexic than for the normal readers. This was found for both words, $\chi^2 (1) = 36.46, p < .001$ and pseudowords, $\chi^2 (1) = 32.44, p < .001$. Also, the length effect from four to five letter for both words, $\chi^2 (1) = 6.49, p < .05$ and pseudowords, $\chi^2 (1) = 13.63, p < .001$ was stronger for the younger and dyslexic readers. We found no significant differences in the effect of length between the younger and dyslexic readers.
The results show that the normal readers have smaller length effects than the dyslexic and younger readers. To examine whether the normal readers were sensitive to word length at all, we tested the contrasts in this group separately. These contrasts revealed that the length effect for words was not significant. The length
effect of the normal readers for pseudowords was significant, \( \chi^2 (1) = 21.66, p < .001 \) (between three and four letters) and \( \chi^2 (1) = 11.25, p < .001 \) (between four and five letters).

**Results Whammy Contrast**

The second contrast concerned the “whammy” effect. In this contrast the mean latency scores of the four and five letter (pseudo)words with a digraph were compared with the mean latencies for the four and five letter (pseudo)words without a digraph. The prediction of the whammy effect is that the mean latencies scores will be faster for four and five letter (pseudo)words without a digraph than for four and five letter (pseudo)words with a digraph (Rastle & Coltheart, 1998).

However, we found that the children responded slower to four and five letter words and pseudowords without a digraph than to four and five letter words and pseudowords with a digraph, \( \chi^2 (1) = 17.16, p < .001 \), and \( \chi^2 (1) = 15.66, p < .001 \), respectively. In addition, we found that this effect was larger for dyslexic and younger readers than for normal reading children for both words and pseudowords, \( \chi^2 (1) = 15.90, p < .001 \), \( \chi^2 (1) = 13.91, p < .001 \). The mean latencies of the younger and dyslexic readers did not differ. Note that these findings contradict the hypothesis that the serial reading strategy of younger and dyslexic readers might interfere with the processing of digraphs. To examine whether four and five-letter (pseudo)words with a digraph were also easier for the normal readers than the four and five-letter (pseudo)words without digraphs, we specified separate contrasts for the normal readers. However, for the normal readers differences between these conditions were not significant.

The only observation that might indicate a whammy effect for the normal readers was found between the four-letter pseudowords with and without digraphs. However, a single contrast for the normal reading children between the four-letter pseudowords with and without a digraph showed that this difference was not significant either, \( \chi^2 (1) = 1.25, p > .10 \).

**Results Phoneme-Length Contrast**

In the previous paragraph it was reported that digraphs do not cause an additional delay on top of the effect of word length. On the contrary, the dyslexic and younger readers were even faster on digraph (pseudo)words as compared to (pseudo)words with the same number of letters but without a digraph. This might be due to the fact that the latter (pseudo)words have the same number of letters but contain one phoneme less. If digraph presence does not cause an additional delay this might suggest that length effects depend on the number of phonemes and not on the number of letters. To test for this possibility we compared (pseudo)words that were equal in their number of phonemes but differed in the number of di-
graphs (one vs. none). In other words, we specified a contrast to compare the three-letter and four-letter (pseudo)words without a digraph with the four-letter and five-letter (pseudo)words with a digraph. Note that the three-letter and four-letter (pseudo)words without a digraph had an equal number of phonemes as, respectively, the four and five letter (pseudo)words with digraphs. This contrast was called the “Phoneme-Length” contrast.

The mean latency scores for words did not differ. However, all reading groups responded faster to three and four-phoneme pseudowords without digraphs than to three and four phoneme pseudowords with digraphs, $\chi^2 (1) = 40.81, p < .001$. Thus, the presence of a digraph had an additional effect, on top of the phoneme-length effect, but only for pseudowords. In addition, we found that this effect was stronger for the dyslexic and younger than for the normal readers, $\chi^2 (1) = 12.16, p < .001$. The effects of the younger and dyslexic readers did not differ significantly. We specified a separate contrast to examine whether the effect was also significant for the normal readers. It was found that the normal readers were also slower on three and four-phoneme pseudowords with a digraph than on comparable pseudowords without digraphs, $\chi^2 (1) = 15.88, p < .001$.

4.3.3 Frequency and phonological neighbourhood size effects

The “Length-Digraph” model was extended by adding frequency and phonological neighbourhood size in order to verify whether the length and digraph-presence effects were still present after controlling for these variables. Before the variables were inserted into the model, we converted the raw frequency and phonological-neighbourhood values into their natural-logarithmic values. As it is impossible to perform logarithmic transformations on zero values, we first added +1 to each phonological neighbourhood size variable (Adelman, Brown, & Quesada, 2006). This was not necessary for the frequency measure since all word frequencies were above zero. Note that we matched all length and digraph conditions for word frequency. Therefore, we did not expect frequency to account for the length and digraph effects. However, we think that it is still relevant to examine whether one of the reading groups responds more strongly to the frequency of the words. To be able to examine and compare the influence of phonological neighbourhood size and frequency for the three groups a separate coefficient was specified for each reading group.

To start with, we examined whether the outcomes of the Length-digraph model were influenced by the addition of the frequency and phonological neighbourhood size variables. The results of the “Length” contrast differed in two respects. Firstly, in the Length-Digraph model, the normal readers did display a length effect for pseudowords. However, when controlling for phonological neighbourhood size, this effect was no longer significant. We will elaborate on this finding in the Discussion. Secondly, the main effect of length between four to five let-
ters for words was no longer significant. In addition, this effect was no longer significantly stronger for dyslexic and younger readers than it was for the normal readers. The results of the “Whammy” and “Phoneme-Length” contrasts did not alter in response to the statistical control for frequency and phonological neighbourhood size.

Next, we examined the frequency and phonological neighbourhood size coefficients per reading group. All reading groups responded significantly faster to more frequent words, $z = 1.70, p < .05$ (normal readers), $z = 6.33, p < .001$ (dyslexics), $z = 3.81, p < .001$ (younger readers). For the word reading of the normal and dyslexic readers no significant effect of phonological neighbourhood size was found. However, we did find a marginal significant effect for the younger readers, $z = -1.44, p = .075$. These readers tended to be faster on words with larger phonological neighbourhood sizes. For pseudoword reading we found phonological neighbourhood size effects for all reading groups. The normal and dyslexic readers were significantly faster on pseudowords with larger phonological neighbourhood sizes, $z = -3.25, p < .001, z = -4.21, p < .001$. A comparable, but marginally significant effect was found for the younger readers, $z = -1.49, p = .068$.

Two orthogonal contrasts were specified in order to test for differences between the frequency and phonological neighbourhood size effects between the three reading groups. First, we compared the normal reading children with the younger and dyslexic readers. We found significant differences for the frequency effect, $\chi^2 (1) = 40.32, p < .001$. The normal reading children were less influenced by the frequency effect than the younger and dyslexic readers. The effect of phonological neighbourhood on word reading did not differ. However, for the pseudowords we found a significant difference in the effect of phonological neighbourhood size, $\chi^2 (1) = 8.64, p < .01$. The younger and dyslexic readers were more strongly affected by the phonological neighbourhood size of pseudowords than the normal reading children. Next, the dyslexic and younger readers were contrasted. It was found that the word frequency effect was more pronounced for the dyslexic than it was for the younger readers, $\chi^2 (1) = 8.01, p < .01$. In addition, we found that the dyslexic readers were also more influenced by phonological neighbourhood size in pseudoword reading than the younger readers, $\chi^2 (1) = 4.84, p < .05$.

4.3.4 Variability differences in normal and dyslexic readers

The mean within-subject standard deviations of the three reading groups for all length and digraph-preservation conditions are presented in Figures 4.3 and 4.4. These standard deviations were derived from the within-subjects variability as estimated in MLwiN. As for the mean latency scores, we specified two orthogonal contrasts to compare differences in variability between the three reading groups. First, we compared the normal reading children with the younger and dyslexic read-
ers. We found a significant difference in the within-subjects variability for words, $\chi^2 (1) = 944.44, p < .001$ and pseudowords, $\chi^2 (1) = 925.78, p < .001$. Accordingly, the dyslexic and younger readers were significantly more variable in word and pseudoword reading than the normal reading children. Second, we compared the dyslexic and younger readers. The within-subject variability of the dyslexic children was significantly larger than that of the younger readers, for both words $\chi^2 (1) = 39.90, p < .001$ and pseudowords, $\chi^2 (1) = 28.08, p < .001$. This means that, although the dyslexic and younger readers were equally fast in general word and pseudoword reading, on the intra-individual level, dyslexic children were significantly more variable in their responses than the younger readers.

In addition, two findings are worth mentioning. Firstly, even for the shortest words in our word sets, the three-letter words and three-letter pseudowords, the variability of the dyslexic children and younger readers was found to be significantly larger than that of the normal reading children $\chi^2 (1) = 183.19, p < .001$, $\chi^2 (1) = 226.98, p < .001$. For the dyslexic readers the within-subjects standard deviation was 3.5 times larger for three-letter words and 4.7 times larger for three-letter pseudowords than the within-subjects standard deviation of the normal readers (see Figure 4.3). The within-subjects standard deviations of the younger readers were 2.1 times larger than the normal readers for three-letter words, and 4.0 times larger for three-letter pseudowords (see Figure 4.4).

Secondly, the within-subjects variability of the dyslexic readers was larger than the variability of the younger readers for both three-letter words and pseudowords, $\chi^2 (1) = 51.79, p < .001$, $\chi^2 (1) = 6.77, p < .01$, respectively. The within-subjects standard deviation for three-letter pseudowords was about 1.2 times larger (an increase of 20%) for the dyslexic children than for the younger readers. Even more impressive was the difference in within-subjects standard deviation for three-letter words, which was 1.6 times larger (an increase of 60%) for the dyslexic children than for the younger readers.

Another interesting finding was that we observed hardly any differences in within-subjects variability for the normal reading children between the five different word and pseudoword conditions (this effect is visible in Figures 4.3 and 4.4). The variance of the dyslexic children, on the other hand, drastically increased in response to additional letters and the presence of a digraph. The younger readers showed another pattern. For three-phoneme words, their variance was quite low, more like that of the normal reading children. However, in response to more challenging word characteristics like consonantal onset clusters (the four-phoneme and five-phoneme words), within-subject variance of the younger readers started to increase as well and their variability on five-letter words resembled those of the dyslexic readers. The younger readers showed an overall high variability for the pseudowords, except for the smallest three-letter pseudowords. For pseudowords the variability of their reading latencies was somewhere between the normal reading and dyslexic children.
4.3. RESULTS

Figure 4.3
Mean within-subjects standard deviations for the letter-length and digraph-presence conditions per reading group for words. Error bars represent standard errors; ‘di’ refers to the presence of a digraph.

Figure 4.4
Mean within-subjects standard deviations for the letter-length and digraph-presence conditions per reading group for pseudowords. Error bars represent standard errors; ‘di’ refers to the presence of a digraph.
4.4 Discussion

The aim of the current study was twofold. Firstly, we examined the separate influences of length and digraph presence on the word reading of normal, dyslexic and younger readers. In line with the findings of earlier studies (Martens & de Jong, 2006; Ziegler et al., 2003; Zoccolotti et al., 2005) we predicted that the younger and dyslexic readers would be more susceptible to the number of letters in a (pseudo)word than the normal reading children. In addition, the letter-by-letter reading strategy of the younger and dyslexic readers was expected to lead to an extra delay in the processing of digraph, on top of the stronger length effect.

The second aim was to examine differences in the variability of reading in normal, dyslexic and younger readers. Although dyslexic children are often assumed to be more variable in reading, intra-individual differences in naming have not been addressed before. We examined to what extent the variability of dyslexic children is due to (pseudo)word length and the presence of a digraph.

4.4.1 Word length and digraph effects

In line with previous studies (Martens, 2006; Ziegler et al., 2003; Zoccolotti et al., 2005), we found that dyslexic and younger readers were more sensitive to the number of letters of words and pseudowords. This probably reflects their stronger reliance on a letter-by-letter reading strategy and less efficient use or a lack of orthographic knowledge. In contrast to earlier studies with dyslexic and younger readers, we controlled for phonological neighbourhood size. Hence, we can safely conclude that the length effects in the dyslexic and younger readers were not caused by differences in neighbourhood size between the various word length conditions. The reading speed of the normal readers was unaffected by length, indicating that they were able to process words in parallel or used more efficient lexical reading strategies (Coltheart et al., 2001). After correction for phonological neighbourhood size, even the normal readers’ length effect for pseudowords disappeared. Apparently, the relatively small length effect for pseudowords in the normal readers was a neighbourhood size effect in disguise. This finding is in line with the assumption of some connectionist models that length effects are primarily caused by differences in neighbourhood size between the word length conditions (Seidenberg & McClelland, 1989). However, although the length effects of the dyslexic and younger readers decreased somewhat after controlling for phonological neighbourhood size, they clearly did not disappear.

Balota et al. (2004) reported that in skilled readers length and neighbourhood independently affected word (and pseudoword) recognition. This is in line with our findings for the dyslexic and younger readers. In contrast, the better readers were not influenced by length after controlling for phonological neighbourhood size. The discrepancy between the normal readers of the current study and the
adult readers in the study by Balota et al. might be due to the strength of the length manipulation. In the current study we only used three to five letter words and in this range length effects might be absent in skilled readers. Similarly, Spinelli et al. (2005) found that proficient readers do not display length effects for words with five or less letters. In contrast, Balota et al. used a larger variety of length including words and pseudowords ranging from two to eight letters. Possibly, length effects in skilled readers can be found if word length exceeds five or six letters (see Spinelli et al.) and therefore Balota et al. might have found independent effects of length and neighbourhood size on (pseudo)word reading.

After the pure length effects were examined, we studied the effect of digraph presence on the length effect. Based on the predictions of the Dual Route Cascaded model (Jackson & Coltheart, 2001b; Rastle & Coltheart, 1998) and the more serial reading strategies that are observed in children (Martens & de Jong, 2006; Ziegler et al., 2003; Zoccolotti et al., 2005), we hypothesized that the presence of a digraph would cause an additional delay in the word and pseudoword reading speed of all groups. We also predicted that this effect would be more pronounced for younger and dyslexic readers as they are known to rely even more strongly on letter-by-letter reading strategies, as reflected by their stronger response to the length of words and pseudowords. Contrary to these predictions, the dyslexic and younger readers appeared to be even faster on four and five-letter (pseudo)words with digraphs than on four and five-letter (pseudo)words without. The normal reading children were equally fast in both conditions. Although the presence of a digraph did not cause an additional delay in the processing of words and pseudowords of the same number of letters, we did find that it caused an additional delay on top of the phoneme length effect for pseudowords in all reading groups. That is words with a digraph were read more slowly than words with an equal number of phonemes but without a digraph. This effect was stronger for the younger and dyslexic readers.

As said, the empirical findings regarding the impact of digraphs on the length effect are not in line with the predictions of the DRC model (Coltheart et al., 2001). However, the observation that there is not an effect of digraph on top of the phoneme length effect in words does fit into the assumptions of the CDP+ framework (Perry et al., 2007). The CDP+ model assumes that the letters of a word are visually parsed into graphemes before they are mapped onto the phoneme units. Therefore, the CDP+ model does not predict that the presence of a digraph causes an additional delay. Actually, the CDP+ model would only predict a phoneme-length effect, as was found for words for the dyslexic and younger readers. In line with these findings, Marinus and de Jong (2008) recently found that normal reading and dyslexic children were equally hampered in their word-naming speed by the visual distortion of a digraph unit. Both results support the conclusion that normal readers and even dyslexic children visually parse digraphs in words before they map them onto their corresponding sounds.
However, this conclusion only explains the results for the words. For *pseudowords*, on the other hand, the digraph did cause an additional delay on top of the phoneme-length effect for all reading groups. In addition, this effect was stronger for the younger and dyslexic readers. Note that these empirical findings cannot be accommodated in the CDP+ model. A possible explanation for this observation might be that, although the children do not have problems with the visual-parsing process, they still might run into a little bit of a delay when the grapheme has to be mapped onto the corresponding phoneme. The fact that vowel digraphs in Dutch are generally less frequent than one-letter graphemes might underlie this pattern as the reading system of the children might still be developing into becoming more efficient in the grapheme to sound mapping process for the more complex graphemes. This hypothesis could also explain why the effect is stronger for the younger readers, who have less experience with grapheme-phoneme mappings, and the dyslexic readers, who are typically found to have difficulties mapping graphemes onto phonemes.

The dyslexic children clearly displayed sensitivity to the effects of length, which is considered a marker effect for the use of a serial-reading strategy. However, we did not find that the dyslexic children had impairments on markers of lexical processing like sensitivity to phonological neighbourhood size or word frequency. These results add to a growing amount of evidence that dyslexic children do not have specific problems in the processing of lexical information (Barca et al., 2006; Ziegler et al., 2008; Ziegler et al., 2003).

### 4.4.2 Variability in reading

We found, as expected, that the intra-individual variability in reading latencies of the dyslexic children was larger than the variability of their normal reading peers. This larger intra-individual variability of the dyslexic readers might, however, be a straightforward function of their slower mean reading speed. In this respect the finding that the dyslexic readers also displayed larger intra-individual variability in reading latencies than the younger readers is important. As the dyslexic and younger readers showed the same effects of length and digraph presence, and had a similar mean reading speed, the larger variability of the dyslexic children cannot be due completely to their overall lower reading speed. According, the present study suggests that, in addition to slow reading, a larger variability in word reading is another characteristic of the dyslexic reading system.

Interestingly, in comparison to the younger readers, the reading of the dyslexic children was relatively most variable for the shortest words. On the three-letter words they were 60% more variable than the younger normal readers and on the four-letter words with a digraph they were 50% more variable. In contrast, the younger readers were somewhat more variable on the longer five-letter words. For pseudowords, the dyslexics were systematically more variable than the younger
that the largest differences in variability between dyslexic and younger readers concern the shortest words seems somewhat surprising. However, from a developmental perspective, stabilization of performance can be considered as a characteristic of the acquisition of a skill. In this respect it might be expected that the reading of shorter words will stabilize more rapidly. As, evidently, younger normal readers are better equipped to acquire reading it might not come as a surprise that differences with dyslexic readers become most visible on words which reading stabilizes relatively early. Indeed, the intra-individual variability in the three-phoneme word reading of the younger readers seemed already to be developing towards the level of variability of the more advanced grade-4 normal readers.

The reading system of these normal readers had stabilized almost completely as their intra-individual word reading variability was hardly affected by length or the presence of a digraph. As said, the variability of the dyslexic readers was already high for three phoneme words but, in addition, increased as word length increased. The younger readers displayed relatively low intra-individual variability in response to the easiest (three phoneme) conditions but their reading system had clearly not stabilized as their variability increased in response to the longer and more complex (pseudo)words.

The finding of a relatively larger intra-individual variability in the reading of dyslexic children is not easy to explain. However, as mentioned in the Introduction, intra-individual variability has been examined in some other developmental disorders, e.g., ADHD (Geurts et al., 2008) and in aging (Williams et al., 2005). For ADHD and aging it has been suggested that the relatively large intra-individual variability might be a nonspecific characteristic of brain pathology. Based on the particular form of the response latency distributions in these groups, the hypothesis has been proposed that their higher intra-individual variability is caused by lapses of attention. This cause, however, does not seem very likely for the dyslexic children in the current study. If the higher intra-individual variability of the dyslexic readers was caused by lapses of attention than the dyslexic readers would be equally more variable than the younger readers in all length and digraph conditions. Instead, we found a highly specific pattern across length and digraph conditions with a main difference in variability between dyslexic and normal readers on short words. Therefore, it seems more likely that the higher variability of the dyslexic readers is caused by difficulties that are specific to reading, such as instability in the retrieval of orthographic knowledge, and is not due to a nonspecific characteristic of brain pathology such as lapses of attention.

In the current study, we demonstrated that the dyslexic readers were more variable in their word reading speed than the younger readers. However, many issues remain to be examined in future research. One issue concerns the nature of
or processes that underlie this larger variability. In this respect it might be fruitful to consider the distribution of the response latencies of dyslexic readers. Given that latency distributions are often positively skewed, especially the size of the right tail of these distributions might be of interest (see for example Geurts et al., 2008, Williams et al., 2005). Another issue is whether the larger intra-individual variability of dyslexics is specific to their reading or concerns a more general characteristic that results in a larger variability in other skill domains as well.