Word-recognition processes in normal and dyslexic readers

Marinus, E.

Citation for published version (APA):
Chapter 6

Size does not matter, frequency does matter: Sensitivity to orthographic neighbours in normal and dyslexic readers

Abstract

The current study examined the influence of the number of orthographically similar candidates, neighbourhood size, on the word and pseudoword naming performance of normal, dyslexic and beginning readers. Participants were 23 Dutch dyslexic grade-4 children, matched to 23 grade-4 chronological age controls and 17 grade-2 reading age controls. Unexpectedly, neighbourhood size had similar effects in all groups: It did not affect word naming and facilitated the naming of pseudowords. However, the presence of a high frequency neighbour had different effects. In contrast to normal readers, beginning and dyslexic readers were slower in naming words with a high-frequent neighbour. These findings suggest a dissociation between global and specific effects of neighbour words. Nevertheless, both findings seem compatible with the view that orthographic representations of beginning and dyslexic children are not (yet) sufficiently specified.

CHAPTER 6. SENSITIVITY TO ORTHOGRAPHIC NEIGHBOURS

6.1 Introduction

The main characteristic of skilled word reading is the ability to recognize words instantly and effortlessly. There is consensus that such fast and efficient processing of words is supported by orthographic knowledge, a system of associations between phonology and orthography (Ehri, 1992; Perfetti, 1992). In contrast to normal readers, the reading of dyslexic children remains slow and effortful. Previous studies have shown that these children need many more exposures to words than normal readers to acquire similar levels of orthographic knowledge (Ehri & Saltmarsh, 1995; Reitsma, 1983; Share & Shalev, 2004).

According to the self-teaching hypothesis (Jorm, Share, Maclean, & Matthews, 1984; Share, 1999), difficulties in building up word specific orthographic knowledge should primarily be attributed to an insufficient ability to phonologically recode words. Phonological recoding refers to the translation of unfamiliar printed words into their spoken equivalents by using letter-sound knowledge. As dyslexic children are known to have specific impairments in the representation, storage and retrieval of speech sounds (Snowling, 2000), it is conceivable that this leads to difficulties with phonological recoding, which in turn might hamper their development of orthographic knowledge.

However, phonological decoding of dyslexic children learning to read transparent orthographies, like Italian, Dutch and German, seems relatively intact as their reading is typically found to be quite accurate (Barca, Burani, di Filippo, & Zoccolotti, 2006; Martens & de Jong, 2006; Ziegler & Goswami, 2005). Apparently, if grapheme-phoneme correspondences are consistent, then even dyslexic children are able to map printed words onto their spoken forms (Landerl, Wimmer, & Frith, 1997). Nevertheless, the reading of these children continues to be slow. One explanation for this overarching slowness is that dyslexic readers persist in using an inefficient, sublexical, decoding strategy (de Luca, Borelli, Judica, Spinelli, & Zoccolotti, 2002; Hutzler & Wimmer, 2004; Spinelli et al., 2005), instead of progressing towards a reliance on more efficient parallel word recognition strategies, as happens in normal reading development. The predominant use of such a sublexical reading procedure is supported by the outcomes of studies that examined word-length effects. The word-length effect refers to the observation that longer words take more time to read than shorter ones. In normal reading development the effect has been found to decrease over the years and is absent, at least for monosyllabic words, in skilled readers (Balota, Cortese, Sergent-Marshall, & Spieler, 2004; Spinelli et al., 2005). However, the reading speed of dyslexic readers remains to be affected by the length of a word (Marinus & de Jong, in press; Spinelli et al., 2005; Ziegler et al., 2003). This seems to reflect continued reliance on decoding.

Thus far, most approaches to illuminate dyslexic’s impairments in reading speed have focused on the (word specific) knowledge and strategies (sublexical
or parallel) that are used to identify a particular (target) word. However, it is generally believed that in the process of identifying a particular word also other orthographically similar words are being activated (Wagenmakers & Raaijmakers, 2006). In the current study, we will examine the effect of orthographically similar words on the recognition of a target word and specifically consider whether such similar words hamper or facilitate the rapid identification of target words in normal and dyslexic readers.

Castles, Davis, Cavalot, and Forster (2007) have recently taken a somewhat similar approach. Using a masked form-priming task they examined the influence of the activation of orthographically similar words on visual word recognition in grade-3 and grade-5 children. In a masked form-priming task the presentation of the target word is preceded by a briefly presented letter string (the prime) that is very similar to the target word (e.g., *rlay* → *PLAY*). It was found that beginning readers in grade 3 were sensitive to form-priming of highly frequent words with a high neighbourhood size. However, two years later the same children no longer showed such priming effects. Castles et al. interpreted these effects as support for their Lexical Tuning Hypothesis. The orthographic representations of the high frequency words in the children in their study had become more finely tuned and as a result the identification speed of these words was no longer sensitive to form-priming effects. Thus, this hypothesis suggests that the priming effect is dependent on the specificity of the orthographic knowledge of the target word.

In the present study we examined the influence of orthographically similar candidates in the orthographic lexicon in word naming. We were particularly interested whether the differences in the number of orthographically similar words had a different effect on normal and dyslexic readers. A useful metric for the number of such candidates is orthographic neighbourhood size, or N-size. The orthographic neighbourhood of a given word represents all the existing words that can be created by replacing one of its letters for another one (Coltheart, Davelaar, Jonasson, & Besner, 1977). Examples of orthographic neighbours of *sand* are *land*, *hand*, but also *send*, *said* and *sang*. In skilled readers, the joint activation of these visually similar words in the orthographic lexicon generally speeds up the reading of a target word (Andrews, 1997).

The influence of orthographic neighbourhood size has been frequently studied in the context of computational modeling of reading. Below we will describe how neighbourhood effects are anchored in two influential classes of models: interactive activation and competition (IAC) and parallel-distributed processing (PDP) models. IAC models of visual word recognition (e.g., McClelland & Rumelhart, 1981) usually assume that there is lateral inhibition among the activated word neighbours. Therefore, the empirical finding that the presence of orthographic neighbours speeds up word recognition might seem somewhat surprising (Lavidor, Johnston, & Snowling, 2006). However, depending on the parameter settings that are used and the larger context in which the interactive activation network is
embedded, the model is also able to produce facilitating effects (Andrews, 1997). For instance, in the Dual Route Cascaded model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), which is partly based on the IAC model of McClelland and Rumelhart, cascaded processing allows a word stimulus to activate the orthographic representations of neighbour words. These orthographic representations in turn facilitate the activation of the phonological representations, which finally have a facilitating effect on the activation of the phonemes in the output system.

In PDP models (Plaut, McClelland, Seidenberg, & Patterson, 1996), there are no word-specific orthographic representations. Instead, knowledge of words is represented as distributed patterns of activity in an associative network. In these models, orthographic N-size effects cannot be interpreted apart from consistency effects. The magnitude of the consistency effect for a given word is the weighted outcome of the summed frequency of friends and enemies of the target word (Jared, McRae, & Seidenberg, 1990). A friend is a word (e.g., lint) that has the same pronunciation as the target word (e.g., mint) whereas an enemy is pronounced differently (e.g., pint). Friends and enemies are both orthographic neighbours of the target word. However, whereas the friends have a facilitating effect on word recognition, an enemy will inhibit the recognition process. This joint interplay of inhibitory and facilitating effects of enemy and friendly neighbours is especially prominent in English. Note, however, that in languages with a more transparent orthography, like Dutch and German, most neighbours will be “friends” with the target word.

Although different underlying mechanisms are assumed, the PDP and DRC models are both able to reproduce the empirical finding that words from large orthographic neighbourhoods are read faster than words from small orthographic neighbourhoods. In addition, both computational models replicate the observation that the orthographic N-size effect is more pronounced for pseudowords than for words and is even absent for high-frequency words (Andrews, 1989, 1997; Coltheart et al., 2001; Peereman & Content, 1995; Plaut et al., 1996). In localist models, like the DRC model, the weaker or even absent orthographic N-size effects for high-frequent words can be interpreted in terms of their higher resting activation levels in the orthographic lexicon. Due to these higher resting levels, the threshold for target-word recognition will be reached more quickly and as a result the facilitating effect of orthographic neighbours will be of less influence (Andrews, 1989; Peereman & Content, 1995). In PDP models it is assumed that if either the frequency or the consistency of a set of words is sufficiently high on its own to produce fast naming speed, increasing the other factor will yield little further improvement. Simply put, if a word is highly frequent, the additional effect of the neighbours will be negligible.

The N-size effect in skilled adult readers and its interactions with lexicality and frequency have been thoroughly investigated. Most studies have used lexical-decision and focused on words only. In general, it was found that N-size facilitates
lexical decisions for low-frequency words (Andrews, 1997; Sears, Hino, & Lupker, 1995). This facilitating effect diminishes as a function of word frequency and is not significant for high-frequent words (Andrews, 1989). The effect of orthographic N-size for lexical decisions about pseudowords is less well studied and the results are less consistent. Until now, both inhibitory effects (Coltheart et al., 1977) and null effects (Arduino & Burani, 2004) have been found.

The effects of N-size on naming are more straightforward. Skilled adult readers respond faster and more accurately to both words and pseudowords with many neighbours than to words and pseudowords with few neighbours (Andrews, 1997; Arduino & Burani, 2004; Peereman & Content, 1995). As for lexical decisions, it has been found that the facilitating effect of orthographic N-size is not present in the naming of high-frequent words (Andrews, 1989; Peereman & Content, 1995).

N-size effects in skilled adults readers were mainly investigated to learn more about lexical access processes in the mature reading system. However, examining sensitivity to N-size and differences in sensitivity to the N-size for words in pseudowords in normal, beginning and dyslexic readers can also be used as a means to track the growth of orthographic knowledge in the developing reading system. For instance, less sensitivity to N-size in dyslexic and beginning readers would indicate that these children have or make less use of lexical knowledge, which in turn might explain their persistent use of a sublexical reading strategy (Barca et al., 2006). In addition, it is also interesting to investigate whether the finding that N-size effects are absent in high-frequency words also holds for children and especially for beginning and dyslexic children. Although both beginning readers and dyslexic children have been found to be sensitive to word frequency (Barca et al., 2006), it is conceivable that the (high-frequent) words in their orthographic lexicon are not as highly pre-activated as in a skilled reading system. As a result, the recognition of these words might still be facilitated by the presence of neighbours. Indeed, in a recent Spanish study, examining the development of sensitivity to orthographic N-size in children from first to sixth grade with a lexical decision task (Duñabeitia & Vidal-Abarca, 2008), evidence was found for a developmental pattern in which sensitivity to neighbourhood size of high-frequent words diminished as a function of reading experience.

In contrast to Duñabeitia and Vidal-Abarca (2008), we used a naming task to investigate sensitivity to N-size in normal and dyslexic readers. In addition, the focus will not be restricted to sensitivity to N-size in words, but we will also examine sensitivity to N-size in pseudowords. Furthermore, most of the earlier naming studies on N-size effects in children have investigated reading accuracy only. In these studies it was found that even children with less than two years of reading experience were already responding more accurately to words and pseudowords with a high N-size as compared to words and pseudowords with a low N-size (Laxon, Gallagher, & Masterson, 2002; Laxon, Masterson, & Moran, 1994; Laxon, Smith, & Masterson, 1995). In addition and in line with skilled adult read-
ers (e.g., Peereman & Content, 1995), Laxon and colleagues (2002) found that the facilitating N-size effect tended to be higher for pseudowords than for words. Interestingly, and converging with our predictions, and the finding of Duñabeitia and Vidal-Abarca (2008) regarding beginning readers, Laxon, Coltheart and Keating (1988) also found poor readers to show larger orthographic N-size effects for words than good readers.

To our knowledge, there are only two studies that have addressed orthographic N-size effects in dyslexic readers. Lavidor, Johnston and Snowling (2006) found that adult dyslexic readers were more sensitive to the facilitating effects of orthographic N-size in a lexical decision task than normal readers. However, this effect only occurred when the words were presented in the left visual field. When the words were presented in the right visual field, orthographic N-size had an inhibitory effect on their lexical decisions, whereas the lexical decision speed and accuracy of normal readers were still facilitated. Ziegler et al. (2003) have been the first and only to compare orthographic N-size effects in normal reading and dyslexic children. In contrast to Laxon and colleagues, Ziegler et al. not only measured reading accuracy but also reading speed. However, in contrast to all the previously described studies, Ziegler et al. focused on sensitivity to orthographic body N-size. Body N-size is a more restricted measure than (general) N-size in that it only counts words that share the same orthographic rime (i.e., body), such as hand, sand and land, with the target word. Ziegler et al. found body N-size to have a facilitating effect on word and pseudoword reading speed in all reading groups. In addition, the body N-size effect was stronger for pseudowords than for words. Ziegler et al. also examined whether the body N-size effect differed between normal reading and dyslexic children. Interestingly, the overall body N-size effect (summed over words and pseudowords) was marginally stronger for the dyslexic readers.

The current study aimed to examine differences in the use of orthographic knowledge between normal, beginning and dyslexic readers. Sensitivity to orthographic neighbourhood size was used as a marker effect. Dyslexic grade-4 children were matched to normal grade-4 and grade-2 readers. The children were learning to read in Dutch, a language with a relatively transparent orthography. All children completed a word and a pseudoword-naming task. Based on the results of earlier studies, we expected all children to respond faster and more accurately to words and pseudowords from a high orthographic neighbourhood than to words and pseudowords from a small orthographic neighbourhood. In addition, we expected the orthographic N-size effect to be more pronounced for pseudowords than for words. Following the results of Duñabeitia and Vidal-Abarca (2008) we expected the N-size effects for words to be stronger in the beginning and dyslexic readers than in the normal grade-4 children. Because only high-frequency words were included, it was anticipated that the normal readers would not display N-size effects in the word-naming task at all. Such
an outcome would indicate that the word-specific orthographic representations of normal reading grade-4 children reading in a transparent orthography have already reached activation levels similar to those in a skilled adult reading system.

6.2 Method

6.2.1 Time frame of data collection

The data were collected within a larger project on word recognition processes in children. This project encompassed three different times of measurement. Screening for dyslexia took place at the beginning of the school year, from October until the beginning of December. The first round of testing was conducted from the second half of January until the beginning of February. The second round of testing was conducted in the second half of February and the beginning of March. The neighbourhood tasks that were used in the current study were administered in the second round.

6.2.2 Screening

Participants of the present study were 240 grade-2 and 501 grade-4 children of 16 different schools for regular education, in the area of Purmerend and Zaanstad (The Netherlands). All children were screened for word-reading speed, passive vocabulary and nonverbal reasoning.

Word reading ability was assessed with the B version of the Dutch One-minute test (Brus & Voeten, 1995), which was administered individually. This test is commonly used in Dutch education to determine children’s reading level. The test consists of 116 words of increasing difficulty. The score was the number of words read correctly. Word reading ability was measured twice, before the matching of the groups and approximately four weeks later.

Receptive vocabulary of the children was measured with the subtest Vocabulary of the RAKIT, a Dutch intelligence test battery for children (Bleichrodt, Drenth, Zaal, & Resing, 1987). This test consists of 60 words of increasing difficulty. 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 Progressive 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.

Based on this screening, 24 dyslexic grade-4 children (9 boys, 15 girls), 24 normal reading grade-4 children (9 boys, 15 girls) and 24 normal reading grade-2 children (9 boys, 15 girls) were selected for participation. All children had normal
or corrected to normal vision. Normal word reading ability was defined to range from three months below to three months above the average reading grade level. The dyslexic children had a reading lag of at least 1.5 years. As a consequence, their reading level equaled the level of normal reading grade-2 children. The dyslexic children were individually matched in receptive vocabulary, non-verbal intelligence, age and gender to the normal reading grade-4 children (the chronological age controls) and in gender to the normal reading grade-2 children (the reading age controls). To anticipate that the reading age controls would progress faster in their reading ability than the dyslexic readers, we selected grade-2 children with a slightly lower reading ability than the dyslexic readers.

### 6.2.3 Participants of the current study

In order to examine whether the reading age controls had caught up with the dyslexic readers, the word reading ability of all children was retested during the first round of testing. Based on these results, it was decided to remove one child from the reading age control group, because her reading level had not progressed between the first and second measurement. The mean raw scores (words per minute) of the normal reading, dyslexic and beginning readers during the first test session halfway January are presented in Table 6.1. The neighbourhood task was administered another three weeks later. The reading of Dutch beginning readers is known to speed up considerably, even within a few weeks and this development of fluent word reading typically emerges halfway through grade 2. Therefore, we calculated an overall mean score for every child across all conditions of the neighbourhood-naming task. The overall mean latency score of five children from the reading age control group fell within one standard deviation of the normal (two year older) readers. Therefore, these younger children were excluded from the analyses.

In addition, we excluded three children (one from each group), because their General Means were three standard deviations higher than those of the other members of their reading group. Hence, for the current study, we used the data of 23 dyslexic, 23 normal reading and 17 reading age controls. Post-hoc analyses (Tukey’s HSD) demonstrated that the reading age controls and dyslexic readers did not differ significantly on their General Mean score, $t(1) = 146.98, p > .10$.

### 6.2.4 Materials

We selected 30 four letter and 30 five letter high frequency words from a word corpus of child literature (Schrooten & Vermeer, 1994). All words consisted of

---

1 The results of the analyses over the complete sample were virtually identical. The only difference was that the overall mean latencies of the dyslexic readers were now significantly slower than the latencies of the beginning readers, which was to be expected because of the faster reading speed latencies of these particular five beginning readers.
Table 6.1

Descriptive statistics of the dyslexic, beginning and normal readers: mean (M) and standard deviation (SD).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dyslexic</th>
<th>Beginning</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Age (years)</td>
<td>9.9</td>
<td>0.4</td>
<td>7.8</td>
</tr>
<tr>
<td>Reading ability (words/min)</td>
<td>39.96</td>
<td>4.61</td>
<td>36.09</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>47.25</td>
<td>3.78</td>
<td>40.61</td>
</tr>
<tr>
<td>Nonverbal reasoning</td>
<td>38.04</td>
<td>4.98</td>
<td>25.39</td>
</tr>
</tbody>
</table>

Four phonemes and had a CCVC, CVCC, CCVVC or CVVCC structure. Half of the words had many (8–16) orthographic neighbours; the other half with few (1–4) orthographic neighbours. The words were matched on word structure, mean word frequency and mean frequency of the neighbours. The orthographic neighbourhood sizes and the mean frequencies of the neighbours were derived from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993).

It was not possible to create a parallel pseudoword set by mixing the onsets, nuclei and codas of the word set because this would result in pseudowords with other neighbourhood sizes than in the original word set. Therefore, the pseudowords were constructed by replacing the letters in the middle of the words by a legitimate Dutch bigram or trigram with the same structure. In a few cases we also had to change the final letter. As a result the words and pseudowords of both N-size conditions were matched on word structure, first letter and, in general, on final letter as well. However, the different N-size conditions could not be matched on first letter. The lists of stimuli are presented in Appendix E and the descriptive statistics of the lists are presented in Table 6.2.

Table 6.2

Descriptive statistics of the words and pseudowords of the neighbourhood task.

<table>
<thead>
<tr>
<th></th>
<th>Words</th>
<th>Pseudowords</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High N-size</td>
<td>Low N-size</td>
</tr>
<tr>
<td>Mean word freq</td>
<td>218 (290)</td>
<td>221 (254)</td>
</tr>
<tr>
<td>Mean n-size</td>
<td>10.5 [8–16]</td>
<td>2.5 [1–4]</td>
</tr>
<tr>
<td>Mean n frequency</td>
<td>937 (742)</td>
<td>930 (960)</td>
</tr>
</tbody>
</table>

Note. Standard deviations are in parentheses. Ranges are in square brackets.

6.2.5 Administration of reading tasks

The children were individually tested in a quiet room at school. The words and pseudowords were presented in two separate blocks and within a block the items were presented in random order. Half of the children were randomly selected to read the word block first; the other half started with the pseudoword block. Each block was preceded by six practice trials and within each block short breaks
were inserted (less than one minute) after every 40th trial. Between the word and pseudoword block was a break of ten minutes, which was filled with an unrelated computer task. The children were instructed to read the (pseudo)words as fast as they could without making errors.

The tasks were programmed in E-Prime version 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 and the onset of the voice key. The (non)words disappeared as soon as the voice key was triggered.

6.2.6 Statistical analyses

In word recognition experiments like the current study, latencies and error percentages can be considered as embedded in a two-level hierarchical structure (Richter, 2006). The repeated word and pseudoword reading observations (level 1) are 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). We specified two separate models for the latency and error data. In each model within-subjects variance of accuracy (or latency) scores was modeled on the first and between-subjects variance was modeled on the second level.

Conceptually, the approach to analyze the data is similar to a multivariate analysis of the mean reaction times with dummy variables denoting group membership, lexicality and orthographic neighbourhood size condition. In the multilevel models accuracy and latency scores were predicted by four dummy variables per reading group. Each dummy variable represented a lexicality by N-size condition, for example, mean reading speed of the normal readers for high N-size words. Due to the separate specification of these dummy variables and in contrast to univariate and multivariate analyses, the variances of the three different reading groups can be separately estimated in the multilevel model. Because dyslexic and beginning readers typically show larger variances in their responses than normal reading children this is an important issue (Quené & van den Bergh, 2004). Another advantage of multilevel analysis is that it can also be applied if the data are not fully complete. This is not possible in a regular multivariate analysis. As some missing

2The subsequent specification of contrasts on these measures to test for differences between the reading groups is comparable to the L-MATRIX that can be used in the General Linear Model approach in SPSS.
data are inevitable in reaction time experiments with children (see data cleaning section below), mean reaction times being usually based on correct and valid trials only, this is also an important advantage.

As the error and latency data are respectively binomially and normally distributed, different estimation and model specifications were needed. For the latency data we estimated a model within the default multilevel procedure for normally distributed data. In such models, the regression parameters of the dummy variables are equivalent to the latency means of each N-size and lexicality condition per reading group. To be able to correctly model and test the binomially distributed accuracy data, we used multilevel logistic regression analyses. In contrast to the latency analyses, the fixed effects in such models do not directly reflect the accuracy scores, but have to be converted first. These converted scores are similar to the descriptive means as obtained in SPSS.

The analyses were conducted with MLwiN 2.02 (Rasbash, Steele, Browne, & Prosser, 2004). In this program differences between parameter estimates (fixed and random effects) can be tested with a Chi-square test.

### 6.3 Results

#### 6.3.1 Data cleaning

Invalid and incorrect responses were not included in the analyses. Latencies were considered invalid when the response was premature (< 325 ms), longer than 6000 ms, in case of a voice key error and when deviating more than three standard deviations from a child’s individual mean score. The deviation scores were calculated per orthographic neighbourhood and lexicality condition. The percentage of invalid latencies for the normal reading children was 6.67% for words and 5.22% for pseudowords, for the dyslexic children 5.29% for words and 5.79% for pseudowords, and for the beginning readers 6.57% for words and 6.57% for pseudowords. The mean error percentages over the valid trials per condition for each reading group are presented in Table 6.3. In general, the normal reading children made fewer errors than the dyslexic and beginning readers.

#### 6.3.2 Effects of neighbourhood density

*Error percentages.* Mean differences between the conditions were tested with multilevel analyses (see Statistical Analysis in the Method section). These analyses are based on all individual word and pseudoword error scores that, in turn, are considered to be nested under individuals.

Two orthogonal contrasts were specified in order to compare differences in mean error percentage between the three reading groups. Firstly, we compared the normal reading children with the beginning and dyslexic readers. We found a
Table 6.3
Mean error percentages and latency scores in the different neighbourhood conditions for words and pseudowords for the dyslexic, beginning and normal readers.

<table>
<thead>
<tr>
<th></th>
<th>Words</th>
<th>Pseudowords</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High N-size</td>
<td>Low N-size</td>
<td>High N-size</td>
<td>Low N-size</td>
</tr>
<tr>
<td>DYS</td>
<td>9.90 (6.77)</td>
<td>7.74 (5.36)</td>
<td>11.63 (6.05)</td>
<td>22.19 (15.52)</td>
</tr>
<tr>
<td>RA</td>
<td>9.41 (6.84)</td>
<td>6.89 (5.23)</td>
<td>8.39 (5.10)</td>
<td>16.93 (11.58)</td>
</tr>
<tr>
<td>CA</td>
<td>1.43 (2.35)</td>
<td>0.53 (1.42)</td>
<td>3.79 (5.41)</td>
<td>6.54 (5.60)</td>
</tr>
</tbody>
</table>

Mean latency scores (ms)

<table>
<thead>
<tr>
<th></th>
<th>High N-size</th>
<th>Low N-size</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS</td>
<td>1092 (309)</td>
<td>1076 (243)</td>
<td>1532 (482)</td>
<td>1670 (556)</td>
</tr>
<tr>
<td>RA</td>
<td>984 (217)</td>
<td>967 (186)</td>
<td>1361 (225)</td>
<td>1470 (393)</td>
</tr>
<tr>
<td>CA</td>
<td>605 (86)</td>
<td>606 (77)</td>
<td>664 (108)</td>
<td>724 (150)</td>
</tr>
</tbody>
</table>

Note. Standard deviations are in parentheses. DYS = dyslexic readers, RA = beginning readers (reading age control), CA = normal readers (chronological age control).

significant difference in accuracy score, $\chi^2 (1) = 60.47, p < .001$. Overall, normal readers made fewer errors compared to the beginning and dyslexic readers. The accuracy of the dyslexic and beginning readers did not differ. In addition, we found a main effect for lexicality, $\chi^2 (1) = 54.15, p < .001$. The responses to pseudowords were less accurate compared to the responses to words and this effect was stronger in the dyslexic and beginning reader groups than in the normal reader group, $\chi^2 (1) = 12.33, p < .001$. The effect did not differ between the beginning and the dyslexic readers. There was no main effect for N-size. However, we did find a significant N-size by lexicality interaction effect, $\chi^2 (1) = 21.42, p < .001$. Follow-up analyses showed significant N-size effects for both words, $\chi^2 (1) = 5.31, p < .05$ and pseudowords, $\chi^2 (1) = 22.50, p < .001$. These effects were similar for all three reading groups. Closer inspection of Table 6.3 reveals that children performed more accurately on high N-size pseudowords than on low N-size pseudowords. Interestingly, the opposite pattern was found for words. Namely, the children responded less accurately to high N-size words than to low N-size words.

**Latencies.** For each child a mean latency score was computed for each lexicality by N-size condition. Mean latency scores were calculated over correct trials only. The mean latency scores per condition and reading group are presented in the lower part of Table 6.3. Mean differences between the conditions were tested with multilevel analyses (see Statistical Analysis in the Method section). These analyses are based on all individual word and pseudoword latency scores that, in turn, are considered to be nested under individuals.

As for the error percentages, we found a significant difference in mean reading speed between the groups, $\chi^2 (1) = 161.11, p < .001$. Overall, the normal readers were significantly faster than the beginning and dyslexic readers. The reading speed of the dyslexic and beginning children did not differ. In addition, we found
6.3. RESULTS

a main effect for lexicality, $\chi^2 (1) = 114.91, p < .001$. The children responded faster to words than to pseudowords. This effect was stronger for the dyslexic and beginning readers than for the normal readers, $\chi^2 (1) = 55.14, p < .001$. The effect did not differ between the beginning and dyslexic readers. The main effect of N-size was also significant, $\chi^2 (1) = 9.14, p < .01$, but was qualified by a significant interaction with lexicality, $\chi^2 (1) = 14.39, p < .001$. Follow-up analyses showed that the children responded significantly faster to high N-size pseudowords than to low N-size pseudowords, $\chi^2 (1) = 15.14, p < .001$, whereas for words the effect of N-size was not significant. The interaction of reader group and N-size was not significant. All groups were similarly influenced by N-size.

In sum, we found that the normal readers were faster and more accurately than the beginning and dyslexic readers. The reading speed and accuracy scores of the beginning and dyslexic readers did not differ. We also found a significant lexicality effect for all groups which was more pronounced for the dyslexic and beginning readers. In addition, all groups responded faster and more accurate to pseudowords with a high N-size as compared to pseudowords with a small N-size. However, none of the groups displayed differences in speed between the low and high N-size words. Moreover, and in contrast to earlier studies investigating orthographic density effects in adults and children, we found that the children were less accurate on high N-size words than on low N-size words.

6.3.3 Supplementary analysis on the effects of high frequent neighbours on word reading

Although we did match our low and high N-size word and pseudoword sets on the mean frequency of the neighbours (see Table 6.2), the word sets were not matched on the presence of a high frequent neighbour (i.e., a neighbour with a higher word frequency than the target word). Carreiras, Perea, and Grainger (1997) found that the presence of a high-frequent neighbour speeds up naming latencies in high N-size words. A closer look at our experimental words revealed that a large proportion of the high N-size words, namely 21 out of 30, had a high-frequent neighbour. Moreover, the opposite was true for the low N-size words of which only 10 out of 30 words had a high-frequent neighbour. Therefore, it could be that the higher error percentage that was found in high N-size words was merely a result of the higher incidence of high-frequency neighbours. To investigate this hypothesis we reanalyzed the word data while controlling for the presence of a high-frequent neighbour. We did not reanalyze the pseudoword data as the concept of “high-frequent neighbour” does not apply for pseudowords as all the neighbours of pseudowords are by default more frequent than the target pseudoword itself.

As predicted, we found that the effect of N-size on word reading accuracy was no longer significant after controlling for the presence of a high-frequent neigh-
bour, $\chi^2 (1) = 1.63, p > .10$. In addition, the presence (or absence) of a high-frequency neighbour was significant for both the accuracy, $z = 4.31, p < .001$ and latency data, $z = 2.98, p < .001$.

To examine whether the presence of a high-frequent neighbour had a different impact on the reading speed and accuracy of the three reading groups, we specified a separate model to disentangle the effects of neighbourhood size and presence of a high-frequent neighbour. We specified four variables per group, representing all N-size and presence/absence-of-a-high-frequent-neighbour conditions.

Obviously, the design was not perfectly balanced in number of items per condition as it concerned a post-hoc analysis to shed light on an unexpected finding. The high N-size by high-frequent neighbour absent condition was calculated over nine items per child, the high N-size by high-frequent neighbour present condition was calculated over 21 items per child, the low-N-size by high-frequent neighbour present condition was calculated over ten items per child and the low-N-size by high-frequent neighbour absent condition was calculated over 20 items per child.

In addition, the high-frequent neighbour present and absent conditions were not matched for word length (four or five letters) and word frequency. Therefore, we added two level-1 predictors to our two-level model (loge of the word frequency and a dummy variable for length) to control for possible confounding effects. Both predictors were not significant for the latency data model, $z = 0.38, p > .10, z = -0.85, p > .10$. However, in the error data model we found a significant effect for word frequency, $z = 3.64, p < .001$, but not for word length, $z = -0.57, p > .10$.

**Error percentages.** The error percentages are presented in the upper part of Table 6.4. We found a main effect for presence of a high-frequency neighbour, $\chi^2 (1) = 4.94, p < .05$. However, the effect disappeared when we controlled for word frequency, $\chi^2 (1) = 0.41, p > .10$. No other effects were significant.

**Latencies.** The mean naming latencies are presented in the lower part of Table 6.4. There was a main effect for presence of a high-frequency neighbour, $\chi^2 (1) = 28.20, p < .001$. Overall, words with a high-frequency neighbour were named slower than words without a high-frequency neighbour. A follow-up contrast showed that this effect was stronger for the beginning and dyslexic readers than for the normal readers, $\chi^2 (1) = 21.28, p < .001$. This effect did not differ for the beginning and dyslexic readers. Interestingly, a single contrast for the normal reading children showed that their reading speed was actually not influenced by the presence of a high-frequent neighbour, $\chi^2 (1) = 0.49, p > .10$. The main effect for N-size in words and the N-size by presence of high-frequency neighbour interaction effect were not significant.

Although the word frequency and length predictors were not significant for the latency data, we specified the same contrasts for a model that included the two control variables to test the robustness of the high-frequent neighbourhood effect. As expected, the outcomes were the same as for the uncontrolled model.
6.4 Discussion

The aim of the present study was to examine the influence of orthographically similar candidates on the word naming accuracy and speed of children. We were particularly interested whether differences in the number of orthographically similar words had a different effect on normal and dyslexic readers as this might provide information about differences in the build-up of orthographic knowledge between these groups. Sensitivity to orthographic neighbourhood size (N-size) was used as a marker effect.

In line with the results of earlier studies, it was found that the children were faster (Ziegler et al., 2003) and more accurate (e.g., Laxon et al., 1995; 2002) in naming pseudowords with a high N-size than in naming pseudowords with a low N-size. In addition, and in line with the outcomes of previous studies with skilled adult readers (e.g., Andrews, 1989; Peereman & Content, 1995), the children responded equally fast to high-frequency words with a high N-size as to high-frequency words with a low N-size. However, it was also found that the children responded less accurate to high N-size words. Finally, and in contrast to our predictions, it was found that dyslexic and normal reading children did not differ in their sensitivity to N-size.

As mentioned above, the children followed the expected pattern of being more accurate in naming pseudowords from a high N-size than from a low N-size. However, the children were also less accurate in naming words from a large compared to words from a small N-size. This result contrasts with findings of earlier naming studies, which typically found facilitating effects of N-size on word reading accuracy (e.g., Laxon et al., 1994; 2002). It should be noted, however, that compared to studies examining N-size effects in English children (e.g., Laxon et al., 1995; 2002)
al., 1994; 2002; Ziegler et al., 2003), the accuracy scores for words in the present study with Dutch children were high: more than 90% correct for the beginning and dyslexic readers and even more than 98% correct for the normal readers. This observation is in accordance with the findings of earlier research that it is easier to learn to read in languages with a transparent orthography (Seymour, Aro, & Erskine, 2003) and the observation that dyslexic children in such languages are mainly characterized by impaired reading speed and not so much by inaccurate reading (de Jong & van der Leij, 2003; Spinelli et al., 2005; Wimmer, 1993). Because of these low error percentages in the word naming conditions, accompanied with large standard deviations, the word accuracy scores must be interpreted with caution. In addition, after controlling for the presence of a high-frequent neighbour (i.e., a neighbour with a higher word frequency than the target word), the word naming accuracy score between the high and low N-size conditions did no longer differ. However, the absence of a neighbourhood size effect for accuracy for words still differs from what was found by Laxon et al. (1994; 2002), but is perfectly in line with our finding of an absence of a neighbourhood size effect for word reading speed.

In addition, there are two other reasons that might account for the different findings of the current study and the earlier studies of Laxon and colleagues (1994; 2002). Firstly, the selected words in the current study were of higher frequency. As earlier studies with skilled adult readers showed (Peereman & Content, 1995) that the N-size effect diminishes as a function of word frequency, the fact that we used words of higher frequency than Laxon and colleagues might account for the finding that we did not find a difference whereas Laxon and colleagues did find facilitating effects. Secondly, the children in the study of Laxon et al. (2002) were on average almost a year younger and had less reading experience than the beginning readers in the current study. Duñabeitia and Vidal-Abarca (2008) recently showed that sensitivity to orthographic neighbourhood size in words diminished as a function of reading experience. It might well be the case that even our beginning readers already had enough reading experience not to be affected by the orthographic neighbourhood size of words.

The outcomes of the current study regarding the facilitating effect of orthographic N-size in pseudoword reading and the absence of a N-size effect for high-frequent words nicely fit into the predictions of both the PDP (Plaut et al., 1996) and DRC model (Coltheart et al., 2001). In the DRC model, these predictions all concern the behaviour of a skilled adult reading system, as there are no explicit predictions about or simulation of the normal and deviant development of orthographic representations. However, based on the PDP framework, Harm and Seidenberg (1999) implemented the development of a normal and dyslexic reading system. Phonological dyslexia was simulated by impairing the representations of phonological information before training the model to read. Results showed that in the normal reading system neighbour words like meat, treat and eat showed
only small differences in their activation patterns, indicating that they are similarly represented in hidden unit space. However, the phonologically impaired reading system had formed overly divergent representations for neighbour words. As a result, it is conceivable that the dyslexic reader would show less sensitivity to orthographic neighbourhood effects. Interestingly, this was neither found in the current study for general neighbourhood size, nor by Ziegler et al. (2003) for orthographic body neighbourhood size.

Although the DRC model does not make explicit predictions about the normal and deviant development of sensitivity to orthographic N-size, differences in sensitivity to orthographic N-size can also be interpreted as a measure of the functioning of the lexical route. It has been proposed that specific deficits in lexical processing might explain dyslexic children’s tendency to persist in a letter-by-letter recoding strategy (Barca et al., 2006; Coltheart et al., 2001). Several marker effects, like sensitivity to word-frequency effects (Barca et al., 2006), sensitivity to body N-size (Ziegler et al., 2003) and the word superiority effect (Ziegler et al., 2008), have already been used to examine whether dyslexic children show specific deficits in the lexical route. However, as in the current study, thus far no differences between normal reading and dyslexic children on such marker effects of lexical processing have been found. In sum, the finding of the current study that dyslexic readers show similar sensitivity to orthographic N-size as normal readers does not seem to be in accordance with the simulations of the PDP model and is not compatible with the contention of the DRC model that dyslexic readers have specific deficits in the lexical route.

Besides its relevance for the modeling of visual word recognition, the results of the present study might also be of interest for the interpretation of the different N-size effects that have been reported for written and spoken word recognition. In contrast to studies on written word recognition, studies on phonological neighbourhood effects in spoken word recognition have typically found inhibitory instead of facilitating effects of neighbourhood size (Allen & Hulme, 2006; Garlock, Walley, & Metsala, 2001). According to Yates, Locker and Simpson (2004) this difference might be explained in terms of the serial nature of the input of spoken words as opposed to the parallel nature of the input of written words. However, if inhibition is caused by the serial nature of the phonological input of the spoken word stimulus, then we would also expect inhibitory instead of facilitating N-size effects for a written word stimulus for dyslexic and beginning readers as the results of earlier studies (Barca et al., 2006; Marinus & de Jong, in press; Zoccolotti et al., 1999) suggest that beginning and dyslexic readers use a serial sublexical processing strategy to decode words. Therefore, the input of both written and spoken words in these readers will be serial in nature, which according to Yates et al. (2004) should have led to inhibitory N-size effects. In contrast, the results of the current study demonstrated that dyslexic and beginning readers, just like normal readers, showed facilitating N-size effects for written pseudowords and Null.
effects for high-frequent words. We therefore think that the opposing N-size effects for written and phonological N-size cannot be totally attributed to the serial versus parallel input formats.

Although orthographic N-size had a similar effect on normal and dyslexic children, we did find differences on a more specific measure, namely in sensitivity to the presence of a high-frequent neighbour. The present study was the first to address the influence of a high-frequent neighbour in normal reading and dyslexic children. Earlier studies with skilled adult readers (e.g., Carreiras et al., 1997; Sears et al., 1995) found that the presence of a high-frequent neighbour facilitated the naming speed for words with a high-neighbourhood size and inhibited the naming speed of low-neighbourhood size words. The normal reading children in the current study appeared to follow the same pattern; however, the effects were not strong enough to yield statistical significance. For the dyslexic and beginning readers another pattern was found. The presence of a high-frequent neighbour slowed down the word reading speed and this effect turned out to be independent of the orthographic neighbourhood size of the word. After correcting for word-frequency effects, the presence of a high-frequent neighbour did not influence word-reading accuracy. However, as discussed earlier, the results regarding the accuracy scores for words needs to be interpreted with caution as the accuracy levels were very high and the corresponding standard deviations as well.

Neither the DRC nor the PDP model makes explicit predictions about the influence of a high-frequent neighbour in word naming. This is not surprising as computational models usually aim to replicate empirical findings and empirical data on effects of high-frequent neighbours is scarce, even for skilled adult readers (see as exceptions Arduino and Burani, 2004 and Carreiras et al, 1997 for the effect of a high-frequent neighbour in naming and Sears et al., 1995, for the effect of a high-frequent neighbour in lexical decision). Regarding high-frequent neighbour effects in lexical decision tasks Coltheart et al. (2001) even concluded that it is premature to consider implementation of the effect of the presence of a high-frequent neighbour, because the empirical evidence is not conclusive yet.

However, the finding that the word reading of beginning and dyslexic children was slowed down by the presence of a high frequent neighbour can easily be accommodated within the general belief that (word specific) orthographic representations become increasingly specified during reading development (e.g., Castles et al., 2007; Ehri, 1998; Share, 1995; 2008) and also within the idea that dyslexic children are hampered in the development of such well-specified representations. If the orthographic representations of these children are not yet sufficiently specified, then during word reading a more frequent neighbour might initially get more activation than the target word, resulting in a delay of the naming of the target word. In contrast, if the word representations are sufficiently detailed, as in normal readers, then the frequency of neighbours, as their number, is no longer relevant. The finding that the normal readers were more accurate (> 98%) than the
dyslexic readers (> 90%) in word reading is also in line with the assumption that their orthographic representations are less well specified.

At first sight, our finding that both dyslexic and beginning readers were unaffected by neighbourhood size during word reading seemed to suggest the presence of well-specified orthographic representations, just like in the normal skilled reading system. For example, both the DRC and the PDP models predict that thresholds of high frequency words in skilled adult readers will be reached so quickly that the concurrent activation of neighbour words has a negligible effect. However, the absence of a neighbourhood size effect for words in dyslexic and beginning readers might also be explained by relatively weaker contribution to the activation of both the less well-specified orthographic representations of the high-frequent target words and their neighbours. Put differently, the threshold of a high-frequency word might be reached more slowly than in normal reading children, but at the same time the activation level of neighbourhood words builds up more slowly as well and therefore the effect of neighbourhood will also be negligible. Note that in this explanation word frequency is not equivalent to the degree of specification of orthographic neighbours although both aspects are probably related. In sum, the current findings suggest a difference in global and specific effects of neighbour words on the word reading of beginning and dyslexic readers. However, both effects can be accommodated within the view that the orthographic representations of these readers are not (yet) sufficiently specified.

Which factors exactly influence the development of well-specified orthographic representations and how differences between well and less well-specified orthographic representations should be defined remain intriguing questions. According to the self-teaching hypothesis (Share, 1995), an important factor in the build-up of orthographic knowledge is the number of encounters with a word. More recently, Castles et al. (2007, see also Castles et al., 1999) proposed that orthographic representations change, or become more finely tuned, as a function of the growing orthographic lexicon. Increasingly specified orthographic representations are needed to accurately and rapidly identify words among a growing number of similar looking candidates. The results of the current study suggest that the fine-tuning of orthographic representations might not only be of relevance for the identification of words from a dense neighbourhood, but also for the identification of words with a high-frequent neighbour.